Doctoral Thesis

A Study of Self-organization Phenomena of Network Structure and Elements under the Constraint of Maximization of Information

（情報量最大化を拘束条件にしたネットワーク構造及びネットワーク素子の自己組織化現象に関する研究）

Hiroshi WATANABE
Department of Mathematics
Hokkaido University

March, 2021.

Contents

1. Introduction

Motivated by biological studies on the mechanisms of neuronal differentiation, we studied a mathematical model of neuronal differentiation by mimicking evolutionary dynamics. In evolutionary dynamics, the fundamental dynamical systems evolve through changes in parameter values of the systems. We consider a network of individual dynamical systems that interact with each other. Hence, an overall system, which is constructed here as coupled dynamical systems, is self-organized in evolvable conditions.

Studies on self-organization have been conducted since the cybernetics movement started, and accelerated in the 1970s and 1980s. Prigogine et al. developed a theory of nonlinear and nonequilibrium thermodynamics and proposed the concept of dissipative structure as a spatiotemporal ordered pattern in macroscopic states (Nicolis and Prigogine, 1977). In far-from-equilibrium systems, Haken founded Synergetics, inspired by the nonequilibrium phase transitions of laser theory. He introduced the slaving principle in which only a few degrees of freedom enslave the other degrees of freedom, and slaving modes represent macroscopic ordered motion (Haken, 1977, 1983). In these studies, self-organization occurred via interactions between microscopic elements, i.e., atoms and/or molecules, which resulted in the appearance of macroscopic ordered motion represented by order parameters. Order parameters can be described by dynamical systems with a few variables.

On the other hand, there are significant organizational processes that cannot be explained within the framework of fixed dynamical systems or even fixed coupled dynamical systems. Among others, functional differentiation of the brain and cell differentiation of embryos are particularly valuable to study for the following reason. Neuro stem cells have been found in the third ventricle, whereby neurogenesis is known to occur in, at least, dentate gyrus and side subventricular zone. Two kinds of medical treatment models for brain injury are now highlighted: one can be performed via the acceleration of differentiation by stimulation in the growth factor of progenitor cells, and the other via nerve grafts with neuro stem cells stemming from fetus brains, stem cells, and even induced pluripotent stem (iPS) cells (H. Okano and S. Yamanaka 2014).

However, few studies on the underlying mechanisms of the neuronal differentiation have been conducted. Motivated by these biological developmental processes, in this study, we focus on the mathematical structure, which is supposed to be embedded in these differentiation processes. Because mathematical modeling of neuronal differentiation should focus on the developing process of the system, conventional self-organization theories are not suitable, and an adequate model should include the changing processes of dynamical system. In this respect, it is plausible to think that an appropriate approach is not based on conventional self-organization theories but on a constrained self-organization theory derived from a variational principle that is a type of optimization formalism.

Considering the biological survival values, the information transmission in multicellular organisms is one of the important functions for survival, and thus it must have been optimized during the evolutionary process. It is known that some chaotic dynamical systems have information transmission capability (Matsumoto and Tsuda, 1985, 1987, 1988; Tsuda and Shimizu, 1985). Then, a question arises: what kind of dynamical systems transmit information more efficiently?

In the present study, each individual dynamical system in a network of dynamical systems develops according to a genetic algorithm under a constraint concerning the transmission of information. In this study, the fitness function for the information transmission was given by the time-dependent mutual information, which was first proposed by Matsumoto and Tsuda (Matsumoto and Tsuda, 1987, 1988).

Concerning the network structure, we adopted a unidirectionally coupled one-dimensional maps such as feed-forward connection networks, random networks, small-world networks, and fully-connected networks. Here, a small-world network is a type of mathematical graph. This network is characterized by a situation in which even when the network has no tight connections, any two nodes are connected via only a few intermediate nodes. Thus, the network represents a "small world."

In each simulation, we fixed the type of network structure. In the small-world networks, the coupling strengths were fixed during the development of each element. In both random and fully-connected networks, the connection strength was also changed according to genetic algorithms, in addition to the change of the individual dynamical systems.

In § 2, we present the details of the mathematical structure of the model. In § 2.1, we describe the one-dimensional map that we used as an individual dynamical system. In § 2.2, we introduce the fitness function. In § 2.3, we describe the genetic algorithm that we used for finding the optimal dynamical system. In § 3, we show the results in the case of unidirectionally coupled one-dimensional maps, which create feed-forward connection networks. In § 4, we treat other typical networks: random networks in § 4.1, small-world networks in § 4.2, and fully-connected networks in § 4.3. In § 4.4, we show the network response to completely randomized inputs. Section 5 devoted to summary and discussion.

2. Neuronal model developed by genetic algorithm

2.1. Function form used for dynamical system

In this study, we treated networks of dynamical systems in terms of coupled maps. We used a one-dimensional continuous map to represent an individual dynamical system, which is defined on R. The map of the k-th element of the network, $f_k(x)$, was represented by the sum of two continuous functions represented by $\tanh(x)$ with six parameters and one shift parameter, as expressed by the following Eq. (1).

$$f_k(x) = a_1\, tanh\left(a_2(x - a_3)\right) - a_4 tanh\left(a_5(x - a_6)\right) + b_k \qquad (1)$$

In neural systems, the values of these functions may represent membrane potentials of neurons or amplitudes of other variables measuring neural fields such as the Local Field Potential (LFP), Electroencephalogram (EEG), Electrocorticogram (ECoG), Magnetoencephalogram (MEG), and bold signals by functional Magnetic Resonance Imaging (fMRI).

Parameters $(a_1, \cdots, a_6)$ were common to all elements of the network. This map was constructed by a combination of two sigmoid functions, where a sigmoid function indicated a bounded and strictly monotone function. Such a sigmoid function has been used as a model for biological neuron in artificial neural network studies. The combination of these functions can represent many functions. The current function's form can represent typical dynamical systems because the form given by Eq. (1) can represent monotonically increasing and decreasing functions, unimodal and bimodal functions, and constant functions, by changing the parameters. For a network with N maps, there were N + 6 parameters $A = (a_1, \cdots, a_6, b_1, \cdots, b_N)$ in the entire system. See Eqs. (4) and (5) for a dynamical system used in this paper.

2.2. Fitness function

To calculate the quantity of information transmitted over the entire system, we used the following fitness function. The amount of information that is shared between the external input and the k-th

map of the network at n time steps can be calculated using the time-dependent mutual information described by Eq. (2).

$$I_n(k) = \sum_{j^{(k)}=1}^{M} p\left(j^{(k)}\right) \log p\left(j^{(k)}\right)^{-1} - \sum_{i=1}^{M} \sum_{j^{(k)}=1}^{M} p(i) p_n\left(j^{(k)} \middle| i\right) \log p_n\left(j^{(k)} \middle| i\right)^{-1} \qquad (2)$$

We calculated the range of the k-th individual map from the recorded time-series data and divided it into M equal parts, where $p(i)$ is the probability that the state of the input is i, $p\left(j^{(k)}\right)$ is the probability that the state of the k-th individual map is j, and $p_n\left(j^{(k)} \middle| i\right)$ is the conditional probability that the state of the k-th individual map is j at n time steps after the state of the input was i.

Here, an overall dynamical system is determined by the parameters $A = (a_1, \cdots, a_6, b_1, \cdots, b_N)$ and a matrix consisting of coupling strength $W = (w_{kl})$. A set of parameters A can be viewed as a gene, which is an object for change.

We defined the capability of information transmission of the present system by the following Eq. (3).

$$J(A, W) = \sum_{k=1}^{N} \max_{n} \{I_n(A, W, k)\} \qquad (3)$$

$I_n(A, W, k)$ indicates the time-dependent mutual information between the external input and the state of each individual dynamical system denoted by k with time steps n. For each individual map, we took the maximum value of this quantity over all time steps. Next, we summed these maximum values over all individual maps, which was regarded as the fitness function.

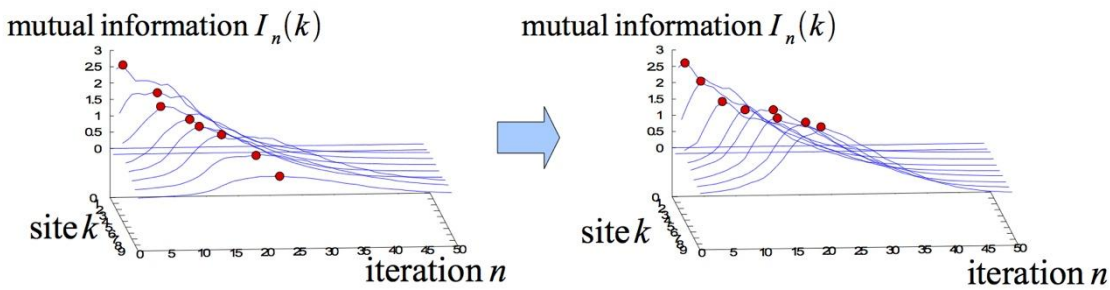

Fig. 1. Fitness function (schematic drawing). We calculated the sum of the maximum value (red circle) of each time-dependent mutual information to evaluate the entire system. The evaluation value is higher at right.

$J(A,W)$ takes 0 as a minimum value when the external input and each individual map are independent, and $J(A,W)$ takes NH_{EX} as a maximum value when information is completely shared between the external input and all individual maps. Here, N is the number of elements, and H_{EX} is the Shannon entropy of the external input. The red circle in Fig. 1 indicates the discrete time step at which each individual map has the maximum value of mutual information. The sum of such maximum values in all individual maps can be an evaluation value of the overall system. In Fig. 1, the evaluation value is higher at right, which indicates the result of evolutionary dynamics.

This fitness function J(A,W) has several important properties. It takes time to transmit information. By taking the maximum value over the time period 1~ n, the system that transmit more information is evaluated regardless of time. The fitness function J(A,W) is based on the amount of information shared between the external input and each individual map. Why did we focus on how efficiently external information is transmitted to the overall system, instead of local information transmission? Had the coupled map network been evaluated solely by local information transmission, it would have evolved to share internal information. In such a case, the system would evolve to be independent of the external information. This provides the different stage from the present purpose.

2.3. Genetic algorithm

Based on the above fitness function, a more efficient dynamical system regarding the information transmission can be found. In the present study, we used the following genetic algorithm as a method of finding better parameters of individual dynamical systems. The genetic algorithm is one of the metaheuristic algorithms that mimics biological evolution (Goldberg, 1989; Holland, 1992). Many genetic algorithms have been proposed with various improvements. Here, we adopted the genetic algorithm called CHC proposed by L. Eshelman (Eshelman, 1991). This algorithm has the following three features: Cross-generational elitist selection, Heterogeneous recombination, and Cataclysmic mutation. Fig. 2 shows the flowchart of the CHC algorithm.

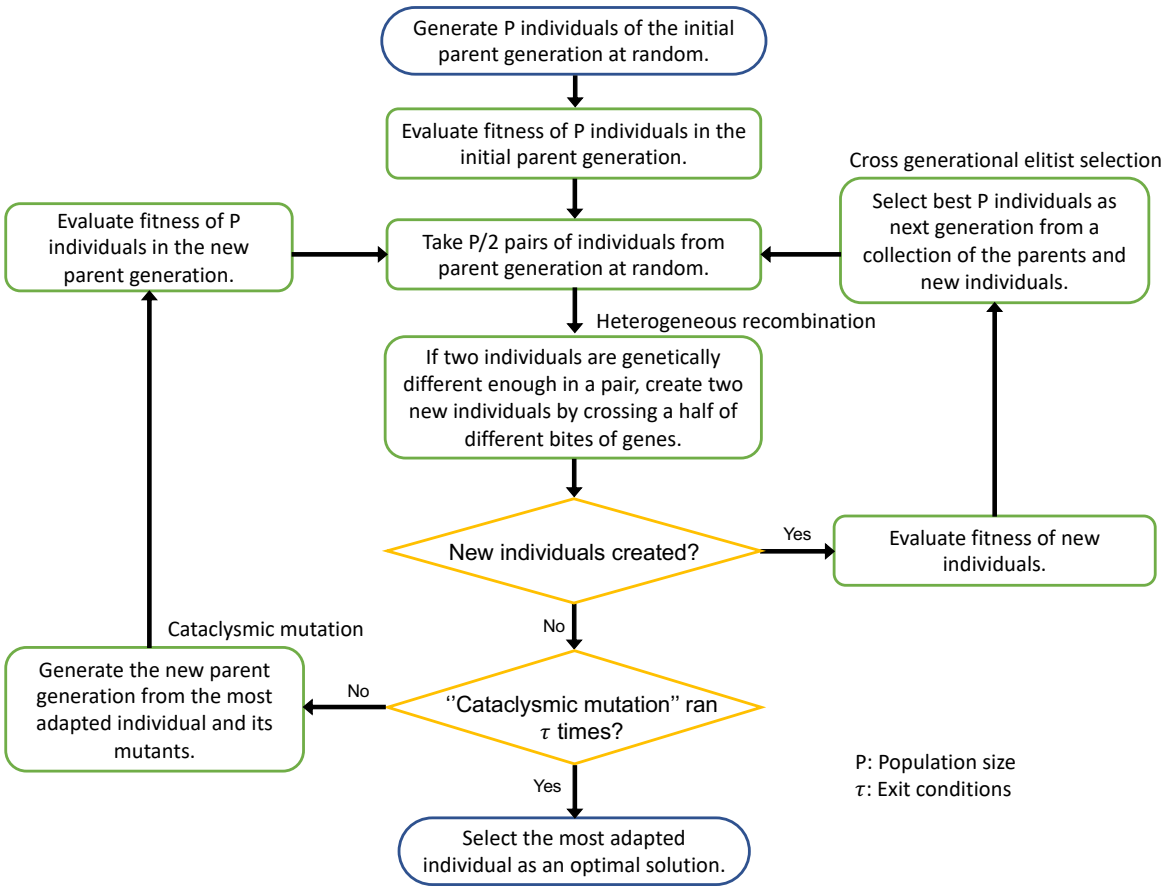

Fig. 2. Flowchart of CHC algorithm from top blue-colored oval, ending with bottom blue-colored oval. Characteristic CHC algorithm appears at upper right (cross-generational elitist selection), center (heterogeneous recombination), and lower left (cataclysmic mutation), indicated by green-colored rounded squares.

3. Unidirectionally coupled one-dimensional maps of feed-forward connections

3.1. Coupled map networks

In the present study, we consider coupled map networks with a discrete time and a continuous state.

In all simulations, the state of the *k*-th individual map at time $n+1$ is expressed by the following Eqs. (4) and (5).

$$x_0(n+1) = f_0\big(x_0(n)\big) + \sum_{l=1}^{N} w_{0l} x_l(n) + DG\big(y(n)\big) + \sigma \tag{4}$$

$$x_k(n+1) = f_k\big(x_k(n)\big) + \sum_{l=1}^{N} w_{kl} x_l(n) + \sigma \tag{5}$$

where $G\big(y(n)\big)$ is the external input at time *n*, w_{kl} is the coupling strength from the *l*-th to the *k*-th individual map, D is the coupling strength from the external input to the 0-th individual map, and σ is an additive noise that was provided by a Gaussian distribution with an average of 0 and variance of 0.0001. The network structure is represented by w_{kl}.

In the present study, we used the following chaotic map as an external input $G\big(y(n)\big)$:

$$y(n+1) = G\big(y(n)\big)$$
$$G(y) = \begin{cases} \left(-(0.125-y)^{\frac{1}{3}} + 0.50607357\right)\cdot e^{-y} + b & (y < 0.125) \\ \left((y-0.125)^{\frac{1}{3}} + 0.50607357\right)\cdot e^{-y} + b & (0.125 \le y < 0.3) \\ 0.121205692 \cdot \left(10ye^{-\frac{10x}{3}}\right)^{19} + b & (0.3 \le y) \end{cases} \tag{6}$$

where b is a shift parameter. Here, $b = 0.023285279$. This chaotic map $G\big(y(n)\big)$ was originally introduced to reproduce the state transitions according to bifurcations in the Belousov-Zhabotinsky (BZ) reaction system (Tomita and Tsuda, 1980).

In the case of simple inputs, for example periodic inputs, dynamical systems with the same period as input period are selected, regardless of the functional form. This is trivial. Therefore, we investigate the case of non-periodic sequences as an external input. We conducted numerical experiments in another various non-periodic inputs such as random sequences consisting of quasi-random number, logistic chaos, and BZ chaos, which represent typical dynamical systems. The present result of evolution, that is, a type of eventually selected map does not change by the choice of the source of non-periodic sequences for feed-forward connections.

3.2. Unidirectionally coupled one-dimensional maps of feed-forward connections

First, we considered unidirectionally coupled one-dimensional maps of feed-forward connections, as shown in Fig. 3. The external input was given only to the first individual map. The network structure was fixed. The individual maps and coupling strengths were optimally changed to maximize the information transmission of the external signal.

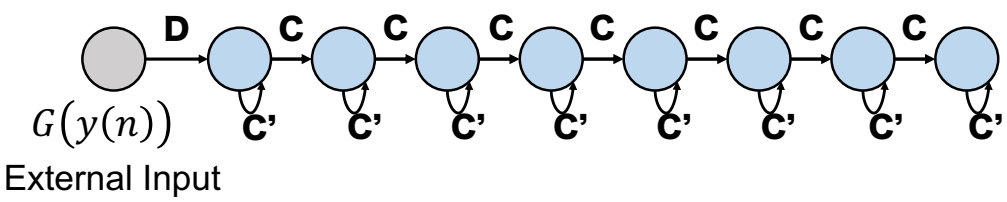

Fig. 3. Unidirectionally coupled one-dimensional maps of feed-forward connections. D is the coupling strength from the external input to the 0-th individual map, and C and C' are the feed-forward connection strength and self-connection strength, respectively. Furthermore, C and C' are the parameters of optimization and are common to all individual maps.

α is defined by the upper limit of the coupling strength. In other words, the coupling strength is optimized in the range of $(0,\ \alpha)$. The genetic algorithm yielded the three types of optimized maps depending on α.

We performed numerical calculations using 10 different seeds of standard random number generators to investigate the effect of α. Changing the seeds of random numbers affects the generation of the initial individuals in the first generation. We used additive Gaussian noise on each individual map, which plays a role in the mutations. Fig. 4 shows a histogram of the evolved individual maps by the optimization algorithm.

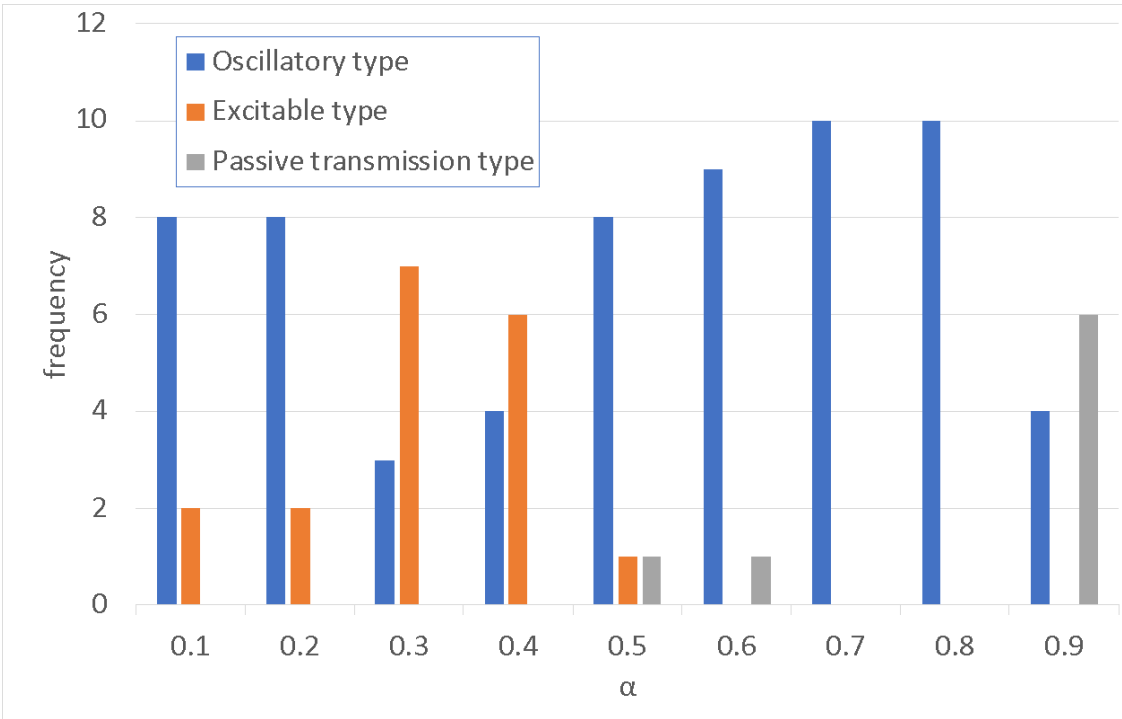

Fig. 4. Histogram of eventually evolved individual maps via optimization algorithm. The number of sites is 10. The additive noise is $\boldsymbol{\sigma \sim N(0, 0.0001)}$

1. Passive transmission type: the case of a large α such that $\alpha = 0.9$ [$N = 10$, $\sigma \sim N(0, 0.0001)$]

When α is significantly large , the individual dynamical systems evolved to simple constant maps. Especially when $\alpha = 1$, which is realized, for instance, via $C = 1$ and $C' = 0$, the individual maps evolved to $f(x) = 0$. From Eqs. (4) and (5),

$$x_0(n+1) = G\big(y(n)\big) \quad x_{k+1}(n+1) = x_k(n) \tag{7}$$

This result indicates that the state of the kth individual map at time n was simply copied to the k+1th individual map at time $n+1$ which is shown in Fig. 5(a). In this case, the external input information can be completely transmitted. This is the trivial case where the solution has obtained maximum information transmission, which we also confirmed via computer simulation.

2. Excitable type: the case of an intermediate α such that $0.4 \geq \alpha \geq 0.3$ ($N = 10$, $\sigma \sim N(0, 0.0001)$)

When α takes intermediate values, the individual maps evolved to excitable maps, as shown in Fig. 5(b).

The map possesses three fixed points. The leftmost fixed point is stable that indicates an equilibrium state, and the other two fixed points are unstable. The middle-fixed point plays a role in the threshold, and the other unstable fixed point plays a role in producing cyclic trajectories. If initial conditions are provided below the threshold, then the dynamical trajectories converge to the equilibrium state, whereas if initial conditions are provided above the threshold, then the dynamical trajectories converge to the equilibrium state after a succession of a large excursions and an undershoot. If there is no external input, all individual maps are attracted to a stable fixed point. Even if there is a small input, each individual map stays around such a fixed point.

However, when the input is sufficiently large to exceed the threshold, the overall dynamics of the network tends to produce chaotic motion, thereby preserving the information quantity of the external signals. This kind of preservation of input information is guaranteed by a large fluctuation of information flow in weakly chaotic systems, where input information can be transmitted before a drastic decay of information owing to the orbital instability of chaos (see also Matsumoto and Tsuda, 1988). This characteristic of the information transmission is not restricted to the unidirectionally coupled maps but is extended to other network architectures as well.

3. Oscillatory type: the case of $0.2 \geq \alpha \geq 0.1$ and $0.8 \geq \alpha \geq 0.5$ ($N = 10$, $\sigma \sim N(0, 0.0001)$)

When α is $0.2 \geq \alpha \geq 0.1$ and $0.8 \geq \alpha \geq 0.5$, the individual maps evolved to a map producing periodic orbits, as shown in Fig. 5(c). The map shown here possesses a stable period two-periodic orbit.

In the presence of an input, the overall network dynamics shows weakly chaotic states, while maintaining the characteristics of period-two periodic motion.

Taking these computation results into account from the aspect of biological evolution, we naturally ask a question of how the neurons evolved, and what is the purpose of such an evolution. Regarding these matters of biological evolution, we propose the following hypothesis.

Hypothesis 1

In biological evolution, neuronal differentiations occurred to achieve a maximum transmission of information of external signals.

As an experimental fact, glial cells produce periodic behaviors of electricity, based on calcium oscillations. Furthermore, glial cells may contribute to DC variations of neuronal activity (A. Ikeda et al., 1999). These observations suggest that one can assign an excitable map and a specific periodic map to a neuron, and both periodic and identity map to glial cells.

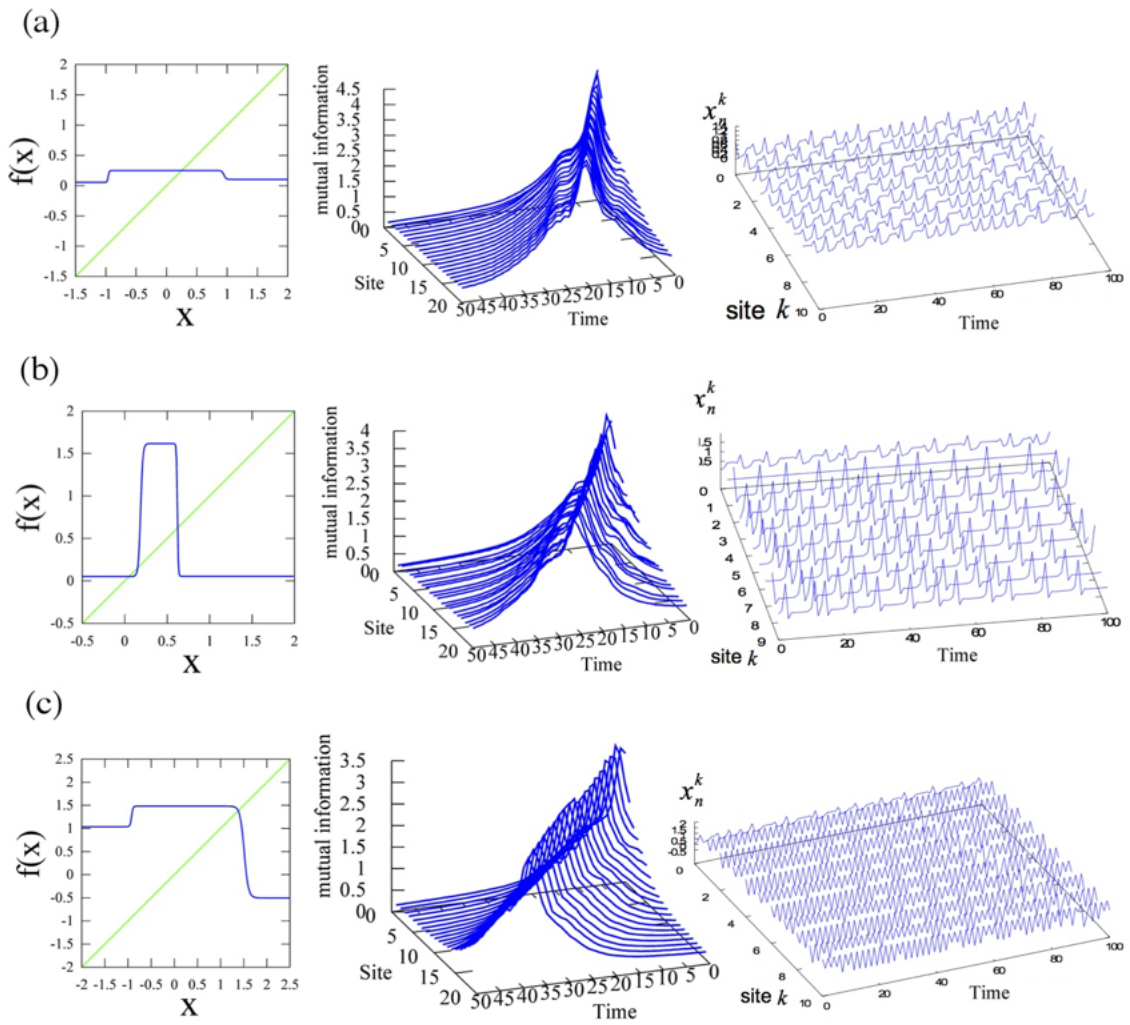

Fig. 5. Three types of optimized maps. From top to bottom: passive transmission type, excitable type, and oscillatory type. Left column shows eventually evolved individual maps via optimization algorithm. Middle column shows temporal change of transmission of input information through network, where network elements are indicated by "site". Right column indicates changes of time-series of amplitudes of individual maps. In neural systems, amplitudes may represent membrane potentials of neurons or amplitudes of other variables measuring neural fields such as LFP, EEG, ECoG, EMG, and fMRI.

4. Further complex networks

We further studied neuronal differentiations owing to evolution dynamics in more complex network architectures.

4.1. Random networks

Erdős and Rényi proposed the Erdős-Rényi model, which generates a random graph by probabilistically combining nodes states (Erdős and Rényi, 1959).In this model, the graph is generated by two nodes in a set of n nodes that are edged with probability p. Fig. 6 shows examples generated with 10 nodes with coupling probability p =0.1, 0.3, and 0.6.

The graphs generated by this model have features such as mean degree $\langle k \rangle$=p(n-1), degree distribution P(deg(v) = k) =$\binom{n-1}{k} p^k (1-p)^{n-1-k}$, and cluster coefficients that converge to zero as n increases.

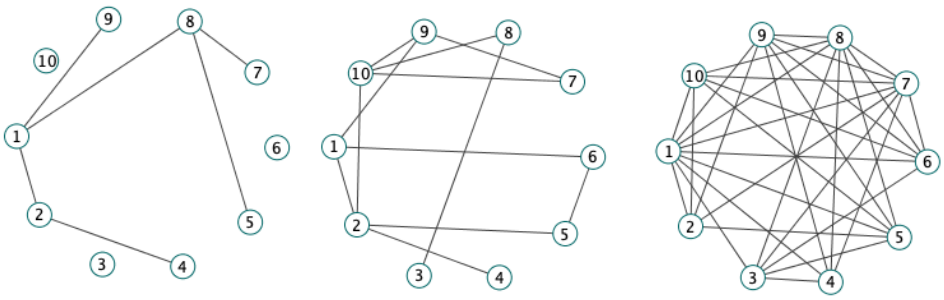

Fig. 6. Examples of construction of random networks (coupling probability = 0.1, 0.3 and 0.6)

We adopted the following networks as directed random networks (see, for example, Fig. 7). In Fig. 7, the number i in the circle indicates the number of sites. Here, the coupling probabilities in 10 individual maps were 0.3 under the constraint of the total number of couplings, which were 30 couplings in the present simulation. In each simulation, we optimized the coupling strengths and individual maps with a fixed network architecture.

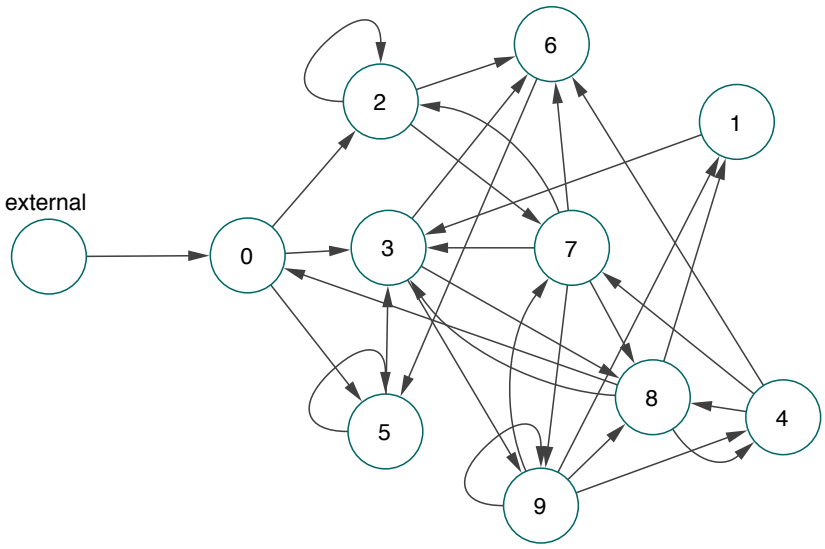

Fig. 7. Example of construction of random networks expressed as weighted directed graph (coupling probability = 0.3).

In the range of coupling strength (-0.1, 0.1), we obtained the optimized network as shown in Fig. 8.

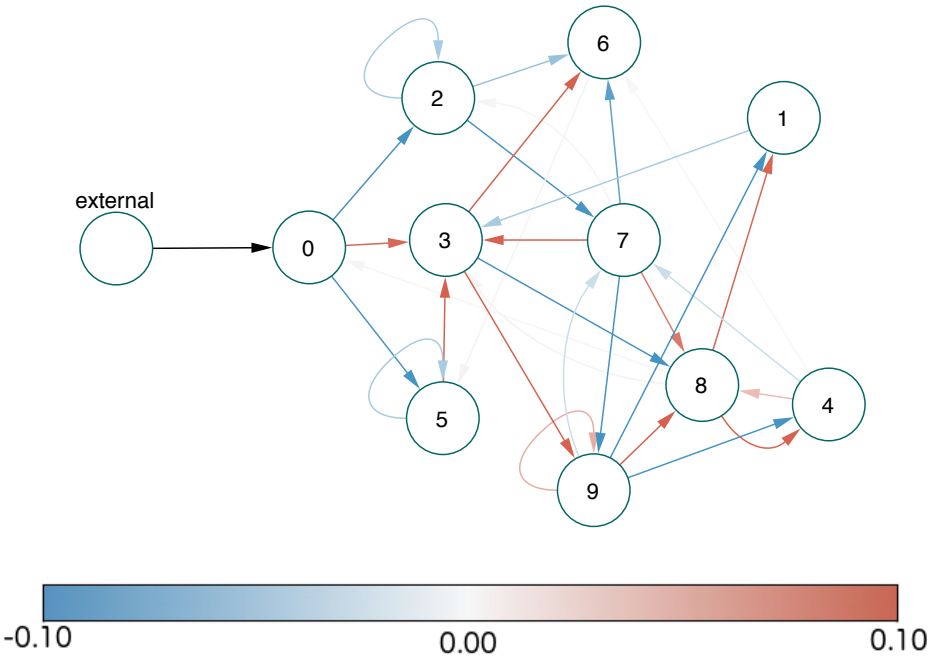

Fig. 8. Optimized random networks in the range of coupling strength (-0.1, 0.1). Red-colored and blue-colored lines and curves indicate excitatory and inhibitory couplings, respectively.

The individual maps eventually developed into the maps shown in Fig. 9. The functional form is common for all sites, but the shift parameter is slightly different: some maps possess a stable fixed point, and others produce period-two periodic orbits.

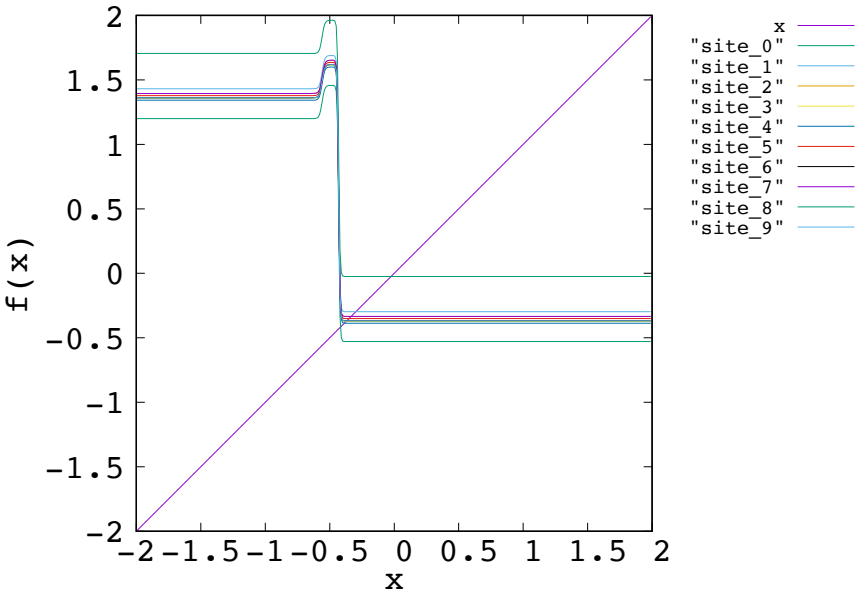

Fig. 9. Optimized map as each individual dynamical system in random networks.

Let site 0 be the zero-th order of the site, which receives input signals directly from the external time-series. Furthermore, let the individual map that directly receives inputs from site 0 be the 1st order of the site. Similarly, the n-th order of the site can be defined by the shortest distance from site 0. The order of each site is listed in Table 1.

Couplings to higher- and lower-ordered sites are called feed-forward connections and feed-back connections, respectively. A coupling to itself is called a self-connection. Coupling to the same order of maps except self-connections is called a homogeneous connection. When couplings are classified as above, the statistical features of the eventually evolved networks are as listed in Table 2. We performed a Wilcoxon-Mann-Whitney test. The absolute values of the strength of the feed-forward and homogeneous connections were significantly larger than those of the feed-back connections (p

< 0.01). The absolute values of the strength of the feed-forward connections were also significantly larger than those of the self-connection strength ($p < 0.01$).

The product of the number of couplings and the mean of the absolute values of the coupling strength suggests a substantial contribution of each coupling type to the overall network architecture.

connections evolved as the greatest contributor for the maximum transmission of the information of input signals. It is interesting to note that the effect of a small number of homogeneous connections was greater than that of a large number of feed-back connections. Furthermore, we classified all couplings into excitatory and inhibitory ones, and further categorized each group into the above four coupling types (see Tables 2 and 3). However, there was no statistically significant difference between the two groups.

The mutual information between the external input and each individual map is shown in Fig. 10. It is observed that the information of the external input is shared with all individual maps with a time delay. Subsequently, we obtained the following proposition.

Proposition 1

Random networks evolved to have a specific character of architecture. Under the constraint of the maximum transmission of input information, the random networks evolved to feed-forward networks including a small number of feed-back connections.

The product of the number of couplings and mean of absolute values of coupling strength defines the effective coupling strength of the network. In this respect, the ratio of feed-forward and feed-back connections is approximately 10 to 2, listed in Table 4. Limited to excitatory and inhibitory connections separately, this ratio was approximately 10 to 3 and 9 to 1, respectively (see Tables 2 and 3). The ratio between the feed-forward and feed-back connections is highly consistent with recent observations of such a ratio in rat and mouse (20:1), and human (10:1) cortical local networks (Seeman et al., 2018).

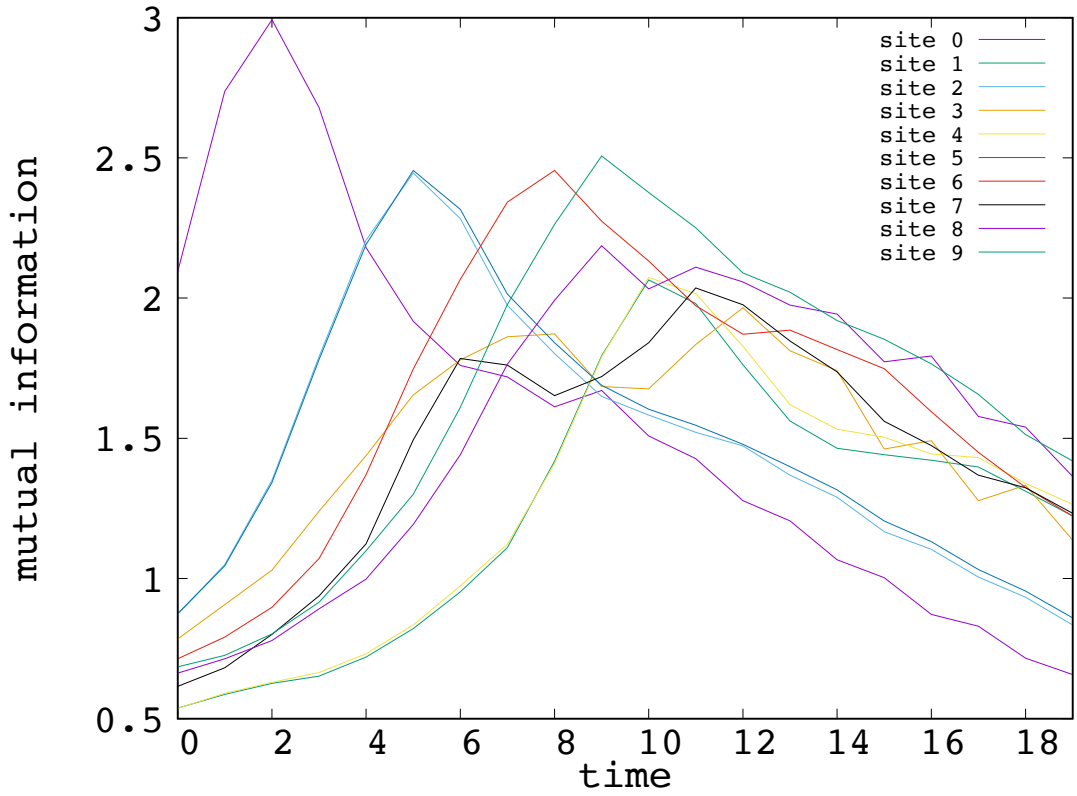

Fig. 10. Mutual information between input signal and each individual map in optimized random networks.

Table 1. Summary of order of the sites, defined by shortest distance between site 0 and each other site.

Order	Site
0	0
1	2, 3, 5
2	6, 7, 8, 9
3	1, 4

Table 2. Summary of the number and strength of excitatory couplings. The numbers in parenthesis in the middle and right most column indicates the standard deviation and proportion, respectively.

Type of couplings	Number of couplings	Mean of absolute value of coupling strength*	Number of couplings × Mean of absolute value of coupling strength**
Feed-forward connections	5	0.098 (0.014)	0.490 (0.51)
Feed-back connections	6	0.024 (0.036)	0.144 (0.15)
Self-connections	1	0.049 (-)	0.049 (0.05)
Homogeneous connections	3	0.094 (0.008)	0.283 (0.29)
Total	15		0.966 (1.00)

Table 3. Summary of the number and strength of inhibitory couplings. The numbers in parenthesis in the middle and the right most column indicates the standard deviation and proportion, respectively.

Type of coupling	Number of couplings	Mean of absolute value of coupling strength*	Number of couplings × Mean of absolute value of coupling strength**
Feed-forward connections	7	0.093 (0.016)	0.649(0.63)
Feed-back connections	3	0.024 (0.016)	0.072(0.07)
Self-connections	2	0.044 (2.75E-05)	0.089(0.09)
Homogeneous connections	3	0.070 (0.033)	0.212(0.21)
Total	15		1.023(1.00)

Table 4. Summary of the number of couplings and absolute values of coupling strength. The number in parenthesis in the middle and the right most column indicates the standard deviation and proportion, respectively.

Type of couplings	Number of Couplings	Mean of absolute value of coupling strength*	Number of couplings × Mean of absolute value of coupling strength**
Feed-forward connection	12	0.095 (0.013)	1.139 (0.57)
Feed-back connection	9	0.024 (0.031)	0.217 (0.11)
Self-connection	3	0.046 (0.0023)	0.138 (0.07)
Homogeneous connection	6	0.082 (0.027)	0.495 (0.25)
Total	30		1.989 (1.00)

*Mean (SD),

**Number of couplings × mean of absolute value of coupling strength (proportion).

4.2. Small-world networks

In sociology, Milgram demonstrated through social experiments the small world phenomenon where any two people are socially connected through a small number of people (Milgram, 1967).

In the Small World model study, Watts and Strogatz proposed the Watts-Strogatz model (Watts and Strogatz, 1998). In this model, we first place the nodes in a circle and connect their k neighborhoods (Fig. 11(a)). In this circular network the clustering coefficient is 1. We select an edge from this network with probability p and randomly cut one of the ends (Fig. 11(b)). We select a new node at random and connect the disconnected side (Fig. 11(c)) If probability p=0, reconnecting does not occur (Fig. 12(a)). When probability p=0.1, the graph is both small-world and clustered, with an average shortest path length of $L \propto \log n$ and a clustering coefficient close to 1 (Fig. 12(b)). If probability p = 1, the graph is random (Fig. 12(c)).

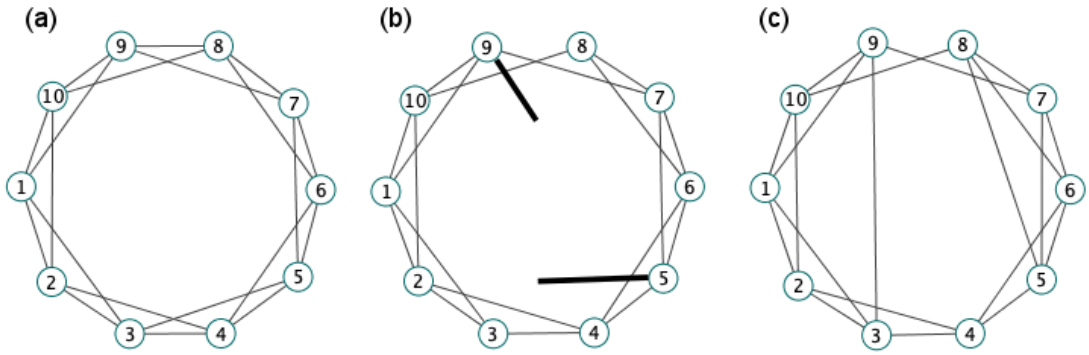

Fig. 11. Example of a small-world network with Watts-Strogatz model (10 nodes, k = 2, reconnection probability p=0.1).

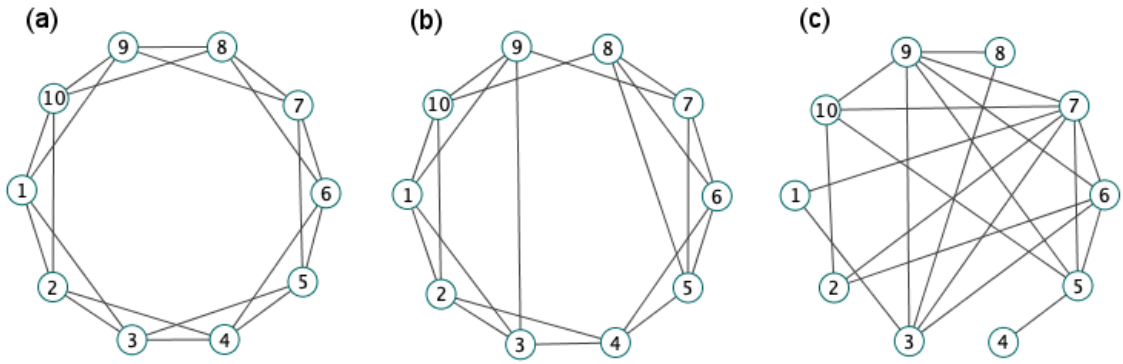

Fig. 12. Effect of reconnection probability p in the Watts- Strogatz model (from left to right, p = 0.0, 0.1, 1.0).

We adopted the following networks as directed small-world networks. The method for the network configuration was based on the Watts-Strogatz model in the following way:

(1) Locate each individual map in a circle to connect with each nearest and next-nearest neighbors.

(2) Reconnect with other maps with a certain probability.

The coupling strength and the network structure were fixed throughout the development of the evolutionary dynamics, and only individual maps were optimized. An external input was fed only to site 0. If the coupling strengths are not fixed, it may become zero via optimization. Here, a zero-coupling strength indicates a disconnection. In a small-world network, disconnections between some sites can significantly change the distance between individual maps. Therefore, the coupling strengths were fixed to maintain the small-world network.

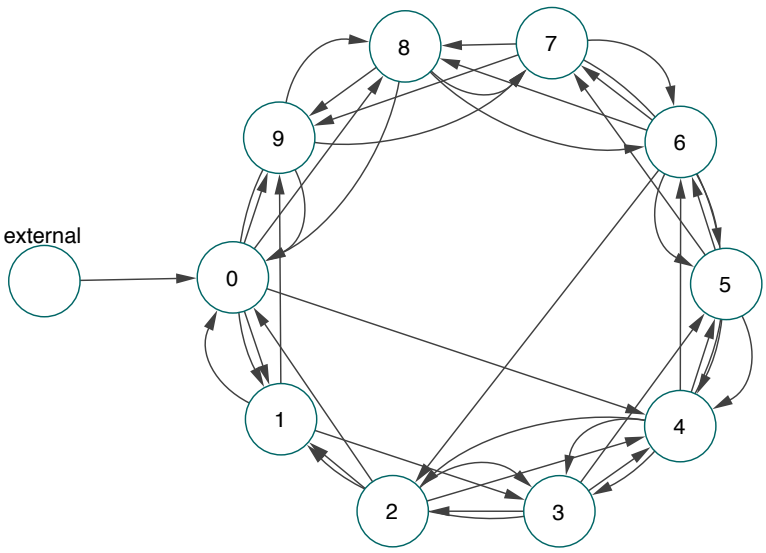

Fig. 13. Example of construction of small-world networks expressed as a weighted directed graph (reconnection probability = 0.1).

We show typical evolved dynamics with some coupling strengths. For the first time, we treat the case in which the coupling strength is fixed to 0.1.

The eventually evolved individual map is shown in Fig. 14. The evolved map possesses one stable fixed point. However, as every map has a similar functional form to the excitable map, which was

obtained in the evolution of unidirectionally coupled maps, the overall dynamics is not restricted to the equilibrium states but produce more dynamic ones such as periodic behaviors, thereby realizing a propagation of pulses (Fig. 15). It can be seen in Fig. 16 that the input information is transmitted in the circular network.

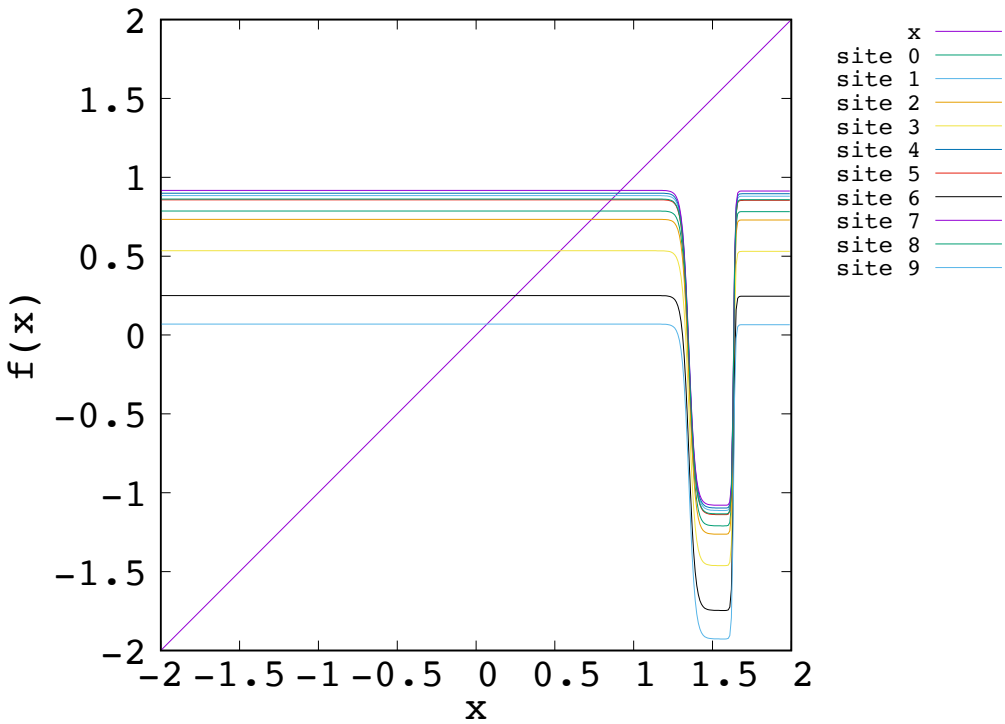

Fig. 14. Optimized individual map in small-world networks (coupling strength = 0.1).

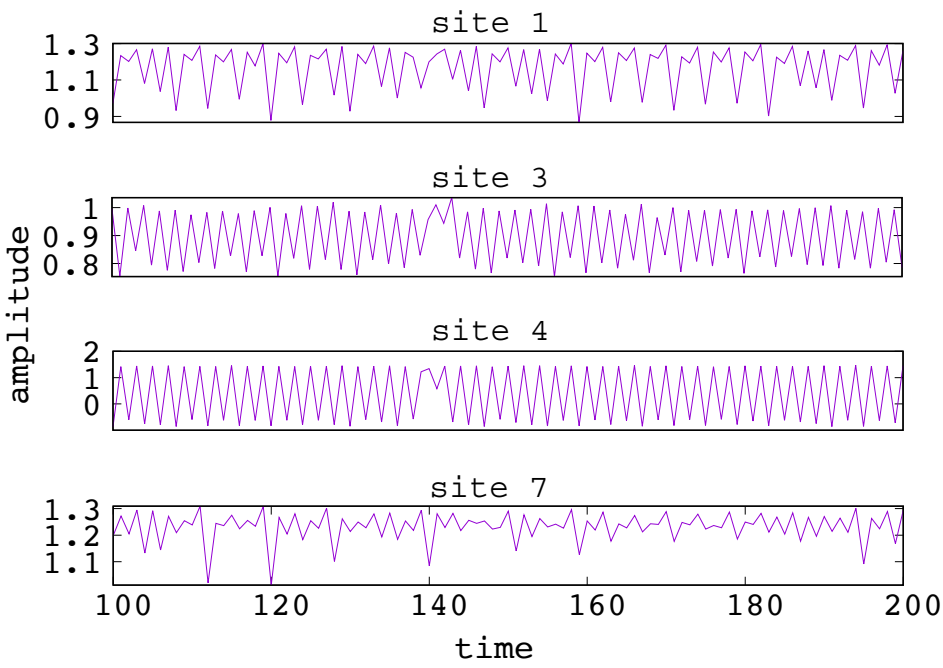

Fig. 15. Time-series of some individual maps in optimized small-world networks (coupling strength = 0.1). The top three figures show intermittent transitions between chaotic states and periodic states with two and four periods. The bottom shows that the oscillations of small amplitudes and the generation of the pulse are repeated.

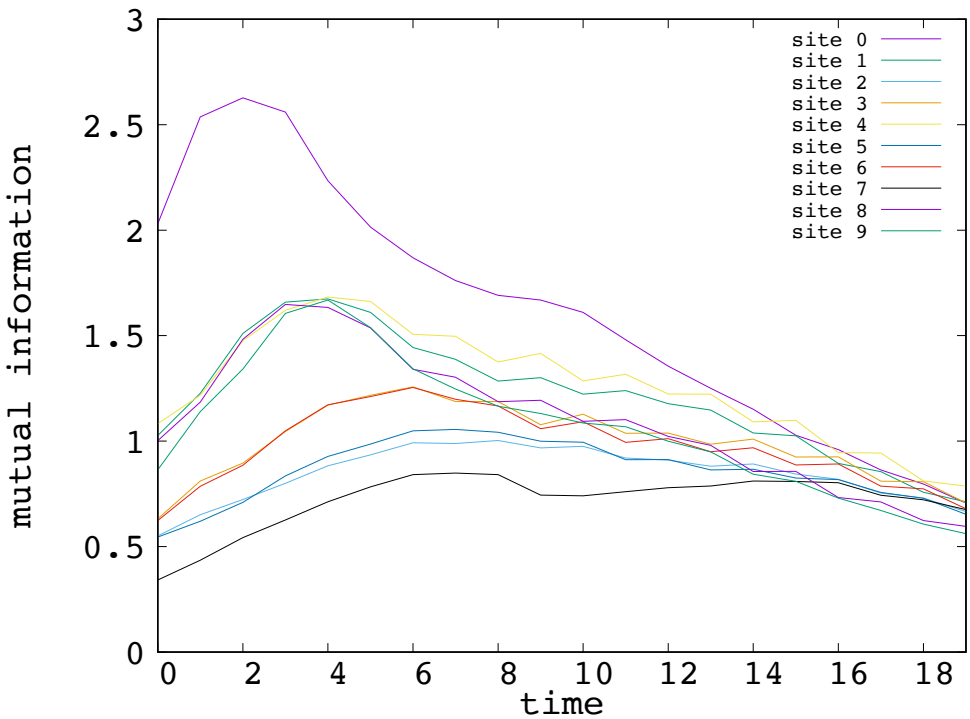

Fig. 16. Mutual information between input and each individual map in optimized small-world networks (coupling strength = 0.1).

We also conducted computer simulations for other coupling strengths and obtained similar results. The varieties of obtained map are the same as in other networks but depending on both the coupling strength and shift parameter, a type of map selected were changed. We obtained the following proposition.

Proposition 2

A small-world network can transmit input information over the entire network. The eventually evolved individual map is modified from the excitable "neuronal" map and tends to produce an equilibrium state directly. The overall network dynamics becomes the intermittent transitions between chaotic and periodic states.

4.3. Fully-connected networks

We used the optimization algorithm for not only individual maps but also the coupling strength in fully-connected networks. These networks include self-connections as shown in Fig. 17.

We simulated two different ranges of coupling strengths whose ranges were (-0.1, 0.1) and (-0.5, 0.5). It should be noted that all simulations included the zero-coupling strength. This implies that computed networks included any network structures.

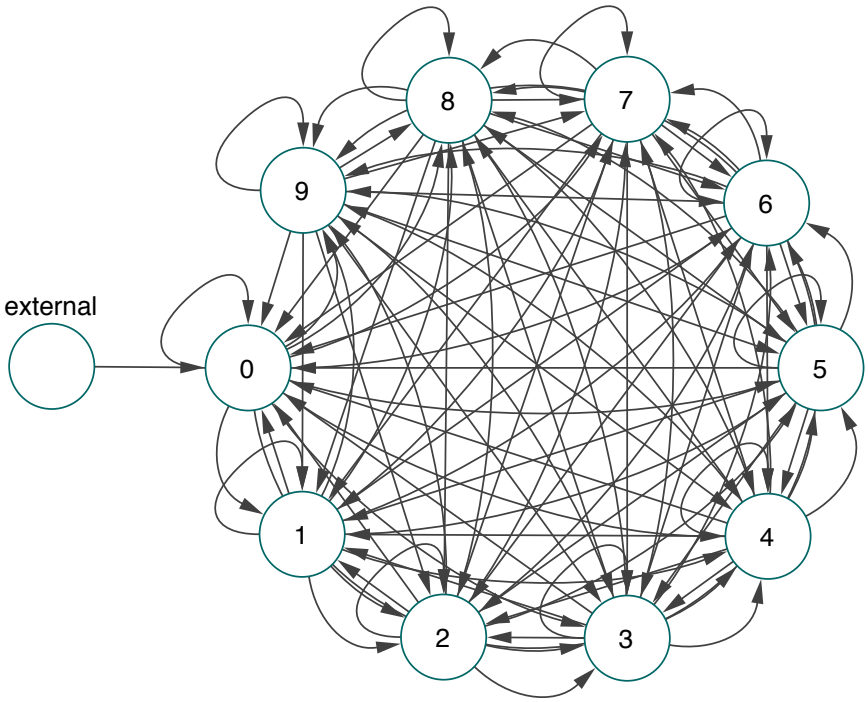

Fig. 17. Fully-connected networks. Each individual map connects to all other nodes including itself.

When the range of coupling strengths was (-0.1, 0.1), we obtained the evolved network as shown in Fig. 18. Furthermore, Fig. 19 and Fig. 20 show individual maps and mutual information with external inputs, respectively. Most individual maps became excitable maps. Because the evolved networks represented any type of network, this computational result implies that the generation of spiking neurons is universal.

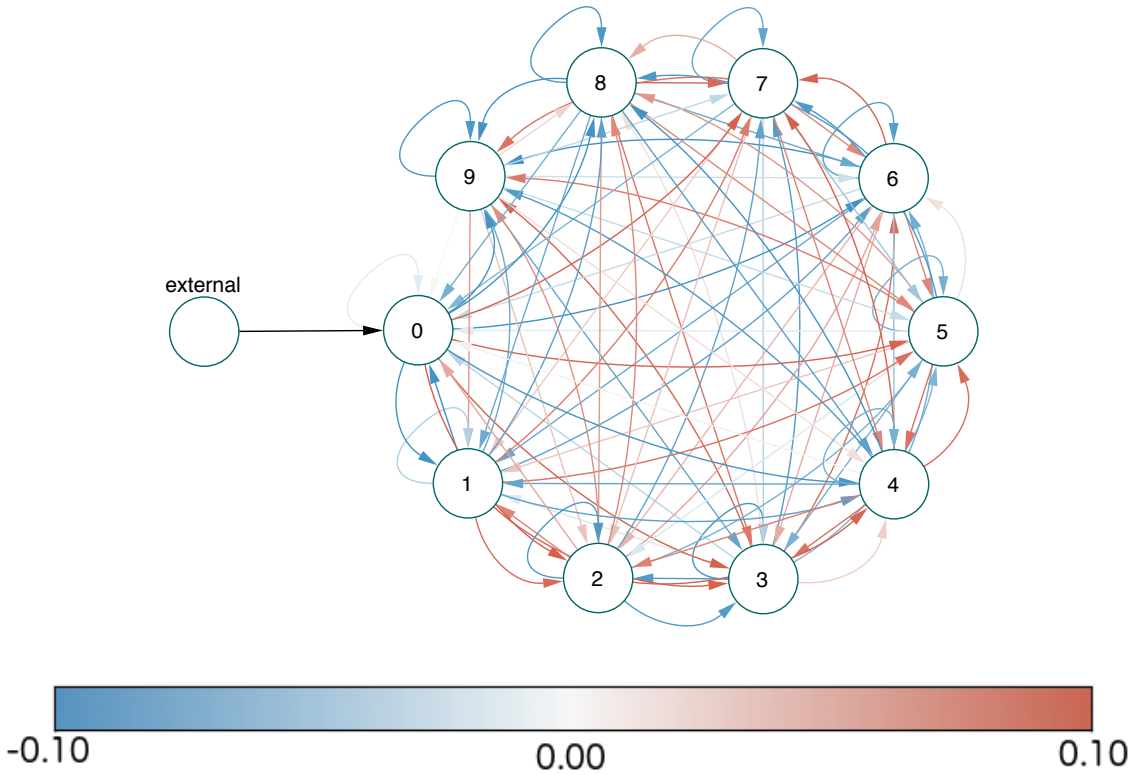

Fig. 18. Optimized fully-connected network under coupling strengths taken over (-0.1, 0.1). Red- and blue-colored curves indicate excitatory and inhibitory coupling, respectively.

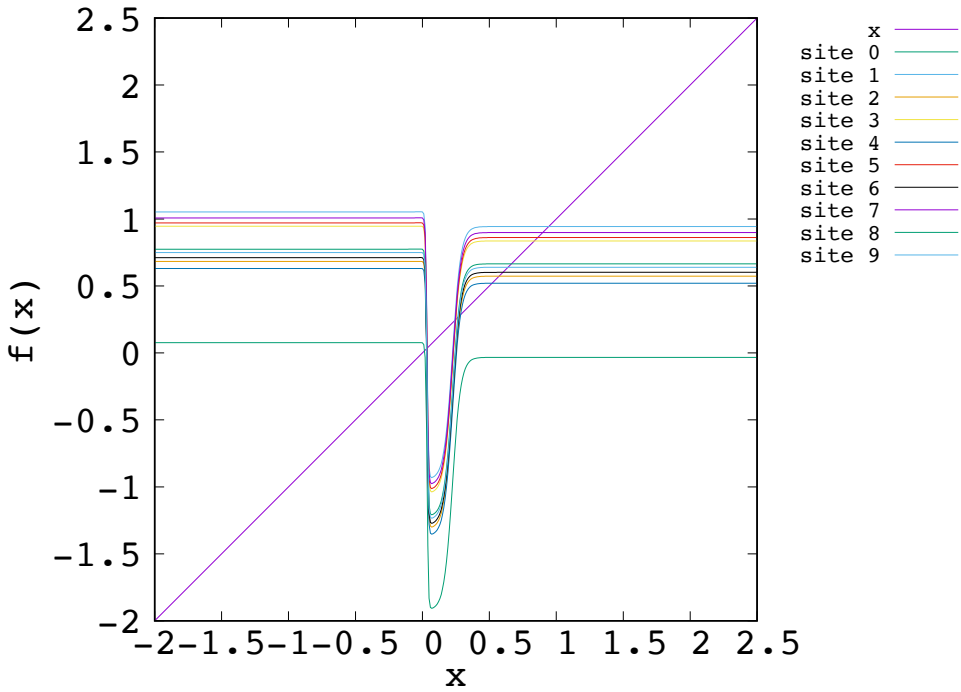

Fig. 19. Optimized individual maps in fully-connected network.

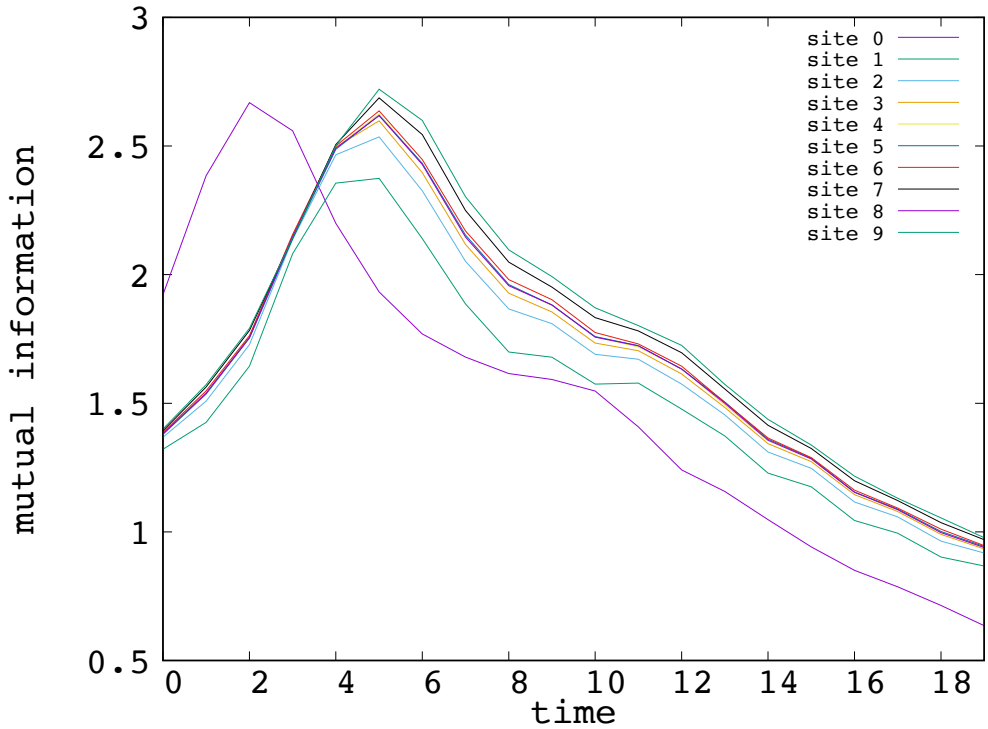

Fig. 20. Mutual information between input signal and each individual map in fully-connected network.

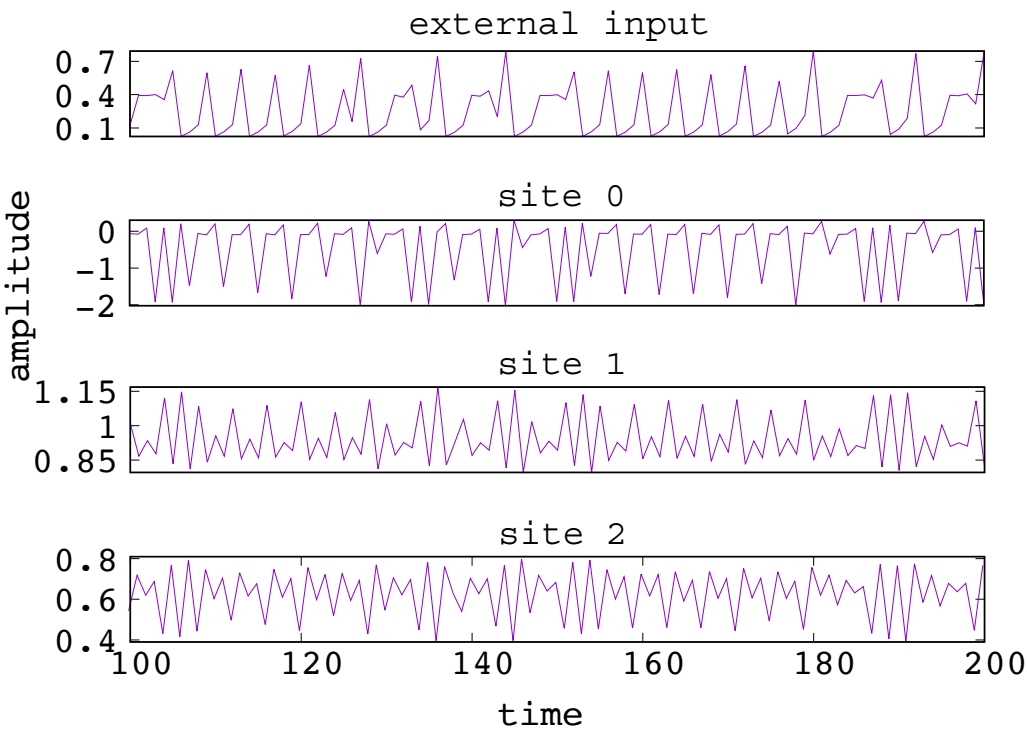

Fig. 21. Time-series of some individual maps in optimized fully-connected network.

Interesting phenomena occurred when the range of coupling strength was (-0.5, 0.5). Fig. 22 shows the eventually evolved network. The evolved individual maps and mutual information between the external input and each individual map are shown in Fig. 23 and 24, respectively. In this case, the mutual information shared between the external input and individual maps except for the "receptor" was larger than the mutual information shared between the external input and the "receptor" that received external input directly. In the network with simple feed-forward connections (Fig. 3), the mutual information with the external input always monotonically decreased in space. Therefore, the mechanism for information transmission must be different. This is discussed in the following subsection.

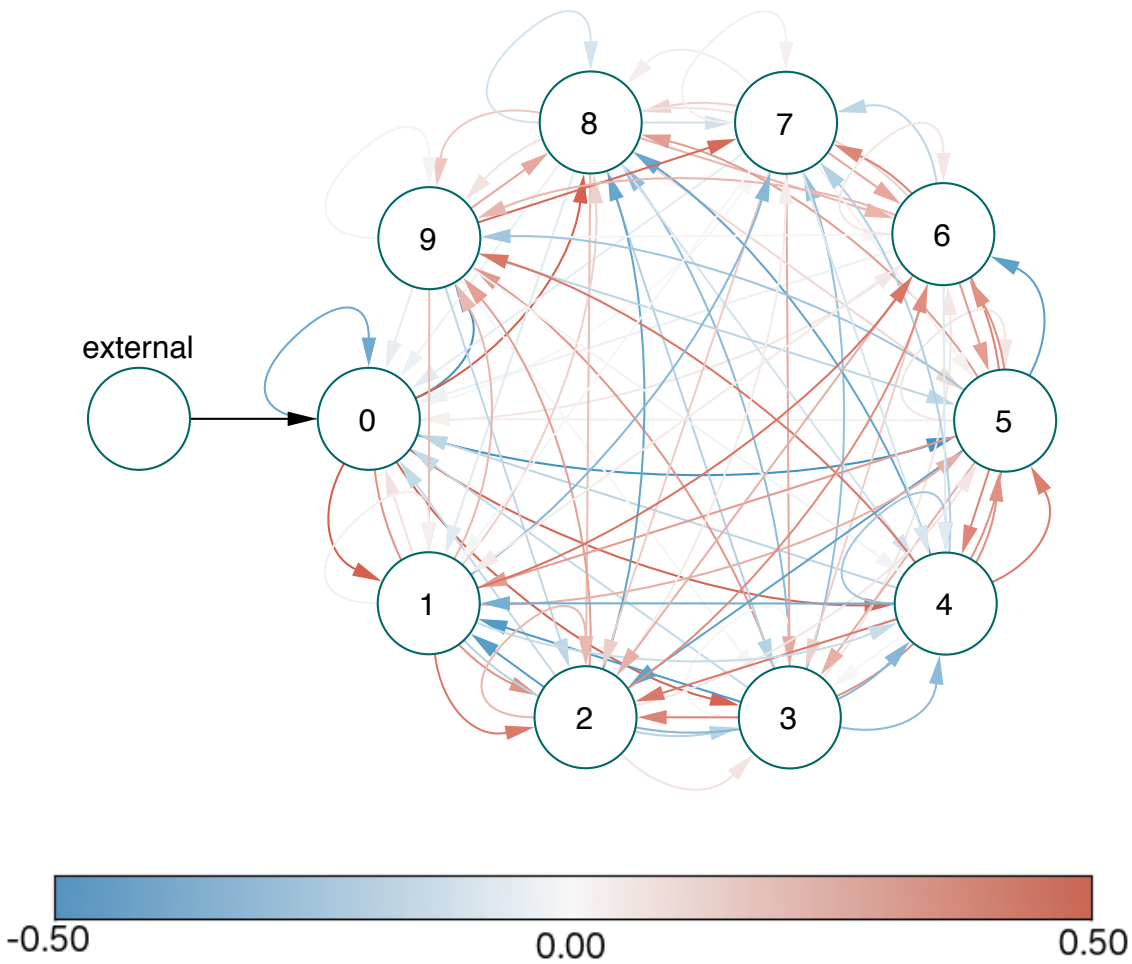

Fig. 22. Optimized fully-connected networks under coupling strength (-0.5, 0.5). Red and blue curves indicate excitatory and inhibitory connections, respectively.

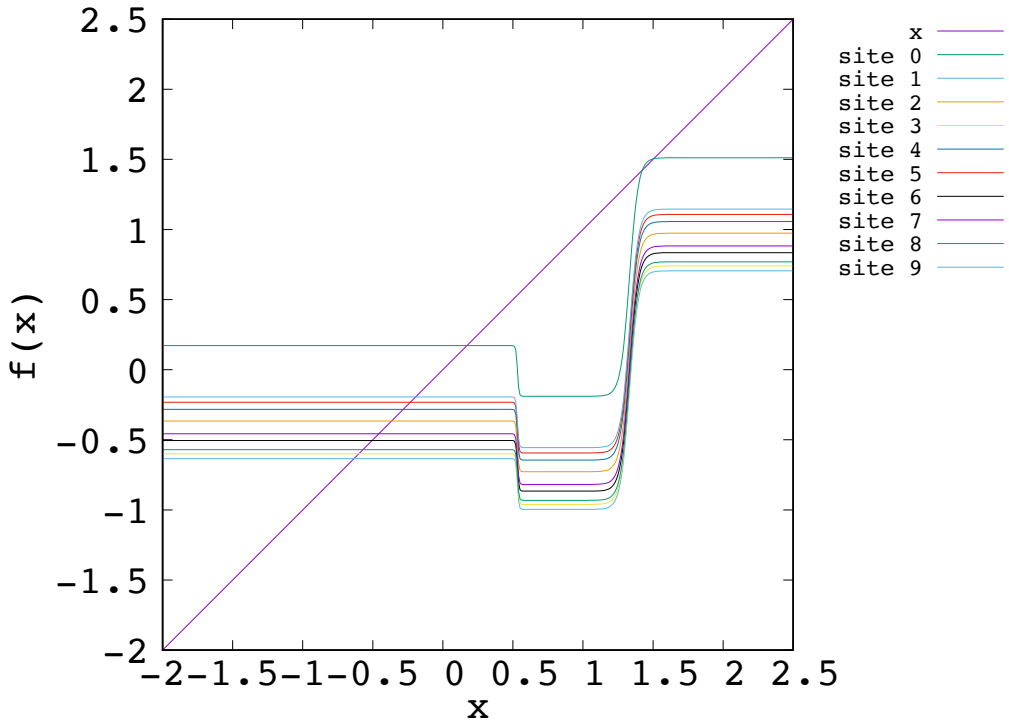

Fig. 23. Optimized map of each element in fully-connected networks.

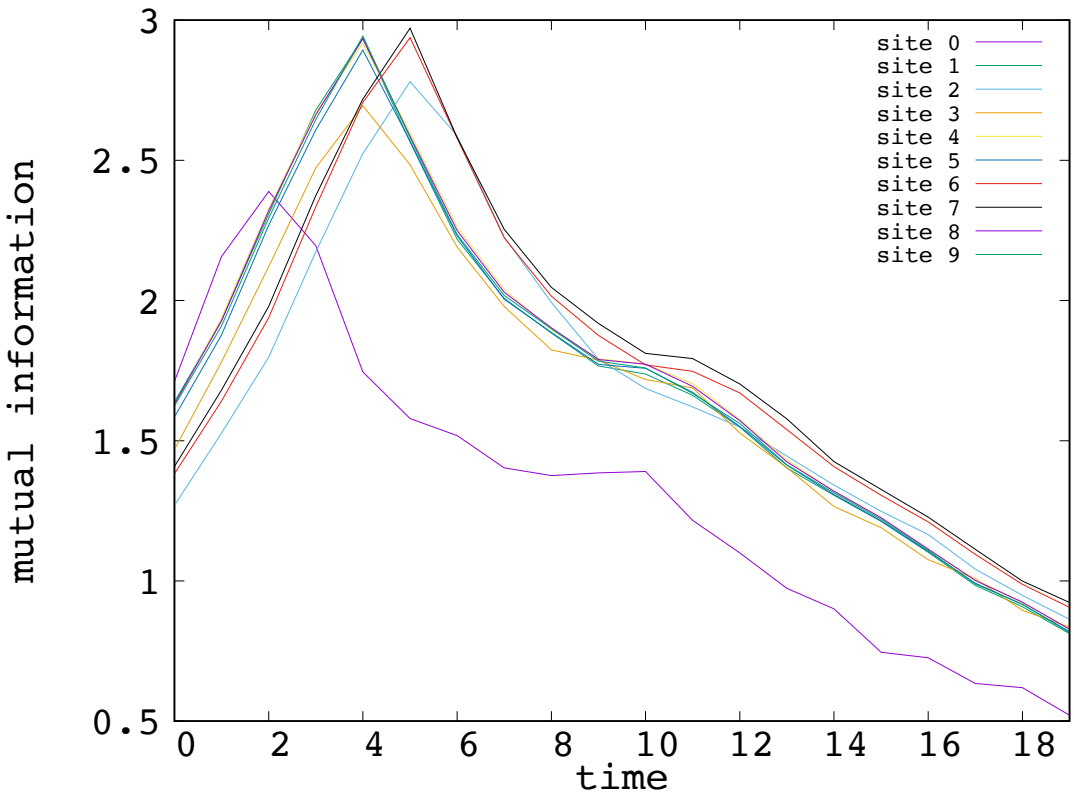

Fig. 24. Mutual information between input and each element in fully-connected networks.

4.4. Randomly shuffled external input data for fully-connected networks

In a fully-connected network developed by the optimization algorithm, the "receptor" did not completely share information with the external input; however, the input information was propagated to other individual maps (Fig. 24). Moreover, much more information from the external input was shared with other internal individual maps than the "receptor."

We hypothesized that this system did not only transfer the information of the external input but also used the temporal structure of the external input to share more mutual information. We generated the following two-input data that differ only in the time structure to confirm this hypothesis.

As original external input data, we generated time-series data according to Eq. (6). As randomly shuffled external input data, we randomly shuffled the original external input data.

We applied these inputs to the evolved network shown in Section 4.3 and measured the mutual information (see Fig. 25). Note that the network structure, coupling strength, and individual maps were fixed.

Furthermore, we have optimized the fully-connected network via a genetic algorithm in a way similar to the previous cases, however, the present case is a noiseless case. In our model, a passive transmission type is always selected when optimized without noise. We applied two types of external inputs (original and randomly shuffled external input data) to the evolved network, and calculated the mutual information between the external input and each individual map (see Fig. 26).

Fig. 25 and Fig. 26 show the change of the information transfer ability in the original input and the randomly shuffled external input data, respectively. It should be noted that the effect is different between the "receptor" and other internal sites. Fig. 26 shows that the evolved network reveals the maximum information transmission ability regardless of the type of time-series data. On the other hand, Fig. 25 shows that the information transmission ability of the evolved network depends on the type of inputs. In Fig. 25., the overall mutual information shared with external inputs in the shuffled data is lower than that of the original data. For the original data, the "receptor" shows the

lowest information transmission, while for the shuffled data, the "receptor" shows the highest information transmission. Destruction of the temporal structure of the external inputs affected the ability of the information transmission of internal elements more significantly than the "receptor". This suggests that for information transmission, the internal elements can use the temporal structure of external input more efficiently than the "receptor". Subsequently, we obtained the following proposition.

Proposition 3

For fully-connected networks, with a selection mechanism, such as satisfying the constraint of the maximum transmission of input information, excitable maps or constant-function maps evolved, depending on the coupling strength. In these networks, information from the external input shared with internal individual maps was much higher than the information shared with the "receptor". In this respect, the fully-connected network transmits external information using the temporal structure of external signals.

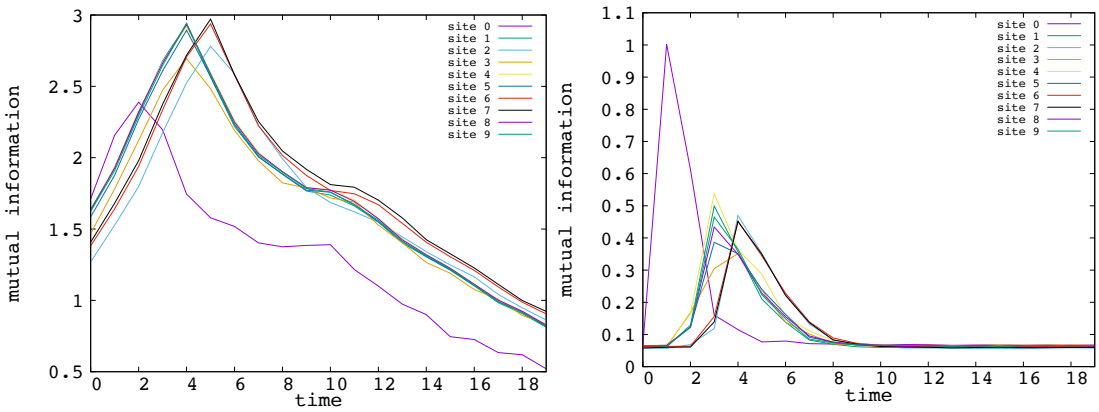

Fig. 25. Mutual information of optimized noisy fully-connected networks. Left figure is the case of original data. The peak of mutual information between external input and "receptor"(site 0) is lower than others. Right figure the case of randomly shuffled external input data. Compared to the left figure, the overall mutual information shared with external inputs is low.

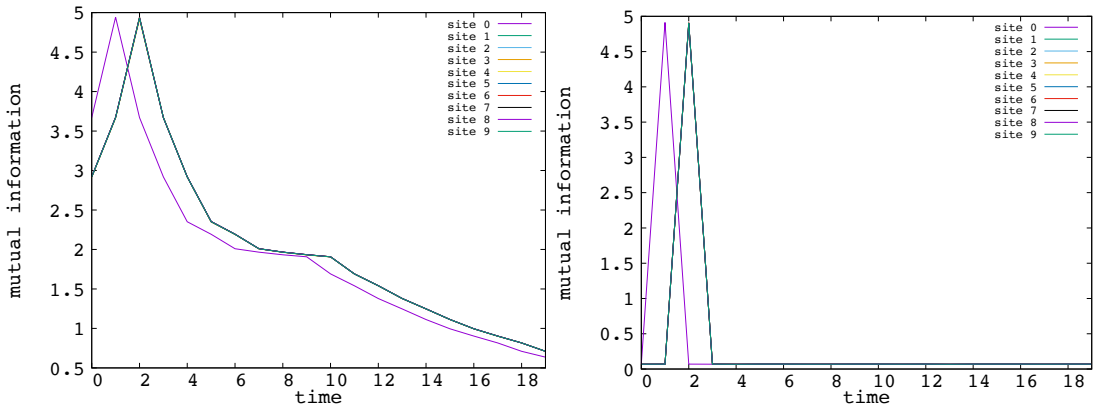

Fig. 26. Mutual information of optimized noiseless fully-connected network. Left figure is the case of original data. Right figure the case of randomly shuffled external input data. All curves are overlapped except for the mutual information between external input and site 0.

5. Summary and Discussion

In the present study, we observed that individual maps of various types of networks evolved to satisfy the constraints: the maximum transmission of external information over the networks. From this observation, we proposed a hypothesis for a mechanism of neuronal differentiations. The computation results were summarized in three propositions. The present computation results suggested that the selective pressure of the efficient transmission of information promoted the functional differentiation of neuronal cells.

We adopted the genetic algorithms to maximize the information transmission capability of coupled networks. In biological evolution, individuals that can use environmental information must have been mostly selected.

We conducted the computer experiments on the evolution of the networks consisting of individual maps, where an individual map is viewed as a biological element. Furthermore, the connections between individual maps can be viewed as the connectivity between subsystems, such as the organs or cells of an individual.

Present network systems have limitations on the number of individual maps, a choice of changeable parameters, and network architectures. However, the computation results seem to show rather universal characters. The results suggest the significance of the selection mechanism of individual units composing the overall network in the process of biological evolution.

Acknowledgments

This work was partially supported by a Grant-in-Aid for Scientific Research on Innovative Areas (Non-Linear Neuro-Oscillology: Towards Integrative Understanding of Human Nature, KAKENHI Grant Number 15H05878) from the Ministry of Education, Culture, Sports, Science and Technology, Japan.

This study was also partially supported by the JST Strategic Basic Research Programs (Symbiotic Interaction: Creation and Development of Core Technologies Interfacing Human and Information Environments, CREST Grant Number JPMJCR17A4).

I would like to express my sincere gratitude to my supervisor, Professor Ichiro Tsuda, for his elaborated guidance, considerable encouragement and invaluable discussions. I would like to thank my co-researcher and friend, Takao Ito, for studying and discussing with me. I would like to thank Yutaka Yamaguti, Satoru Tadokoro, and Takao Namiki for their helpful discussions.

Finally, I would like to express the deepest appreciation to my parents and Aya Nakamura for their support and understanding over a very very very long time.

References

Erdős, P., Rényi, A., 1959. On random Graphs I. Publ. Math. 6, 290–297.

Eshelman, L.J., 1991. The CHC adaptive search algorithm: How to have safe search when engaging in nontraditional genetic recombination, in: Found. Genet. Algor. Morgan Kaufmann Publishers, 265–283.

Goldberg, D.E., 1989. Genetic Algorithms in Search, Optimization and Machine Learning. Addison-Wesley.

Haken, H., 1977. Synergetics. Springer-Verlag, Berlin, Germany.

Haken, H., 1983. Advanced Synergetics. Springer-Verlag, Berlin, Germany.

Holland, J.H., 1992. Adaptation in Natural and Artificial Systems. MIT Press.

Ikeda, A., Taki, W., Kunieda, T., Terada, K., Mikuni, N., Nagamine, T., et al., 1999. Focal ictal direct current shifts in human epilepsy as studied by subdural and scalp recording. Brain 122, 827-838.

Matsumoto, K., Tsuda, I., 1985. Information theoretical approach to noisy dynamics. J. Phys. A: Math. Gen. 18, 3561–3566.

Matsumoto, K., Tsuda, I., 1987. Extended information in one-dimensional maps. Phys. D 26, 347–357.

Matsumoto, K., Tsuda, I., 1988. Calculation of information flow rate from mutual information. J. Phys. A: Math. Gen. 21, 1405–1414.

Milgram, S., 1967. The Small World Problem. Psychol. Today 2, 60–67.

Nicolis, G., Prigogine, I., 1977. Self-Organization in Nonequilibrium Systems. Wiley, New York, NY, USA.

Okano, H., Yamanaka, S., 2014. iPS cell technologies: significance and applications to CNS regeneration and disease. Molecular Brain, 7, 22.

Seeman, S.C., Campagnola, L., Davoudian, P.A., Hoggarth, A., Hage, T.A., Bosma-Moody, A., Baker, C.A., Lee, J.H., Mihalas, S., Teeter, C., Ko, A.L., Ojemann, J.G., Gwinn, R.P., Silbergeld, D.L., Cobbs, C., Phillips, J., Lein, E., Murphy, G., Koch, C., Zeng, H., Jarsky, T., 2018. Sparse

recurrent excitatory connectivity in the microcircuit of the adult mouse and human cortex. ELife 7. https://doi.org/10.7554/eLife.37349.

Tomita, K., Tsuda, I., 1980. Towards the interpretation of Hudson's experiment on the Belousov-Zhabotinsky reaction: Chaos due to delocalization. Prog. Theor. Phys. 64, 1138–1160.

Tsuda, I., Shimizu, H., 1985. Self-organization of the dynamical channel, in: Haken, H. (Ed.), Complex Systems: Operational Approaches in Neurobiology, Physics, and Computers. Springer-Verlag, pp. 240–251.

Watts, D., Strogatz, S., 1998. Collective dynamics of small - world networks. Nature 393, 440–442.