Storage capacity and retrieval time of small-world neural networks

Hiraku Oshima and Takashi Odagaki
Department of Physics, Kyushu University, Fukuoka 812-8581, Japan

To understand the influence of structure on the function of neural networks, we study the storage capacity and the retrieval time of Hopfield-type neural networks for four network structures: regular, small world, random networks generated by the Watts-Strogatz (WS) model, and the same network as the neural network of the nematode *Caenorhabditis elegans*. Using computer simulations, we find that (1) as the randomness of network is increased, its storage capacity is enhanced; (2) the retrieval time of WS networks does not depend on the network structure, but the retrieval time of *C. elegans*'s neural network is longer than that of WS networks; (3) the storage capacity of the *C. elegans* network is smaller than that of networks generated by the WS model, though the neural network of *C. elegans* is considered to be a small-world network.

PACS number(s): 84.35.+i

I. INTRODUCTION

The brain has various functions such as memory, learning, awareness, thinking, and so on. These functions are produced by the activity of neurons that are connected to each other in the brain. There are many models to reproduce the memory of the brain, and the Hopfield model is one of the most studied [1]. The Hopfield model was proposed to reproduce associative memory, and it has been studied extensively by physicists because this model is similar to the Ising model of spin glasses. This model was studied circumstantially, for example, the storage capacity was analyzed by the replica method [2,3]. However, in these studies, the neural networks are completely connected, i.e., each neuron is connected to all other neurons. It was not clear how the properties of the model depend on the connections of neurons until recently.

In recent years the study of complex networks has been paid much attention. A network consists of vertices and edges. A vertex is a site or point on the network such as a neuron; the vertices are connected by edges such as an axon or synapse of a neuron. Several characteristic network structures have been proposed, and the small-world and the scale-free networks have been studied heavily. Small-world networks have the properties that the characteristic path length is very short, and simultaneously the clustering coefficient is large [4]. Here, the characteristic path length is the average path length between any two vertices on the network, and the clustering coefficient represents the degree of its connectivity among the neighbors of a given vertex. Scale-free networks have the distribution of degrees following a power law, where the degree is the number of neighbors of each vertex [5]. These characteristic networks have been identified in various real networks such as acquaintance networks, the World Wide Web, power grids, and neural networks, etc. [6,7]. A model reproducing the small-world networks was proposed by Watts and Strogatz (WS) [4]. A model reproducing scale-free networks was proposed by Barabási and Albert [8]. These models are often used to study complex networks.

In the brain it is believed that the neural network also forms a small-world network [9,10]. For example, the neural network of the nematode *Caenorhabditis elegans* is small world. It is an important question why the neural network is

a small-world network. It is possible that neural networks are small world in order to optimize some of their functions. To answer this question, it is necessary to understand first how the network structure affects the functions. Some researchers focus on the associative memory of neural networks to understand the effects of the network structure, and they have investigated either small-world or scale-free networks [9,11–15]. According to these studies, the stability of stored patterns depends on the controlling parameter of the network structure in the WS model [9] and on the clustering coefficient of the network [13]. In these investigations, the effects of network structure on memory have been studied through the stability of the stored patterns on several networks. However, it is not clear how many patterns the system can store in various networks, and how the number of patterns that can be stored depends on the characteristic path length and the clustering coefficient. This upper limit of the number of patterns that can be stored is called the storage capacity, and this quantity is important because the system changes from the storable phase to a spin-glassy phase at this point. In addition, the effects of the network structure on the time required for the system to retrieve patterns have not been studied yet. We define this time as the retrieval time, and it is worth studying the retrieval time because it is not useful for the neural network to store many patterns if retrieving them takes a long time.

In this paper, we investigate how the storage capacity and the retrieval time of the associative memory of a Hopfield-type model depend on the characteristics of the network. The connection of neurons is determined by the structure of the network, in contrast to the Hopfield model where all neurons are connected mutually. We focus on the small-world network generated by the WS algorithm and the neural network of the nematode *C. elegans*. *C. elegans* is a living organism, the neural network of which has been investigated experimentally. We show that (1) as the randomness of network is increased, its storage capacity is enhanced; (2) the retrieval time of WS networks does not depend on the network structure, but the retrieval time of *C. elegans*'s neural network is longer than that of WS networks; (3) the storage capacity of the *C. elegans* network is smaller than that of networks generated by the WS model with the same number of neurons and average degree.

In Sec. II, we explain the model in detail, and we show the results of computer simulation in Sec. III. In Sec. IV, the results are analyzed. We discuss these results in Sec. V.

II. MODEL

A. Hopfield model on a network

We use McCulloch-Pitts neurons [16], each element of which can have only two states, $S_i = 1$ (firing) and -1 (not firing). Each neuron is connected to some other neurons via synaptic connections, which are characterized by their synaptic weights J_{ij} determined by stored patterns. To represent connectivity among neurons, we use the adjacency matrix $\{a_{ij}\}$, where $a_{ij} = 1$ when there is a connection between i and j, $a_{ij} = 0$ otherwise. The adjacency matrix $\{a_{ij}\}$ uniquely determines the structure of the network.

The stored patterns in the neural network are represented by network states $\boldsymbol{\xi}^{\mu}$, where $\mu = 1, \ldots, p$, p being the number of stored patterns. These patterns are randomly generated with probabilities $P(\xi_i^{\mu} = \pm 1) = 1/2$. Using these patterns, the synaptic weights J_{ij} are given by the Hebbian rule

$$J_{ij} = \frac{1}{\langle k \rangle} a_{ij} \sum_{\mu=1}^{p} \xi_i^{\mu} \xi_j^{\mu} \tag{1}$$

for $i, j = 1, \ldots, N$. Here, N is the number of neurons in the network, and $\langle k \rangle$ is the average degree of the network. The degree is defined by the number of neighbors of a vertex. We set $J_{ii} = 0$, so that no connection within a neuron is allowed.

The state of neurons is updated asynchronously by the Monte Carlo method. A neuron is chosen at random, and the state of each neuron is updated according to

$$S_i = \text{sgn}\left(\sum_{j=1}^{N} J_{ij} S_j\right). \tag{2}$$

The Monte Carlo time is advanced by 1 after N neurons have been accessed by the updating process.

B. Storage capacity

To quantify pattern retrieval, we use the overlap order parameters [2]

$$m^{\mu} = \frac{1}{N} \sum_{i=1}^{N} S_i \xi_i^{\mu}, \tag{3}$$

where μ is the pattern number. The overlap is a parameter which represents the degree of similarity between the state of the system S and the stored pattern $\boldsymbol{\xi}^{\mu}$. For example, $m^1 = 1$ if the state is the same as the pattern $\boldsymbol{\xi}^1$, while $m^1 = 0$ if the state is far from the pattern $\boldsymbol{\xi}^1$.

We measure the stability of the pattern $\boldsymbol{\xi}^1$ to check whether it is stored or not. First, we construct the structure of the network $\{a_{ij}\}$. Second, we store p random patterns in the network by setting J_{ij} according to Eq. (1). We change p from 1 to 50 to judge the storage capacity. Third, we initialize the system to the pattern $\boldsymbol{\xi}^1$ at $t = 0$, and evolve the state in time until the dynamics converges. Finally, to check whether

the pattern $\boldsymbol{\xi}^1$ is stored or not, we measure the overlap m^1 between the final state and $\boldsymbol{\xi}^1$. The pattern $\boldsymbol{\xi}^1$ is considered as stored if the overlap m^1 is no less than 0.9. When $m^1 \geq 0.9$, the number of error bits defining the difference between the final state of network and the pattern $\boldsymbol{\xi}^1$ is no more than 5% of all bits. It is practically reasonable that a pattern within 5% error is considered as being stored.

Repeating the above procedure several times by using different samples of networks and stored patterns, we count up the number of times that the pattern is successfully stored, and calculate the fraction of the number of trials that is successfully stored. These calculations are the same as those used for all-to-all connection networks in Ref. [17]. We define this fraction as the retention rate. The retention rate is unity when p is small, while it is 0 when p is large. There exists a transition point where the retention rate changes from 1 to 0 sharply when p is increased. Therefore, we define the storage capacity $\alpha_c \equiv p_c / N$ as the number of patterns when the retention rate becomes 0.5.

C. Retrieval time

The retrieval time is defined as the time required for the system to retrieve a stored pattern. The time t is measured in Monte Carlo time steps. First, we initialize the system to the state of the pattern $\boldsymbol{\xi}^1$ with noise. To add noise to $\boldsymbol{\xi}^1$, we select some neurons of $\boldsymbol{\xi}^1$ randomly and flip them, and $m^1(t=0)$ is set as 0.9, 0.8, 0.7, 0.6 in this way. We do not use smaller $m^1(0)$ than these values, because the system tends to go to a different stable point from the pattern $\boldsymbol{\xi}^1$, if $m^1(0)$ is small. Next, we measure $m^1(t)$ at each time step until the dynamics converges. This convergence time is denoted as t_f. After repeating this procedure several times for other samples, we scale the overlap $m^1(t)$ and calculate its average according to

$$f(t) = \left\langle \frac{m^1(t) - m^1(t_f)}{m^1(0) - m^1(t_f)} \right\rangle, \tag{4}$$

where $f(0) = 1$ and $f(t > t_f) = 0$. We define the retrieval time as the reciprocal of the slope of $-\ln f(t)$, i.e.,

$$\frac{1}{\tau} = -\frac{d \ln f(t)}{dt} \tag{5}$$

because $f(t)$ decays exponentially.

III. RESULTS

First, we compared regular, small-world, random networks which are generated by the Watts-Strogatz model [4] in order to understand the effects of network structure in this model. Next, we compared the neural network of the nematode *C. elegans* [18] and WS networks in order to clarify differences between the real and model networks.

A. Results for the WS model

We carried out a Monte Carlo simulation of the Hopfield model with several networks generated by the WS model with $N = 1000$ and average degree $\langle k \rangle = 100$. First, we con-
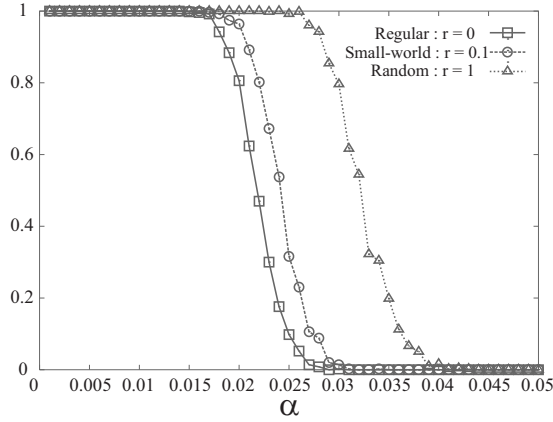
FIG. 1. Retention rate as a function of $\alpha = p/N$. The squares, circles, and triangles represent regular ($r=0$), small-world ($r=0.1$), and random networks ($r=1$), respectively. All networks have $N = 1000$ and $\langle k \rangle = 100$.

structed a network with a probability r, where r is the rewiring probability in the WS model. The edges on the network are selected by r, and they are rewired randomly. The network is regular, random, and small world when $r=0$, $r=1$, and $0 < r \ll 1$, respectively. Second, we stored p random patterns, from 1 up to 50, by setting J_{ij} according to Eq. (1). Third, we evolved the system in time and measured the overlap m^1 after the dynamics converged. We repeated this procedure 500 times using different samples of networks and patterns with the same r and p.

We used $r=0$, 0.1, and 1 each of which corresponds to a regular, small-world, and random network, respectively. Figure 1 shows the retention rate as a function of $\alpha = p/N$. The storage capacity α_c is defined by α when the retention rate becomes 0.5. From Fig. 1, we find

$$\alpha_c = \begin{cases} 0.022 & \text{for regular, } r=0, \\ 0.024 & \text{for small world, } r=0.1, \\ 0.032 & \text{for random, } r=1. \end{cases}$$

We got these values by fitting the lines with $\{1 + \tanh[a(\alpha_c - \alpha)]\}/2$, where a and α_c are fitting parameters. We find that the storage capacity is an increasing function of r. α_c depends on N and $\langle k \rangle$, but its qualitative behavior that the storage capacity is an increasing function of r is invariant.

Next, we measured the retrieval time of the model, which is shown in Fig. 2 for four different initial states, i.e., $m^1(t=0) = 0.9, 0.8, 0.7, 0.6$, when three patterns were stored. These data were calculated using 10 000 different samples of networks and patterns. We find that the retrieval time of all networks is approximately one Monte Carlo step because the slope of $\ln f(t)$ is unity. Therefore, it does not strongly depend on network structure. We confirmed that this result is independent of N and $\langle k \rangle$.

B. Results for *C. elegans*

C. elegans is a living organism used frequently in biological experiments. Although its neural network consists of only

302 neurons, it has abilities of learning and memory. The connections in this network have been investigated experimentally, and its data are available from the database on the web [18].

We reconstructed *C. elegans*'s network by using the connection data from the database [18]. A connection between two neurons by either a synapse or a gap junction is assumed to be an undirected edge. If there are two or more connections between two neurons, we regarded them as one connection. *C. elegans*'s network is a small-world network with $N = 251$ and $\langle k \rangle \cong 14$, because it has about the same L as but larger C than the random network (Table I).

We carried out the same procedure as in the previous section for the neural network of *C. elegans*. We compared its storage capacity with that of three WS networks, $r = 0, 0.3, 1$, which have the same number of neurons ($N = 251$) and approximately the same average degree ($\langle k \rangle = 14$). The WS network with $r=0.3$ has roughly the same characteristic path length L and clustering coefficient C as *C. elegans*'s network. Figure 3 shows the retention rate as a function of α. The storage capacity of *C. elegans*'s network is the smallest among the networks examined,

$$\alpha_c = \begin{cases} 0.022 & \text{for regular, } r=0, \\ 0.024 & \text{for small world, } r=0.3, \\ 0.026 & \text{for random, } r=1, \\ 0.018 & \text{for } C. \text{ elegans}. \end{cases}$$

The retrieval time is shown in Fig. 4 for different initial states $m^1(t=0) = 0.9, 0.8, 0.7, 0.6$, when one pattern was stored. The retrieval time of WS networks is one Monte Carlo step, as in the previous section. The neural network of *C. elegans* retrieved a stored pattern more slowly than other networks; its retrieval time is approximately 1.1 Monte Carlo steps. We conclude that the retrieval time of WS networks does not depend on the structure of the network, but the retrieval time of *C. elegans* is longer than that of WS networks.

IV. ANALYSIS

We analyzed the relations between the storage capacity and network properties in order to understand what characteristics of the structure of the network determine the storage capacity. We calculated the storage capacities of the neural networks generated by the WS model with several probabilities, and plotted them against the rewiring probability r (Fig. 5), the characteristic path length L (Fig. 6), and the clustering coefficient C of each network (Fig. 7). Each set of data is fitted by a curve as a guide for the eyes. According to Fig. 6, the characteristic path length may have a large effect on the storage capacity because the storage capacity changes significantly in the area of $L \approx 2$. The data are fitted well by $\alpha_c(L) = a_L(L - b_L)^{-c_L} + d_L$. Here the fitting parameters are $a_L = 0.0005$, $b_L = 2 c_L = 0.8$, $d_L = 0.02$. On the other hand, according to Fig. 7, the storage capacity depends on the clustering coefficient linearly. The data are fitted well by $\alpha_c(C) = a_C C + b_C$, and the fitting parameters are $a_C = -0.02$, $b_C = 0.03$.
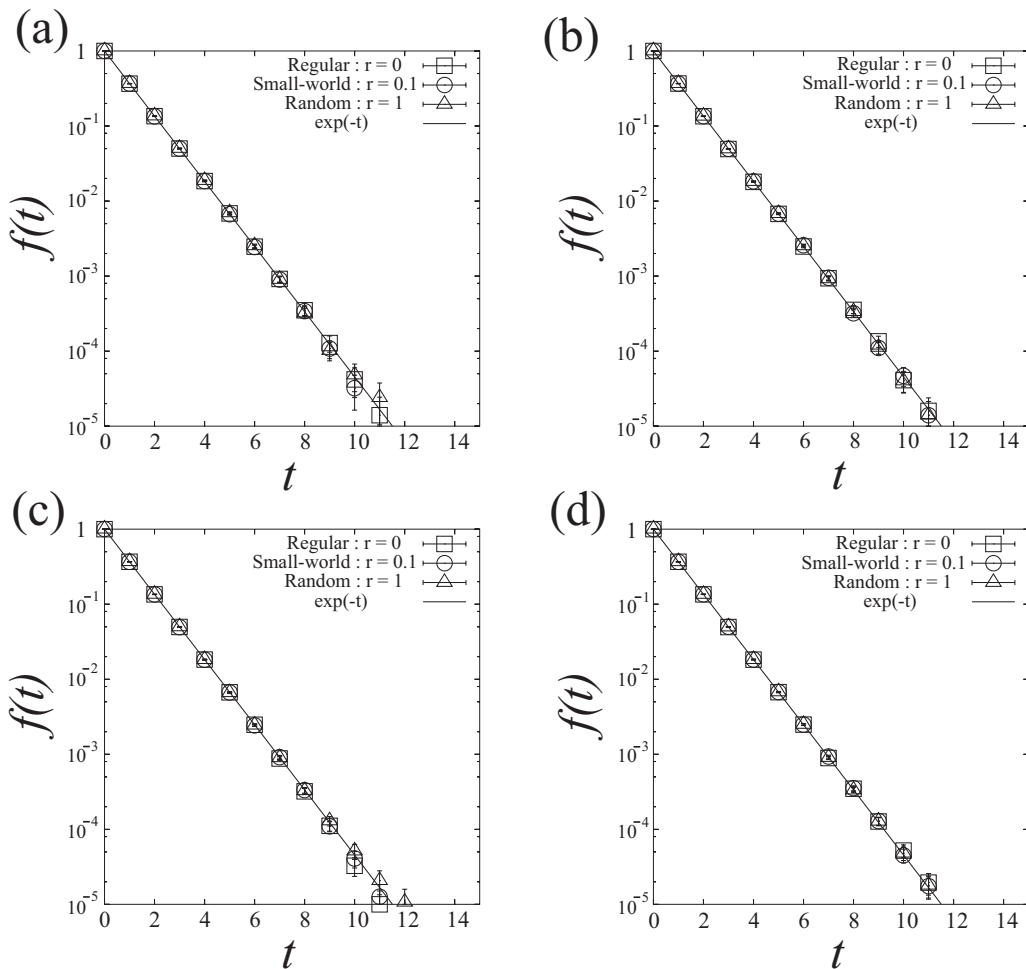
FIG. 2. Process of retrieval in different networks, regular (squares), small-world (circles), and random networks (triangles), when $p=3$. All networks have $N=1000$ and $\langle k \rangle = 100$. The initial state of the system is $m^1(0) = $ (a) 0.9, (b) 0.8, (c) 0.7, and (d) 0.6.

Therefore the storage capacity may depend on both L and C. In Fig. 5, α_c is saturated at large r because C is saturated at large r. The data are fitted well by $\alpha_c(r) = a_r \tanh(b_r r) + 0.02$, and the fitting parameters are $a_r = -0.01$, $b_r = 2$.

Next, in order to investigate the dependence of the degree distribution of neural networks, we compared the storage capacities of the random network generated by the WS model and the random network in which all neurons have the same degree, called a regular-random network. The regular-random network is a random network whose edges are ran-

TABLE I. The characteristic path length L and clustering coefficient C of the *C. elegans*, the random ($r=1$), and the small-world network ($r=0.3$) generated by the WS model. *C. elegans*'s network has about the same L as the random network, while it has larger C. The small-world network with $r=0.3$ has roughly the same L and C as *C. elegans*'s network.

Network	L	C	$\langle k \rangle$
C. elegans	2.65	0.245	$\cong 14$
Random ($r=1$)	2.38	0.0525	14
Small world ($r=0.3$)	2.54	0.261	14

domly connected, but each neuron has the same degree. We generated this network by connecting two neurons selected randomly under the condition that each neuron has the same
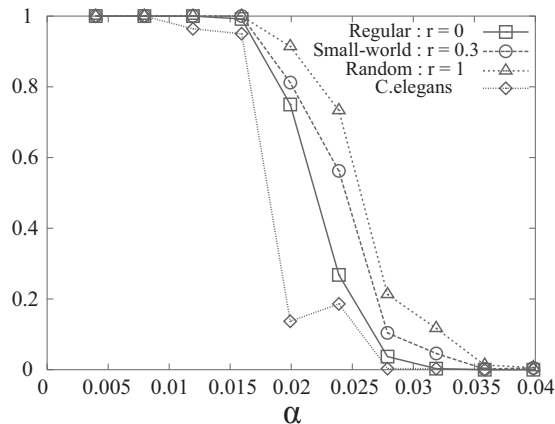

FIG. 3. Retention rate in *C. elegans* neural network (diamonds). Data for three WS networks are also shown for comparison. These networks are regular (squares), small world (circles), and random (triangles), which have the rewiring probability $r=0$, 0.3, and 1, respectively. All networks have $N=251$ and $\langle k \rangle \cong 14$.
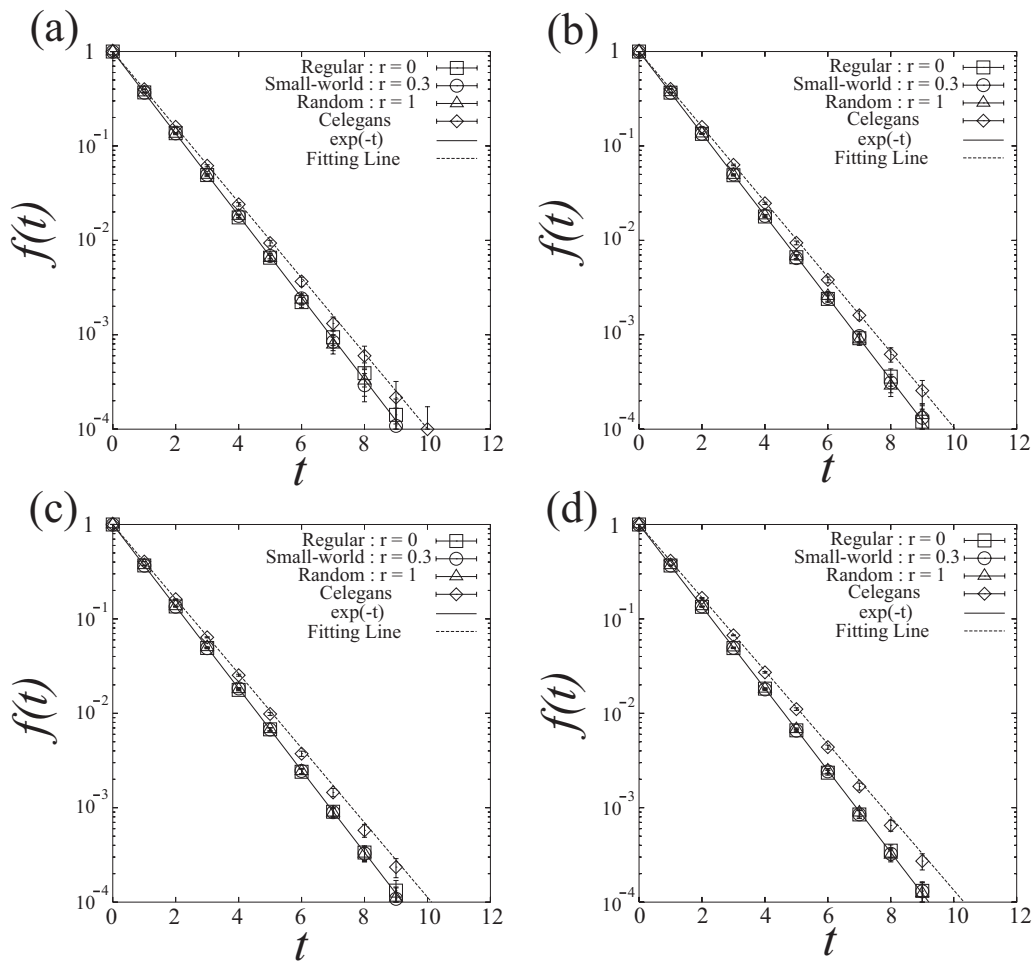
FIG. 4. Process of retrieval in different networks, i.e., regular (squares), small world (circles), random (triangles), and *C. elegans* (diamonds), when $p=1$. All networks have $N=251$ and $\langle k \rangle \cong 14$. The data of *C. elegans* are fitted with a line of slope approximately $1/1.1$. The initial state of the system is $m^1(0)=(a)$ 0.9, (b) 0.8, (c) 0.7, and (d) 0.6.

number of neighbors. On the other hand, the degree distribution of the random network generated by the WS model with $r=1$ obeys the Poisson distribution approximately. By comparing the results for these networks, we can understand the effects of a neuron that has lower or higher degree than average. On the other hand, the degree distribution of the random network generated by the WS model with $r=1$ obeys the Poisson distribution approximately. By comparing the results for these networks, we can understand the effects of a neuron that has lower or higher degree than av-

erage. The relation between the storage capacity and the network properties of this network are plotted in Figs. 6 and 7. There is little difference between the storage capacities of the
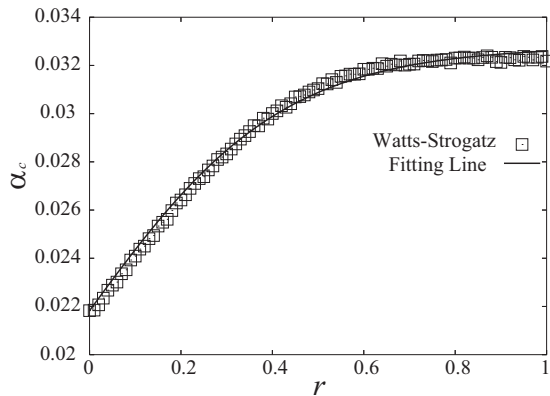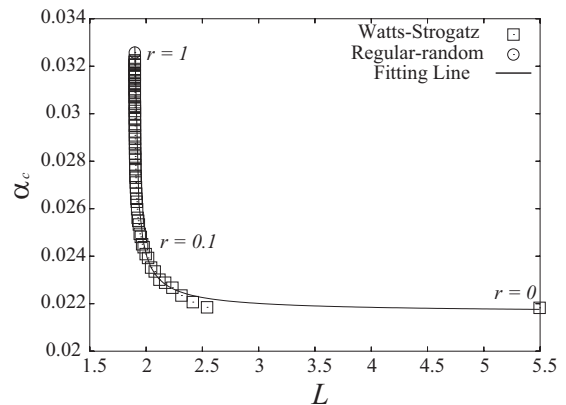

FIG. 6. Storage capacity as a function of the characteristic path length of different networks: that generated by the Watts-Strogatz model (squares) and the regular-random network (circles). The parameters are $N=1000$ and $\langle k \rangle=100$. The curve is a guide for the eyes.

FIG. 5. Storage capacity as a function of the rewiring probability r. The parameters are $N=1000$ and $\langle k \rangle=100$. The curve is a guide for the eyes.
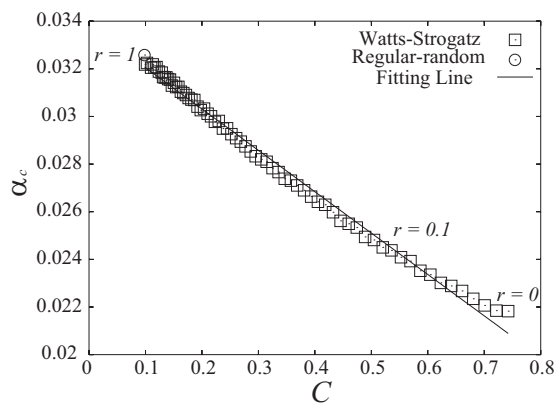
FIG. 7. Storage capacity as a function of the clustering coefficient of different networks: that generated by the Watts-Strogatz model (squares) and a regular-random network (circles). The parameters are $N=1000$ and $\langle k \rangle=100$. The curve is a guide for the eyes.

regular-random network and random network. We conclude that the storage capacity does not depend much on the degree distribution.

V. DISCUSSION

We investigated the storage capacity and the retrieval time of the Hopfield model on several different networks. We showed that the storage capacity depends on the randomness of networks. We found that the retrieval time of WS networks does not depend on the network structure, but the retrieval time of *C. elegans*'s neural network is longer than that of WS networks. The storage capacity may depend on both the characteristic path length and the clustering coefficient. However, it is unclear which effect is larger because L and C cannot be changed independently in the WS model. The effect of C was studied in Ref. [13], but the effect of L still has not been studied independently. It is a topic for future research to clarify the effects of network properties independently. Furthermore, the storage capacity of *C. elegans*'s network is smaller than that of the network generated by the WS model. We consider that *C. elegans*'s network has a small storage capacity because it has many neurons whose degree is 1. The error often occurs on a neuron with $k=1$, because its state becomes different from a stored pattern whenever its neighborhood goes to the different state from stored.

The neural network in the brain is thought to be small world. However, the results in this paper show that more randomized networks have more storage capacity, and the real neural network of *C. elegans* has a longer retrieval time than WS networks. Our work implies that the neural network of the brain is not optimized only for maximizing the storage capacity and minimizing the retrieval time. The reason why the brain's network is not random but small world may be that making long-distance connections as in a random network costs more energy or proteins than making short-distance connections [9]. The question why the neural network is small world is still unresolved. The other possibility is that the Hopfield model does not explain the memory of the brain well. Then we need a better model than the Hopfield model in order to reproduce the memory of the brain.

[1] J. J. Hopfield, Proc. Natl. Acad. Sci. U.S.A. **79**, 2554 (1982).

[2] D. J. Amit, *Modeling Brain Function: The World of Attractor Neural Networks* (Cambridge University Press, New York, 1989).

[3] J. Hertz, A. Krogh, and R. G. Palmer, *Introduction to the Theory of Neural Computation* (Addison-Wesley, Redwood City, CA, 1991).

[4] D. J. Watts and S. H. Strogatz, Nature (London) **393**, 440 (1998).

[5] A. L. Barabasi and R. Albert, Science **286**, 509 (1999).

[6] M. E. J. Newman, SIAM Rev. **45**, 167 (2003).

[7] R. Albert and A. L. Barabasi, Rev. Mod. Phys. **74**, 47 (2002).

[8] A. L. Barabasi, R. Albert, and H. Jeong, Physica A **272**, 173 (1999).

[9] P. N. McGraw and M. Menzinger, Phys. Rev. E **68**, 047102 (2003).

[10] S. H. Strogatz, *Sync: The Emerging Science of Spontaneous Order* (Hyperion, New York, 2003).

[11] J. W. Bohland and A. A. Minai, Neurocomputing **38**, 489 (2001).

[12] J. Q. Lu, J. He, J. D. Cao, and Z. Q. Gao, Phys. Lett. A **354**, 335 (2006).

[13] B. J. Kim, Phys. Rev. E **69**, 045101(R) (2004).

[14] D. Stauffer, A. Aharony, L. D. Costa, and J. Adler, Eur. Phys. J. B **32**, 395 (2003).

[15] L. G. Morelli, G. Abramson, and M. N. Kuperman, Eur. Phys. J. B **38**, 495 (2004).

[16] W. S. McCulloch and W. Pitts, Bull. Math. Biol. **52**, 99 (1990).

[17] B. M. Forrest and D. J. Wallace, in *Models of Neural Networks*, edited by E. Domany, J. L. van Hemmen, and K. Schulten (Springer-Verlag, Berlin, 1991), pp. 121–148.

[18] K. Oshio, Y. Iwasaki, S. Morita, Y. Osana, S. Gomi, E. Akiyama, K. Omata, K. Oka, and K. Kawamura, Keio University Technical Report, 2003 (unpublished), URL http://www.bio.keio.ac.jp/ccep/