Estimation of network structures from multivariate point processes

(多変量点過程時系列データを用いたネットワーク構造推定)

2014 年 3 月

埼玉大学大学院理工学研究科 (博士後期課程)

理工学専攻 (主指導教員　池口 徹 教授)

黒田 佳織

Abstract

Nonlinear dynamical systems often produce complicated behavior due to interaction between the systems. Such complicated behavior usually depends on connectivity in networks, that is, network topology. Thus, to analyze, model or predict complicated behavior produced from the networks, it is essential to understand the network structures as well as the nonlinear dynamics. Although it is not so easy to investigate the interactions directly, recent developments in measurement technique makes it possible to observe multivariate time series. Then, it is possible to estimate the network structure through the multivariate time series. If an observed time series is continuous and smooth, and sampled by a fixed interval, the network structures can be estimated through statistical measures applied to the continuous time series. However, the nonlinear dynamical systems are often observed as event sequences, for example, firing of neurons, occurrence timing in seismic tremors, transaction timing in stock markets, lightning strike, and so on. Such event sequences are often referred as point processes. It is difficult to directly apply the conventional statistical measures to such point processes. Then, it is an important issue to develop a method to estimate network structures in case that the point processes are observed. In this thesis, we proposed estimation methods of network structures only from the point processes.

In the proposed methods, we introduced three strategies: (1) transformation of point processes into continuous time series, (2) using normalized distance between point processes and (3) using multi-dimensional scaling with the distance between point processes.

In the first strategy, we applied the method of transforming point processes into continuous time series to detect the connectivity between elements. As the transformation method, we used a kernel density estimator. In the kernel density estimator, we have to select the optimal kernel bandwidth because the transformed time series depends on the bandwidth. Then we proposed two selection methods of the kernel bandwidth: a kernel bandwidth optimization method for estimating firing rates and a selection method of an optimal time delay in the attractor reconstruction which is a basic technique in chaotic time series analysis. By using these bandwidth selection methods in the kernel density estimator, we could reconstruct input information applied neurons from the point processes, and estimate network structures from the point processes. In addition, we treat point processes which have the information of the amplitude and event timing as marked point processes.

If we use not only the information of the event timing but also the amplitude information, we can estimate more precisely the network structures. Then we extended the transformation method and the estimation method of network structure to marked point processes.

Although these methods work well, one should be careful to apply these methods, because it is possible to lose essential information of point processes by transforming point processes into continuous time series. Then, as the second strategy, we proposed new methods for estimating network structures from point processes by using normalized distance between point processes. Using the distance between point processes, we proposed two measures, a spike time metric coefficient and a partial spike time metric coefficient. The spike time metric coefficient defined by using the normalized distance is a similar measure to the correlation coefficient. Then we applied partialization analysis to the spike time metric coefficient. We experimentally confirmed that these measures can estimate true connectivities from point processes by removing spurious correlations. However, these measures are heuristically defined. Then, to calculate theoretically, we proposed another estimation method by using multi-dimensional scaling with the distance between point processes as the third strategy. We also proposed estimation methods of network structures from marked point processes by using distance between marked point processes. Furthermore, we proposed methods of estimating evolving network structures by dividing the point processes into small temporal epochs and applying the method of estimating static network structures. As a result, the proposed method can estimate the evolving neural network structures and the direction of couplings with high estimation accuracy.

Contents

Chapter 1

Introduction

Interactions among elements in nonlinear dynamical systems often produce complicated behavior. Such complicated behavior usually depends on connectivity of elements in networks, that is, network topology. Thus, to analyze, model or predict the behavior, it is inevitable and essential to understand the network structures as well as the nonlinear dynamics. Although it is not so easy to investigate the interactions between elements directly, recent developments in measurement technologies make it possible to observe multivariate time series data. Then it is possible to estimate the network topology through the multivariate time series data.

On the other hand, complex phenomena are ubiquitous in the real world and often observed as event sequences, for example, firing of neurons, occurrence timing in seismic tremors, transaction timing in stock markets, lightning strike, and so on. Such event sequences are often referred as point processes. In the point processes, a point process which has information of only event timing, such as a spike sequence, is called a simple point process. However, the amplitude information with event sequences can be essential for several phenomena, for example, financial systems, seismic events, and so on [1, 2]. Such sequences are often referred as marked point processes [3]. If the observed time series from systems are smooth and continuous, the connectivity of the system can be identified with several statistical measures [4, 5, 6, 7]. One of the most popular measures in the partialization analysis is a partial correlation coefficient which is based on a correlation coefficient and the partialization analysis. Frenzel et al. proposed partial mutual information which is based on mutual information and partialization analysis [4]. Schelter et al. proposed a partial phase synchronization index which is based on phase synchronization and the partialization analysis [5].

Sameshima et al. proposed partial directed coherence which is based on the Granger causality [6, 7] and the partialization analysis. Although the methods described in Refs. [4, 5, 6, 7] work well for the smooth and continuous time series, the application of these point processes remain unclear. Then it is an important issue to develop methods for estimating network structures in case that point processes are observed. From this point of view, in this thesis, we proposed estimation methods of network structures only from point processes.

As a conventional method, Sameshima et al. proposed an estimation method of connectivity between neurons by transforming spike sequences into continuous time series and using the partial directed coherence[7]. Although this conventional method has good performance to estimate connectivity between neurons from spike sequences, this method has an issue that the method how to select the kernel bandwidth has not proposed. To transforming the spike sequence into the continuous time series, it is important to select an optimal kernel bandwidth because the transformed time series depend on the optimal kernel bandwidth. To resolve this issue, first, we proposed two selection methods of the kernel bandwidth: a kernel bandwidth optimization method for estimating firing rates and a selection method of an optimal time delay in the attractor reconstruction which is a basic technique in chaotic time series analysis. By using these bandwidth selection methods in the kernel density estimator, we could reconstruct input information applied neurons from the point processes and estimate network structures from the point processes. In addition, we treat point processes which have the information of the amplitude and event timing as marked point processes. If we use not only the information of the event timing but also the amplitude information, we can estimate more precisely the network structures. Then we extended the transformation method and the estimation method of network structure to marked point processes. Second, we introduced a new strategy to estimate network structures from the point processes: we used distance between point processes. Victor et al. proposed a spike time metric which is one of the statistic to quantify a distance between two point processes [13]. By using the spike time metric, we proposed a spike time metric coefficient which is similar measure to a correlation coefficient and applied the partialization analysis to the spike time metric coefficient. We checked the performance of the proposed method in numerical simulations, and the results show that the proposed method exhibits high performance. However, we have not theoretically proven why the measure works well. Then we proposed a new method to estimate network structures from point processes by using multi-dimensional scaling with the distance between point processes. The multi-dimensional scaling is a classical method for representing arrangement in the Euclidean space from a distance relation [8]. Using the distance between point processes which calculated by the spike time metric, we can obtain the position vectors which represent the linear relationship among spike sequences in the Euclidean space, and applied a linear regression model to the obtained position vectors. Then

we could use the statistical measures of the partialization analysis such as the partial correlation coefficient and the partial directed coherence because these measures are based on the linear regression model. We also proposed methods of estimating network structures from marked point processes by using distance between marked point processes. In addition, we proposed estimation methods of not only static network structures but also evolving network structures. It is important to estimate dynamic structures, or to detect how the neural network structure changes, because one of the intrinsic properties in neural networks is learning. When the neural networks accept external stimulation, neural networks change their structures by learning. In the proposed method, we divided the observed multi-spike sequences into small temporal epochs. Then, we applied the estimation method of the static network structure to the temporally divided multi-spike sequences and estimated their connectivities.

Figure 1.1 shows relation to chapters in this thesis. In Chapter 2, we reviewd the partialization analysis and explained several statistical measures: the partial correlation coefficient, the partial mutual information and the partial directed coherence. In Chapter 3, the methods of estimating the network structures from point processes by transforming point processes into continuous time series was proposed. First, we proposed selection methods of the optimal kernel bandwidth. By using the selection methods of the optimal kernel bandwidth, we reconstructed input information applied to a neuron from a point process. Next, we estimated network structures only from simple point processes using the transforming method. Further, we extended the transforming method to the marked point process and proposed the estimation method of network structures from marked point processes. In Chapter 4, we proposed the method of estimating network structures by using normalized distance between point processes. As the distance between point proccesses, the spike time metric proposed by Victor et al. [13] is introduced. We proposed two measures, a spike time metric coefficient and a partial spike time metric coefficient, using the spike time metric. In Chapter 5, the estimation method of network structures by using distance between point processes and the multi-dimensional scaling. We proposed the estimation methods of network structures from not only simple point process but also marked point process. In addition, the estimation method of evolving network structures is also proposed. Finally, this thesis is concluded with Chapter 6.

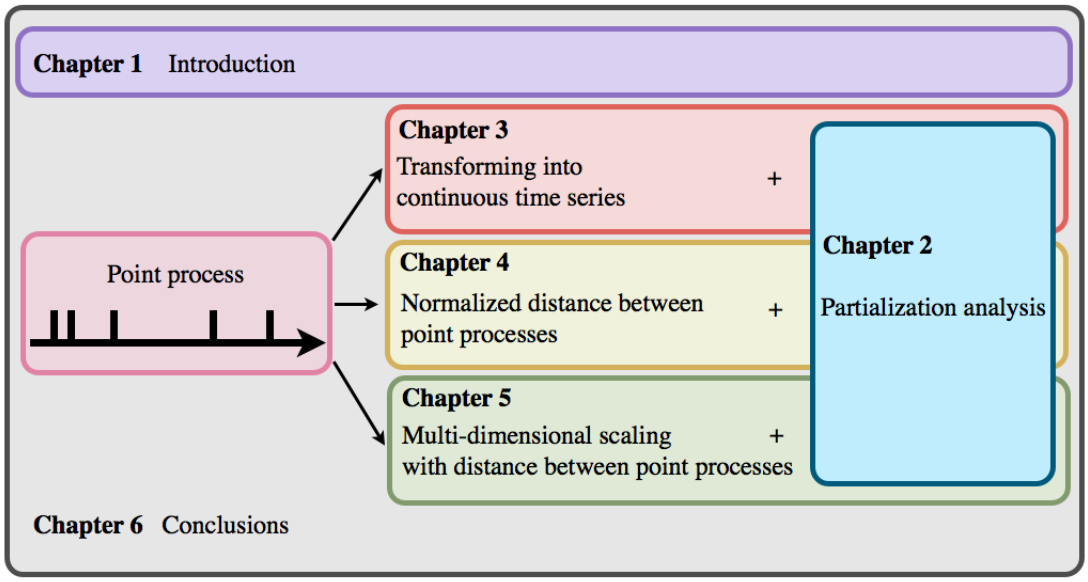


Figure 1.1: Relation to chapters in this thesis.

Chapter 2

Partialization Analysis

In recent years, the several partialization analysis for continuous time series are proposed [4, 5, 6, 7]. In this chapter, we explain the partialization analysis and several statistical measures.

2.1 Spurious Correlation

To estimate the connectivity between elements from time series, we quantify the similarity between two time series. However, we cannot estimate the connectivity between elements because of the spurious correlation.

For example, if two elements, i and j, are not connected, the correlation between two time series takes zero because these time series are different so much (Fig. 2.1(a)). However, if two elements are driven by a common input, or indirectly connected but not directly connected, their time series can have spurious correlation (Fig. 2.1(b), (c)). To remove such a common influence, the partialization analysis is effective. In the next section, we explain several statistical measures.

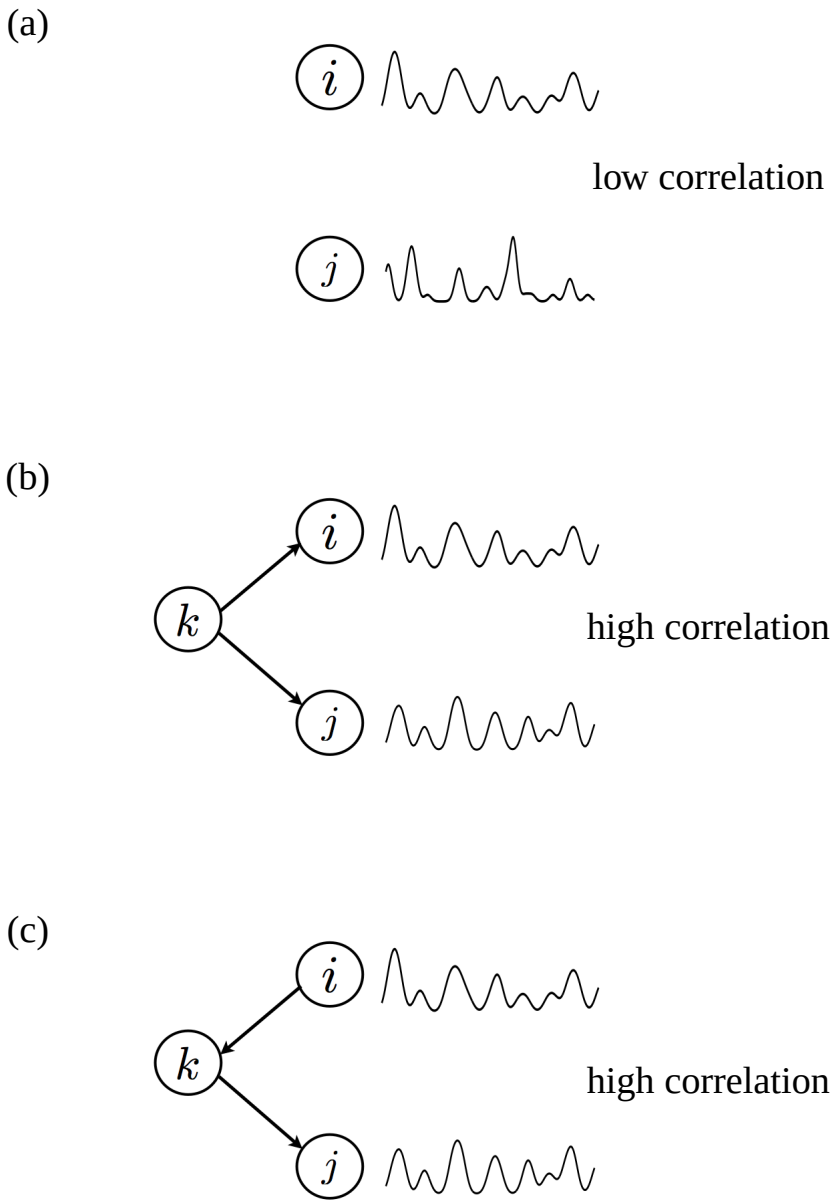


Figure 2.1: Three coupling ways and the correlation: (a) element *i* and element *j* are independent, (b) element *i* and element *j* are driven by a common input from another element *k*, and (c) element *i* and element *j* are indirectly connected.

2.2 Statistical Measures

2.2.1 Partial Correlation Coefficient

The most general measure in the partialization analysis is the partial correlation coefficient. Let us consider time series data $y_1(t), y_2(t), \ldots, y_m(t), (t = 1, 2, \ldots, n)$ are observed. $\bar{y}_i = \frac{1}{n}\sum_{t=1}^{n} y_i(t)$, $s_{ij} = \frac{1}{n}\sum_{t=1}^{n}(y_i(t) - \bar{y}_i)(y_j(t) - \bar{y}_j)$. Then, a correlation coefficient is defined as

$$r_{12} = \frac{s_{12}}{\sqrt{s_{11}s_{22}}}. \tag{2.1}$$

Let us define linear multiple regression models as follows:

$$y_1(t) = a_0 + a_3y_3(t) + a_4y_4(t) + \cdots + a_my_m(t) + e_1(t), \tag{2.2}$$
$$y_2(t) = b_0 + b_3y_3(t) + b_4y_4(t) + \cdots + b_my_m(t) + e_2(t), \tag{2.3}$$

where a_i, b_i are regression coefficients, and e_1, e_2 are residuals. The linear multiple regression models can be written as:

$$y_1(t) = \hat{a}_0 + \hat{a}_3y_3(t) + \hat{a}_4y_4(t) + \cdots + \hat{a}_my_m(t), \tag{2.4}$$
$$y_2(t) = \hat{b}_0 + \hat{b}_3y_3(t) + \hat{b}_4y_4(t) + \cdots + \hat{b}_my_m(t). \tag{2.5}$$

Next, we estimate the regression coefficients of linear multiple regression models as follows:

$$\sum_{t=1}^{n} e_1^2(t) = \sum_{t=1}^{n}(y_1(t) - (a_0 + a_3y_3(t) + a_4y_4(t) + \cdots + a_my_m(t)))^2, \tag{2.6}$$
$$\sum_{t=1}^{n} e_2^2(t) = \sum_{t=1}^{n}(y_2(t) - (b_0 + b_3y_3(t) + b_4y_4(t) + \cdots + b_my_m(t)))^2. \tag{2.7}$$

Let us denote that $F(a_0, a_3, \cdots, a_m) = \sum_{t=1}^{n}(y_1(t) - (a_0 + a_3y_3(t) + a_4y_4(t) + \cdots + a_my_m(t)))^2$. Then we calculate

$$\frac{\partial F}{\partial a_0} = 0, \quad \frac{\partial F}{\partial a_3} = 0, \quad \frac{\partial F}{\partial a_4} = 0, \quad \cdots, \quad \frac{\partial F}{\partial a_m} = 0.$$

From these calculations, the estimated regression coefficients are obtained as:

$$\hat{a}_i = -\frac{S_{22,1(i+1)}}{S_{22,11}}. \tag{2.8}$$

where S is defined by the equation

$$S = \begin{vmatrix} s_{11} & s_{12} & \cdots & s_{1m} \\ s_{21} & s_{22} & \cdots & s_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ s_{m1} & s_{m2} & \cdots & s_{mm} \end{vmatrix}. \tag{2.9}$$

Here, S_{ij} is the (i, j)th-cofactor matrix of S, and $S_{ij,kl}$ is the (k, l)th-cofactor matrix of S_{ij}. In the same way, the regression coefficient b_i are estimated as follows:

$$\hat{b}_i = -\frac{S_{11,2(i+1)}}{S_{11,22}}. \tag{2.10}$$

From the Eqs. (2.23) and (2.10), we can denote $u(t)$ and $v(t)$ as follows:

$$\begin{aligned} u(t) &= y_1(t) - (\hat{a}_0 + \hat{a}_3 y_3(t) + \cdots + \hat{a}_m y_m(t)), \\ v(t) &= y_2(t) - (\hat{b}_0 + \hat{b}_3 y_3(t) + \cdots + \hat{b}_m y_m(t)). \end{aligned}$$

The correllation coefficient between $u(t)$ and $v(t)$ is defined by the equation

$$r_{uv} = \frac{s_{uv}}{\sqrt{s_{uu} s_{vv}}}, \tag{2.11}$$

and r_{uv} is the partial correlation coefficient. Therefore, partial correlation coefficient between $y_1(t)$ and $y_2(t)$ is the correlation coefficient between $y_1(t)$ and $y_2(t)$ which removed regression from $y_3(t), \ldots, y_m(t)$. The partial correlation coefficient can be rewritten by using elements of the inverse matrix of the correlation matrix. Because s_{uu}, s_{vv} and s_{uv} are rewitten as follows:

$$\begin{aligned} s_{uu} &= \frac{1}{n}\sum_{t=1}^{n}(y_u(t) - \overline{y_u})^2 \\ &= \frac{1}{n}\sum_{t=1}^{n}((y_i(t) - \overline{y_i}) - (\hat{a_1}(y_k(t) - \overline{y_k})))^2 \\ &= \frac{1}{n}\sum_{t=1}^{n}((y_i(t) - \overline{y_i})^2 - 2\hat{a_1}(y_i(t) - \overline{y_i})(y_k(t) - \overline{y_k}) + \hat{a_1}(y_k(t) - \overline{y_k})^2) \\ &= s_{ii} - 2\hat{a_1} s_{ik} + \hat{a_1}^2 s_{kk} \\ &= s_{ii} - \frac{2}{S_{22,11}}(s_{ii}S_{22,11} - s_{ik}S_{22,13} - s_{ii}S_{22,11}) + \frac{s_{ki}}{s_{kk}}\hat{a_1} s_{kk} \\ &= s_{ii} - \frac{2}{S_{22,11}}(s_{ii}S_{22,11} - s_{ik}S_{22,13} - s_{ii}S_{22,11}) + \frac{s_{ki}}{s_{kk}}\hat{a_1} s_{kk} \\ &= s_{ii} - \frac{2}{S_{22,11}}S_{22} - 2s_{ii} - \frac{S_{22}}{S_{22,11}} + s_{ii} \\ &= \frac{S_{22}}{S_{22,11}}, \end{aligned} \tag{2.12}$$

$$s_{vv} = \frac{S_{11}}{S_{22,11}}, \tag{2.13}$$

$$\begin{aligned}
s_{uv} &= \frac{1}{n}\sum_{t=1}^{n}(y_u(t) - \overline{y_u})(y_v(t) - \overline{y_v}) \\
&= \frac{1}{n}\sum_{t=1}^{n}((y_i(t) - \overline{y_i}) - (\hat{a_1}(y_k(t) - \overline{y_k})))((y_j(t) - \overline{y_j}) - (\hat{b_1}(y_k(t) - \overline{y_k}))) \\
&= \frac{1}{n}\sum_{t=1}^{n}((y_i(t) - \overline{y_i})(y_j(t) - \overline{y_j}) - \hat{a_1}((y_j(t) - \overline{y_j}) - \hat{b_1}(y_k(t) - \overline{y_k})) - \cdots \\
&\quad \hat{b_1}(y_i(t) - \overline{y_i})(y_k(t) - \overline{y_k}) + \hat{a_1}\hat{b_1}(y_k(t) - \overline{y_k})^2) \\
&= s_{ij} - \hat{a_1}s_{jk} - \hat{b_1}s_{ik} + \hat{a_1}\hat{b_1}s_{kk} \\
&= s_{ij} - \hat{a_1}s_{jk} - \frac{s_{kj}}{s_{kk}}s_{ik} + \hat{a_1}\hat{b_1}s_{kk} \\
&= s_{ij} - 2\hat{a_1}s_{jk} + \hat{a_1}\hat{b_1}s_{kk} \\
&= s_{ij} - \frac{2}{S_{22,11}}(s_{ij}S_{22,11} - s_{jk}S_{22,13} + s_{ij}S_{22,11}) + \hat{a_1}s_{kj} \\
&= s_{ij} - \frac{2}{S_{22,11}}S_{12} - 2s_{ij} + \frac{S_{12}}{S_{22,11}} + s_{ij} \\
&= -\frac{S_{12}}{S_{22,11}}.
\end{aligned} \tag{2.14}$$

From Eqs. (2.12),(2.13),(2.14), the partial correlation coefficient is rewritten as follows:

$$\begin{aligned}
\gamma_{uv} &= \frac{s_{uv}}{\sqrt{s_{uu}s_{vv}}} \\
&= \frac{-\frac{S_{12}}{S_{22,11}}}{\sqrt{\frac{S_{22}}{S_{22,11}}\frac{S_{11}}{S_{22,11}}}} \\
&= -\frac{S_{12}}{\sqrt{S_{11}S_{22}}} \\
&= -\frac{s_{ij}s_{kk} - s_{ik}s_{jk}}{\sqrt{(s_{jj}s_{kk} - s_{jk}^2)(s_{ii}s_{kk} - s_{ik}^2)}} \\
&= -\frac{(s_{ij}s_{kk} - s_{ik}s_{jk})/\sqrt{s_{ii}s_{jj}s_{kk}^2}}{\sqrt{(s_{jj}s_{kk} - s_{jk}^2)(s_{ii}s_{kk} - s_{ik}^2)}/\sqrt{s_{ii}s_{jj}s_{kk}^2}} \\
&= -\frac{(s_{ij}/\sqrt{s_{ii}s_{jj}}) - (s_{ik}/\sqrt{s_{ii}s_{kk}})(s_{jk}/\sqrt{s_{jj}s_{kk}})}{\sqrt{(1 - s_{jk}^2/s_{jj}s_{kk})(1 - s_{ik}^2/s_{ii}s_{kk})}} \\
&= -\frac{V_{12}}{\sqrt{V_{11}V_{22}}} \\
&= -\frac{\sigma_{12}}{\sqrt{\sigma_{11}\sigma_{22}}}, \qquad (\frac{V_{ij}}{|V|} = \sigma_{ij})
\end{aligned}$$

where σ_{ij} is the (i, j)th element of inverse matrix of the correlation matrix V which is defined as follows:

$$V = \begin{vmatrix} r_{ii} & r_{ji} & r_{ki} \\ r_{ij} & r_{jj} & r_{kj} \\ r_{ik} & r_{jk} & r_{kk} \end{vmatrix}. \tag{2.15}$$

The partial correlation coefficient measures the degree of association between two time series with removing spurious linear influences. However, in the real systems, it is often the case that the influences are not only linear but also nonlinear. In this section, we proposed a nonlinear partial correlation coefficient which removes spurious nonlinear influences. We derived the nonlinear partial correlation coefficient from a multivariable nonlinear regression model.

We defined the nonlinear multivariable regression model of $y_i(t)$ with a set of $2(n-2)$ controlling variables $\mathbf{Y} = \{y_k(t), y_k^2(t)\}, (k \neq i \neq j, k = 1, 2, \ldots, n)$ and that of $y_j(t)$ with a set of $2(n-2)$

controlling variables as follows:

$$y_i(t) = a_0 + \sum_{\substack{k=1 \\ k\neq i\neq j}}^{n} (a_{2k-1}y_k(t) + a_{2k}y_k^2(t)) + e_i(t), \tag{2.16}$$

$$y_j(t) = b_0 + \sum_{\substack{k=1 \\ k\neq i\neq j}}^{n} (b_{2k-1}y_k(t) + b_{2k}y_k^2(t)) + e_j(t), \tag{2.17}$$

where $e_i(t)$ and $e_j(t)$ are residuals.

The averaged y_i is described by $\bar{y}_i = \frac{1}{T}\sum_{t=1}^{T} y_i(t)$, and the covariance between $y_i(t)$ and $y_j(t)$ is described by $s_{ij} = \frac{1}{T}\sum_{t=1}^{T}(y_i(t) - \bar{y}_i)(y_j(t) - \bar{y}_j)$. To estimate the unknown parameters a_k and b_k of multivariable regression model, we use the method of least squares. The sum of squares to be minimized are

$$\sum_{t=1}^{T} e_i^2(t) = \sum_{t=1}^{T}\{y_i(t) - (a_0 + \sum_{\substack{k=1 \\ k\neq i\neq j}}^{n} (a_{2k-1}y_k(t) + a_{2k}y_k^2(t)))\}^2, \tag{2.18}$$

$$\sum_{t=1}^{T} e_j^2(t) = \sum_{t=1}^{T}\{y_j(t) - (b_0 + \sum_{\substack{k=1 \\ k\neq i\neq j}}^{n} (b_{2k-1}y_k(t) + b_{2k}y_k^2(t)))\}^2. \tag{2.19}$$

Let us denote that $F(a_0, a_1, \cdots, a_{2n}) = \sum_{t=1}^{T}\{y_i(t) - (a_0 + \sum_{\substack{k=1 \\ k\neq i\neq j}}^{n} (a_{2k-1}y_k(t) + a_{2k}y_k^2(t)))\}^2$, then we calculate

$$\frac{\partial F}{\partial a_0} = -2\sum_{t=1}^{T}\{y_i(t) - (a_0 + \sum_{\substack{k=1 \\ k\neq i\neq j}}^{n} (a_{2k-1}y_k(t) + a_{2k}y_k^2(t)))\} = 0, \tag{2.20}$$

$$\frac{\partial F}{\partial a_{2l-1}} = -2\sum_{t=1}^{T}\{y_i(t) - (a_0 + \sum_{\substack{k=1 \\ k\neq i\neq j}}^{n} (a_{2k-1}y_k(t) + a_{2k}y_k^2(t)))\}y_l(t) = 0, \tag{2.21}$$

$$\frac{\partial F}{\partial a_{2l}} = -2\sum_{t=1}^{T}\{y_i(t) - (a_0 + \sum_{\substack{k=1 \\ k\neq i\neq j}}^{n} (a_{2k-1}y_k(t) + a_{2k}y_k^2(t)))\}y_l^2(t) = 0, \tag{2.22}$$

for $l \neq i \neq j, l = 1, 2, \cdots, n$. For Eq.(2.20),

$$\sum_{t=1}^{T} y_i(t) - a_0 T - \sum_{\substack{k=1 \\ k\neq i\neq j}}^{n} (a_{2k-1}\sum_{t=1}^{T} y_k(t) + a_{2k}\sum_{t=1}^{T} y_k^2(t)) = 0, \tag{2.23}$$

$$a_0 = \bar{y}_i - \sum_{\substack{k=1 \\ k\neq i\neq j}}^{n} (a_{2k-1}\bar{y}_k + a_{2k}\bar{y}_k^2). \tag{2.24}$$

We substitute a_0 for Eqs.(2.20), (2.21) and (2.22), and obtain the following equations for $l \neq i \neq j, l = 1, 2, \cdots, n$.

$$\sum_{t=1}^{T}\{(y_i(t) - \bar{y}_i) - \sum_{\substack{k=1 \\ k\neq i\neq j}}^{n}(a_{2k-1}(y_k(t) - \bar{y}_k) + a_{2k}(y_k^2(t) - \bar{y}_k^2))\} = 0, \tag{2.25}$$

$$\sum_{t=1}^{T}\{(y_i(t) - \bar{y}_i) - \sum_{\substack{k=1 \\ k\neq i\neq j}}^{n}(a_{2k-1}(y_k(t) - \bar{y}_k) + a_{2k}(y_k^2(t) - \bar{y}_k^2))\}y_l(t) = 0, \tag{2.26}$$

$$\sum_{t=1}^{T}\{(y_i(t) - \bar{y}_i) - \sum_{\substack{k=1 \\ k\neq i\neq j}}^{n}(a_{2k-1}(y_k(t) - \bar{y}_k) + a_{2k}(y_k^2(t) - \bar{y}_k^2))\}y_l^2(t) = 0. \tag{2.27}$$

Subtracting Eq.(2.25) times $\bar{y}_l$ from Eq.(2.25) for $l \neq i \neq j, l = 1, 2, \cdots, n$, we obtain

$$\sum_{\substack{k=1 \\ k\neq i\neq j}}^{n}(a_{2k-1}s_{kl} + a_{2k}s_{k^2l}) = s_{li}. \tag{2.28}$$

In the same way, we also obtain from Eqs. (2.25) and (2.27),

$$\sum_{\substack{k=1 \\ k\neq i\neq j}}^{n}(a_{2k-1}s_{kl^2} + a_{2k}s_{k^2l^2}) = s_{l^2i}, \tag{2.29}$$

for $l = 1, 2, \cdots, n \neq i \neq j$. Here, let us describe the variance-covariance matrix S as

$$S = \begin{bmatrix} s_{11} & s_{11^2} & \cdots & s_{1k} & s_{1k^2} & \cdots & s_{1n} & s_{1n^2} \\ s_{1^21} & s_{1^21^2} & \cdots & s_{1^2k} & s_{1^2k^2} & \cdots & s_{1^2n} & s_{1^2n^2} \\ \vdots & \cdots & \cdots & \cdots & \cdots & \cdots & \vdots & \vdots \\ s_{k1} & s_{k1^2} & \cdots & s_{kk} & s_{kk^2} & \cdots & s_{kn} & s_{kn^2} \\ s_{k^21} & s_{k^21^2} & \cdots & s_{k^2k} & s_{k^2k^2} & \cdots & s_{k^2n} & s_{k^2n^2} \\ \vdots & \cdots & \cdots & \cdots & \cdots & \cdots & \vdots & \vdots \\ s_{n1} & s_{n1^2} & \cdots & s_{nk} & s_{nk^2} & \cdots & s_{nn} & s_{nn^2} \\ s_{n^21} & s_{n^21^2} & \cdots & s_{n^2k} & s_{n^2k^2} & \cdots & s_{n^2n} & s_{n^2n^2} \end{bmatrix}. \tag{2.30}$$

Then Eqs. (2.28) and Eq. (2.29) are solved with the Cramer formula:

$$\hat{a}_l = \frac{\begin{vmatrix} S_{11} & S_{11^2} & \cdots & S_{1i} & \cdots & S_{1n} & S_{1n^2} \\ S_{1^2 1} & S_{1^2 1^2} & \cdots & S_{1^2 i} & \cdots & S_{1^2 n} & S_{1^2 n^2} \\ \vdots & \cdots & \cdots & \cdots & \cdots & \vdots & \vdots \\ S_{l1} & S_{l1^2} & \cdots & S_{ki} & \cdots & S_{ln} & S_{ln^2} \\ S_{l^2 1} & S_{l^2 1^2} & \cdots & S_{k^2 i} & \cdots & S_{l^2 n} & S_{l^2 n^2} \\ \vdots & \cdots & \cdots & \cdots & \cdots & \vdots & \vdots \\ S_{n1} & S_{n1^2} & \cdots & S_{ni} & \cdots & S_{nn} & S_{nn^2} \\ S_{n^2 1} & S_{n^2 1^2} & \cdots & S_{n^2 i} & \cdots & S_{n^2 n} & S_{n^2 n^2} \end{vmatrix}}{\begin{vmatrix} S_{11} & S_{11^2} & \cdots & S_{1k} & S_{1k^2} & \cdots & S_{1n} & S_{1n^2} \\ S_{1^2 1} & S_{1^2 1^2} & \cdots & S_{1^2 k} & S_{1^2 k^2} & \cdots & S_{1^2 n} & S_{1^2 n^2} \\ \vdots & \cdots & \cdots & \cdots & \cdots & \cdots & \vdots & \vdots \\ S_{k1} & S_{k1^2} & \cdots & S_{kk} & S_{kk^2} & \cdots & S_{kn} & S_{kn^2} \\ S_{k^2 1} & S_{k^2 1^2} & \cdots & S_{k^2 k} & S_{k^2 k^2} & \cdots & S_{k^2 n} & S_{k^2 n^2} \\ \vdots & \cdots & \cdots & \cdots & \cdots & \cdots & \vdots & \vdots \\ S_{n1} & S_{n1^2} & \cdots & S_{nk} & S_{nk^2} & \cdots & S_{nn} & S_{nn^2} \\ S_{n^2 1} & S_{n^2 1^2} & \cdots & S_{n^2 k} & S_{n^2 k^2} & \cdots & S_{n^2 n} & S_{n^2 n^2} \end{vmatrix}}$$

$$= \frac{S_{(2j-2)(2j-2),(2i-1)l}}{S_{(2j-2)(2j-2),(2i-1)(2i-1)}}, \tag{2.31}$$

where $S_{ij,kl}$ is the (k, l)th-cofactor matrix of the (i, j) cofactor matrix of S.

In the same way, we obtain the following equation:

$$\hat{b}_l = \frac{\begin{vmatrix} s_{11} & s_{11^2} & \cdots & s_{1j} & \cdots & s_{1n} & s_{1n^2} \\ s_{1^2 1} & s_{1^2 1^2} & \cdots & s_{1^2 j} & \cdots & s_{1^2 n} & s_{1^2 n^2} \\ \vdots & \cdots & \cdots & \cdots & \cdots & \vdots & \vdots \\ s_{l1} & s_{l1^2} & \cdots & s_{kj} & \cdots & s_{ln} & s_{ln^2} \\ s_{l^2 1} & s_{l^2 1^2} & \cdots & s_{k^2 j} & \cdots & s_{l^2 n} & s_{l^2 n^2} \\ \vdots & \cdots & \cdots & \cdots & \cdots & \vdots & \vdots \\ s_{n1} & s_{n1^2} & \cdots & s_{nj} & \cdots & s_{nn} & s_{nn^2} \\ s_{n^2 1} & s_{n^2 1^2} & \cdots & s_{n^2 j} & \cdots & s_{n^2 n} & s_{n^2 n^2} \end{vmatrix}}{\begin{vmatrix} s_{11} & s_{11^2} & \cdots & s_{1k} & s_{1k^2} & \cdots & s_{1n} & s_{1n^2} \\ s_{1^2 1} & s_{1^2 1^2} & \cdots & s_{1^2 k} & s_{1^2 k^2} & \cdots & s_{1^2 n} & s_{1^2 n^2} \\ \vdots & \cdots & \cdots & \cdots & \cdots & \cdots & \vdots & \vdots \\ s_{k1} & s_{k1^2} & \cdots & s_{kk} & s_{kk^2} & \cdots & s_{kn} & s_{kn^2} \\ s_{k^2 1} & s_{k^2 1^2} & \cdots & s_{k^2 k} & s_{k^2 k^2} & \cdots & s_{k^2 n} & s_{k^2 n^2} \\ \vdots & \cdots & \cdots & \cdots & \cdots & \cdots & \vdots & \vdots \\ s_{n1} & s_{n1^2} & \cdots & s_{nk} & s_{nk^2} & \cdots & s_{nn} & s_{nn^2} \\ s_{n^2 1} & s_{n^2 1^2} & \cdots & s_{n^2 k} & s_{n^2 k^2} & \cdots & s_{n^2 n} & s_{n^2 n^2} \end{vmatrix}}$$

$$= -\frac{S_{(2i-1)(2i-1),(2j-2)l}}{S_{(2i-1)(2i-1),(2j-2)(2j-2)}}. \tag{2.32}$$

We substitute $\hat{a}_l$ and $\hat{b}_l$ for Eqs. (2.16) and (2.17) respectively and obtain the following equations,

$$u_i(t) = y_i(t) - (\hat{a}_0 + \sum_{\substack{k=1 \\ k\neq i\neq j}}^{n} (\hat{a}_{2k-1} y_k(t) + \hat{a}_{2k} y_k^2(t))), \tag{2.33}$$

$$u_j(t) = y_j(t) - (\hat{b}_0 + \sum_{\substack{k=1 \\ k\neq i\neq j}}^{n} (\hat{b}_{2k-1} y_k(t) + \hat{b}_{2k} y_k^2(t))). \tag{2.34}$$

The variance of the residual $u_i(t)$ and covariance between the residuals $u_i(t)$ and $u_j(t)$ are denoted as

$$v_{ii} = \frac{1}{T}\sum_{t=1}^{T}(u_i(t) - \overline{u}_i)^2, \tag{2.35}$$

$$v_{jj} = \frac{1}{T}\sum_{t=1}^{T}(u_j(t) - \overline{u}_j)^2, \tag{2.36}$$

$$v_{ij} = \frac{1}{T}\sum_{t=1}^{T}(u_i(t) - \overline{u}_i)(u_j(t) - \overline{u}_j). \tag{2.37}$$

The correlation coefficient between $u_i(t)$ and $u_j(t)$ is defined as

$$Pr_2(i,j) = \frac{v_{ij}}{\sqrt{v_{ii}v_{jj}}}. \tag{2.38}$$

Eq. (2.38) is the nonlinear partial correlation coefficient between $y_i(t)$ and $y_j(t)$ resulting from the nonlinear regression of $y_i(t)$ with $\mathbf{Y}$ and that of $y_j(t)$ with $\mathbf{Y}$. Because

$$\begin{aligned}
v_{ii} &= \frac{1}{T}\sum_{t=1}^{T}(u_i(t)-\bar{u}_i)^2 && (2.39)\\
&= \frac{S_{(2j-2)(2j-2)}}{S_{(2j-2)(2j-2),(2i-1)(2i-1)}}, && (2.40)\\
v_{jj} &= \frac{1}{T}\sum_{t=1}^{T}(u_j(t)-\bar{u}_j)^2 && (2.41)\\
&= \frac{S_{(2i-1)(2i-1)}}{S_{(2i-1)(2i-1),(2j-2)(2j-2)}}, && (2.42)\\
v_{ij} &= \frac{1}{T}\sum_{t=1}^{T}(u_i(t)-\bar{u}_i)(u_j(t)-\bar{u}_j) && (2.43)\\
&= -\frac{S_{(2i-1)(2j-2)}}{S_{(2i-1)(2i-1),(2j-2)(2j-2)}}, && (2.44)
\end{aligned}$$

then we have

$$\begin{aligned}
Pr_2(i,j) &= \frac{v_{ij}}{\sqrt{v_{ii}v_{jj}}} && (2.45)\\
&= \frac{-\frac{S_{(2i-1)(2j-2)}}{S_{(2i-1)(2i-1),(2j-2)(2j-2)}}}{\sqrt{\frac{S_{(2j-2)(2j-2)}}{S_{(2j-2)(2j-2),(2i-1)(2i-1)}}\frac{S_{(2i-1)(2i-1)}}{S_{(2i-1)(2i-1),(2j-2)(2j-2)}}}} && (2.46)\\
&= \frac{-S_{(2i-1)(2j-2)}}{\sqrt{(S_{(2j-2)(2j-2)})(S_{(2i-1)(2i-1)})}} && (2.47)\\
&= \frac{-S_{(2i-1)(2j-2)}/\sqrt{s_{ii}s_{jj}\prod_{k=1,k\neq i\neq j}^{n}s_{kk}^2 s_{k^2k^2}^2}}{\sqrt{(S_{(2j-2)(2j-2)})(S_{(2i-1)(2i-1)})}/\sqrt{s_{ii}s_{jj}\prod_{k=1,k\neq i\neq j}^{n}s_{kk}^2 s_{k^2k^2}^2}} && (2.48)\\
&= \frac{-R_{(2i-1)(2j-2)}}{\sqrt{(R_{(2j-2)(2j-2)})(R_{(2i-1)(2i-1)})}} && (2.49)\\
&= \frac{-\sigma_{(2i-1)(2j-2)}}{\sqrt{\sigma_{(2i-1)(2i-1)}\sigma_{(2j-2)(2j-2)}}}, && (2.50)
\end{aligned}$$

where σ_{ij} is the (i, j)th element of inverse matrix of the correlation matrix of Eq. (2.51).

$$R = \begin{bmatrix}
r_{11} & r_{1^2 1} & \cdots & r_{i1} & r_{(i+1)1} & \cdots & r_{j1} & r_{(j+1)1} & \cdots & r_{n1} & r_{n^2 1} \\
r_{11^2} & r_{1^2 1^2} & \cdots & r_{i1^2} & r_{(i+1)1^2} & \cdots & r_{j1^2} & r_{(j+1)1^2} & \cdots & r_{n1^2} & r_{n^2 1^2} \\
\vdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \vdots & \vdots \\
r_{1i} & r_{1^2 i} & \cdots & r_{ii} & r_{(i+1)i} & \cdots & r_{ji} & r_{(j+1)i} & \cdot & r_{ni} & r_{n^2 i} \\
r_{1(i+1)} & r_{1^2(i+1)} & \cdots & r_{i(i+1)} & r_{(i+1)(i+1)} & \cdots & r_{j(i+1)} & r_{(j+1)(i+1)} & \cdots & r_{n(i+1)} & r_{n^2(i+1)} \\
\vdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \vdots & \vdots \\
r_{1j} & r_{1^2 j} & \cdots & r_{ij} & r_{(i+1)j} & \cdots & r_{jj} & r_{(j+1)j} & \cdot & r_{nj} & r_{n^2 j} \\
r_{1(j+1)} & r_{1^2(j+1)} & \cdots & r_{i(j+1)} & r_{(i+1)(j+1)} & \cdots & r_{j(j+1)} & r_{(j+1)(j+1)} & \cdots & r_{n(j+1)} & r_{n^2(j+1)} \\
\vdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \vdots & \vdots \\
r_{1n} & r_{1^2 n} & \cdots & r_{in} & r_{(i+1)n} & \cdots & r_{jn} & r_{(j+1)n} & \cdots & r_{nn} & r_{n^2 n} \\
r_{1n^2} & r_{1^2 n^2} & \cdots & r_{in^2} & r_{(i+1)n^2} & \cdots & r_{jn^2} & r_{(j+1)n^2} & \cdots & r_{nn^2} & r_{n^2 n^2}
\end{bmatrix} \tag{2.51}$$

2.2.2 Partial Mutual Information

The mutual information can measure nonlinear dependencies between two random variables. Let us define a random variable X with probabilities p_x of outcomes x. Then entropy is defined by $H(X) = -\sum_x p_x \ln p_x$. The mutual information between two random variables X and Y is described by:

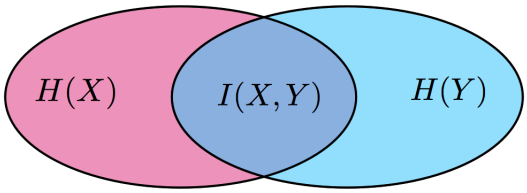
$$I(X, Y) = H(X) + H(Y) - H(X, Y) \tag{2.52}$$

where $H(X, Y)$ is the joint entropy between X and Y, defined by $H(X, Y) = -\sum_{xy} p_{xy} \ln p_{xy}$. The mutual information is non-negative ($I(X, Y) \geq 0$), and bounded $0 \leq I(X, Y) \leq \min\{H(X), H(Y)\}$. If X and Y are independent, the mutual information takes zero.

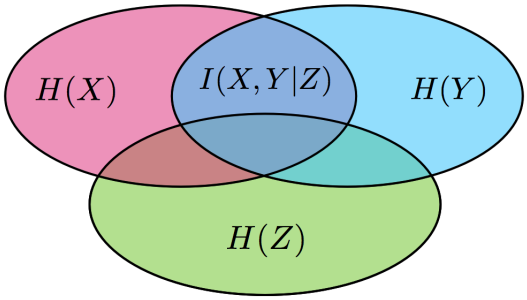
However, the mutual information can be spuriously biased if two elements are indirectly connected. Then, partial mutual information is effective to remove such spurious bias [4]. The part of $I(X, Y)$ that is not in Z is the partial mutual information $I(X, Y|Z)$ (Fig. 2.2(b)) given by:

$$I(X, Y|Z) = H(X, Z) + H(Y, Z) - H(X, Y, Z) - H(Z). \tag{2.53}$$

This partial mutual information can measure the mutual dependence of the two time series with removing spurious bias.



(a)



(b)

Figure 2.2: Concepts of (a) the mutual information and (b) the partial mutual information. The partial mutual information $I(X, Y|Z)$ selects the part of mutual information $I(X, Y)$ which is not in Z.

2.2.3 Partial Directed Coherence

The partial directed coherence (PDC) [6] is a frequency domain counterpart of the Granger causality [38] which is based on a multiple linear regression model.

The PDC [6] is defined by

$$|\bar{\pi}_{ij}(f)| = \left| \frac{\bar{A}_{ij}(f)}{\sqrt{\bar{\boldsymbol{a}}_j(f)^H \bar{\boldsymbol{a}}_j(f)}} \right|, \tag{2.54}$$

where

$$\bar{A}_{ij}(f) = \begin{cases} 1 - a_{ij}\mathrm{e}^{-i2\pi fr}, & \text{if } \; i = j; \\ -a_{ij}\mathrm{e}^{-i2\pi fr}, & \text{otherwise;} \end{cases} \tag{2.55}$$

and $\bar{\boldsymbol{a}}_j(f) = (\bar{A}_{j1}, \bar{A}_{j2}, \ldots, \bar{A}_{jN})^{\mathrm{T}}$, and H stands for the Hermitian transpose.

The PDC takes values between 0 and 1. The PDC can quantify a similarity between the two time series with removing spurious correlations between them, for example, in a case that two elements are driven by a common input from another elements.

Chapter 3

Estimation Method of Network Structure by Transforming Point Processes into Continuous Time Series

To estimate network structures from point processes, we proposed methods of transforming point processes into continuous time series in this chapter.

First, we proposed a method for reconstructing hidden input information applied to neurons by transforming point processes into continuous time series. Next, we proposed estimation methods of network structures by the transformation method from simple point processes into continuous time series. Third, we also proposed methods of estimating network structures from marked point processes. In this method, we proposed a method of transforming marked point processes into continuous time series.

3.1 Reconstruction of Input Information from Point Process

Neurons code input information and generate spike trains. The generated spike trains reflect the input information. To understand how the information is coded by the neurons, it is an important issue to reconstruct input information from spike trains.

Some methods for reconstructing input information from spike sequences are proposed. Sauer proved that there is one-to-one correspondence between the system states and interspike interval (ISI) vectors of sufficiently large dimension, and show that attractors of the original dynamical systems can be reconstructed only from an output spike train of an integrate-and-fire neuron model [26, 27]. Suzuki et al. examined fundamental characteristics of ISI reconstruction with a leaky integrated-and-fire neuron model, and applied ISI reconstruction to output of cricket wind receptor cells[28]. Although these study are approaches to reconstruct input information from spike trains which use information of ISI directly, Hori et al. proposed a method for reconstructing input information from spike trains as a different approach [31]. In the Ref. [31], the method estimate continuous input information from interspike intervals by using a sampling function. This method can reconstruct input information with broader case than the Sauer's method. However, they assumed that a membrane time constant and a threshold of a neuron are known. Then it is difficult to reconstruct input information only from spike trains.

Then, in this section, we proposed a method for reconstructing hidden input information applied to neurons. The information we used is only observed spike sequences. In the proposed method, we used a kernel density function to transform the observed spike trains to an instantaneous mean-firing-frequency time-series. Furthermore, we proposed selection methods of an optimal kernel bandwidth. Then we reconstructed the hidden input information through this transformation method. Consequently, we found that high correlation coefficients and low normalized mean square errors are evaluated between the input time-series and its reconstructed time-series by the proposed method.

3.1.1 Methods

Transformation from a spike sequence to a continuous time series

To transform an observed spike train to a continuous time series, we used a kernel density function. Using the kernel function, we can transform from a spike train to an instantaneous mean-

Figure 3.1: Transformation from a spike train to an instantaneous mean-firing-rate time series using a kernel function.

firing-rate time series (Fig. 3.1). Let us define a spike train as $t_i, (t_1 < t_2 < \cdots < t_N)$. The transformed time series $f(t)$ is described by the following equation:

$$f(t) = \frac{1}{NT} \sum_{i=1}^{N} k(\frac{t - t_i}{T}), \tag{3.1}$$

where $k(\cdot)$ is a kernel function and T is a bandwidth. As the kernel function, we used the Gaussian function defined as follows:

$$k(t) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{t^2}{2}) \tag{3.2}$$

Selection Method of Optimal Bandwidth

We have to select the optimal bandwidth because the transformed time series depends on the bandwidth. To optimize the kernel bandwidth, we used a kernel bandwidth optimization method which is proposed by Shimazaki et al. [25]. The method is used in spike rate estimation.

Method 1 The optimal kernel bandwidth is obtained by minimizing the mean integrated squared error (MISE) between the underlying rate λ_t and the kernel density estimate $\hat{\lambda}_t$. The MISE is described by the following equation:

$$\text{MISE} = \int_a^b E[(\hat{\lambda}_t - \lambda_t)^2] dt, \tag{3.3}$$

where $E[\dot{]}$ is ensemble mean and interval of spike train is $[a, b]$. A cost function is defined by subtracting the term of the MISE does not depend on the choice of a kernel and assumption

of Poisson point process as follows:

$$C(w) = \sum_{i,j} \psi_w(t_i, t_j) - 2 \sum_{i \neq j} k_w(t_i - t_j), \tag{3.4}$$

where

$$\psi_w(t_i, t_j) = \int_a^b k_w(t - t_i) k_w(t - t_j) dt.$$

The optimal fixed kernel bandwidth is obtained as a minimizer of the cost function.

Method 2 Attractor reconstruction[34] is a basic technique in chaotic time series analysis. In the attractor reconstruction, mth dimensional state space is constructed for a time series $x(t)$ as follows:

$$\mathbf{v}(t) = \{x(t), x(t + \tau), \cdots, x(t + (m - 1)\tau)\}, \tag{3.5}$$

where τ is delay.

It is known fact that states of finite dimensional dynamical systems correspond in a one-to-one manner with delay-coordinate vectors [29, 30].

To select an optimal time delay, the selection method using mutual information between $x(t)$ and $x(t+\tau)$ is used [35]. In this method, the time delay that produces the first local minimum of mutual information between $x(t)$ and $x(t + \tau)$ is selected as an optimal time delay.

3.1.2 Numerical Simulation

To evaluate our method, we used a leaky integrate-and-fire (LIF) model [34]. The LIF model is described by following equation:

$$\dot{v} = -\frac{v}{\tau_0} + I(t), \tag{3.6}$$

where v is the membrane potential, τ_0 is a membrane time constant, and $I(t)$ is the input at time t. After the membrane potential v reaches a threshold θ, the membrane potential v is reset to v_0. In the simulations, we set $\theta = 1, v_0 = 0$, and $\tau_0 = 5$[ms].

In the simulations, we used periodic time series, quasi-periodic time series, chaotic time series, and real time series of Japanese vowel sound /a/ as input time series $I(t)$.

In the simulations, we normalized time series data to a range between 0 and 1. In accordance with this, we set $\theta = 1$ and $v_0 = 0$.

Case 1 Periodic Time Series

As the periodic time series, we used the sinusoidal function which is described by:

$$I(t) = \frac{1}{2}\left(\sin(2\pi f t) + 1\right). \tag{3.7}$$

We set the frequency $f = 100$[Hz].

Case 2 Quasi-Periodic Time Series

As the quasi-periodic time series, we set $I(t)$ to the variable x of the Langford systems [36] which are described by the following equations:

$$\begin{cases} \dot{x} &= (z-\beta)x - \omega y, \\ \dot{y} &= \omega x + (z-\beta)y, \\ \dot{z} &= \lambda + \alpha z - \dfrac{x^3}{3} - (x^2+y^2)(1.0+\rho z) + \epsilon z x^3. \end{cases} \tag{3.8}$$

We set the parameters $\alpha = 1, \beta = 0.7, \lambda = 0.6, \omega = 3.5, \rho = 0.25$, and $\epsilon = 0$.

Case 3 Chaotic time series

We set the input time series $I(t)$ to the variable x of the Lorenz systems [37] as the chaotic time series. The Lorenz equations are described by the following equations:

$$\begin{cases} \dot{x} &= -\sigma x + \sigma y, \\ \dot{y} &= -xz + rx - y, \\ \dot{z} &= xy - bz. \end{cases} \tag{3.9}$$

We set the parameters $\sigma = 10, b = 8/3$, and $r = 28$. These parameters lead to chaotic behavior of the systems.

To calculate Eqs. (3.8), (3.9), we applied a fourth order Runge-Kutta method with integration step 0.001. We used $100,000$ data after a period of transition.

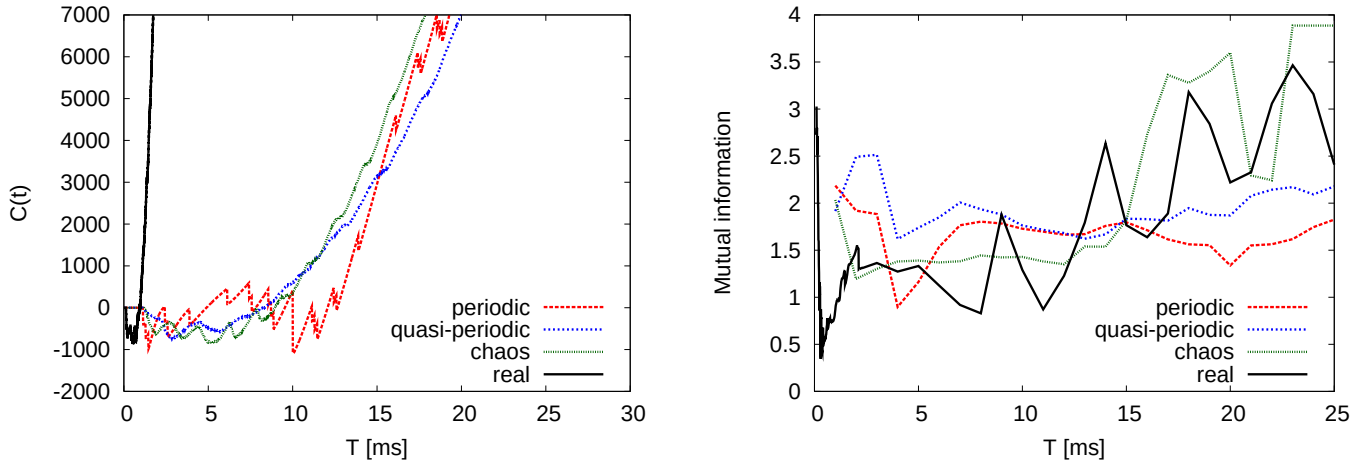
Case 4 Real time series

The above mentioned time series are generated from numerical models. Then we set $I(t)$ to a time series of Japanese vowel sounds /a/ (a male subject, sampled at 44[kHz]) as a real time series.

3.1.3 Results

We show the results of the bandwidths calculated by the methods 1 and 2 for each input time series in Fig. 3.2. Fig. 3.2(a) shows results of the costs when the bandwidth is changed in the

method 1. Fig. 3.2(b) shows results of the mutual information when the bandwidth is changed in the method 2. The optimal bandwidth is obtained when the cost $C(t)$ and the mutual information are minimum in the method 1 and the method 2, respectively. From the results, the optimal bandwidths are calculated as shown in Table 3.1.



(a) (b)

Figure 3.2: (a) Costs and (b) mutual information when the bandwidths are changed.

Table 3.1: The optimal bandwidths calculated by the methods 1 and 2.

	method 1	method 2
periodic	10.02 [ms]	3.80[ms]
quasi-periodic	2.85 [ms]	4.81 [ms]
chaos	5.27 [ms]	3.92[ms]
real	0.51 [ms]	0.36 [ms]

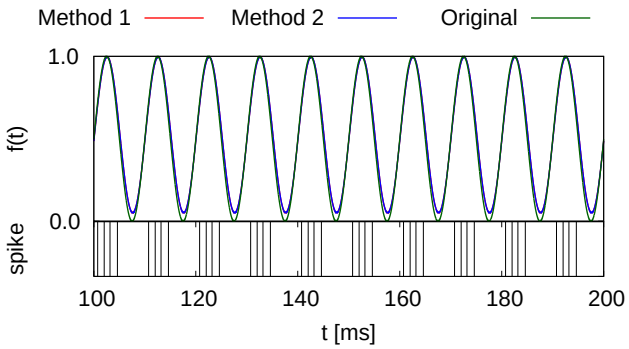
Next, in Fig. 3.3, we show results of transformed time series from each input time series using the obtained bandwidth. In Fig. 3.3, upper figures show input time series and transformed time series, and bottom figures show spike trains obtained from the LIF model. In the case of periodic time series (Fig. 3.3(a)), we can see that the input time series and the transformed time series with

the method 1 and 2 are overlapped. However, in the case of the quasi-periodic time series (Fig. 3.3(b)), the transformed time series are not similar to the input time series because the bandwidths are fixed. On the other hand, in the case of the chaotic time series and real time series (Fig. 3.3(c), (d)), the transformed time series are similar to the input time series in spite of complicated behavior. To evaluate the calculated bandwidth quantitatively, we investigated correlation coefficients and normalized mean square error (NMSE) between the input time series and transformed time series when the bandwidth T is changed (Fig. 3.4). From the results. the correlation coefficients are low and the NMSEs are large when the bandwidth is too small or too large However, when the bandwidths are calculated by the method 1 and method 2, the correlation coefficients are close to unity, and the NMSEs are close to zero.

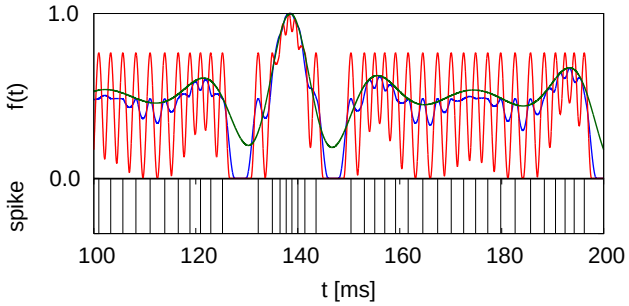
We also show correlation diagrams between the input time series and the transformed time series and the correlation coefficients in Fig. 3.5. From the results, we can identify that positive correlations exist.

In addition, we compared reconstructed attractors in delay coordinates [29, 30, 34] from the input time series with these of the transformed time series (Fig.3.6). In Fig. 3.6(a), (b), we set time delay $\tau = 1.0$[ms]. In Fig. 3.6(c), we set the time delay $\tau = 1.9$[ms]. In Fig. 3.6(d), we set the time delay $\tau = 1.2$[ms].

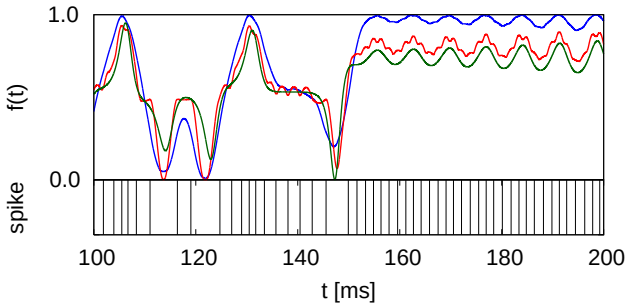
On the other hand, the results of quasi-periodic time series are low accuracy. In the simulations, we set the low firing rates. We confirmed that, if we set the high firing rates, we can reconstruct the input time series with high accuracy.



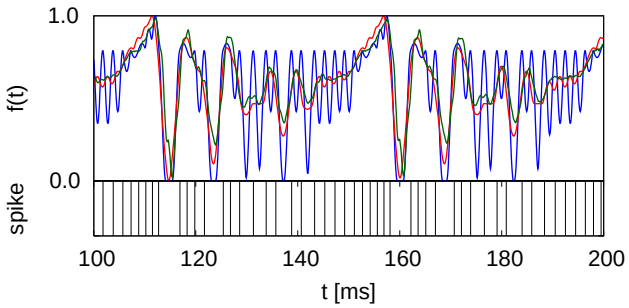
(a) periodic



(b) quasi-periodic

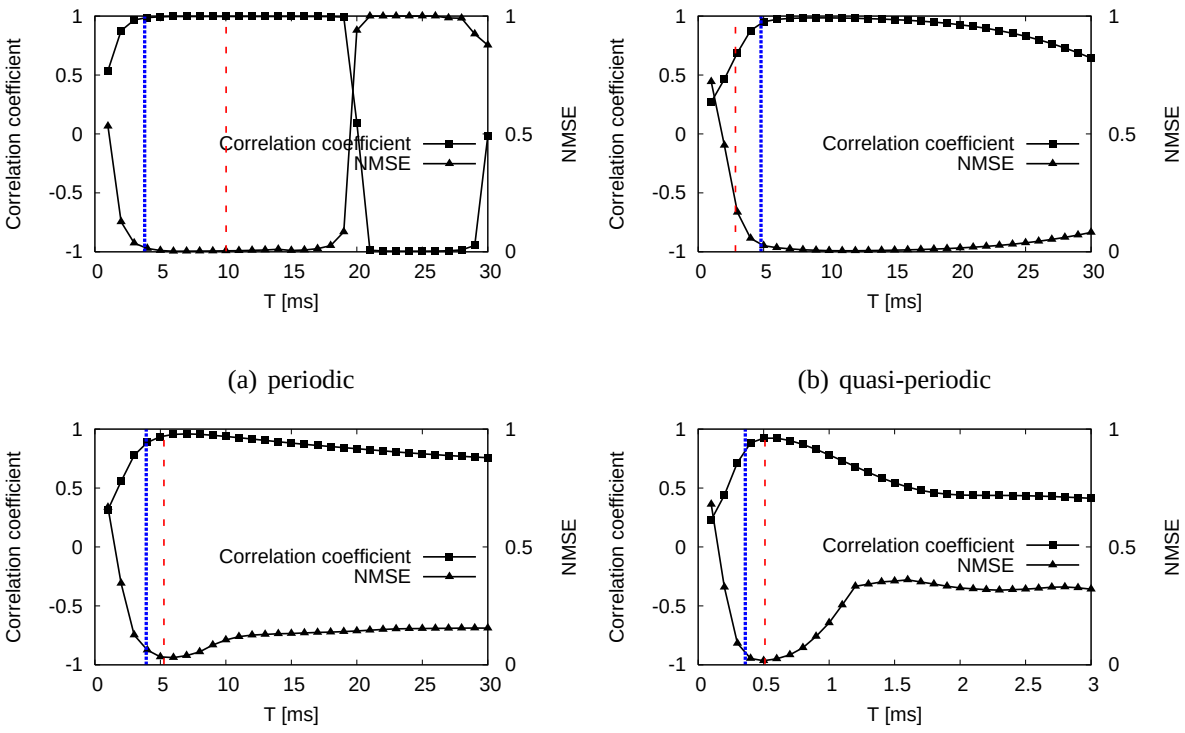


(c) chaos



(d) real

Figure 3.3: Input time series and transformed time series by the methods 1 and 2. Bottom figures show spike trains obtained from the LIF model.



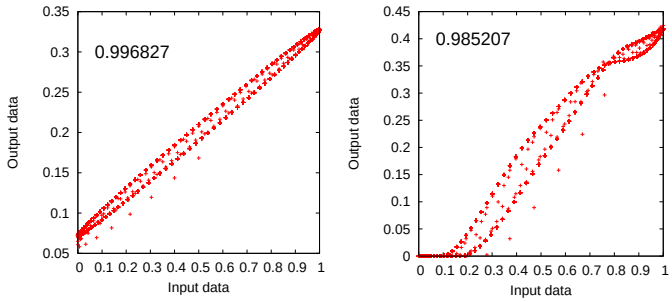
(a) periodic

(b) quasi-periodic

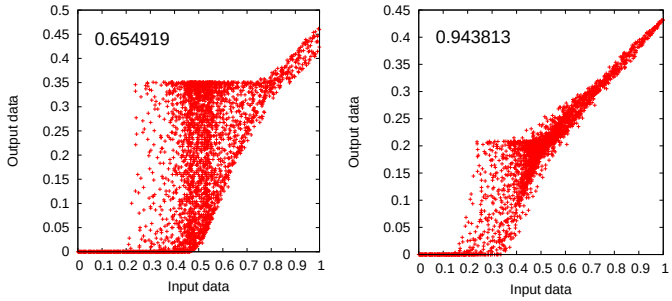
(c) chaos

(d) real

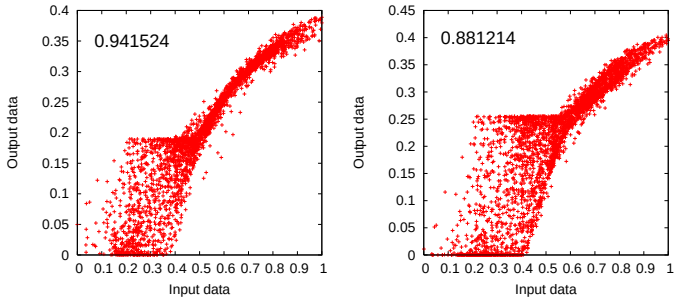
Figure 3.4: Cross correlation coefficients and normalized mean square errors between input time series and transformed time series for the band width T. The red line shows the bandwidth calculated by the method 1, and the blue line shows the bandwidth calculated by the method 2.



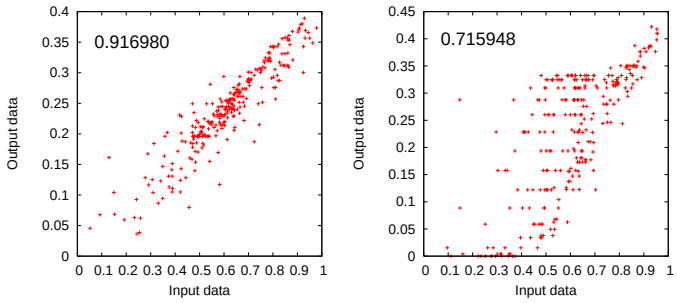
(a)periodic



(b)quasi-periodic

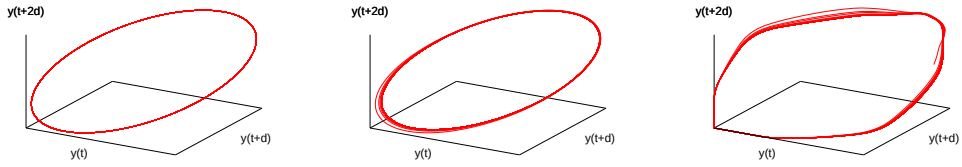


(c)chaos

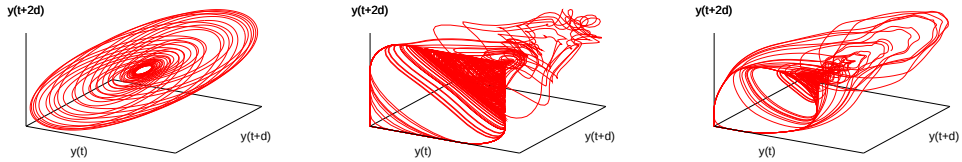


(d)real

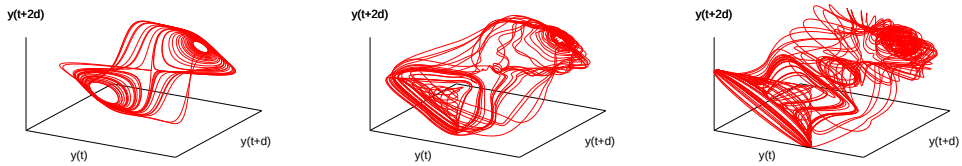
Figure 3.5: Correlation diagrams between input time series and transformed time series by the method 1 (left) and the method 2 (right) with $\tau_0 = 5.0$[ms].



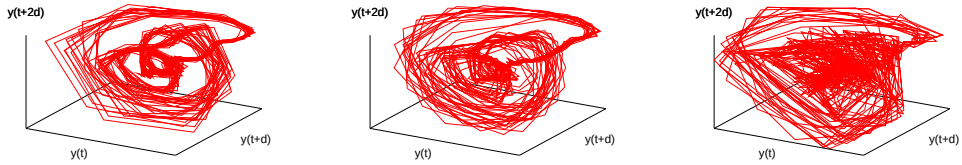
(a) periodic



(b) quasi-periodic



(c) chaotic



(d) real

Figure 3.6: Reconstructed attractors in delay coordinates from input time series (left), transformed time series by the method 1 (middle), and the transformed time series by the method 2 (right) with $\tau_0 = 5.0$.

3.2 Estimation Method of Network Structure from Simple Point Processes

In the previous section, we used the transforming method from a spike sequence to a continuous time series. Using the transforming method from a spike sequence to a continuous time series, we can use the time series analysis method for continuous time series. Then, in this section, we use the transforming method to estimate network structures from spike sequences. We applied two methods to transform spike sequences into continuous time series. The first one is an interpolation of inter-spike intervals by sinusoidal waves. The second one is a kernel density estimator [25] with theoretical optimization of its kernel bandwidth. In addition, we applied the partial correlation analysis to the transformed continuous time series.

3.2.1 Methods of Transforming Point Processes into Continuous Time Series

In order to transform the spike sequences into continuous time series, we applied two transforming methods to the spike sequences.

Let us describe the jth spike in the ith spike sequence as t_j^i. Then, the ith spike sequence is described as $s_i = \{t_1^i, t_2^i, \ldots, t_{l_i}^i\}$ where l_i is the last index of s_i. Firstly, we interpolate the jth segment (inter-spike interval, ISI) bounded by two adjacent spikes, t_j^i and t_{j+1}^i, by the following equation:

$$x_j^i(t) = \frac{1}{2} + \frac{1}{2}\cos\frac{2\pi}{T_j^i}(t - t_j^i),\ (t_j^i \le t \le t_{j+1}^i), \tag{3.10}$$

where $T_j^i = t_{j+1}^i - t_j^i$. Then, $x_j^i(t)$ $(j = 1, 2, \ldots, l_{i-1})$ is concatenated to produce a transformed continuous time series from the ith spike sequence. In Fig. 3.7, a spike sequence and corresponding transformed continuous time series are shown. In method 2, we transformed continuous time series from spike sequences by using the kernel density estimator.

We used the Gaussian function as the kernel function. The kernel density function is described

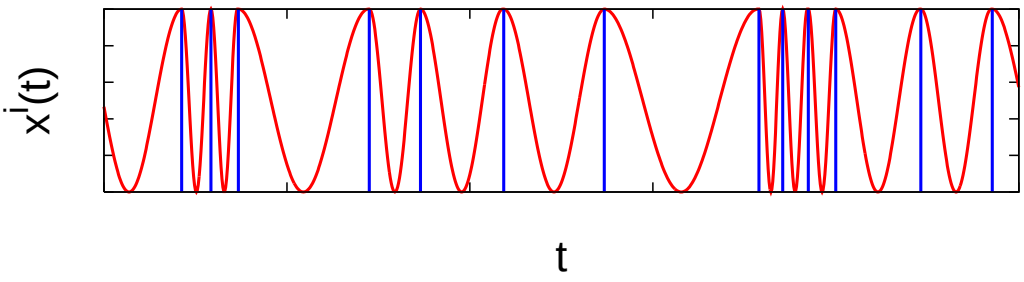


Figure 3.7: Example of transformation of a spike sequence into a continuous time series by the method 1. Blue lines indicate spike timings and the red curve indicates transformed continuous time series.

by the following equation:

$$f_K^i(t) = \frac{1}{l_i w} \sum_{j=1}^{l_i} K\left(\frac{t - t_j^i}{w}\right). \tag{3.11}$$

where w is the band width and $K(\cdot)$ is the kernel function. The Gaussian function with mean zero and unit variance as a kernel function is defined as:

$$K(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}. \tag{3.12}$$

In Fig. 3.8, an observed spike sequence and corresponding transformed continuous time series are shown. Here, the bandwidth w used for the kernel estimation is optimized by the method for selecting a fixed kernel bandwidth [25] (refer to the method 1 in Sec. A.2).

3.2.2 Simulations

We examined the validity of the proposed method using a mathematical model. In our numerical experiments, we assumed that we can only observe asynchronous spike sequences while the true network structure is unknown.

In our experiments, to evaluate the validity of the proposed measure, we used the Izhikevich neuron model [14]. The dynamics of the ith neuron in the neural network is described by the

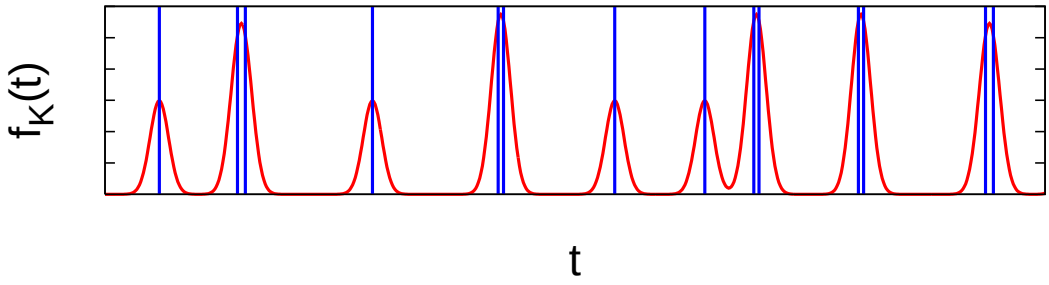


Figure 3.8: Example of transformation of a spike sequence into a continuous time series by the method 2. Blue lines indicate spike timings and the red curve indicates transformed continuous time series.

following equations:

$$\begin{aligned} \frac{dv_i(t)}{dt} &= 0.04v_i^2(t) + 5v_i(t) + 140 - u_i(t) + I_i(t), \\ \frac{du_i(t)}{dt} &= a(bv_i(t) - u_i(t)), \\ \text{if } v_i(t) &\geq 30 \text{ [mV], then} \begin{cases} v_i(t) \leftarrow c, \\ u_i(t) \leftarrow u_i(t) + d, \end{cases} \end{aligned} \tag{3.13}$$

where $v_i(t)$ is the membrane potential, $u_i(t)$ is the membrane recovery variable; and a, b, c, and d are dimensionless parameters. The parameters are set to $a = 0.02$, $b = 0.2$, $c = -65$, and $d = 8$. If $v_i(t) \geq 30$[mV], v_i and u_i are reset according to Eq. (3.13). The variable $I_i(t)$ is the sum of the external and synaptic inputs from coupled neurons which is defined as follows:

$$I_i(t) = \sum_{j=1}^{n} \sum_{k=1}^{m_j} g_{ij} w_{ij} \delta(t - t_j^k - \tau_{ij}) + 5\eta_i(t) \tag{3.14}$$

where g_{ij} is a coupling strength from j to i, τ_{ij} is a delay time between i and j which is randomly set to a value between 1 [ms] to 4 [ms], and w_{ij} is the (i, j)th element of the connection matrix of the network structure. If the neurons are coupled from j to i, w_{ij} takes unity. Otherwise, w_{ij} takes zero. The amplitude of the external inputs is set to 5 times $\eta_i(t)$, where $\eta_i(t)$ is a Gaussian random number with a mean value and standard deviation of zero and unity, respectively.

We generated a complex network structure having a regular ring topology with 100 neurons by a random rewiring of the synaptic connections between neurons in the same manner as that

described in Ref. [15]. We set the parameters k (the number of edges in the regular network) to four and p (rewiring probability) to 0.1.

We conducted numerical experiments according to the following procedures. First, to generate multi-spike sequences, we constructed a neural network using the Izhikevich simple neuron model and applied external inputs to the neural network. Second, we transformed the spike sequences into continuous time series. Third, we calculated the partial correlation coefficient of the transformed continuous time series. If the ith and the jth neurons are coupled, the partial correlation coefficient becomes large. On the other hand, if these neurons are not coupled, it becomes small. Finally, we classify coupled pairs and uncoupled pairs by the Otsu thresholding [17] which is based on a linear discriminant analysis.

To evaluate an overall estimation accuracy, we compared the estimated network structure with the true network structure. For this evaluation, we used the following index:

$$E = \frac{\sum_{i,j}^{N}(\alpha_{ij}\tilde{\alpha}_{ij} + (1-\alpha_{ij})(1-\tilde{\alpha}_{ij}))}{N(N-1)} \tag{3.15}$$

where α_{ij} and $\tilde{\alpha}_{i,j}$ are the directional connectivity of the true network structure and the estimated network structure, respectively. If the ith and jth elements are (estimated to be) coupled, α_{ij} and $\tilde{\alpha}_{ij}$ take a value of unity. If they are not, α_{ij} and $\tilde{\alpha}_{ij}$ take a value of zero. If E approaches unity, the estimation accuracy increases.

3.2.3 Results

To compare the proposed methods which transform spike sequences into continuous time series and the conventional method which uses APSTMC, we also apply the APSTMC [12] to the same network structure.

Figure 3.9 shows histograms of APSTMC and partial correlation coefficients. Although the conventional method and the method 1 can separate coupled and uncoupled pairs of neurons, a few uncoupled pairs are estimated as coupled pairs. However, coupled and uncoupled pairs of neurons are more clearly distinguished by the method 2.

In Fig. 4.7, we show the results when the rewiring probability is changed. In the method 1, the estimation accuracy for the random network ($p = 1.0$) is low. However, the conventional method and the method 2 show high estimation accuracy for all rewiring probabilities.

We examined how the estimation accuracy depends on the temporal epoch for observing spikes

(Fig. 5.6). The conventional method exhibits higher accuracy than other methods when the temporal epoch of the spike sequences is shorter than 20 [s]. However, in the method 2, the estimation accuracy is higher than the other methods when the temporal epoch of the spike sequences is longer than 20 [s].

In Fig. 3.12, we show the results when the coupling strength is changed. If the coupling strength becomes larger than four, the estimation accuracy of the conventional method is high. If the coupling strength becomes larger than five, the estimation accuracy becomes higher in the method 2. If the coupling strength becomes larger than six, the estimation accuracy of the method 1 is the highest.

In addition, we show the results for different network sizes (Fig. 4.9). In the method 1, the estimation accuracy worsens as the network size increases. However, the estimation accuracy in the method 2 and the conventional method are still high even if the network size is large.

From these results, the estimation accuracy in the method 2 is better than the other methods. In this neural network model, we used different dynamics of neurons; regular spiking, intrinsically bursting, and chattering neurons. If two neurons of different dynamics are coupled, the correlation between the two transformed continuous time series decreases in the conventional method [12] and the method 1. However, in the method 2, true relation can be identified if observed spike sequences of two neurons are transformed with an optimal bandwidth, even though the dynamics of the two neurons differs (Fig. 3.14).

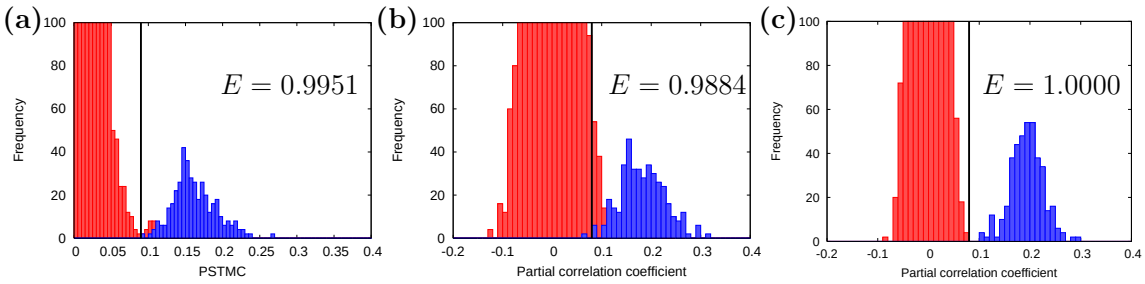


Figure 3.9: Histograms of (a) PSTMC in the conventional method, partial correlation coefficients by (b) the method 1 and (c) the method 2. The number of neurons is 100. The temporal epoch of spike sequences used for estimation is 50 [s]. Histograms of all of the PSTMC and the partial correlation coefficients are indicated in red, and those of the coupled elements are superimposed by blue. The black vertical lines show thresholds decided by the Otsu thresholding. If the PSTMC and the partial correlation coefficients are less than the threshold, corresponding neurons are classified into the uncoupled class.

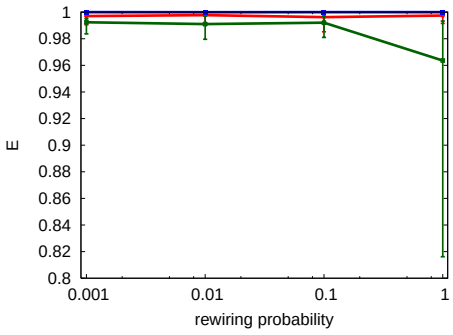


Figure 3.10: Estimation accuracy E of the network structure for rewiring probabilities p. The number of neurons, the temporal epoch and the coupling strength are 100, 50 [s] and 6. The red line indicates the conventional method, the green line indicates the method 1 and the blue line indicates the method 2. Error bars with 20 trials are also provided.

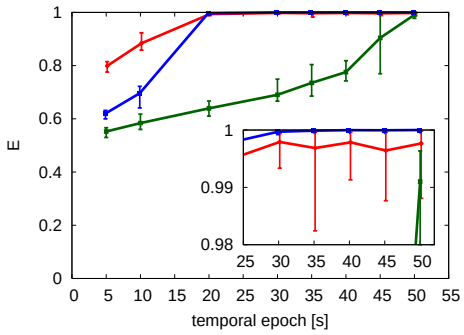


Figure 3.11: Estimation accuracy E of the network structure for several temporal epochs. The number of neurons and the coupling strength are 100 and 6. Red line indicates the conventional method, green line indicates the method 1 and blue line indicates the method 2. Error bars with 20 trials are also provided. The inset shows an enlargement.

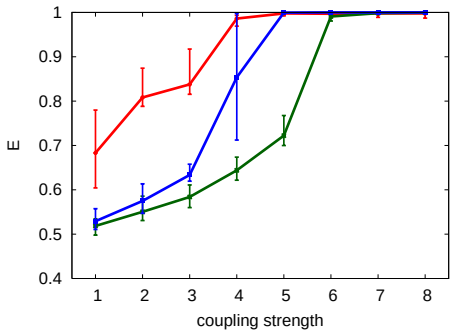


Figure 3.12: Estimation accuracy E of the network structure for several coupling strength. The number of neurons and the temporal epoch are 100 and 50 [s]. Red line indicates the conventional method, green line indicates the method 1 and blue line indicates the method 2. Error bars with 20 trials are also provided.

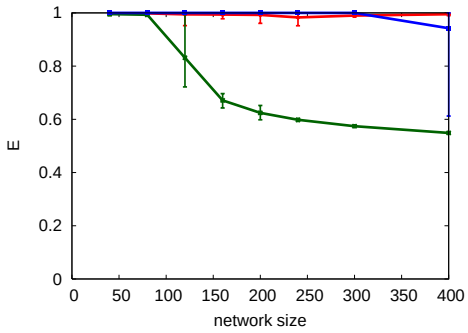
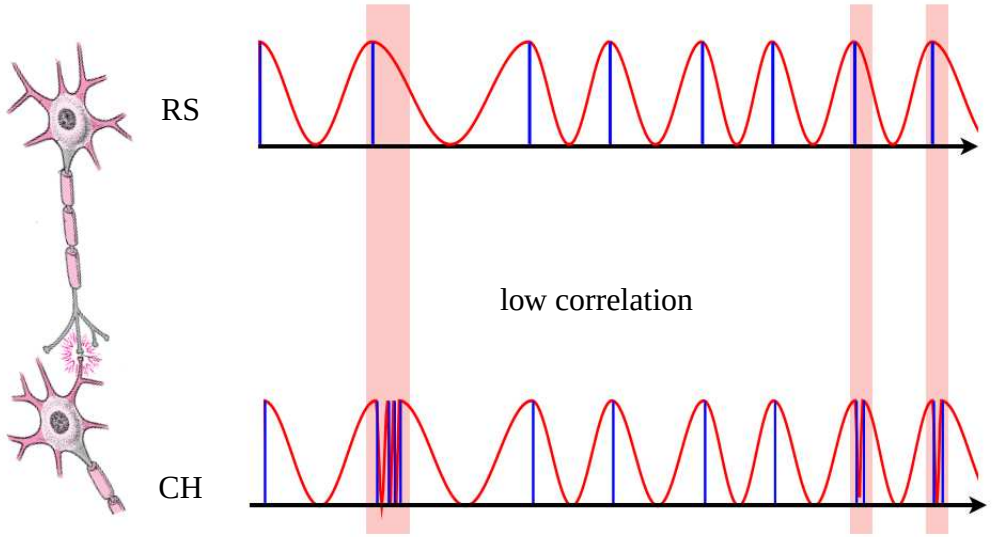
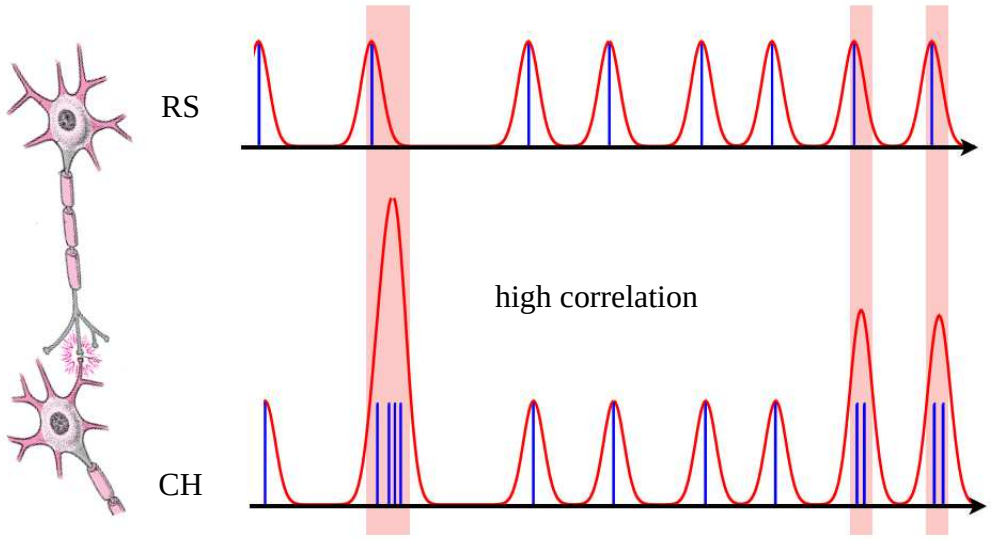


Figure 3.13: Estimation accuracy E of the network structure for several network sizes. The temporal epoch and the coupling strength are 50 [s] and 6. The red line indicates the conventional method, the green line indicates the method 1 and the blue line indicates the method 2. Error bars with 20 trials are also provided.



(a)Method 1



(b)Method 2

Figure 3.14: Observed spike sequences (blue) and transformed times series (red) in case that the regular spiking (RS) neuron and the chattering (CH) neuron are coupling.

3.3 Estimation Method of Network Structures from Marked Point Processes

In the Sec. 3.2, we have treated observed event sequences as a point process, which means that the observed event sequences do not have amplitude information but only event timing. However, the amplitude information can be essential for several phenomena, for example, financial systems and seismic events. Such sequences are often referred as a marked point process. If we use not only the information of the event timing but also the amplitude information, we can estimate more precisely the network structures. In this section, we proposed a method for transforming a marked point process into a continuous time series. In addition, we applied the partialization analysis to the transformed continuous time series and estimated connectivity of nonlinear dynamical systems: the coupled Lorenz systems [37].

3.3.1 Method of Transforming Marked Point Processes into Continuous Time Series

We used a kernel density estimator for transforming an event sequence into a continuous time series in Sec. 4.2. To apply this method to a marked point process, we modified the kernel density estimator by using the amplitude of marks. Let us define the lth event timing of the ith marked point process data as $t_i^l (l = 1, 2, \ldots, N)$, and the amplitude at t_i^l as $h(t_i^l)$. Then, we use the following equations to transform the marked point process data into a continuous time series:

$$f_i(t) = \sum_{\substack{l=1, \\ t-\frac{T}{2} \le t_i^l \le t+\frac{T}{2}}}^{N} K\Big(\frac{t - t_i^l}{T}\Big) h(t_i^l), \quad K(t) = \frac{1}{2}(1 + \cos 2\pi t), \tag{3.16}$$

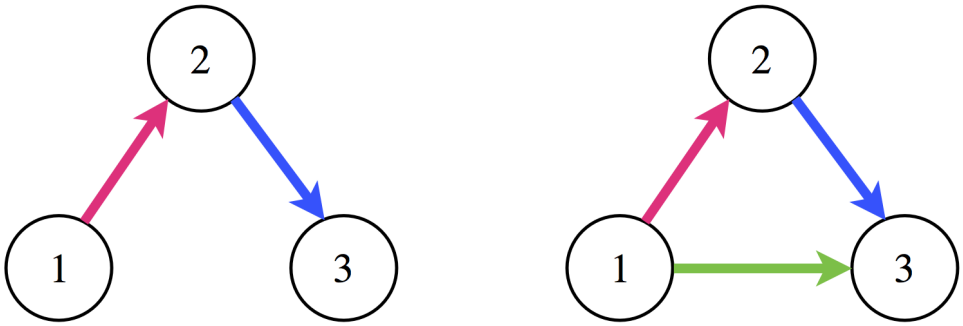
where T is the bandwidth and $K(t)$ is the Hann window function. Here, the bandwidth T is optimized by a method for selecting a fixed kernel bandwidth [25].

3.3.2 Simulations

To confirm the effectiveness of our method, we produced marked point process data from the three coupled Lorenz systems [37] described by the following equations:

$$\begin{cases} \dot{x}_i(t) &= -\rho x_i(t) + \rho y_i(t), \\ \dot{y}_i(t) &= -x_i(t)z_i(t) + r x_i(t) - y_i(t) + \sum_{j\neq i} k_{ij} y_j(t-\tau_{ij}), \\ \dot{z}_i(t) &= x_i(t)y_i(t) - b z_i(t), \end{cases} \tag{3.17}$$

with $i, j = 1, 2, 3$. We set the parameters $\rho = 10, r = 28$, and $b = 8/3$. The delays are set to $\tau_{12} = 5$ and $\tau_{23} = 15$. The coupling strengths are set to $k_{12} = k_{23} = 8$ and $k_{13} = 0$. The coupling structure of the systems is shown in Fig. 3.15. We defined the lth event timing t_i^l as the time when $|y_i(t)|$ takes the lth local maxima and the amplitude of the lth event as $|y_i(t_i^l)|$.



(a) (b)

Figure 3.15: Coupling structures: (a) true structure and (b) misestimated results for the three coupled Lorenz systems.

3.3.3 Results

First, we examined how the mutual information and the correlation coefficients between the original continuous time series y_1 and the transformed time series depend on the band width T (Fig. 3.16). When the bandwidth is equal to 7.0, which is calculated by the method for selecting a fixed kernel bandwidth [25], the correlation coefficients and the mutual information are high.

Second, we estimated the connectivity of systems only from event timings (we set $h(t_i^l) = 1$ for $i = 1, 2, 3$ and $l = 1, 2, \cdots, N$). We show the results of the cross mutual information (Fig. 3.18 (a)–(c)) and partial mutual information (Fig. 3.18 (d)–(f)). From the results, it is difficult to detect coupling because no peaks are found. In contrast, if we use not only event timings t_i^l but also the additional information $h(t_i^l)$, the results show clear peaks (Fig. 3.19). From the results of the cross mutual information (Fig. 3.19 (a)–(c)), we can identify sharp peaks. These results indicate the connectivity of the coupled Lorenz systems as Fig. 3.15(b). However, the connectivity between the systems 1 and 3 is misestimated because of a spurious bias. To remove the spurious bias, we used the partial mutual information. In Fig. 3.19(d)–(f), we show the results of the partial mutual information. The results shown in Fig. 3.19(e) indicate that no connection exists between the systems 1 and 3. These results indicate that we can estimate the connectivity of the systems as Fig. 3.15(a) which is the true connectivity of the systems. In addition, we can clearly identify the delays $\tau_{12} = 5$ and $\tau_{23} = 15$ (Fig. 3.19(d)–(f)), because the peaks occur at these values.

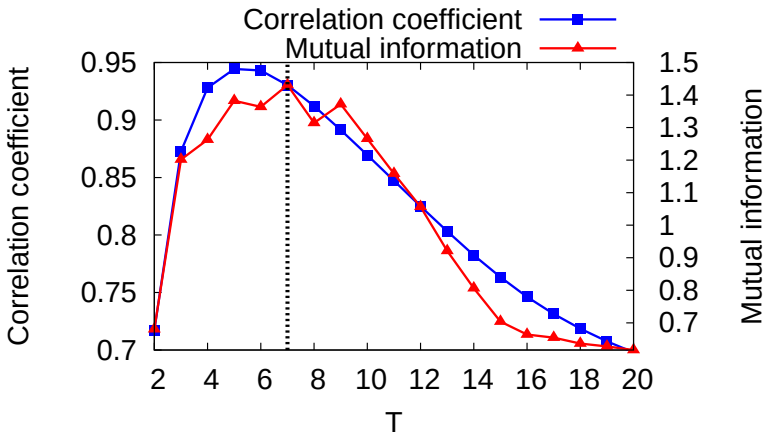


Figure 3.16: Correlation coefficients and mutual information between the original continuous time series and the transformed time series with the bandwidth T.

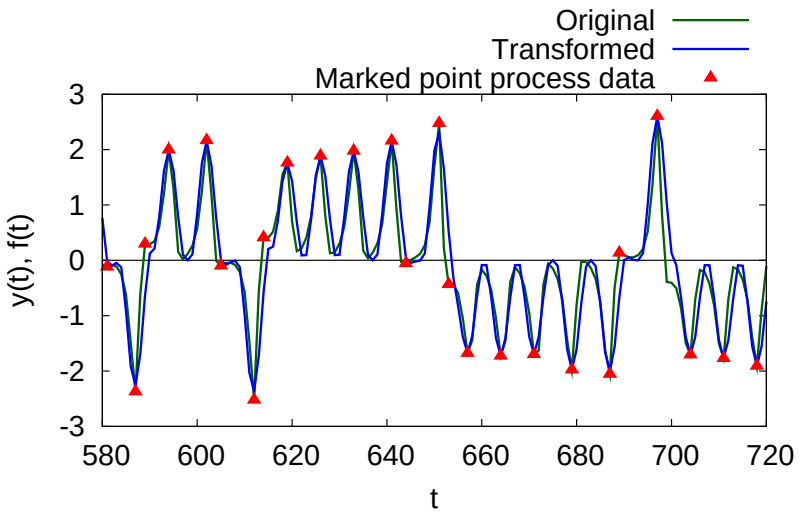


Figure 3.17: Example of the marked point process data generated from the coupled Lorenz systems. The transformed time series by using Eq. (3.16) with the optimal bandwidth ($T = 1.5$ from Fig. 3.16) and the original continuous time series.

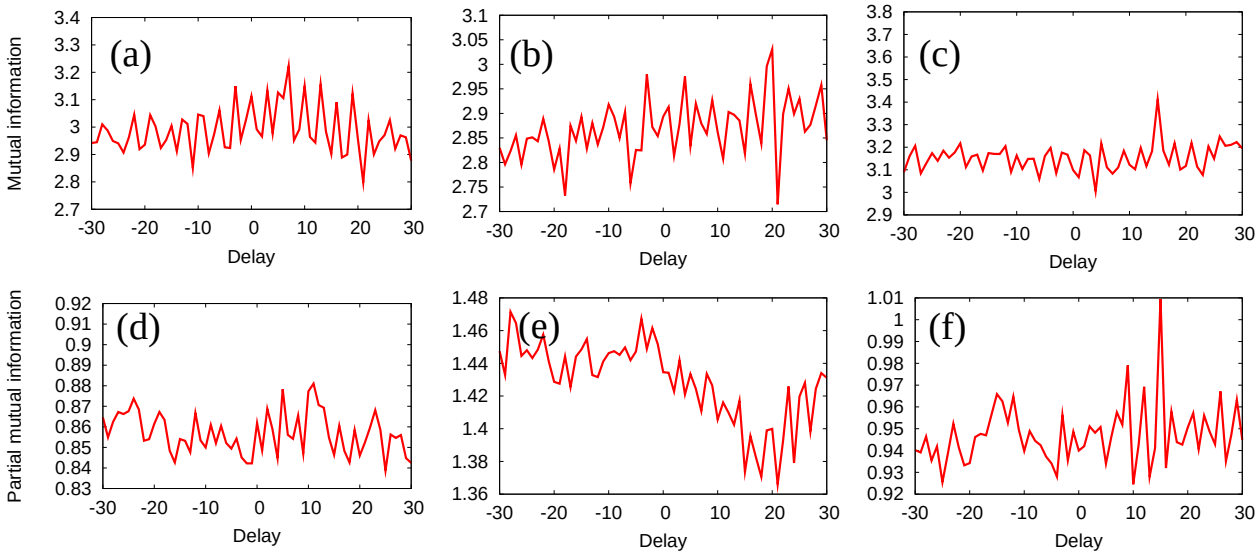


Figure 3.18: Averaged mutual information for 20 trials: (a) $I(f_1, f_2)$, (b) $I(f_1, f_3)$ and (c) $I(f_2, f_3)$ and averaged partial mutual information for 20 trials: (d) $I(f_1, f_2|f_3)$, (e) $I(f_1, f_3|f_2)$ and (f) $I(f_2, f_3|f_1)$ when we use only event timings t_i^l for the coupled Lorenz systems.

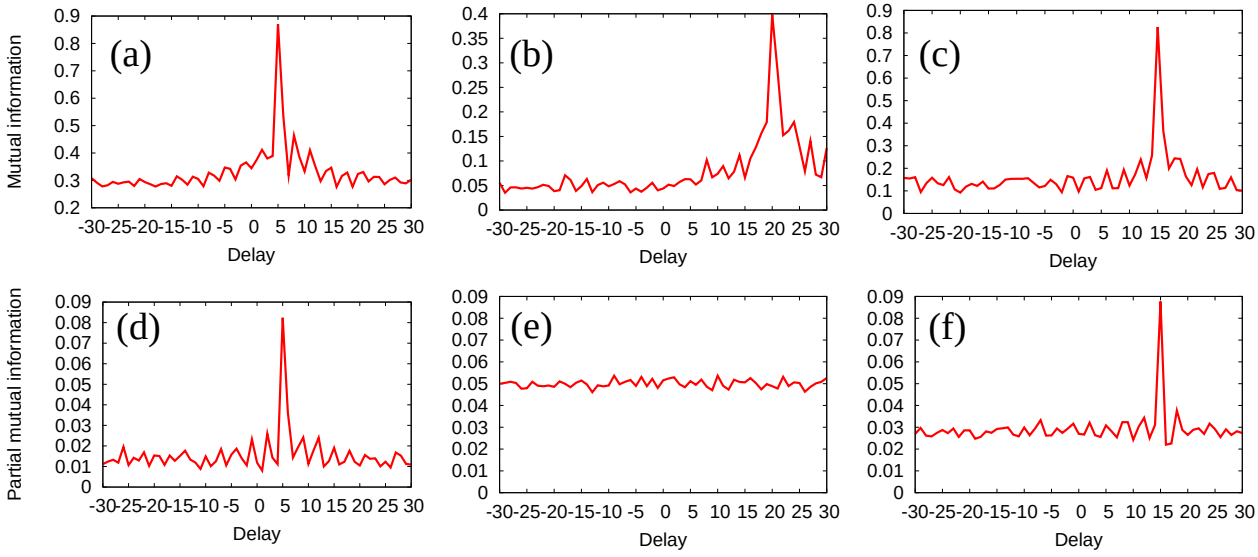


Figure 3.19: Averaged mutual information for 20 trials: (a) $I(f_1, f_2)$, (b) $I(f_1, f_3)$ and (c) $I(f_2, f_3)$, and averaged partial mutual information for 20 trials: (d) $I(f_1, f_2|f_3)$, (e) $I(f_1, f_3|f_2)$ and (f) $I(f_2, f_3|f_1)$ when we use not only event timings t_i^l but also the amplitude information $h(t_i^l)$ for the coupled Lorenz systems.

3.4 Summary of Chapter 3

In this chapter, we proposed method of estimating network structures by transforming point processes into continuous time series.

To transform point processes into continuous time series, we used the kernel density function. When the kernel function is used, we have to select the optimal bandwidth because the transformed time series depends on the bandwidth. Then we used two selection methods of the kernel bandwidth. As the first method, we used a kernel bandwidth optimization method which is proposed by Shimazaki et al. [25]. The method is used in spike rate estimation. As the second method, we also used the attractor reconstruction[34] which is a basic technique in chaotic time series analysis. By using these methods, we found that high correlation coefficients and low normalized mean square errors are evaluated between the input time-series and its reconstructed time-series.

Using the transformation method, we estimated network structures from simple point processes. To evaluate the validity of the proposed method, we used a neural network constructed from a mathematical model of the Izhikevich simple neuron model [14] and generated multi-spike sequences. The results show that our method exhibits high performance.

In additon, for marked point processes, we extended the proposed method. We proposed a method of transforming marked point processes into continuous time series which is based on the kernel density estimator. Then, we applied the partialization analysis to the transformed time series. As a result, we can estimate the connectivity of the coupled Lorenz systems from marked point processes.

Chapter 4

Estimation Method of Network Structure Using Distance between Point Processes and Partialization Analysis

In Chapter 3, we proposed estimation methods of network structures by transforming point processes into continuous time series. Although these methods work well, one should be careful to apply these methods, because it is possible to lose essential information of spike sequences by transforming spike sequences into continuous time series. From this point of view, we proposed methods for estimating network structures only from point process data without transforming point process into continuous time series. In the proposed methods, we use distance between point processes.

4.1 Estimation Method of Network Structure from Simple Point Processes

In Section 4.2, we propose an estimation method of network structures from spike sequences as simple point processes. In the proposed method, we use a spike time metric which is one of the statistic to quantify a distance between two spike sequences and apply the partialization analysis to the spike time metric..

4.1.1 Spike Time Metric

The spike time metric was proposed by Victor and Purpura [13]. In the spike time metric, two operations are used to quantify the distance between two spike sequences; then, the costs incurred to carry out these operations are defined to measure the dissimilarity between spike sequences. The first operation is deletion or insertion of a single spike, the cost of both of which is unity. The second operation is a movement or temporal shift of a single spike, the cost of which is proportional to the interval for which the single spike is moved. For example, if a spike sequence A is the same as a sequence A' except for single spikes that occur at t_A^a in A and $t_{A'}^{a'}$ in A', the cost $c_q(A, A')$ corresponds to $q|t_A^a - t_{A'}^{a'}|$. The parameter q determines which operation takes precedence in the deletion and insertion, or the movement. In the spike time metric [13], a metric distance between two spike sequences Z and Z' is defined as

$$D_q(Z, Z') = \min\left\{\sum_{k=1}^{N-1} c_q(V_k, V_{k+1})\right\}, \tag{4.1}$$

where $V_1, V_2, \ldots, V_N$ are elementary steps from Z to Z' [13]. The metric distance between the two spike sequences is the minimum total cost of a set of elementary steps that transforms a spike sequence into another spike sequence. An example of transforming Z to Z' is shown in Fig. 4.1. From V_1 (=Z) to V_3, we move two spikes. From V_3 to V_4, we insert one spike (the fifth one). From V_4 to V_5 (=Z'), we move a single spike to reach Z'. Consequently, the cost of transforming Z into Z' is $1 + q(\Delta t_1 + \Delta t_2 + \Delta t_3)$ in this example.

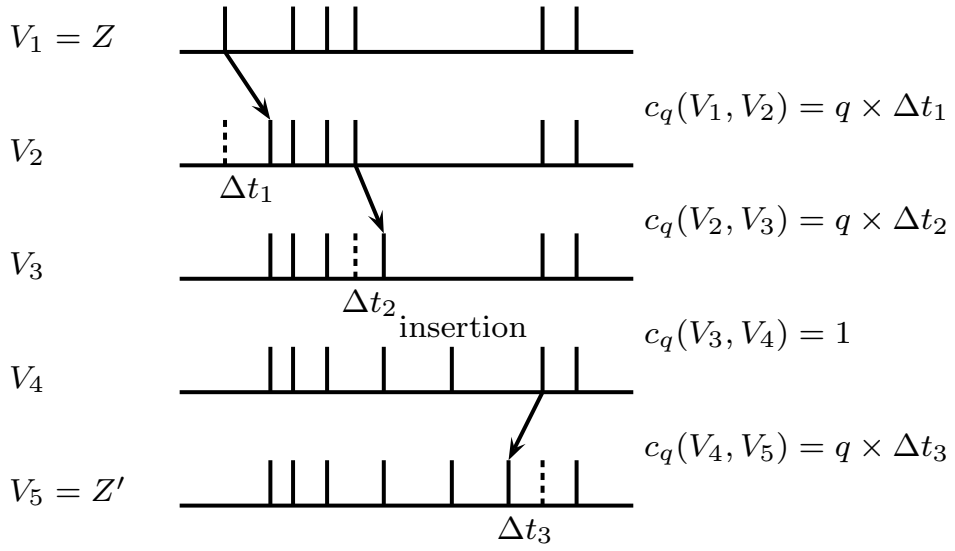


Figure 4.1: Example of transforming the spike sequence Z into Z'.

Spike time metric coefficient and partial spike time metric coefficient

Let us denote the ith and jth spike sequences observed from a neural system as X_i and X_j, respectively. Using Eq. (5.8), the spike time metric coefficient (STMC) between the two spike sequences X_i and X_j is defined as:

$$S_q(X_i, X_j) = 1 - \frac{D_q(X_i, X_j)}{\max_{i,j}\{D_q(X_i, X_j)\}}. \tag{4.2}$$

The STMC is a normalized measure of the spike time metric, and it takes a value between 0 and 1. The STMC is similar to a function of the correlation coefficient between X_i and X_j whereas the spike time metric (Eq. (5.8)) is a function of the distance between X_i and X_j. If two neurons that produce X_i and X_j are coupled, $S_q(X_i, X_j)$ is expected to become greater than that in the case in which the two neurons are uncoupled, because the distance between X_i and X_j in the coupled cases becomes less than that in the uncoupled cases.

Although the STMC is a measure similar to the correlation coefficient, it can be spuriously biased if the two neurons are driven by a common input from other neurons. To avoid such a bias, as in the case of deriving a partial correlation coefficient from the correlation coefficient, the partial

spike time metric (PSTMC) between the two spike sequences X_i and X_j is defined as

$$P_q(X_i, X_j) = \left| \frac{\alpha(i, j)}{\sqrt{\alpha(i, i)\alpha(j, j)}} \right|,$$

where $\alpha(i, j)$ is the (i, j)th entry of the inverse matrix of $S_q(X_i, X_j)$. The PSTMC can reveal the unbiased correlation between the two spike sequences by removing any spurious correlation as the partialization analysis does. In other words, the PSTMC works as a partial correlation coefficient between the two spike sequences based on the spike time metric. Using the PSTMC, we can find hidden relations between neurons and estimate the network structure.

4.1.2 Simulations

We examined the validity of the proposed measures using a mathematical model. In our numerical experiments, we assumed that we can only observe asynchronous spike sequences while the true network structure is unknown. To evaluate the validity of the proposed measures, we used the Izhikevich neuron model.

We checked the validity of our method under the following two conditions. The first condition is that the network is homogeneous; the neurons are only regular spiking (RS) neurons whose parameters are $a = 0.02, b = 0.2, c_i = -65$, and $d_i = 8$. The second condition is that the network is heterogeneous; the neurons are RS, intrinsically bursting (IB), and chattering (CH) neurons. The parameters were set to $a = 0.02, b = 0.2, c_i = -65 + 15r_i^2$, and $d_i = 8 - 6r_i^2$ where r_i are uniform random numbers between $[0, 1]$. The synaptic weight was set to six and the amplitude of the external inputs was five times G, where G was a Gaussian random number with a mean value and standard deviation of zero and unity, respectively. For the sake of simplicity, we did not consider conduction delays.

To apply our proposed measures to the spike sequences, it is important to decide q appropriately, because it determines a relative weight between the two operations in the spike time metric: deletion and insertion, or the movement. Then, we pre-examined $S_q(X_i, X_j)$ and $P_q(X_i, X_j)$ by changing q, as shown in Fig. 4.2. From Fig. 4.2(b), if we set q between 50 and 100, the network structure can be accurately estimated by P_q, because the disparity between the coupled and uncoupled elements is clearly distinguished by P_q. Then, we experimentally decide q only with the observed asynchronous spike sequences in the following manner.

First, let us assume that we have two spike sequences A and A' that are identical except for a single spike that occurs at t_A^q in A and $t_{A'}^{q'}$ in A'. To transform A into A', we have two operations.

The first one is the insertion and deletion whose cost is two (the cost of each is unity). The second one is the movement whose cost is $q|t_A^a - t_{A'}^{a'}|$. Then, solving the equation $2 = q|t_A^a - t_{A'}^{a'}|$ (in the case in which the costs are same), we obtain a critical value of q (if $2 < q|t_A^a - t_{A'}^{a'}|$, the insertion and deletion are selected, otherwise the movement is selected). To decide q appropriately, we have to define a possible range for the movement of a single spike, $t_A^a - t_{A'}^{a'}$, because it decides the critical values of q. Then, we evaluate a maximum value of the average minimum time difference by

$$\frac{1}{N_i}\sum_{k=1}^{N_i}\min_l\left|t_i^k - t_j^l\right|,$$

where N_i is the number of spikes in the ith sequence and t_i^k is the kth spike timing in the ith spike sequence. Figure 4.3 shows an example. From Fig. 4.3, the maximum value of the average minimum time difference is 0.025 [s]. To obtain the critical value of q, we solve the equation $2 = q|t_A^a - t_{A'}^{a'}|$ by substituting 0.025 in $t_A^a - t_{A'}^{a'}$. Then, we have $q = 80$; this agrees well with the fact that if we set q between 50 and 100, the disparity between the coupled and uncoupled elements is clear (Fig. 4.2(b)). Although this determination procedure is heuristic, we have confirmed that it works well in other cases.

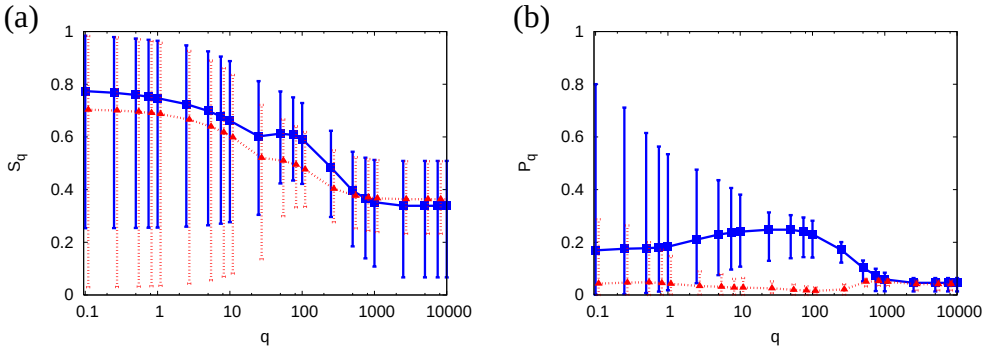


Figure 4.2: Relation between q and (a) S_q and (b) P_q for the 100 regular ring topology of a homogeneous network. Blue solid lines indicate S_q and P_q between coupled elements. Red dotted lines indicate S_q and P_q between uncoupled elements. Error bars which indicate minimum and maximum values with 30 trials are also provided.

If the ith and the jth neurons are coupled, $S_q(X_i, X_j)$ and $P_q(X_i, X_j)$ might increase. On the other hand, if these neurons are not coupled, these measures might decrease. Thus, to find coupled neuron pairs, we extracted higher values of these measures by discriminating the coupled and

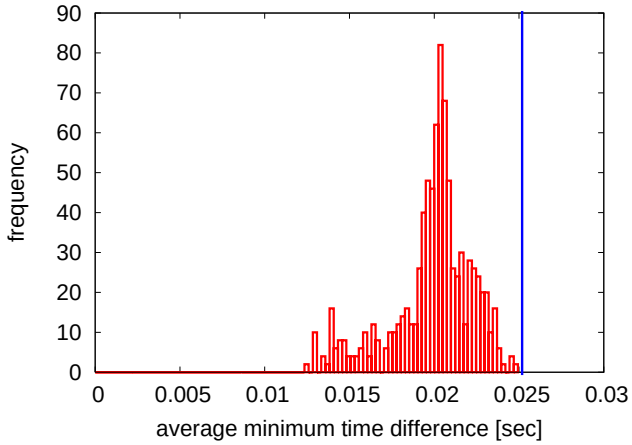


Figure 4.3: Frequency histogram of an average minimum time distance between two spikes for the 100 regular ring topology of a homogeneous network. Blue vertical line indicates a boundary between the deletion and the insertion or the movement. It also indicates the maximum value of the histogram.

uncoupled pairs by calculating a threshold. To exclude any subjective discrimination, the threshold is decided by the Otsu thresholding [17], which is based on a linear discriminant analysis.

To evaluate the overall estimation accuracy, we compared the estimated network structure with the true network structure. For this evaluation, we introduced two evaluation indices, C–$\tilde{\text{C}}$ and U–$\tilde{\text{U}}$, defined as follows:

$$\begin{aligned} \text{C}-\tilde{\text{C}} &= \frac{\sum_{i,j} w_{ij}\tilde{w}_{ij}}{\sum_{i,j} w_{ij}}, \text{and} \\ \text{U}-\tilde{\text{U}} &= \frac{\sum_{i,j}(1-w_{ij})(1-\tilde{w}_{ij})}{\sum_{i,j}(1-w_{ij})}, \end{aligned}$$

where w_{ij} and $\tilde{w}_{ij}$ are the connectivity of the true network structure and the estimated network structure, respectively. If the ith and the jth elements are (estimated to be) coupled, w_{ij} and $\tilde{w}_{ij}$ take unity. If they are not, w_{ij} and $\tilde{w}_{ij}$ take a value of zero. If C–$\tilde{\text{C}}$ and U–$\tilde{\text{U}}$ approach unity, the estimation accuracy increases.

4.1.3 Results

We checked the validity of our method in simple networks. As simple networks, we used 3-neuron and 4-neuron network structures (Fig. 4.4).

We compared the values of S_q and P_q (Fig. 4.5, blue and red). In the case of three neuron network, the STMC of uncoupled pairs is zero (Fig. 4.5(a), uncoupled, red), because the uncoupled pair is the most dissimilar in all neuron pairs. However, for uncoupled pairs of neurons, the PSTMC is larger than the STMC. In the case of four neuron network, the STMC of the uncoupled pairs is as large as that of the coupled pairs by spurious correlations (Fig. 4.5(b)). However, the large STMC becomes small in the PSTMC. As same as the case of the three neuron network, the PSTMC is larger than the STMC in the case that the STMC is small. Therefore, we used the STMC and the PSTMC adaptively. In the adaptive measure, if $P_q(X_i, X_j)$ is larger than $S_q(X_i, X_j)$, we used $S_q(X_i, X_j)$. The adaptive measure, called the adaptive partial spike time metric (APSTMC), is defined as:

$$A_q(X_i, X_j) = \min(S_q(X_i, X_j), P_q(X_i, X_j)), \tag{4.3}$$

where $\min(\cdot, \cdot)$ takes the value of smaller one, $S_q(X_i, X_j)$ or $P_q(X_i, X_j)$. In the three neuron network, $A_q(X_i, X_j)$ of uncoupled pairs is zero (Fig. 4.5(a), pink). In the four neuron network, the APSTMCs of uncoupled pairs are distributed within the range where the value is small (Fig. 4.5(b), pink). In the followin

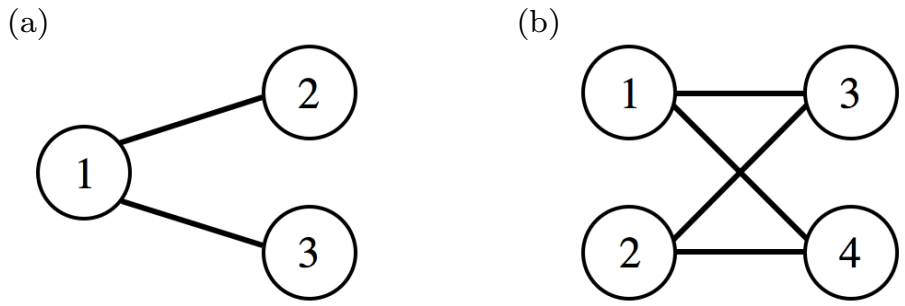


Figure 4.4: Network structures which are composed of (a) three and (b) four neurons. The neurons are mutually connected.

Next, to evaluate the performance of the proposed measure, we generated a complex network structure having a regular ring topology with 100 neurons by a random rewiring of the synaptic

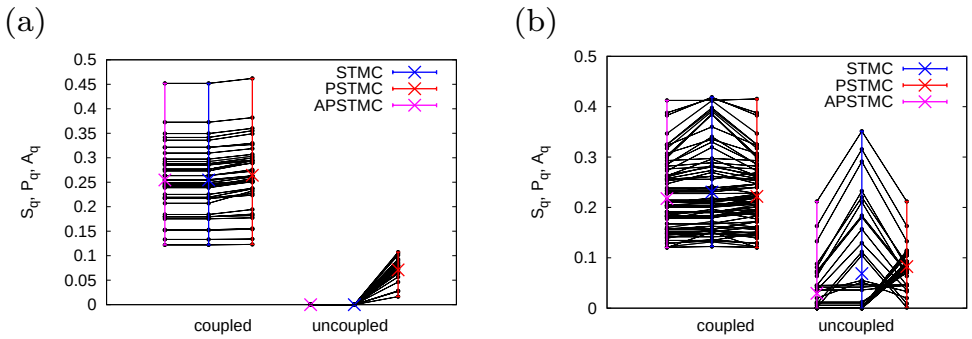


Figure 4.5: Transition from S_q to P_q and A_q for (a) three and (b) four neuron networks. The blue points indicate average value of S_q, the red points, P_q, and the pink points, A_q with 20 trials. Error bars which indicate minimum and maximum values are also provided.

connections between neurons in the same manner as that described in Ref. [15]. We also introduced another complex network structure, namely, a scale-free network [16]. We set the parameters m_0 (number of nodes in an initial network) to four and m (number of edges attached to the network) to two, as in the case of Ref. [16]. In these networks, the neurons are mutually connected.

For a small-world network whose components are homogeneous, although our method can separate coupled (blue) and uncoupled (red) pairs of neurons with S_q, a few uncoupled pairs are estimated as coupled pairs (Fig. 4.6(a)). However, coupled and uncoupled pairs are more clearly distinguished when we use A_q (Fig. 4.6(b)). In the heterogeneous case, the distribution of uncoupled and coupled pairs widely overlap (Fig. 4.6(c)). As shown in Fig. 4.6(d), coupled and uncoupled pairs are effectively discriminated. In the homogeneous case of a scale-free network, the discrimination of coupled and uncoupled pairs is improved as in the case of the small-world network when we use A_q (Figs. 4.6(e) and (f)). However, as compared to the small-world case, coupled pairs are misestimated (Figs. 4.6(b) and (f)). Similarly, the use of A_q realizes more clear separation than the use of S_q (Figs. 4.6(g) and (h)). Then, the estimation accuracy of the scale-free network becomes lower than that of the small-world network.

In Fig. 4.7, we show the results of the small-world network when the rewiring probability is changed. The results show that the estimation of the small-world structure with the STMC depends on the rewiring probability p (Fig. 4.7). In particular, the estimation accuracy for the random network ($p = 1.0$) is low. However, if we use APSTMC, the estimation accuracy improves. In addition, there is significant difference in C–$\tilde{\text{C}}$ and U–$\tilde{\text{U}}$ between the STMC and the APSTMC (P

< 0.05; two-sample t-test) (Fig. 4.7).

In Fig. 5.6, we examined how the estimation accuracy depends on the temporal epoch for observing spikes. C–$\tilde{\mathrm{C}}$ of APSTMC for a small-world network is similar to that of STMC, and C–$\tilde{\mathrm{C}}$ in both STMC and APSTMC is high even if the temporal epoch of the spike sequences is short (Fig. 5.6(a)). However, the accuracy by U–$\tilde{\mathrm{U}}$ exhibits a significant difference (Fig. 5.6(b)). The estimation accuracy with APSTMC is high, whereas that with STMC is low. If the temporal epoch becomes longer than 20 [s], the estimation accuracy with APSTMC is high. On the other hand, for the scale-free network, we obtain the same tendency with the small-world case; however, at least 30 [s] are required to obtain high accuracy of U$-\tilde{\mathrm{U}}$ (Fig. 5.6(d)). The estimation accuracy of C$-\tilde{\mathrm{C}}$ is not improved (the value is less than 0.9) (Fig. 5.6(c)).

In Fig. 4.9, we show the results for different network sizes. For all the network sizes, even though the network is small-world or scale-free, the APSTMC exhibits a high estimation accuracy of C$-\tilde{\mathrm{C}}$ and U$-\tilde{\mathrm{U}}$ (Figs. 4.9(a)–(d)). However, the estimation accuracy of C$-\tilde{\mathrm{C}}$ worsens slightly as the network size increases even if APSTMC is used.

We also examined how stable our method under noise influence (Fig. 4.10). In this simulation, the noise is classification error in spike sorting, and the noise rate μ is the ratio of the number of spikes replaced with other spike sequences randomly to all the number of spikes in a spike sequence. For the small-world network, the estimation accuracy of C$-\tilde{\mathrm{C}}$ and U$-\tilde{\mathrm{U}}$ with APSTMC is high if the noise rate μ is smaller than 0.2. On the other hand, for the scale-free network, the APSTMC exhibits a high estimation accuracy of U$-\tilde{\mathrm{U}}$ when the noise rate μ is smaller than 0.1.

Fig. 4.11 shows correlation diagrams between PSTMC and coupling strength. From the results, there is a correlation between PSTMC and coupling strength. This results indicate that we can see a distribution of coupling strength from the distribution of the PSTMC.

In this thesis, we evaluated the performance of the estimation method for network structures by the proposed measure using only spike sequences. Our results show that the estimation accuracy of the small-world network is much better than that of the scale-free network. This result is attributable to the large difference between the firing rates of neurons in the scale-free network (Fig. 4.12). In Fig. 4.12, we show the results of local estimation accuracy $\mathrm{C}_i-\tilde{\mathrm{C}}_i$ and $\mathrm{U}_i-\tilde{\mathrm{U}}_i$, respectively defined as

$$\begin{aligned} \mathrm{C}_i-\tilde{\mathrm{C}}_i &= \frac{\sum_j (w_{ij}\tilde{w}_{ij})}{\sum_j w_{ij}}, \\ \mathrm{U}_i-\tilde{\mathrm{U}}_i &= \frac{\sum_j ((1-w_{ij})(1-\tilde{w}_{ij}))}{\sum_j (1-w_{ij})}. \end{aligned}$$

Even though a neuron has small connectivity or small degree, its estimation accuracy of $\mathrm{C}_i-\tilde{\mathrm{C}}_i$

decreases if the average degree of its adjacent neurons is high (Fig. 4.12(a)). If the difference between the firing rates is large, the value of the spike time metric increases because the spike time metric determines the insertion and deletion of spikes. Therefore, the reduction in the estimation accuracy of the scale-free network originates from the large difference between the firing rates of neurons.

We also compared estimation accuracy which estimated from membrane potentials with that which estimated from spike sequences. In case of membrane potentials, we estimated the network structure by using a partial correlation coefficient. The results of estimation accuracy for several temporal epochs are shown in Fig. 4.13. When the temporal epoch is 10[s], the estimation accuracy is high in case of membrane potentials. However, when the temporal epoch is longer than 20[s], the result of PSTMC is almost same estimation accuracy as that of membrane potentials. Fig. 4.14 shows the results of estimation accuracy for several coupling strengths. From the results, when the coupling strength is smaller than 3, PSTMC exhibits higher estimation accuracy than partial correlation coefficient between membrane potentials. Although the estimation accuracy in case of PSTMC is lower than the partial correlation coefficient between membrane potentials, the result of PSTMC is almost same estimation accuracy as that of the partial correlation coefficient between membrane potentials. The results of estimation accuracy for several sizes are shown in Fig. 4.15. Similarly, the result of PSTMC is almost same estimation accuracy as that of the partial correlation coefficient between membrane potentials.

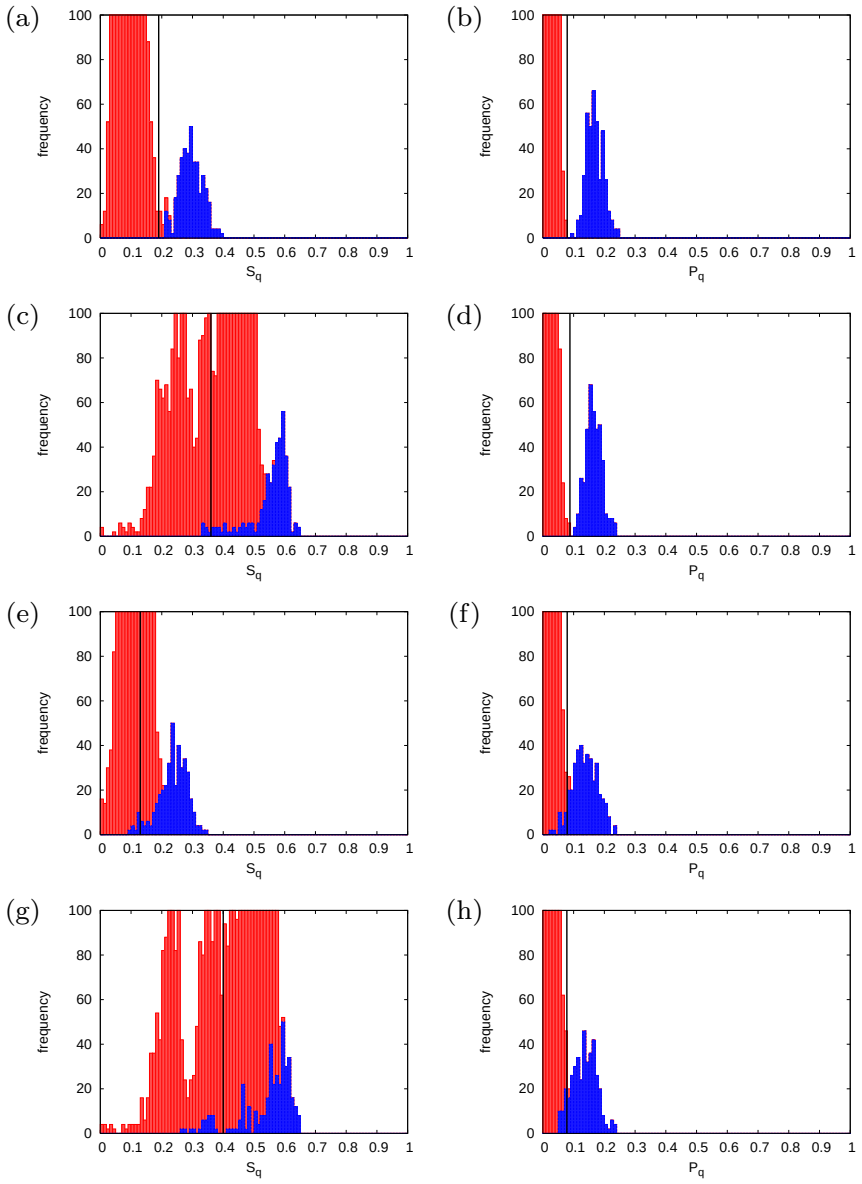


Figure 4.6: Histograms of (a) S_q and (b) A_q for the small-world network, and (e) S_q and (f) A_q for the scale-free network comprising only RS neurons, and (c) S_q and (d) A_q for the small-world network, and (g) S_q and (h) A_q for the scale-free network comprising RS, IB, and CH neurons. The number of neurons is 100. The temporal epoch of spike sequences used for estimation is 50 [s]. Histograms of all of S_q and A_q are indicated in red, and those of the coupled elements are superimposed by blue. The vertical lines indicated a threshold decided by Otsu thresholding. If S_q and A_q are less than the threshold, corresponding neurons are classified into the uncoupled class.

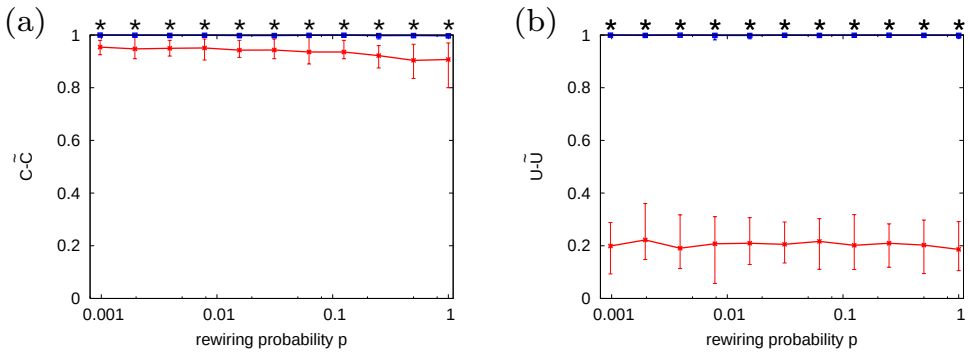


Figure 4.7: Estimation accuracy of the small-world network structure: (a) C–$\tilde{C}$ and (b) U–$\tilde{U}$ for rewiring probabilities p. The number of neurons is 100. The temporal epoch of spike sequences used for estimation is 50 [s]. The network is comprising RS, IB, and CH neurons. Red lines indicate S_q and blue lines, A_q. * indicates a significant difference between the STMC and the APSTMC ($P < 0.05$, two-sample t-test). Error bars which indicate minimum and maximum values with 20 trials are also provided.

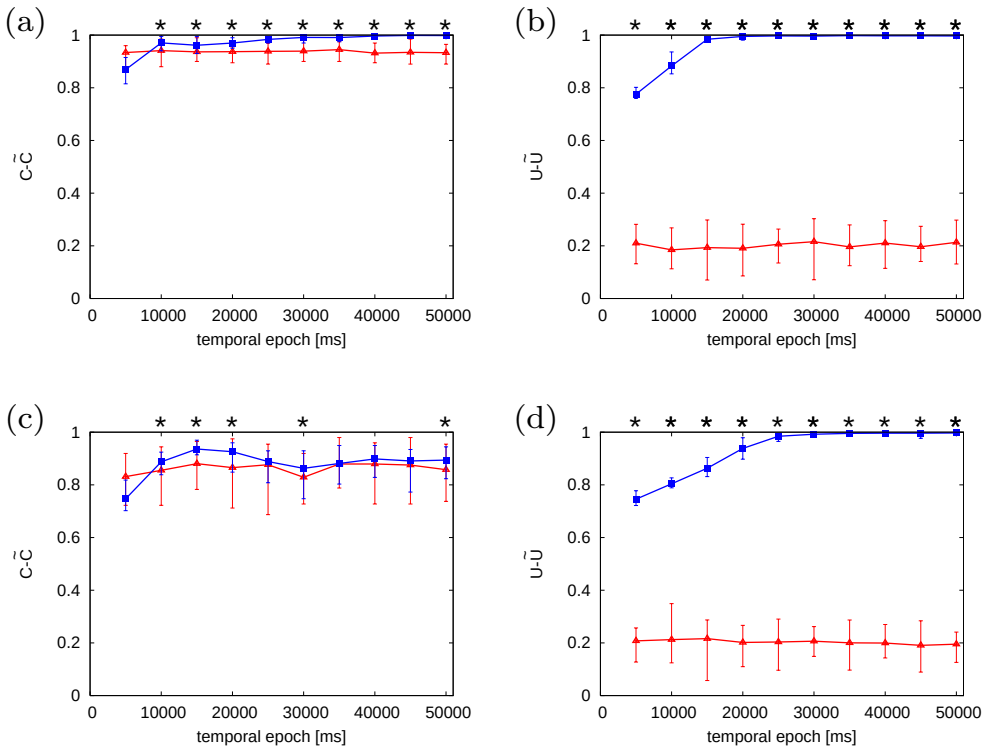


Figure 4.8: Estimation accuracy of the network structure for several temporal epochs: (a) C–$\tilde{\text{C}}$ and (b) U–$\tilde{\text{U}}$ for small-world network, (c) C–$\tilde{\text{C}}$ and (d) U–$\tilde{\text{U}}$ for scale-free network. The number of neurons is 100. The network is comprising RS, IB, and CH neurons. Red lines indicate S_q and blue lines, A_q. * indicates a significant difference between the STMC and the APSTMC ($P < 0.05$, two-sample t-test). Error bars which indicate minimum and maximum values with 20 trials are also provided.

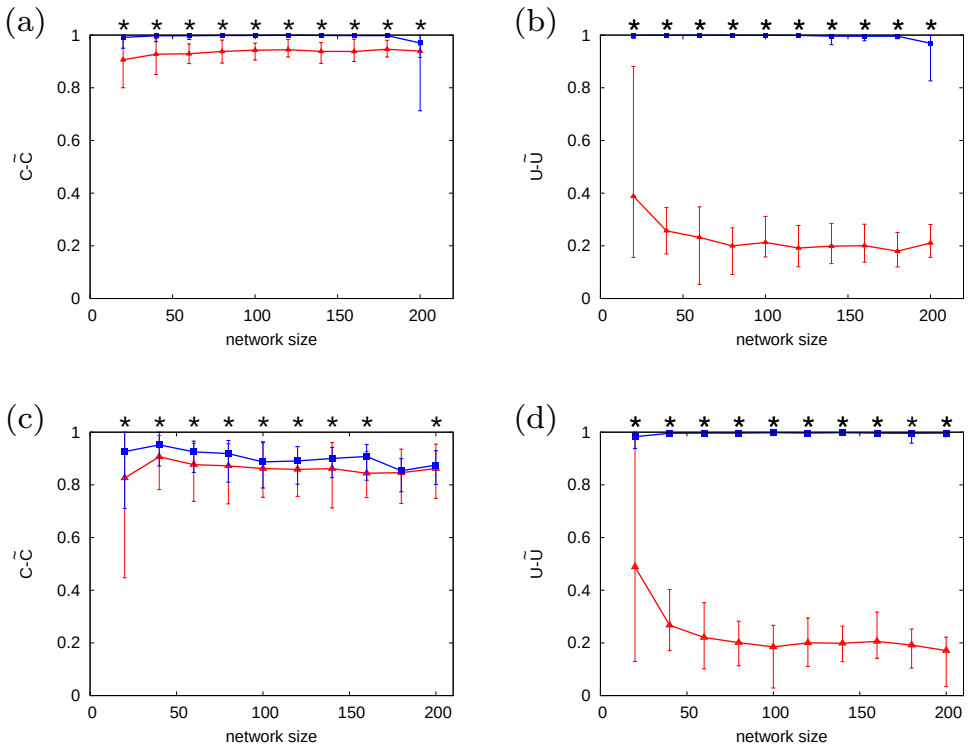


Figure 4.9: Estimation accuracy of the network structure for several network sizes: (a) C–$\tilde{\text{C}}$ and (b) U–$\tilde{\text{U}}$ for small-world networks, (c) C–$\tilde{\text{C}}$ and (d) U–$\tilde{\text{U}}$ for scale-free networks. The temporal epoch of spike sequences used for estimation is 50 [s]. The network is comprising RS, IB, and CH neurons. Red lines indicate S_q and blue lines, A_q. * indicates a significant difference between the STMC and the APSTMC ($P < 0.05$, two-sample t-test). Error bars which indicate minimum and maximum values with 20 trials are also provided.

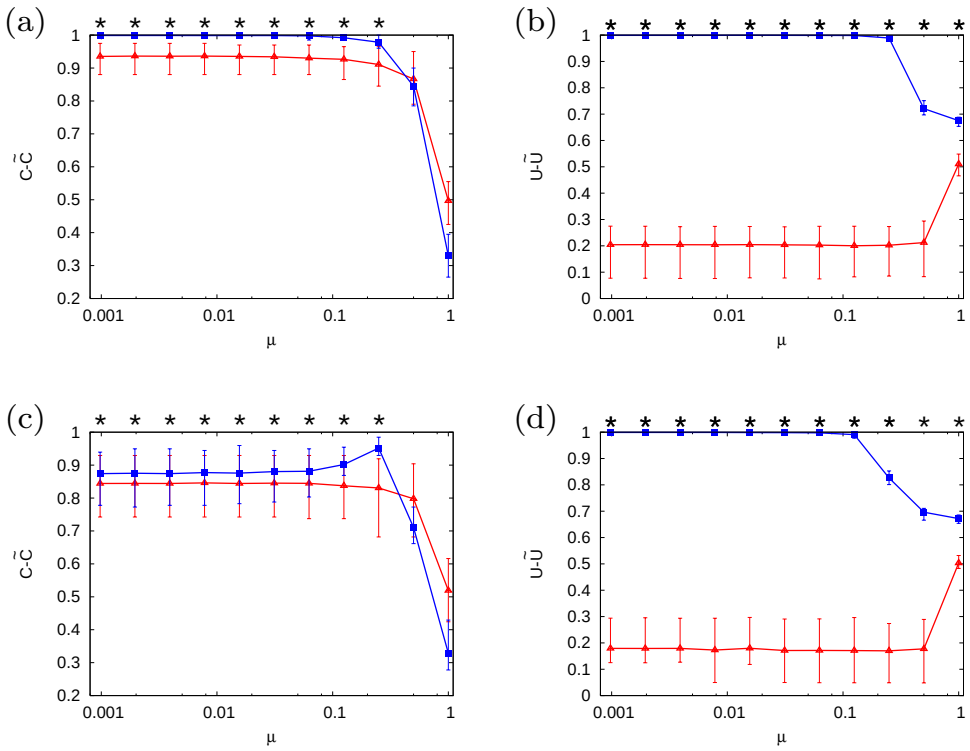
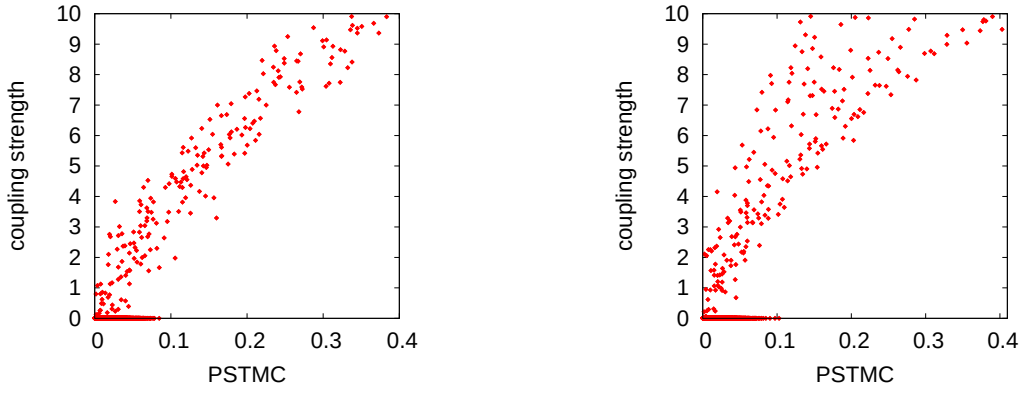


Figure 4.10: Estimation accuracy of the network structure under noise influence: (a) C–$\tilde{\text{C}}$ and (b) U–$\tilde{\text{U}}$ for small-world networks, (c) C–$\tilde{\text{C}}$ and (d) U–$\tilde{\text{U}}$ for scale-free networks. The number of neurons is 100. The temporal epoch of spike sequences used for estimation is 50 [s]. The network is comprising RS, IB, and CH neurons. Red lines indicate S_q and blue lines, A_q. * indicates a significant difference between the STMC and the APSTMC ($P < 0.05$, two-sample t-test). Error bars which indicate minimum and maximum values with 20 trials are also provided.



(a) (b)

Figure 4.11: Correlation diagrams between PSTMC and coupling strength for (a) the small-world network and (b) scale-free network. The number of neurons is 100. The temporal epoch of spike sequences used for estimation is 50 [s]. The network is comprising RS neurons.

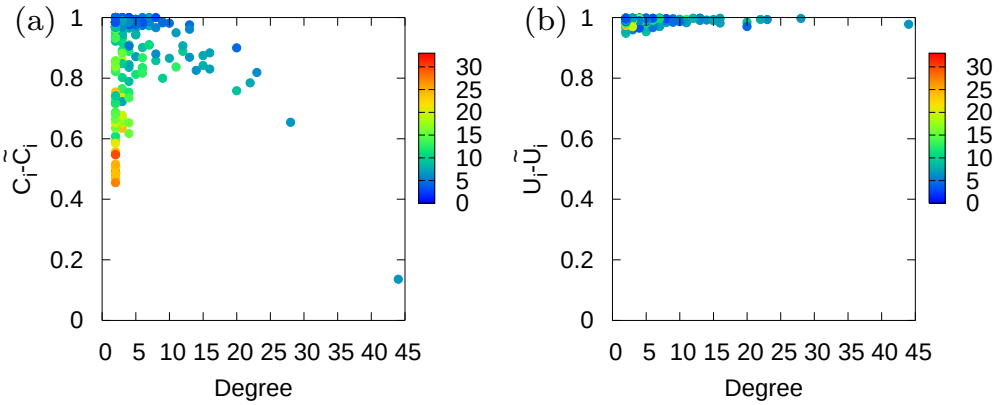
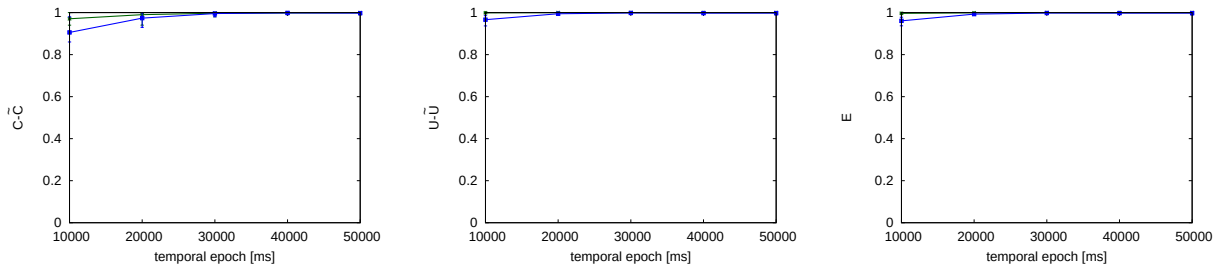
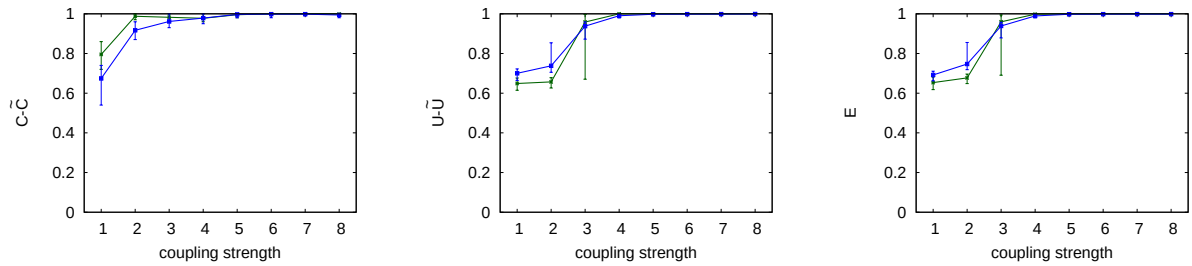


Figure 4.12: Estimation accuracy when we use A_q and connectivity of the ith neuron for the scale-free network: (a) $C_i - \tilde{C}_i$ and (b) $U_i - \tilde{U}_i$. The number of neurons is 300. The temporal epoch of spike sequences used for estimation is 50 [s]. The network is comprising RS, IB, and CH neurons. Color bars indicate the average degree of adjacent neurons.



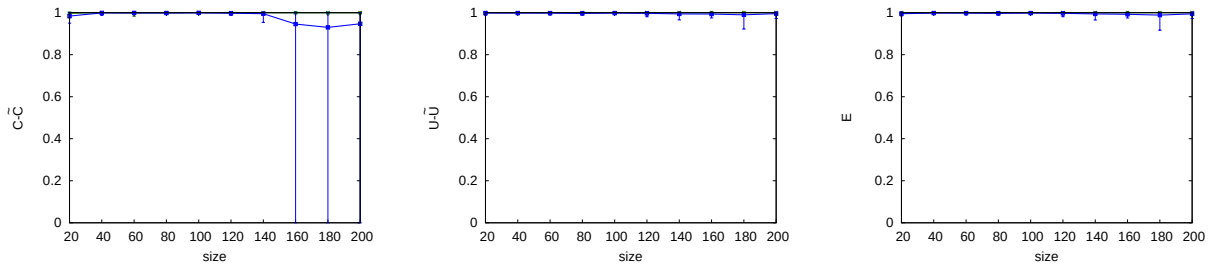
(a) (b) (c)

Figure 4.13: Estimation accuracy of the network structure for several temporal epochs: (a) $C_i - \tilde{C}_i$, (b) $U_i - \tilde{U}_i$ and (c) E. Green lines indicate partial correlation coefficients between membrane potentials and blue lines, P_q. The number of neurons is 50. The coupling strength is 6. The network is comprising RS, IB, and CH neurons.



(a) (b) (c)

Figure 4.14: Estimation accuracy of the network structure for several coupling strengths: (a) $C_i - \tilde{C}_i$, (b) $U_i - \tilde{U}_i$ and (c) E. Green lines indicate partial correlation coefficients between membrane potentials and blue lines, P_q. The number of neurons is 50. The temporal epoch of spike sequences used for estimation is 50 [s]. The network is comprising RS, IB, and CH neurons.



(a) (b) (c)

Figure 4.15: Estimation accuracy of the network structure for several network sizes: (a) $C_i - \tilde{C}_i$, (b) $U_i - \tilde{U}_i$ and (c) E. Green lines indicate partial correlation coefficients between membrane potentials and blue lines, P_q. The temporal epoch of spike sequences used for estimation is 50 [s]. The coupling strength is 6. The network is comprising RS, IB, and CH neurons.

4.2 Estimation Method of Network Structures and direction of couplings from Simple Point Processes

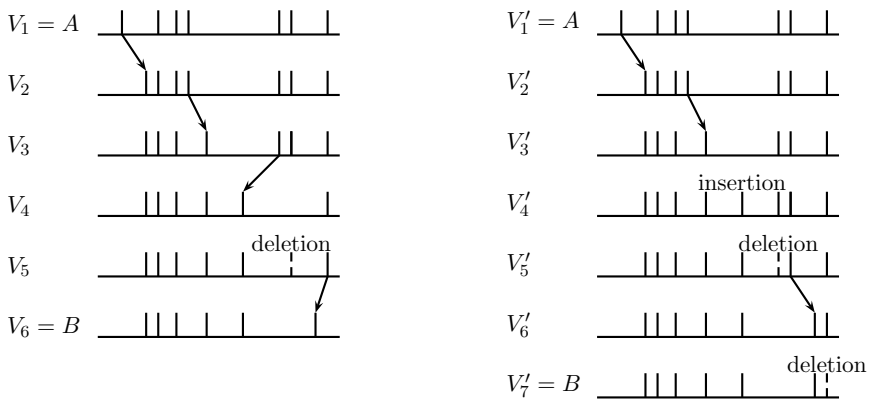
In this section, we propose a new estimation method of direction of connectivity between neurons in neural networks only from multiple spike sequences. The proposed method is based on a spike time metric, or a statistical measure to quantify a degree of dissimilarity between two spike sequences, and the partialization analysis. Although the method of the Section can estimate network structures, the direction of connectivity cannot be estimated. To resolve this issue, we modify the definition of the conventional cost in the spike time metric. Then the proposed method can effectively estimate direction of connectivity between neurons. To check the validity, we applied the proposed method to multiple spike sequences that are produced by a mathematical neural network model. As a result, our method could estimate the neural network structure and the direction of couplings with high accuracy.

4.2.1 Proposed measure

We introduce a new measure for estimating the direction of coupling between neurons. If we use the spike time metric, the direction of coupling cannot be detected although we can estimate connection between two neurons, because the spike time metric holds a condition of symmetry: $D(A, B) = D(B, A)$. Then, we adapt the cost of the movement of a single spike for estimating the direction of coupling. The movement of a single spike in the spike time metric, the spike can be shifted both forward and backward. Here, we consider that a spike is shifted only forward, and define the cost of the movement as follows:

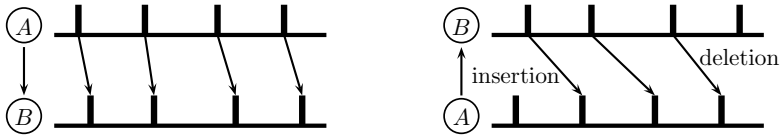
$$c(V, V') = q(t_V^i - t_{V'}^j), \quad t_V^i > t_{V'}^j. \tag{4.4}$$

Figure 4.16(b) shows an example of transforming a spike sequence A to B by the proposed measure. By using the cost of Eq.(4.4), if a neuron A is unidirectionally coupled to a neuron B, $D(A, B)$ exhibits a smaller value than $D(B, A)$ (Fig. 4.17). Then, we can determine the direction of coupling by using this proposed measure.



(a) (b)

Figure 4.16: Examples of transforming a spike sequence A into B in (a) the spike time metric and (b) the proposed measure.



(a) (b)

Figure 4.17: The difference of the total cost between (a) $D(A, B)$ and (b) $D(B, A)$ in the case that a neuron A is unidirectionally connected to a neuron B. (a) $D(A, B)$ takes a small value. (b) $D(B, A)$ takes a larger value than $D(A, B)$.

4.2.2 Simulations

To evaluate the validity of our method, we used a neural network constructed from a mathematical model of the Izhikevich simple neuron model [14] and generated multiple spike sequences.

The dynamics of the ith neuron in the neural network is described by the following equations:

$$\begin{aligned} \dot{v}_i &= 0.04v_i^2 + 5v_i + 140 - u_i + I_i, \\ \dot{u}_i &= a(bv_i - u_i), \\ \text{if } v_i &\geq 30[\text{mV}], \text{ then} \begin{cases} v_i \leftarrow c, \\ u_i \leftarrow u_i + d, \end{cases} \end{aligned}$$

where v_i is the membrane potential, u_i is the membrane recovery variable; and a, b, c, and d are dimensionless parameters. The parameters were set to $a = 0.02$, $b = 0.2$, $c = -65$, and $d = 8$. The variable I_i is the sum of the external and synaptic inputs from coupled neurons. The synaptic weight is set to 8 and the amplitude of the external inputs to 5 times G, where G is a Gaussian random number with a mean value and standard deviation of zero and unity, respectively. The neural network is composed of only excitatory neurons which are regular spiking neurons. The neurons are unidirectly connected. We set delays between neurons to 2[ms] to 4[ms] randomly.

To apply our proposed measures to the spike sequences, it is important to decide q appropriately, because it determines a relative weight between the two operations in the spike time metric: deletion and insertion, or the movement. We experimentally decided q only with the observed asynchronous spike sequences in the following manner. First, let us assume that we have two spike sequences V and V' that are identical except for a single spike that occurs at t_V^i in V and $t_{V'}^j$ in V' under the condition that $t_V^i < t_{V'}^j$. To transform V into V', we have two operations. The first operation is the insertion and deletion. Its cost is two (each cost is unity). The second operation is the movement. Its cost is $q(t_{V'}^j - t_V^i)$. Then, if we solve the equation $2 = q(t_{V'}^j - t_V^i)$ (in the case that both costs are same), we obtain a critical value of q (if $2 > q(t_{V'}^j - t_V^i)$, the movement is selected, otherwise the insertion and deletion are selected). To decide q appropriately, we have to define a possible range for the movement of a single spike, $t_{V'}^j - t_V^i$, because it decides the critical values of q. Then, we evaluate an average minimum time difference by

$$\overline{\Delta T} = \frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{j=1, i \neq j}^{N} \overline{\Delta T}_{ij}, \tag{4.5}$$

where $\overline{\Delta T}_{ij} = \dfrac{1}{N_i} \displaystyle\sum_{k=1}^{N_i} \min_l (t_i^k - t_j^l), (0 < t_i^k - t_j^l < \dfrac{\overline{ISI}}{2})$, N_i is the number of spikes in the ith sequence, N is the number of spike sequences, t_i^k is the kth spike timing in the ith spike sequence, and $\overline{ISI}$ is the mean interspike interval for all the multiple spike sequences. To exclude long time difference, we applied the condition that $t_i^k - t_j^l < \frac{\overline{ISI}}{2}$.

To obtain the critical value of q, we solve the equation $2 = q\overline{\Delta T}$. Although this determination procedure is heuristic, we have confirmed that it works well in other cases.

Using q decided by the abovementioned method, we applied the STMC and PSTMC to the multiple spike sequences. If two neurons are coupled, S and P might be large. Then, we calculate a threshold that classifies the coupled and the uncoupled pairs. The threshold is decided by the Otsu thresholding method [17] which is based on a linear discriminant analysis.

To evaluate the estimation accuracy, we compared the estimated network structure with the true network structure. For this evaluation, we define an index as follows:

$$E = \frac{\sum_{i,j}^{N}(\alpha_{ij}\tilde{\alpha}_{ij} + (1-\alpha_{ij})(1-\tilde{\alpha}_{ij}))}{N(N-1)}, \tag{4.6}$$

where N is the number of elements, α_{ij} and $\tilde{\alpha}_{ij}$ are the directional connectivity of the true network structure and the estimated network structure, respectively. If the ith element is (estimated to be) connected to the jth element, α_{ij} and $\tilde{\alpha}_{ij}$ take a value of unity. If they are not, α_{ij} and $\tilde{\alpha}_{ij}$ take a value of zero. If E approaches unity, the estimation accuracy increases.

4.2.3 Results

First, we checked the validity of our proposed method and compared with the conventional method. As a network structure, we generated a small-world network structure [**?**], which has initially ring lattice topology with 50 neurons and degree 4 of each neurons. The temporal epoch of spike sequences is 50[s], and the coupling strength is set to 8. We examined the case that the neurons are unidirectly coupling. Figure 4.18(a) and (b) shows histograms of S and P in the conventional method, respectively. Figure 4.18(c) and (d) shows histograms of S and P in the proposed method. In Fig. 4.18(a) and (b), some pairs coupled with an opposite direction are misestimated as a coupled class, because $S(i, j)$ equals $S(j, i)$. Thus, the direction of couplings cannot be well estimated in the conventional method. On the other hand, S in the proposed method can estimate the direction of couplings although the threshold is not optimal (Fig. 4.18(c)). However, by using P in the proposed method, we can estimate the coupled pairs and the direction of couplings. In addition, the threshold is optimally calculated (Fig. 4.18(d)).

To check the validity of the proposed method, we evaluated the performance of the proposed method in case of changing rewiring probability in the small-world network model [**?**], network sizes, temporal epochs for observed spikes, and coupling strengths. We show the results when the rewiring probability is changed (Fig. 4.19(a)). If we use S, the estimation accuracy is low. In particular, when the rewiring probability is 1.0, the estimation accuracy is low. On the other hand,

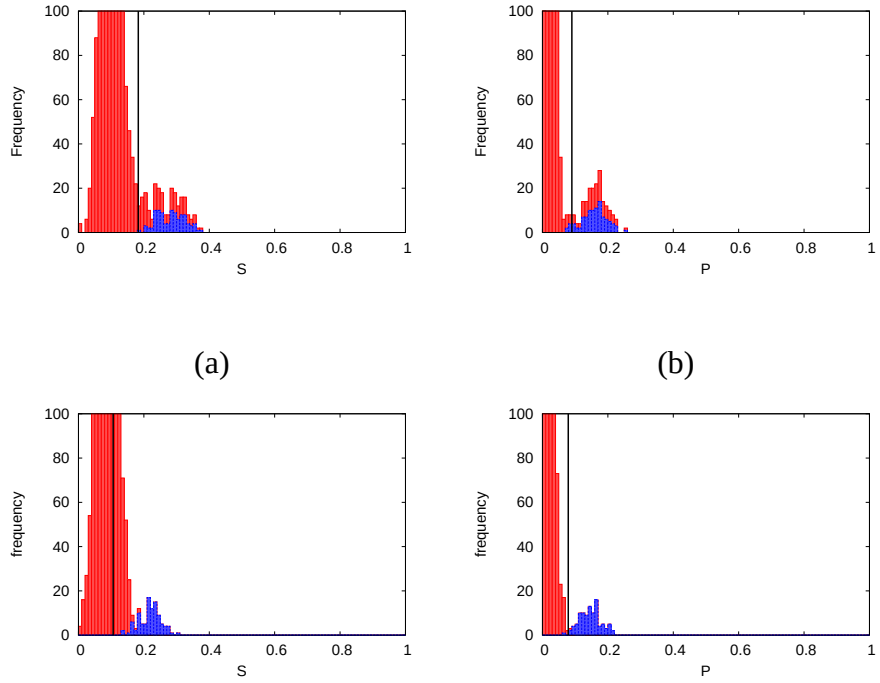
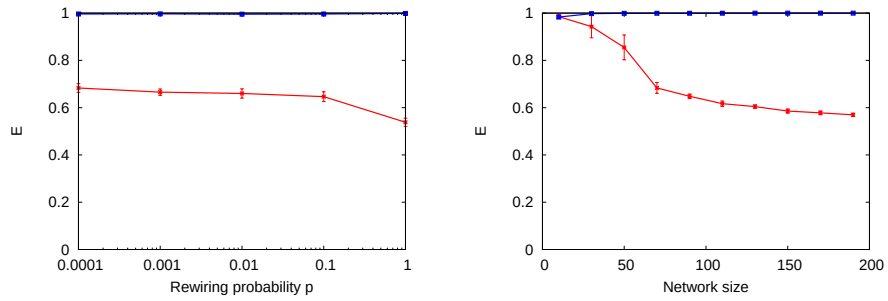


Figure 4.18: Histograms of (a) S and (b) P in the conventional method, and those of (c) S and (d) P in the proposed method. The number of neurons is 50. Histograms of all of S and P are shown by red bars, and those of the coupled elements are superimposed by blue. The vertical lines indicate a threshold decided by the Otsu thresholding. If S and P are larger than the threshold, corresponding neurons are classified into the coupled class.

if we use P, the estimation accuracy is high for all the rewiring probability.

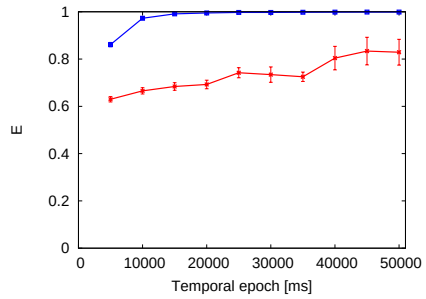
In Fig. 4.19(b), we show the results when the network size is changed. The estimation accuracy in P is high although the estimation accuracy in S decreases as the network size increases.

We also examined how the estimation accuracy depends on the temporal epochs for observed spikes (Fig. 4.19(c)). The estimation accuracies in both S and P become higher as the temporal epoch becomes longer. If the temporal epoch is longer than 15[s], the estimation accuracy in P is almost unity.

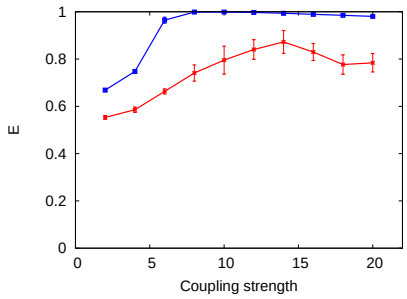


(a)

(b)



(c)



(d)

Figure 4.19: Estimation accuracy of the network structure for (a) rewiring probabilities, (b) network sizes, (c) temporal epochs, and (d) coupling strength. Red and blue lines represent S and P, respectively. Error bars which indicate standard errors with 20 trials are also provided.

In addition, we show the results in case that values of coupling strength are changed (Fig. 4.19(d)). If the coupling strength is larger than 8, the estimation accuracy is high. However, the estimation accuracy becomes slightly worse as the coupling strength becomes larger.

4.3 Summary of Chapter 4

In this chapter, we proposed two measures, spike time metric coefficient and partial correlation coefficient, to estimate network structures only from the information of observed asynchronous spike sequences. We applied the proposed measures to simple networks and complex networks [15, 16]. The results show that network structures can be estimated by adaptive application of

spike time metric coefficients and partial spike time metric coefficients with high performance.

In Eq. (4.3), we heuristically defined the PSTMC from association of the conventional partialization analysis. Then, we experimentally confirmed that the PSTMC can measure true relation between spike sequences by removing spurious correlations. However, we do not have theoretically proven why the PSTMC works well yet. As a future work, we have to prove that why the PSTMC works well.

Chapter 5

Estimation Method of Network Structure Using Distance and Multi-Dimensional Scaling

In Chapter 4, we proposed two measures, the spike time metric coefficient and the partial correlation coefficient. By using these measures, we could estimate network structures only from point processes. However, we do not have theoretically proven why the PSTMC works well. Then, in this chapter, we propose a new method to estimate network structures from point processes by using the distance between point processes and multi-dimensional scaling. The multi-dimensional scaling is a classical method for representing arrangement in the Euclidean space from a distance relation [8]. Using the distance between point processes which calculated by the spike time metric, we can obtain the position vectors which represents the linear relationship among spike sequences in the Euclidean space and apply a linear regression model to the obtained position vector. Then we can use the statistical measures of the partialization analysis such as the partial correlation coefficient and the partial directed coherence because these measures are based on the linear regression model.

5.1 Estimation Method of Network Structure from Simple Point Process

In this section, we propose a method of estimating network structures from spike sequences as simple spike processes. In the proposed method, first, we measure a distance between spike sequences by using the spike time metric that is proposed by Victor and Purpula [13]. Next, we apply the multi-dimensional scaling [8] to the distance between spike sequences to represent a relationship among spike sequences in the Euclidean space. By using the multi-dimensional scaling [8], we can obtain position vectors of each spike sequence in the Euclidean space. Then we can consider that the obtained position vectors represent a linear relationship among the spike sequences. Then, the linear regression model can be applied to the obtained position vectors. To estimate connectivity between neurons, we use the partial directed coherence [6] which is based on the linear regression model and partialization analysis. To check the validity of the proposed method, we apply the proposed method to observed multiple spike sequences from a mathematical neuron model. In numerical simulations, we show that our method can estimate neural network structure even though the neurons have many common, or spurious, inputs from the other neurons.

5.1.1 Multi-Dimensional Scaling

To represent the relationship between point processes, we used the multi-dimensional scaling (MDS) [8] which is a classical method for representing arrangement in the Euclidean space from a distance relation. In the proposed method, we used the STM as the distance relation.

Then a distance matrix $\boldsymbol{S}$ is defined as $\boldsymbol{S} = \{s_{ij}\}$, where s_{ij} is described as follows:

$$s_{ij} = \frac{1}{2}(D(i,m)^2 + D(j,m)^2 - D(i,j)^2), \tag{5.1}$$

where $D(i, j)$ is the STM between the ith and the jth point processes.

Next, by applying the eigenvector decomposition to $\boldsymbol{S}$, we obtain a coordinate matrix X as follows:

$$\boldsymbol{S} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^{\mathrm{T}} = (\boldsymbol{V}\boldsymbol{\Lambda}^{\frac{1}{2}})(\boldsymbol{V}\boldsymbol{\Lambda}^{\frac{1}{2}})^{\mathrm{T}} \equiv \boldsymbol{X}\boldsymbol{X}^{\mathrm{T}}, \tag{5.2}$$

where $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$, $\boldsymbol{V} = (\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n)$, n is the number of nonzero eigenvalues of $\boldsymbol{S}$, and λ_i and $\boldsymbol{v}_i$ are the ith eigenvalue and eigenvector of the matrix $\boldsymbol{S}$, respectively. The matrix $\boldsymbol{X}$ is a coordinate matrix described by $\boldsymbol{X} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N)^{\mathrm{T}}$, where $\boldsymbol{x}_k = (x_{k1}, x_{k2}, \ldots, x_{kn})^{\mathrm{T}}$.

Using the STM as the distance, we can obtain the position vector $\boldsymbol{x}_i$ which represents the linear relationship between corresponding marked point processes in the Euclidean space. Then we can apply a linear regression model to the obtained position vector $\boldsymbol{x}_i$ as $\boldsymbol{x}_i = \sum_{j=1}^{N} a_{ij}\boldsymbol{x}_j + \boldsymbol{w}_i$,where $\boldsymbol{w}_i$ is a multivariate Gaussian white noise process and a_{ij} is a regression coefficient from $\boldsymbol{x}_j$ to $\boldsymbol{x}_i$.

Using the STM as the distance relation, we can obtain the position vector $\boldsymbol{x}_i$ which represents the linear relationship among spike sequences in the Euclidean space. Then we can apply a linear regression model to the obtained position vector $\boldsymbol{x}_i$ as follows:

$$\boldsymbol{x}_k = \sum_{r=1}^{N} a_{kr}\boldsymbol{x}_r + \boldsymbol{w}_k, \tag{5.3}$$

where $\boldsymbol{w}_i$ is multivariate Gaussian white noise process and a_{kr} is a regression coefficient from $\boldsymbol{x}_r$ to $\boldsymbol{x}_k$.

5.1.2 Numerical Simulations

To evaluate the validity of our method, we used a neural network constructed from a mathematical model, or the Izhikevich simple neuron model [14]. We conducted numerical experiments according to the following procedures.

1. We constructed the neural network whose elements are the Izhikevich simple neuron model, and observed multiple point processes.
2. We calculated the distance D between point processes by using the STM.
3. We applied the MDS to the distance D.
4. Using observed position vectors in the MDS, we calculated the PDC.
5. If $|\overline{\pi}_{ij}|$ (Eq. (2.54)) is larger than 0.1, we decided the neuron pair of i and j as the coupled pair [6].

To confirm the estimation accuracy, we compared the structure of an estimated network with that of the original network. We used two measures defined by

$$\mathrm{C}\text{–}\tilde{\mathrm{C}} = \frac{\sum_{i,j}(\alpha_{ij}\tilde{\alpha}_{ij})}{\sum_{i,j}\alpha_{ij}}, \tag{5.4}$$

$$\mathrm{U}\text{–}\tilde{\mathrm{U}} = \frac{\sum_{i.j}((1-\alpha_{ij})(1-\tilde{\alpha}_{ij}))}{\sum_{i,j}(1-\alpha_{ij})}, \tag{5.5}$$

where α_{ij} ($\tilde{\alpha}_{ij}$) is the (i, j)th element of the adjacency matrix of the original (estimated) network structure. If the ith and jth neurons are coupled, α_{ij} and $\tilde{\alpha}_{ij}$ take unity. If they are not coupled, α_{ij} and $\tilde{\alpha}_{ij}$ take zero. If C$-\tilde{\text{C}}$ and U$-\tilde{\text{U}}$ are close to unity, our method estimates the original network structure well.

5.1.3 Results

First, we used a simple network structure shown in Fig. 5.1. In the network structure of Fig. 5.1, spurious couplings exist between neurons, and the connected neurons are mutually connected. A sample of point processes observed from this network is shown in Fig. 5.2. Figure 5.3 shows the result of the value of the PDC. From the result, coupled pairs and uncoupled pairs are clearly distinguished. Furthermore, $|\overline{\pi}_{ij}|$ of uncoupled pairs take smaller than the threshold (0.1).

We also investigated the estimation accuracy when the network size is changed. In the following simulations, we generated a ring lattice network structure in which the number of edges is four. Figure 5.4 shows the C$-\tilde{\text{C}}$ and U$-\tilde{\text{U}}$ when the network size is changed. The estimation accuracy is high (both of C$-\tilde{\text{C}}$ and U$-\tilde{\text{U}}$ are higher than 0.8) when the network size is smaller than 100. However, the estimation accuracy C$-\tilde{\text{C}}$ becomes worse as the network size becomes larger than 100.

We also examined how the estimation accuracy depends on the coupling strength. The results are shown in Fig. 5.5. Both of C$-\tilde{\text{C}}$ and U$-\tilde{\text{U}}$ are high when the coupling strength is between 4 and 7. On the other hand, the estimation accuracy of C$-\tilde{\text{C}}$ is worse when the coupling strength is weaker than 3. If the coupling strength is stronger than 9, both of C$-\tilde{\text{C}}$ and U$-\tilde{\text{U}}$ decrease. In this situation, the STMs take almost same values because the point processes are almost identical due to the strong connection. Then, in this method, it is considered that it is difficult to estimate the connectivity in the case as spike synchronizes.

We finally examined how the estimation accuracy depends on the temporal epoch for observed spikes. The results are shown in Fig. 5.6. The estimation accuracy is close to unity as the temporal epoch is long. The results show that both of C$-\tilde{\text{C}}$ and U$-\tilde{\text{U}}$ is high (> 0.8) when the temporal epoch for observed multiple spikes is longer than 35[s]. Then, to keep the high estimation accuracy for both of C$-\tilde{\text{C}}$ and U$-\tilde{\text{U}}$, our method needs point processes for 35[s] at least.

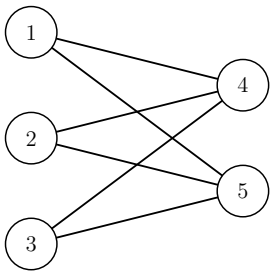


Figure 5.1: A network structure which is composed of five neurons. In this structure, the connected neurons are mutually connected.

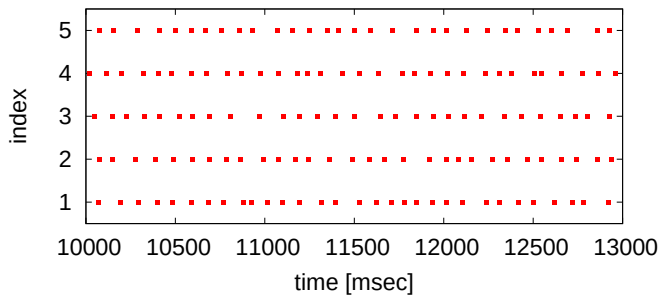


Figure 5.2: A sample of spike sequences observed from the neural network shown in Fig. 5.1.

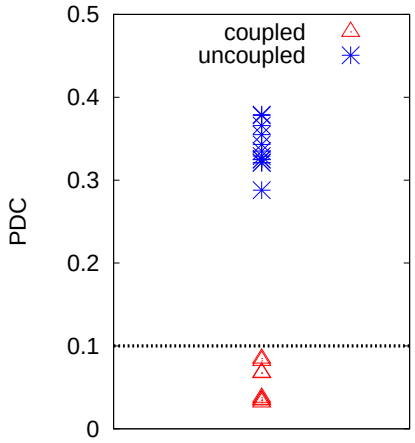


Figure 5.3: The result of applying the PDC. Red triangles and blue asterisks indicate the PDCs for uncoupled pairs and coupled pairs, respectively. The dotted line shows a threshold for $|\overline{\pi}_{ij}|$ (0.1).

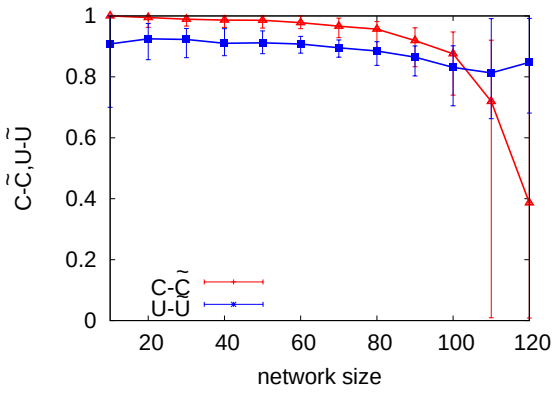


Figure 5.4: Estimation accuracy of the network structure in case of changing the network size. We set the coupling strength is 6, and the temporal epoch is 50[s]. Red and blue lines represent C$-\tilde{\text{C}}$ and U$-\tilde{\text{U}}$, respectively. Error bars which indicate minimum and maximum values with 20 trials are also provided.

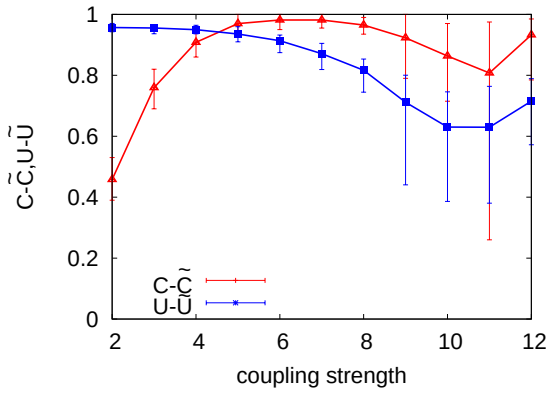


Figure 5.5: Estimation accuracy of the network structure in case of changing the coupling strength. Red and blue lines represent C−C̃ and U−Ũ, respectively. We set the number of neurons is 50, and the temporal epoch is 50[s]. Error bars which indicate minimum and maximum values with 20 trials are also provided.

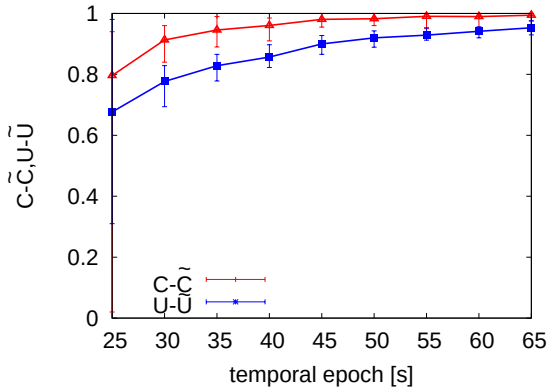


Figure 5.6: Estimation accuracy of the network structure in case of changing the temporal epoch. Red and blue lines represent C−C̃ and U−Ũ, respectively. We set the number of neurons is 50, and the coupling strength is 6. Error bars which indicate minimum and maximum values with 20 trials are also provided.

5.2 Estimation Method of Network Structures from Marked Point Process

In this section, we modified the conventional spike time metric to be applicable to marked point process.

5.2.1 Spike Time Metric for Marked Point Process

The spike time metric [13] is one of the statistics to quantify a distance between two point processes. The statistic consists of two operations. The first operation is deletion or insertion of an event: their cost is unity. The second operation is a movement of an event: the cost is $q\Delta t$ where q is a parameter and Δt is the temporal duration. Suzuki et al. proposed a metric for marked point processes which is extended the spike time metric [28]. The cost for marked point processes is modified as follows:

$$c_{p,q}(A, A') = \begin{cases} 1, & \text{(deletion)} \\ 1, & \text{(insertion)} \\ p|h(t_i^A) - h(t_{i'}^{A'})| + q|t_i^A - t_{i'}^{A'}|, & \text{(movement)} \end{cases} \tag{5.6}$$

where p and q are parameters. The cost of deletion or insertion of an event is same as that of the spike time metric. However, the cost of movement is modified. If marked point process data A is the same as data A' except for an event that occur at t_i^A and its amplitude $h(t_i^A)$ in A and $t_{i'}^{A'}$ and its amplitude $h(t_{i'}^{A'})$ in A', the cost of movement is $p|h(t_i^A) - h(t_{i'}^{A'})| + q|t_i^A - t_{i'}^{A'}|$.

However, it is also important to consider an amplitude of marks when an event is deleted or inserted. Then we proposed new costs of deletion and insertion as follows:

$$c_{p,q}(A, A') = \begin{cases} p|h(t_k^A)|, & \text{(deletion)} \\ p|h(t_l^{A'})|, & \text{(insertion)} \\ p|h(t_i^A) - h(t_{i'}^{A'})| + q|t_i^A - t_{i'}^{A'}|, & \text{(movement)} \end{cases} \tag{5.7}$$

where $h(t_k^A)$ is a mark of a deleted event, and $h(t_l^{A'})$ is a mark of an inserted event. The cost of movement is same as Eq. (5.6).

Then, a metric between two marked point process data Z and Z' is defined as follows:

$$D(Z, Z') = \min \left\{ \sum_{k=1}^{N-1} c_{p,q}(V_k, V_{k+1}) \right\}, \tag{5.8}$$

where $V_1 = Z, V_2, \cdots, V_N = Z'$ are elementary steps from Z to Z'. The metric between the two marked point process data is the minimum total cost of a set of elementary steps that transforms marked point process data into another data.

5.2.2 Numerical Simulations and Results

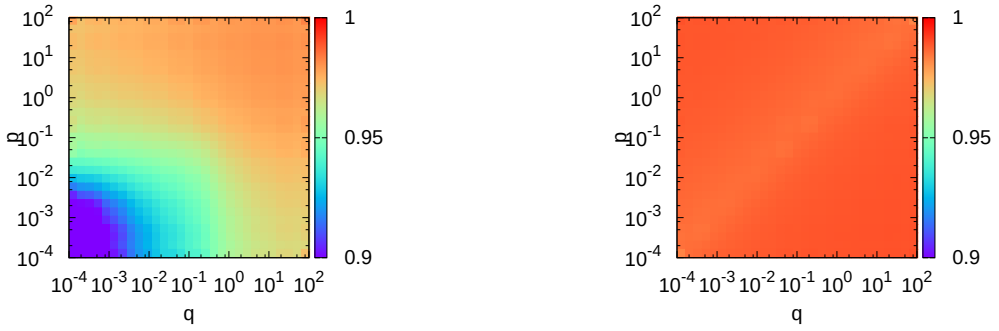
Evaluation of the proposed method with dynamical noise

To investigate the validity of the proposed method, we used a mathematical model. We observed marked point process data from Rössler systems [24] with dynamical noise, which are described by the following equations:

$$\begin{cases} \dot{x}_i &= -\omega y_i - z_i + \sigma\eta_i, \\ \dot{y}_i &= \omega x_i + a y_i, \\ \dot{z}_i &= b + x_i z_i - c z_i. \end{cases} \tag{5.9}$$

with $l = 1, 2$, where σ is the noise strength and η_i is Gaussian noise with a mean value and standard deviation of zero and unity, respectively. The Gaussian noises $\eta_1(t)$ and $\eta_2(t)$ are uncorrelated. We define the ith event timing t_i as the time when $|\tilde{x}_l(t)|$ takes the ith local maxima and the amplitude h_i of the ith event as $|\tilde{x}_l(t_i)|$. The corresponding marked point process is represented by $M_l = (t_i, h_i)$. To evaluate the validity of the proposed measure, we investigated correlation coefficients between noise strength σ and the metric $D(M_1, M_2)$ when the parameters p and q are changed.

In Fig. 5.7(a), the result of the conventional method shows low correlation (< 0.9) when the parameters p and q are small. In Fig. 5.7(b), the result of proposed method shows high correlation (> 0.95) in all parameters. In addition, the results of relations between distance and noise strength are shown in Fig. 5.8. When the parameters p and q are small (Fig. 5.8(a)), the distance of the conventional method andthe noise strength have not correlation. The distance of the proposed method and the noise strength have correlation. Thus the proposed measures can quantify the degree of dissimilarity between two marked point process data using broader parameters than the conventional method.



(a) (b)

Figure 5.7: Correlation coefficients between the noise strength σ and the proposed measure $D(M_1, M_2)$ when the parameters p and q are changed.

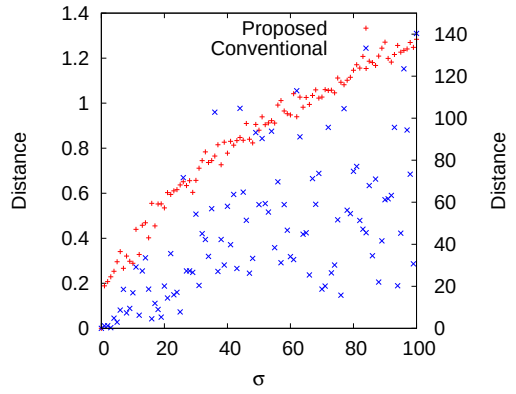
Evaluation of the proposed method with observation noise

We also investigate the validity of the proposed method using observed marked point process data from Rössler systems [24] with observation noise. In this simulation, we set $\sigma = 0$ in Eq. (5.9) (no dynamical noise) and used the first variable x of Eq. (5.9). The observed time series is defined as follows:

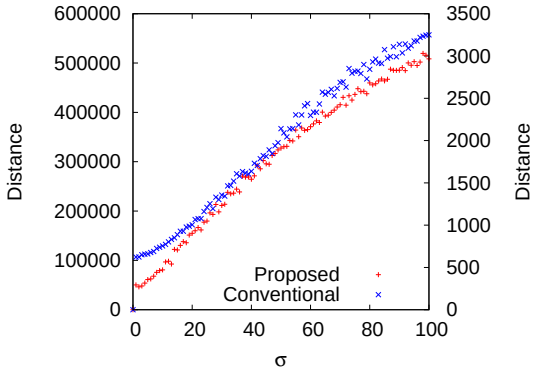
$$\tilde{x}_l = x + s\eta_l, \tag{5.10}$$

with $l = 1, 2$, where s is the noise strength and η_l is Gaussian noise with a mean value and standard deviation of zero and unity, respectively. The Gaussian noises η_1 and η_2 are uncorrelated. We define the ith event timing t_i as the time when $|\tilde{x}_l|$ takes the ith local maxima and the amplitude h_i of the ith event as $|\tilde{x}_l(t_i)|$. The corresponding marked point process is represented by $M_l = (t_i, h_i)$.

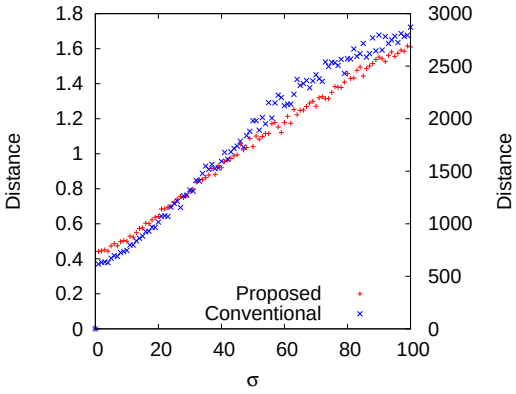
The results of the correlation coefficients between noise strength s and the metric $D(M_1, M_2)$ when the parameters p and q are changed are shown in Fig. 5.9. From the results, correlation coefficients of the proposed method (Fig. 5.9(b)) are higher than these of the conventional method (Fig 5.9(a)).



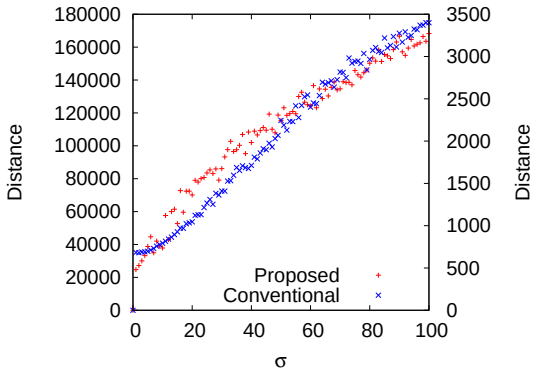
(a)$(p, q) = (0.0001, 0.0001)$



(b)$(p, q) = (0.0001, 104.8576)$



(c)$(p, q) = (104.8576, 0.0001)$



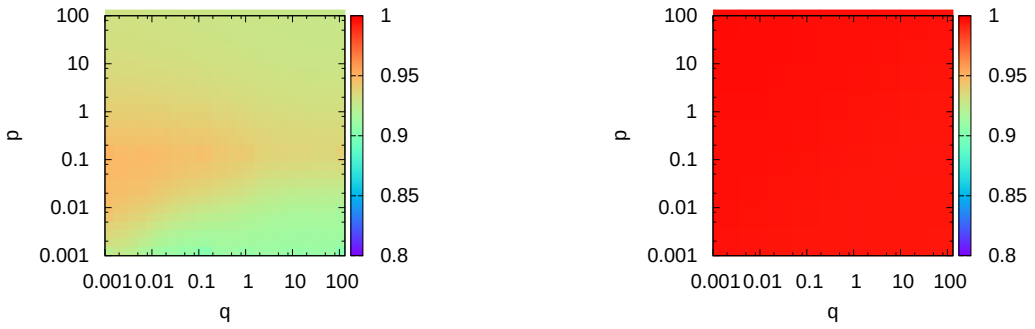
(d)$(p, q) = (13.1072, 13.1072)$

Figure 5.8: Relations between distance and noise strength. Left vertical axis indicates distance calculated by the proposed measure. Right vertical axis indicates distance calculated by the conventional measure.

Estimation of network structures

Using the proposed measure, we estimated network structures from marked point process data. In the simulation, we used coupled Rössler systems:

$$\begin{cases} \dot{x}_i &= -\omega y_i - z_i + \sum_j k_{ij} x_j, \\ \dot{y}_i &= \omega x_i + a y_i, \\ \dot{z}_i &= b + x_i z_i - c z_i. \end{cases} \tag{5.11}$$



(a) (b)

Figure 5.9: Correlation coefficients between the noise strength s and the proposed measure $D(M_1, M_2)$ when the parameters p and q are changed.

The parameters are set to $a = 0.15, b = 0.2, c = 10$. These parameters lead to chaotic behaviors of the systems. In Eq. (5.11), ω is Gaussian random number with a mean value and standard deviation of unity and 0.01, respectively. The coupling strengths are set to $k_{ij} = 0.2$ if the ith and the jth systems are coupled, otherwise these are set to $k_{ij} = 0$. Then we translated continuous time series x_i into marked point processes. We defined the lth event timing t_i^l as the time when $x_i(t)$ takes the lth local maxima or local minima and the amplitude of the lth event as $x_i(t_i^l)$.

We conducted numerical experiments according to the following procedures. First, we generated marked point processes from $x_i(t)$ of the coupled Rössler systems. Second, we calculated the STM for marked point processes. Here, although the parameters p and q in the STM have to been appropriately determined, we have not estimated optimal values of parameters because it is clarified that the proposed measure does not depend on the parameter values of p and q. Thus, in this paper, we heuristically decided the parameters values. Third, we obtained position vectors by applying the MDS to the STM. Then we calculated the partial correlation coefficients between the position vectors. If two systems are coupled, the partial correlation coefficient is close to unity. Otherwise, it is close to zero. Then we classified whether pairs of systems are coupled or not using histogram of the partial correlation coefficient by the Otsu thresholding method [**?**]. If the partial correlation coefficient is larger than a threshold, we defined the pair as a coupled pair. Otherwise, we defined it as an uncoupled pair.

We checked the proposed method with a ring lattice network structure which has 10 nodes and

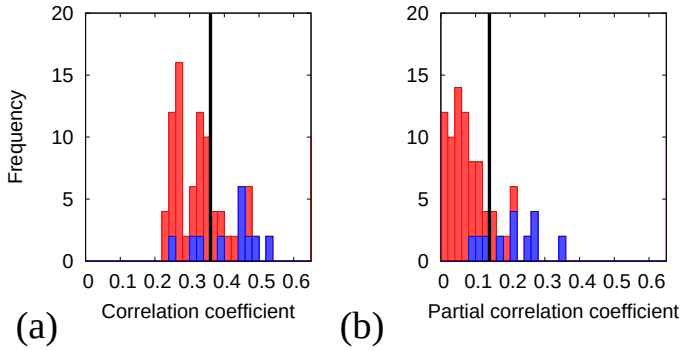


Figure 5.10: Histograms of absolute values of (a) the correlation coefficient and (b) the partial correlation coefficient in case that only event timings are used. Histograms of all coefficients are shown in red, and those of coupled pairs are superimposed by blue. The vertical lines indicate thresholds.

degree 2 of each node. First, in Fig. 5.10, we show histograms of absolute values of the correlation coefficients and the partial correlation coefficients in case that only event timings are used. In Fig. 5.10(a) and (b), some pairs are misestimated with correlation coefficients, however the partial correlation coefficients can more clearly distinguish coupled pairs and uncoupled pairs

Next, we show histograms of the absolute values of the correlation coefficients and the partial correlation coefficients in case that both event timings and amplitude information of marked point processes are used. In Fig. 5.11(a), the correlation coefficients of uncoupled pairs take large values. Thus distributions of coupled pairs and uncoupled pairs overlap, and consequently coupled pairs and uncoupled pairs are not clearly distinguished. However, coupled pairs and uncoupled pairs are clearly distinguished by using the partial correlation coefficient (Fig. 5.11(b)). The results also indicate that the proposed measure using mark information is naturally appropriate for marked point processes.

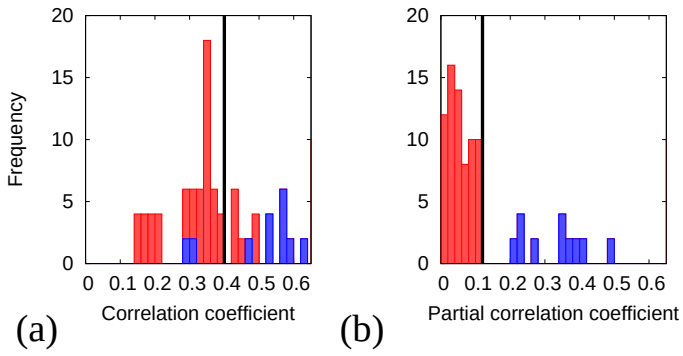


Figure 5.11: Histograms of absolute values of (a) the correlation coefficient and (b) the partial correlation coefficient in case that both event timings and amplitude information are used. Histograms of all coefficients are shown in red, and those of coupled pairs are superimposed by blue. The vertical lines indicate thresholds.

5.3 Estimation of Evolving Network Structure

In section , we estimated static network structures. However, it is also important to estimate dynamic structures, or to detect how the neural network structure changes, because one of the intrinsic properties in neural networks is learning. When the neural networks accept external stimulation, neural networks change their structures by learning.

In this thesis, we proposed a method for estimating evolution of neural network structures by applying the methods of estimating static neural network structures to the case that neural networks dynamically evolve. To check the validity of the proposed method, we conducted numerical experiments by using a neural network model with a learning rule of spike-timing-dependent plasticity [43]. In the experiments, we first observed the multi-spike sequences from the neural network with the STDP learning [43]. Next, we divided the observed multi-spike sequences into small temporal epochs. Then, we applied the method of section 5.1 to the temporally divided multi-spike sequences and estimated their connectivities. As a result, we could estimate the evolving neural network structure with high estimation accuracy.

5.3.1 Method

To estimate connectivity of neurons only from spike sequences, we have already proposed a spike time metric coefficient (STMC) and a partial spike time metric coefficient (PSTMC).

To estimate the direction of couplings between neurons, we defined a directional spike time metric. We calculated the spike time metric between two spike sequences $X_i(t)$ and $X_j(t+\tau)$ defined as

$$D_{ij}(\tau) = D(X_i(t), X_j(t+\tau)), \tag{5.12}$$

where τ is a temporal difference between two spike sequences X_i and X_j. We distinguish the direction of couplings by whether the difference τ at the minimum value of $D_{ij}(\tau)$ is positive or not. If the difference τ at the minimum value of $D_{ij}(\tau)$ is positive, we judged that the direction of coupling is from the ith neuron to the jth neuron.

5.3.2 Numerical Simulation

To evaluate the validity of our method, we used a neural network constructed from a mathematical model of Izhikevich's simple neuron model [14] and generated multi-spike sequences. The number of neurons is 100. The neural network is composed of only excitatory neurons which are regular spiking neurons. Each neuron connects to 10 postsynaptic neurons.

We use an STDP function proposed by Song et al. [43] which is defined by

$$\Delta g = \begin{cases} A_p e^{-|\Delta t|/\tau_p} & \text{if } \Delta t > 0, \\ -A_d e^{-|\Delta t|/\tau_d} & \text{otherwise,} \end{cases} \tag{5.13}$$

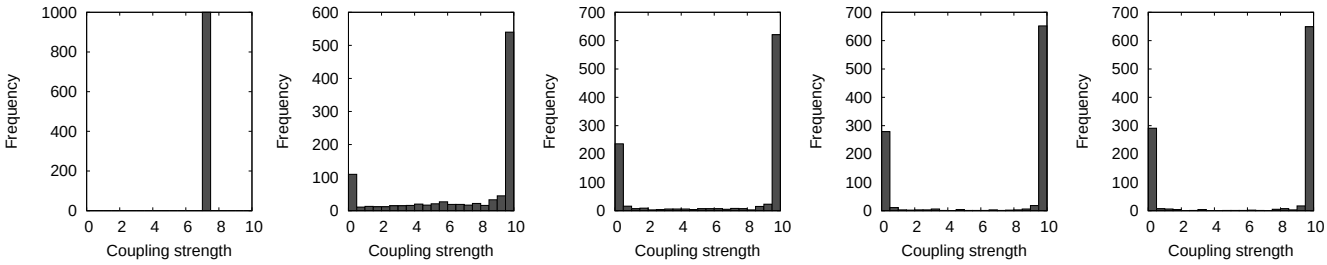
where A_p and A_d are the learning rates of the long-term potentiation (LTP) and depression (LTD), and τ_p and τ_d are the time constants that determine the exponential decays of the LTP and the LTD. We set the parameters $A_p = 0.01$, $A_d = 0.012$, and $\tau_p = \tau_d = 20$ [ms]. In Eq. (5.13), the variable Δt represents a relative spike timing between a presynaptic and a postsynaptic neuron. The coupling strength is updated as $g \leftarrow g + \Delta g$ at every second. The coupling strength is limited between 0 to 10. At an initial condition, we set that the coupling strength is 7 and an initial network structure is a random network.

We conducted numerical experiments in the following way. First, we generated multi-spike sequences by constructing a neural network using Izhikevich's simple neuron model and the STDP rule of Eq. (5.13). Next, we calculated the PSTMC between spike sequences for 1,000 [s]. Then, we divided this total temporal length of 1,000 [s] into small temporal windows. The length of the small temporal window is 100 [s]. Using multi-spike sequences in this 100 [s] small windows, we applied the methods of Section 5.1. We classified coupled pairs and uncoupled pairs by the Otsu thresholding [17]. Then, we estimated the direction of couplings by calculating $D_{ij}(\tau)$ for the estimated coupled pairs. Finally, we evaluated the estimation accuracy.

5.3.3 Results

Figure 5.12 shows results of histograms of coupling strength. From these results, we can see that the distribution of the coupling strength changes, namely, the network structure evolves. We also show results of estimation accuracy of the evolving neural network structures in Fig. 5.13. From the results, the estimation accuracy $\mathrm{C}-\tilde{\mathrm{C}}$ takes relatively a low value at 0 [s]. This reason is that the coupling strength changes rapidly at the initial stage with the STDP. Namely, it is relatively hard to estimate the structure because the evolution of the network structure has fast dynamics.

However, the value of both of U−Ũ takes relatively a high value. The temporal epoch proceeds, the estimation accuracy converges to higher values. It means that the proposed method can detect the evolution of STDP neural network structures.



(a) 0 [s] (b) 250 [s] (c) 500 [s] (d) 750 [s] (e) 1000 [s]

Figure 5.12: Histograms of coupling strength at each time.

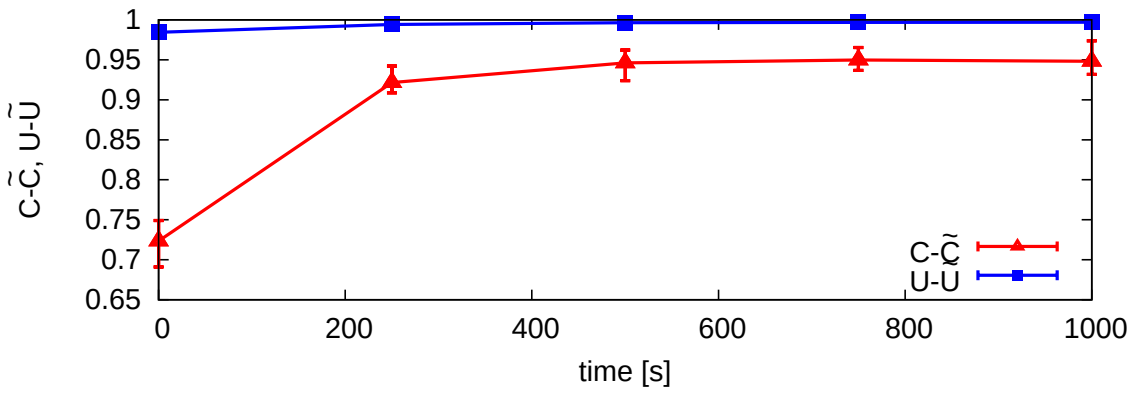


Figure 5.13: Estimation accuracy of the evolving neural network structures. Red line shows C−C̃ and blue line shows U−Ũ. Error bars indicate minimum and maximum values with 20 trials are also provided.

5.4 Summary of Chapter 5

In Chapter 5, we proposed the estimation method of network structures only from point processes by using distance between point processes and the MDS. First, we proposed a new method for estimating connectivity between neurons only from observed multiple spike sequences. In our method, using the MDS, we represented a relationship among spike sequences in the Euclidean space from the STM to adapt observed position vectors to a linear regression model. By calculating the PDC from the position vectors, we estimated connectivities between neurons. As a result, we could estimate the connectivity between neurons using our method.

In addition, we extended the method for marked point processes. As a result, using not only event timings but also corresponding amplitude information, we could estimate more precisely the connectivity between systems than using only event timings.

Furthermore, we proposed a method for estimating evolution of neural network structures through the spike-timing-dependent plasticity by applying the methods of estimating static neural network structures to the case that neural networks dynamically evolve. As a result, the proposed method can estimate the evolving neural network structures and the direction of couplings with high estimation accuracy.

Chapter 6

Conclusions

In this thesis, we proposed estimation methods of network structures only from point processes. In the proposed methods, we introduced three strategy: transforming point processes into continuous time series, using normalized distance between point processes, and using multi-dimensional scaling with distance between point processes.

In Chapter 3, we proposed a method of estimating network structures by transforming point processes into continuous time series. To transform point processes into continuous time series, we used the kernel density function. When the kernel function is used, we have to select an optimal bandwidth because the transformed time series depends on the bandwidth. Then we used two selection methods of the kernel bandwidth. As the first method, we used a kernel bandwidth optimization method which is proposed by Shimazaki et al. [25]. The method is used in spike rate estimation. As the second method, we used a selection method of an optimal time delay in the attractor reconstruction [34] which is a basic technique in chaotic time series analysis. By using these methods, we found that high correlation coefficients and low normalized mean square errors are evaluated between the input time-series and its reconstructed time-series.

Using the transformation method, we estimated network structures from simple point processes. To evaluate the validity of the proposed method, we used neural networks constructed from a mathematical model of the Izhikevich simple neuron model [14] and generated multi-spike sequences. The results show that our method exhibits high performance.

In additon, for marked point processes, we extended the proposed method. We proposed a

method of transforming marked point processes into continuous time series which is based on the kernel density estimator. Then we applied the partialization analysis to the transformed time series. As a result, we can estimate the connectivity of the coupled Lorenz systems from marked point processes.

In Chapter 4, we proposed two measures, spike time metric coefficient and partial spike time metric coefficient, to estimate network structures only from the information of observed asynchronous spike sequences. We applied the proposed measures to simple networks and complex networks [15, 16]. The results show that network structures can be estimated by adaptive application of spike time metric coefficients and partial spike time metric coefficients with high performance. We heuristically defined the PSTMC from association of the conventional partialization analysis. Then we experimentally confirmed that the PSTMC can measure true relations between spike sequences by removing spurious correlations. However, we have not theoretically proven why the PSTMC works well yet. As a future work, we have to prove that why the PSTMC works well.

In Chapter 5, we proposed the estimation method of network structures only from point processes by using distance between point processes and the MDS. First, we proposed a new method for estimating connectivity between neurons only from observed multiple spike sequences. In our method, using the MDS, we represented a relationship among spike sequences in the Euclidean space from the STM to adapt observed position vectors to a linear regression model. By calculating the PDC from the position vectors, we estimated connectivities between neurons. As a result, we could estimate the connectivity between neurons using our method. In addition, we extended the method for marked point processes. As a result, using not only event timings but also corresponding amplitude information, we could estimate more precisely the connectivity between systems than using only event timings. Furthermore, we proposed a method for estimating evolution of neural network structures through the spike-timing-dependent plasticity by applying the methods of estimating static neural network structures to the case that neural networks dynamically evolve. As a result, the proposed method can estimate the evolving neural network structures and the direction of couplings with high estimation accuracy.

Bibliography

[1] Y. Ogata, “Space-Time Point-Process Models for Earthquake Occurrences,” *Ann. Inst. Statist. Math.* **50**(2):379, 1998 .

[2] J.H. Dshalalow, “On multivariate antagonistic marked point processes” *Math. Comput. Model.* **49**:432, 2009.

[3] D. J. Daley and D. Vere-Jones, “An Introduction to the Theory of Point Processes,” *Springer-Verlag*, 2008.

[4] S. Frenzel and B. Pompe, “Partial Mutual Information for Coupling Analysis of Multivariate Time Series,” *Phys. Rev. Lett.*, **99** 204101, 2007.

[5] B. Schelter, M. Winterhalder, R. Dahlhaus, J. Kurths and J. Timmer, “Partial Phase Synchronization for Multivariate Synchronizing Systems,” *Phys. Rev. Lett.*, **96** 208103, 2006.

[6] L.A. Baccala and K. Sameshima, “Partial directed coherence: a new concept in neural structure determination,” Biological Cybernetics, **84**:463, 2001.

[7] K. Sameshima and L.A. Baccala, “Using partial directed coherence to describe neuronal ensemble interactions,” Journal of Neuroscience Methods, Vol. 94, pp. 93–103, 1999.

[8] T.F. Cox and M.A.A. Cox, HAPMAN&HALL/CRC, 2000.

[9] J. Zhou and J. Lu, “Topology Identification of Weighted Complex Dynamical Networks,” *Physica A*, **386**:481, 2007.

[10] D. Yu, M. Righero and L. Kocarev, “Estimating Topology of Networks,” *Phys. Rev. Lett.*, **97** 188701, 2006.

[11] M. Eichler, R. Dahlhaus and J. Sandkuhler, “Partial Correlation Analysis for the Identification of Synaptic Connections,” *Biological Cybernetics*, **89**:289, 2003.

[12] T. Ashizawa et al., “New measures for estimating neural network only from multi-spike sequences,” *Proc. of NOLTA’07*, 417, 2007.

[13] J. D. Victor and K. P. Purpura, “Metric-space Analysis of Spike Trains: Theory, Algorithms and Application,” *Network: Comput. Neural Syst.*, **8**:127, 1997.

[14] E. M. Izhikevich, “Simple Model of Spiking Neurons,” *IEEE Trans. on Neural Networks*, **14**:1569, 2003.

[15] D. J. Watts and S. H. Strogatz, “Collective Dynamics of ’Small-world’ Networks,” *Nature*, **393**:440, 1998.

[16] A. L. Barabasi and R. Albert, “Emergence of Scaling in Random Networks,” *Science*, **286**:509, 1999.

[17] N. Otsu, “A Threshold Selection Method from Gray Level Histograms,” *IEEE Trans. on Syst., Man and Cybernetics*, **9**:62, 1979.

[18] W. M. Kistler, W. Gerstner and J. L. Hemmen, “Reduction of the Hodgkin-Huxley Equations to a Single-Variable Threshold Model,” *Neural Comput.* **9**:1015, 1997.

[19] 田中豊，脇本和昌，“多変量統計解析法,” 現代数学社，1983.

[20] M. Hénon, “A two-dimensional mapping with a strange attractor,” *Commun. Math. Phys.* **50**:69, 1976.

[21] K. Aihara, T. Takabe and M. Toyoda, “Chaotic neural networks,” *Phys. Lett. A.* **144**:333, 1990.

[22] Y. Ueda, “Randomly transitional phenomena in the system governed by Duffing’s equation,” *J. Stat. Phys.* **20**:2, 1979.

[23] H. Fujii, H. Ito, K. Aihara, N. Ichinose and M. Tsukada, “Dynamical Cell Assembly Hypothesis-Theoretical Possibility of Spatio-Temporal Coding in the Cortex,” *Neural Networks*, **9**:1303, 1996.

[24] O. E. Rössler, “An equation for continuous chaos,” *Phys. Lett. A.* **57**:397, 1976.

[25] H. Shimazaki and S. Shinomoto, “Kernel Bandwidth Optimization in Spike Rate Estimation,” *J. Comput. Neurosci*, **29**:171, 2009.

[26] T. Sauer, “Reconstruction of Dynamical Systems from Interspike Intervals,” Physical Review Letters, Vol. 72, No. 24, pp. 3811–3814, 1994.

[27] T. Sauer, “Interspike interval embedding of chaotic signals,” Chaos, Vol. 5, No. 1, p. 127, 1995.

[28] H. Suzuki, K. Aihara, J. Murakami and T.Shimozawa, “Analysis of Neural Spike Trains with Interspike Interval Reconstruction”, Biological Cybernetics, Vol.82, pp. 305–311, 2000.

[29] F. Takens, “Detecting strange attractors in turbulence,” In D. A. Rand and B. S. Young, editors, “Dynamical Systems of Turbulence,” Vol. 898 of Lecture Notes in Mathematics, pp. 366–381, Berlin, 1981. Springer-Verlag.

[30] T. Sauer, J. A. Yorke and M. Casdagli, “Embedology,” Journal of Statistical Physics, Vol. 65, pp.579–616, 1991.

[31] 堀玄，岡田真人，深井朋樹，“スパイク間隔による埋め込みの拡張について,” 電子情報通信学会技術研究報告, Vol. 105, No. 139, pp. 1–5, 2005.

[32] J. J. Collins, C. C. Chow and T. T. Imhoff, “Aperiodic stochastic resonance in excitable systems,” Physical Review E, Vol. 52, No. 4, pp. R3321–R3324, 1995.

[33] C. J. de Luca, R. S. LeFever, M. P. McCue and A. P. Xenakis, “Behaviour of human motor units in different muscles during linearly varying contractions,” The Journal of Physiology, Vol. 329, pp. 113–128, 1982.

[34] 池口徹, 山田泰司, 小室元政,“カオス時系列解析の基礎と応用,” 合原一幸 編, 産業図書, 2000.

[35] Andrew M. Fraser and Harry L. Swinney, “Independent coordinates for strange attractors from mutual information,” *Phys. Rev. A*, **33**(2):1134, 1986.

[36] W. F. Langford, “Numerical Studies of torus bifurcations,” *International Series of Numerical Mathematics*, **70**:285, 1984.

[37] E. N. Lorenz, “Deterministic Nonperiodic Flow,” *Journal of the atmospheric sciences*, **20**:130, 1963.

[38] C.W.J. Granger, Econometrica **37**:424, 1969.

[39] S. Suzuki et al., “Definition of Distance for Marked Point Process Data and Its Application to Recurrence Plot-Based Analysis of Exchange Tick Data fo Foreigh Currencies,” *IJBC*, **20**:3699, 2010.

[40] L. Zhu, Y. C. Lai, F. C. Hoppensteadt and J. He: Neural Computation, Vol. 15, pp. 2359–2377, 2003.

[41] A. Nedungadi, G. Rangarajan, N. Jain, and M. Ding: Journal of Computational Neuroscience, Vol. 27, “Analyzing multiple spike trains with nonparametric granger causality,” pp.55–64, 2009.

[42] R. G. Andrzejak and T. Kreuz: Europhysics Letters, Vol. 96, 50012, 2011.

[43] S. Song, K.D. Miller, L.F. Abbott: “Competitive Hebbian learning through spike-timing-dependent synaptic plasticity,” Nature Neuroscience, Vol. 3, No. 9, pp. 919–926, 2000.

Publications

Journal Papers

- 黒田佳織，島田裕，鈴木麻衣，池口徹, “スパイク列から瞬時平均発火時系列への変換を用いたニューロンへの入力情報の再構成,” 電子情報通信学会論文誌, Vol. J94-A, No. 2, 2011.
- Kaori Kuroda, Tohru Ashizawa, and Tohru Ikeguchi, “Estimation of Network Structures only from Spike Sequences,” Physica A, Vol.390, pp.4002–4011, 2011.
- Kaori Kuroda, Kantaro Fujiwara, and Tohru Ikeguchi, “Identification of neural network structure from multiple spike sequences,” Lecture Notes in Computer Science, Vol.7664, pp.184–191, 2012.

International Conference Proceedings

- Kaori Kuroda and Tohru Ikeguchi,”Reconstructing scale-free structure of neural networks from multi-spike sequences,” Proceedings of the International Symposium on Nonlinear Theory and its Applications, pp.54–57, Oct. 2009.
- Kaori Kuroda and Tohru Ikeguchi,”Estimation of Neural Network Structure by Transforming Spike Sequences to Continuous Time Series,” Proceedings of the International Symposium on Nonlinear Theory and its Applications, pp. 123–126, Sep. 2010.
- Kaori Kuroda and Tohru Ikeguchi,”Reconstructing connectivity of elements in nonlinear dynamical systems from marked point processes,” Proceedings of the International Symposium on Nonlinear Theory and its Applications, pp.697–700, Sep. 2011.
- Kaori Kuroda and Tohru Ikeguchi, “Estimation of Input Information Applied to Neurons by Local Adaptive Kernel Density Function,” International Joint Conference on Neural Networks, Aug. 2011.
- Kaori Kuroda and Tohru Ikeguchi, “Estimation of Connectivities of Neurons Using Multi-Dimensional Scaling,” Proceedings of International Symposium on Nonlinear

Theory and its Applications, pp.382–385, Oct. 2012.

- Kaori Kuroda and Tohru Ikeguchi, “Estimation of neural network structures and direction of couplings using spike time metric,” Proceedings of the International Workshop on Nonlinear Circuits and Signal Processing, pp. 504–507, Mar. 2013.
- Kaori Kuroda, Kantaro Fujiwara and Tohru Ikeguchi, “Detection of Learning in Neural Networks Only from Spike Sequences,” Proceedings of International Symposium on Nonlinear Theory and its Applications, pp.142-145, Sep. 2013.
- Yong Gao, Kaori Kuroda, Yutaka Shimada, Kantaro Fujiwara, Tohru Ikeguchi, “Classification of Real Networks by Using Classical Multidimensional Scaling,” Proceedings of International Symposium on Nonlinear Theory and its Applications, pp. 126–129, Sep. 2013.
- Akio Watanabe, Kaori Kuroda, Kantaro Fujiwara, Tohru Ikeguchi, “A Chaotic Local Search Algorithm with Adaptive Exchange of Elements for Quadratic Assignment Problems,” Proceedings of International Symposium on Nonlinear Theory and its Applications, pp. 362–365, Sep. 2013.

National Conference Proceedings

- Kaori Kuroda and Tohru Ikeguchi, “Partialization Analysis for Nonlinear Connections of Second Order”, 信学技報, Vol.109, No.269, pp.115-120, 2009 年 11 月.
- Kaori Kuroda, Atsushi Uchida and Tohru Ikeguchi, “Estimation of connectivity of nonlinear dynamical systems by partial correlation analysis”, 信学技報, Vol.109, No.366, pp.1–5, 2010 年 1 月.
- 鈴木麻衣，黒田佳織，島田裕，池口徹, “スパイク列から連続時系列への変換を用いた入力情報の再構成”, 信学技報, Vol.109, No.458, pp.9–14, 2010 年 3 月.
- Kaori Kuroda and Tohru Ikeguchi, “Transforming spike sequence to continuous time series to estimate neural network structure,” 2010 年電子情報通信学会 総合大会, A-2-21, p.65, 2010 年 3 月.
- Kaori Kuroda and Tohru Ikeguchi, “A method for transforming marked point process to continuous time series,” 2010 年電子情報通信学会ソサイエティ大会, A-2-3，2010 年 9 月.
- Kaori Kuroda and Tohru Ikeguchi, “Adaptive Reconstruction of Input Information Applied to Neurons from Spike Trains,” 2011 年電子情報通信学会 総合大会, A-2-4, 2011 年 3 月.
- Yong Gao, Kaori Kuroda, Yutaka Shimada, Kantaro Fujiwara, Tohru Ikeguchi,”Analysis on real networks by classical multidimensional scaling,” 信学技報, Vol.112, No.487, NLP2012-162, pp.91–96, 2013 年 3 月.

- Suguru Yaginuma, Kaori Kuroda, Yutaka Shimada, Kantaro Fujiwara, Tohru Ikeguchi, ”Prediction of growth of complex networks,” 信学技報, Vol. 112, No. 487, NLP2012-163, pp. 97–102, 2013 月 3 月.
- Tsubasa Kawai, Ryoya Kusumi, Kaori Kuroda, Hideyuki Kato, Kantaro Fujiwara, Kenya Jin’no and Tohru Ikeguchi, “Experimental observation of common noise-induced synchronization in chaotic oscillators,” 電子情報通信学会技術研究報告, Vol. 112, No. 487, NLP2012-172, pp. 149–154, 2013 年 3 月.
- 渡辺明生，黒田佳織，藤原寛太郎，池口徹，“二次割当問題の交換要素を適応的に決定するカオス的局所探索法”， 信学技報, Vol. 112, No. 487, NLP2012-145, pp. 7–12, 2013 年 3 月.
- Kaori Kuroda, Kantaro Fujiwara, and Tohru Ikeguchi, “Estimation of connectivity of systems from distance relation between marked point processes,” 2013 年電子情報通信学会ソサイエティ大会，AS-1-7，2013 年 9 月.
- Akio Watanabe, Kaori Kuroda, Kantaro Fujiwara, Tohru Ikeguchi, “Improvement by Chaotic Neuro Dynamics of an Adaptive Local search Algorithm for solving Combinatorial Optimization Problems,” 2013 年電子情報通信学会ソサイエティ大会，AS-1-2，2013 年 9 月.

Acknowledgments

I would particularly like to thank Prof. Tohru Ikeguchi. Prof. Tohru Ikeguchi gives me fruitful comments and advices not only for my studies but also laboratory life, thesis writing, presentation skills, and so on. I am deeply grateful to Prof. Tohru Ikeguchi.

I would like to thank Professor Yutaka Ohsawa, Professor Takaomi Shigehara, Associate Professor Atsushi Uchida, Associate Professor Takashi Komuro for judging my thesis.

I want to thank Prof. Yoshihiko Horio (Tokyo Denki University), Prof. Masaharu Adachi (Tokyo Denki University), Prof. Jin'no (Nippon Institute of Technology), Prof. Mikio Hasegawa (Tokyo University of Science) and Prof. Hiroo Sekiya (Chiba University) who also gave me insightful comments in academic meetings.

I express my appreciation to Dr. Takafumi Matsuura, Dr. Hideyuki Kato, Dr. Yutaka Shimada who gave me constructive comments and suggestions.

I also thank all members of Ikeguchi Laboratory for having many opportunities to discuss studies.

Finally, my special thanks go to my parents and sisters who gave warm encouragements and invaluable supports.