

## Graph structure modeling for multi-neuronal spike data

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2016 J. Phys.: Conf. Ser. 699 012012

(<http://iopscience.iop.org/1742-6596/699/1/012012>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 142.66.184.160

This content was downloaded on 17/03/2017 at 18:23

Please note that [terms and conditions apply](#).

You may also be interested in:

[Effect of Topology Structures on Synchronization Transition in Coupled Neuron Cells System](#)

Liang Li-Si, Zhang Ji-Qian, Xu Gui-Xia et al.

# Graph structure modeling for multi-neuronal spike data

Shotaro Akaho<sup>1</sup>, Sho Higuchi<sup>2</sup>, Taishi Iwasaki<sup>3</sup>, Hideitsu Hino<sup>4</sup>,  
Masami Tatsuno<sup>5</sup> and Noboru Murata<sup>3</sup>

<sup>1</sup> The National Institute of Advanced Industrial Science and Technology, 1-1-1 Umezono, Tsukuba, Ibaraki 305-8568 Japan

<sup>2</sup> NTT DATA Corporation, 3-9 Toyosu 3-chome, Koto-ku, Tokyo 135-8671 Japan

<sup>3</sup> Waseda University, 3-4-1 Okubo, Shinjuku-ku, Tokyo 169-8555 Japan

<sup>4</sup> University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8577 Japan

<sup>5</sup> University of Lethbridge, Lethbridge, AB, T1K3M4, Canada

E-mail: s.akaho@aist.go.jp

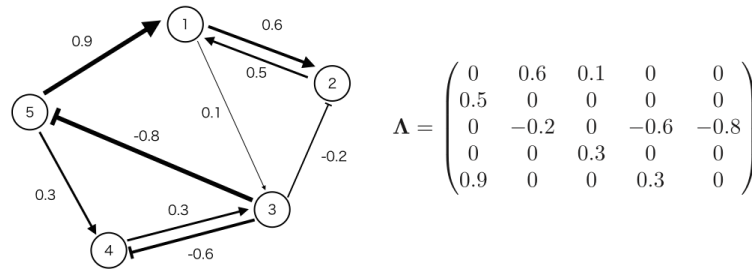
**Abstract.** We propose a method to extract connectivity between neurons for extracellularly recorded multiple spike trains. The method removes pseudo-correlation caused by propagation of information along an indirect pathway, and is also robust against the influence from unobserved neurons. The estimation algorithm consists of iterations of a simple matrix inversion, which is scalable to large data sets. The performance is examined by synthetic spike data.

## 1. Introduction

Due to recent advances in experimental technologies, we can record activities of a large number of neurons simultaneously. It has been suggested that cell assemblies can be formed dynamically in the course of motor and cognitive functions [1]. In order to understand how the cell assemblies contribute to brain functions, it is important to clarify the relationship between neurons. For example, Wilson and McNaughton estimated a functional connection in the hippocampal CA1 network using a cross-correlation function [2]. Recently, researchers have also started investigating higher-order interactions using information-theoretic approaches [3–5]. However, many of the existing methods have the following difficulties:

- (i) Pseudo-correlation: Neuron activities can propagate to other neurons via indirect pathways. This causes pseudo correlation.
- (ii) Direction of connection: Functional connectivity estimated by correlation-based methods is often symmetric. It is not possible to extract information about the direction of connections.
- (iii) Openness of the system: In neurophysiological experiments, only a limited number of neurons can be recorded. The activity of measured neurons is also affected by other neurons that are not observed.
- (iv) Positive and negative connectivity: The previously proposed method based on diffusion process [6] was developed to analyze a network of excitatory (positive) connections. However, there are also inhibitory (negative) connections in the brain.





**Figure 1.** An example of a graph structure representing direct connectivity  $\Lambda$

To overcome these difficulties, we propose a novel method to extract dependency from multivariate time series data.

## 2. Model

Suppose there are  $N$  neurons measured, membrane potential  $v_i(t + \Delta)$  of the neuron  $i$  at time  $t + \Delta$  is affected by the neuron  $j$  at time  $t$  by

$$v_i(t + \Delta) = b_i(t) + \theta_{ij} Y_j(t), \quad (1)$$

where  $\Delta$  is a small period of time,  $Y_j(t)$  is a binary random variable representing spike of the neuron  $j$ , and  $b_i(t)$  is input from all other neurons including unmeasured ones except the neuron  $j$ . The coefficient  $\theta_{ij}$  represents the connection from the neuron  $j$  to the neuron  $i$ , including not only a direct connection but also indirect connection through two-steps, three-steps and longer pathways through other neurons. Let  $\lambda_{ij}$  be the direct connection from the neuron  $j$  to the neuron  $i$ . We assume that indirect connection propagates in a multiplicative way. For instance, the strength of connection from the neuron  $j$  to the neuron  $i$  through the neuron  $k$  is equal to  $\xi_k^{(1)} \lambda_{ik} \lambda_{kj}$ , where  $\xi_k^{(1)}$  denotes a decaying factor. Therefore  $\theta_{ij}$  is decomposed as the sum of multi-step connections,

$$\theta_{ij} = \lambda_{ij} + \sum_k \xi_k^{(1)} \lambda_{ik} \lambda_{kj} + \sum_{k,l} \xi_{kl}^{(2)} \lambda_{ik} \lambda_{kl} \lambda_{lj} + \dots, \quad (2)$$

where in this paper we assume there are no direct self-connection, hence  $\lambda_{ii} = 0$ . For simplicity, we further assume decaying factors  $\xi_k^{(1)}$ ,  $\xi_{kl}^{(2)}$ ,  $\dots$  are all equal to one by constraining  $\lambda_{ij} < 1$  to incorporate the decaying effect in  $\lambda_{ij}$ . As a result, we have a simple relation between the accumulated total connection  $\theta_{ij}$  and the direct connection  $\lambda_{ij}$ ,

$$\Theta = \Lambda + \Lambda^2 + \dots = (I - \Lambda)^{-1} - I, \quad (3)$$

where  $\Theta$  and  $\Lambda$  are matrices whose  $ij$  elements are  $\theta_{ij}$  and  $\lambda_{ij}$  respectively.

In this paper, we assume that we can construct a matrix  $X$  whose off-diagonal element is related to  $\Theta$  as

$$x_{ij} = \alpha \theta_{ij} + \epsilon_{ij}, \quad i \neq j \quad (4)$$

where  $x_{ij}$  is the  $ij$  element of  $X$ ,  $\alpha$  is a scalar parameter, and  $\epsilon_{ij}$  is a random noise. We assume the diagonal elements of  $X$  are not obtained, and let  $\text{diag} X = O$ , then we have

$$X + D \simeq \alpha \Theta = \alpha \{ (I - \Lambda)^{-1} - I \}, \quad (5)$$

where  $D$  is a diagonal matrix representing unknown diagonal elements.

As a consequence, the problem is to obtain the connectivity matrix  $\Lambda$  for given  $X$  by estimating parameters  $D$ ,  $\alpha$  and solving eq. (5). An example of  $\Lambda$  is shown in figure 1, which can be regarded as weighted edges of a graphical structure.

In the rest of the paper, first we propose an iterative algorithm to find  $\alpha$ ,  $D$ ,  $\Lambda$  for a given  $X$ . Next, we give a way of constructing  $X$  from measured neural spike data, and show some simulation results.

### 3. Optimization algorithm

The objective function to be minimized is the squared loss defined from eq. (5),

$$J(\alpha, D, \Lambda) = \|X + D - \alpha\{(I - \Lambda)^{-1} - I\}\|_F^2 \quad (6)$$

where  $\|\cdot\|_F$  denotes the Frobenius norm.

We further introduce *sparsity* of the connectivity matrix  $\Lambda$  as a constraint. Otherwise, most neurons are connected each other, which makes it difficult to interpret the results. The sparsity constraint is also consistent with a neuroscientific finding that approximately 10% of excitatory neurons are connected [7]. Here we use  $L_0$  norm constraint as sparsity, then the problem is to minimize the objective function  $J(\alpha, D, \Lambda)$  under the constraint

$$\|\Lambda\|_0 \leq M. \quad (7)$$

In this paper, the hyperparameter  $M$ , which represents an upperbound of the number of nonzero components of  $\Lambda$ , is estimated by StARS criterion as described later.

#### 3.1. Objective minimization

We apply alternating iterative procedures to minimize the objective function  $J(\alpha, D, \Lambda)$ , i.e., each parameter  $\alpha$ ,  $D$  or  $\Lambda$  is optimized by fixing the other parameters. For sparsity, we apply a greedy approach.

First,  $D$  is initialized by  $rI$ , where  $r$  is a constant so that  $\det(X + D) \neq 0$ .  $\alpha$  is randomly initialized from uniform distribution on  $(0, 1]$ . Fixing  $D$  and  $\alpha$ , we obtain  $\hat{\Lambda}'$ , which is an estimator of  $\Lambda$ , by solving eq. (5) and taking  $\lambda_{ii} = 0$  into account,

$$\hat{\Lambda}' = \text{offdiag} \left[ - \left( \frac{1}{\alpha}(X + D) + I \right)^{-1} \right], \quad (8)$$

where  $\text{offdiag}[A]$  is an operator that sets  $A$ 's diagonal elements to zero. Since  $\hat{\Lambda}'$  is not a sparse matrix, we obtain a sparse estimate of  $\Lambda$  by cutting off the small values of  $\hat{\Lambda}'$ ,

$$\hat{\lambda}_{ij} = \begin{cases} \hat{\lambda}'_{ij} & \text{if } |\hat{\lambda}'_{ij}| \geq \zeta \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

where  $\zeta$  is the  $M$ -th largest value among  $\{|\hat{\lambda}'_{ij}|\}$ .

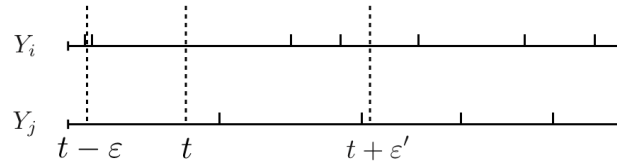
In order to optimize  $\alpha$  for a fixed  $\hat{\lambda}_{ij}$ , by letting  $D = \text{diag}(\alpha(I - \hat{\Lambda})^{-1} - I)$  in  $J(\alpha, D, \hat{\Lambda})$ , we have a quadratic optimization function, and the solution is given by

$$\hat{\alpha} = \frac{\sum_{i \neq j} x_{ij} \left[ (I - \hat{\Lambda})^{-1} - I \right]_{ij}}{\sum_{i \neq j} \left[ (I - \hat{\Lambda})^{-1} - I \right]_{ij}^2}. \quad (10)$$

Then  $D$  is updated by

$$D = \text{diag} \left( \hat{\alpha} \left\{ (I - \hat{\Lambda})^{-1} - I \right\} \right), \quad (11)$$

and the procedures above are repeated for a predefined number of times.



**Figure 2.** Two time-windows for measuring spikes

### 3.2. Hyperparameter optimization

In order to determine the number of nonzero elements  $M$  in  $\Lambda$ , AIC and BIC are often used. However, in this paper, we apply StARS (Stability Approach to Regularization Selection) [8] because neural data is noisy and thus the criterion is required to be robust against noise. StARS tends to give rise to less sparse graphs than AIC and BIC. Thus it is more suitable if we want to avoid relevant connections are omitted.

Since the original proposal of StARS is formulated for undirected graph, in the following, we describe the method extended to the case of directed graph.

Suppose we have multidimensional neural spike time series ( $N$  neurons  $\times$  time length  $T$ ). We randomly extract  $K$  subsequences  $S_1, S_2, \dots, S_K$ , where each  $S_i$  has time length  $b$  (i.e.,  $S_i$  is  $N \times b$  matrix), and  $S_i$  and  $S_j$  can be overlapped.

We apply the algorithm described in the subsection above for a subsequence  $S_k$  with a fixed number of connections  $M$ , and we define the indicator variable,

$$\psi_{ij}^M(S_k) = \begin{cases} 1 & \text{if a connection from } j \text{ to } i \text{ exists,} \\ 0 & \text{otherwise.} \end{cases} \quad (12)$$

Collecting those values for all subsequences, we estimate the probability of existence of the connection,

$$\phi_{ij}^b(M) = \frac{1}{K} \sum_k \psi_{ij}^M(S_k). \quad (13)$$

We also calculate

$$\xi_{ij}^b(M) = 2\phi_{ij}^b(M)(1 - \phi_{ij}^b(M)) \quad (14)$$

which represents instability of the estimation. This is twice as large as the variance of Bernoulli probability of existence of the connection. Let us define average of  $\xi_{ij}^b(M)$  by

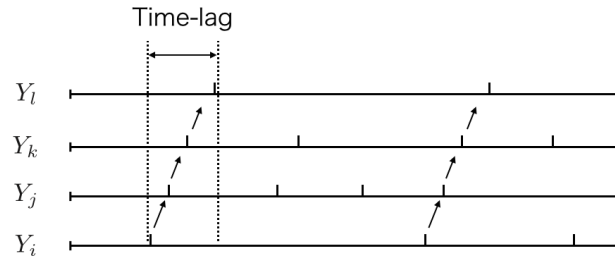
$$D_b(M) = \frac{\sum_{i,j} \xi_{ij}^b(M)}{N(N-1)}. \quad (15)$$

We choose a maximal number of  $M$  so that the value of  $D_b(M')$  is less than a certain threshold  $\gamma$  for all  $M' \leq M$ .

## 4. Construction of statistics

The algorithm described in the section above assumes that off-diagonal elements of  $X$  is given. In this section, we propose to estimate those values from observed neural spike data.

In eq. (1), we formulated a generative model of neuron connectivity. However, since neuron spike trains are point processes, we have to observe them within a certain width of time window. Here we consider two time windows  $[t - \epsilon, t]$  and  $[t, t + \epsilon']$  as shown in figure 2, where the former corresponds to the firing  $Y_j(t)$  in eq. (1) and the latter corresponds to the time for propagation



**Figure 3.** Time lag for multi-step propagation

to the  $i$ -th neuron  $v_i(t + \Delta)$ . For simplicity, we use the notation  $Y_i(t) = 1$  for the event  $\{Y_i(\tau) = 1, \exists \tau \in [t - \epsilon, t]\}$ , and  $Y_i(t + \Delta) = 1$  for the event  $\{Y_i(\tau) = 1, \exists \tau \in [t, t + \epsilon']\}$ .

The values of  $\epsilon$  and  $\epsilon'$  are determined based on the characteristics of neuron's behavior.  $\epsilon$  should be taken a little longer than refractory period. On the other hand, for  $\epsilon'$ , since multi-step propagation through other neurons requires more time as shown in figure 3, it should be long enough to capture the time-lag of the propagation.

In eq. (1), the  $i$ -th neuron fires at time  $t + \Delta$  randomly depending on the membrane potential  $v_i(t + \Delta)$ . Although this relation is typically modelled by a sigmoidal function, for simplicity, we assume it is a linear function, thus eq. (1) can be written as

$$\Pr[Y_i(t + \Delta) = 1 | Y_j(t)] \propto v_i(t) = b_i(t) + \theta_{ij}Y_j(t). \quad (16)$$

For  $Y_j(t) = 0$  and 1, the expected values of eq. (16) are

$$E[Y_i(t + \Delta) | Y_j(t) = 0] \propto E[b_i(t)], \quad (17)$$

$$E[Y_i(t + \Delta) | Y_j(t) = 1] \propto E[b_i(t)] + \theta_{ij}. \quad (18)$$

By taking difference between eq. (17) and eq. (18), we have

$$E[Y_i(t + \Delta) | Y_j(t) = 1] - E[Y_i(t + \Delta) | Y_j(t) = 0] \propto \theta_{ij}. \quad (19)$$

The left hand side of eq. (19) can be estimated from observed time sequence, and can be used as an estimator of  $x_{ij}$  that is proportional to  $\theta_{ij}$ . Remark that the influence  $b_i(t)$  from other neurons including unmeasured neurons is cancelled out by taking difference.

## 5. Simulations

### 5.1. Synthetic data

We apply the proposed method to synthetic data generated by neural models proposed by Izhikevich [9]. The Izhikevich model is known to be able to simulate various types of realistic neuron behaviors even though the model can be formulated by small number of parameters.

Let  $v_i(t)$ [mV] ( $i = 1, \dots, N'$ ) be the membrane potential of the neuron  $i$ ,  $u_i(t)$  be the inactivator giving negative feedback,  $I_i(t)$  be the post-synaptic current with random external input  $I_i^{\text{ext}}$ ,  $W_{ij}$  be the connection weight from the neuron  $j$  to the neuron  $i$ ,  $Y_i(t)$  be the binary state (0 or 1) of neuron  $i$ . Izhikevich model is described by the following difference equations:

For  $v_i(t) \geq 30[\text{mV}]$ ,

$$Y_i(t) = 1 \quad (20)$$

$$I_i(t) = I_i^{\text{ext}}(t) + \sum_j W_{ij} Y_j(t) \quad (21)$$

$$v_i(t + dt) = c_i \quad (22)$$

$$u_i(t + dt) = u_i(t) + d_i \quad (23)$$

$$Y_i(t + dt) = 0, \quad (24)$$

and for  $v_i(t) < 30[\text{mV}]$ ,

$$v_i(t + dt) = v_i(t) + 0.04v_i^2(t) + 5v_i(t) + 140 - u_i(t) + I_i(t) \quad (25)$$

$$u_i(t + dt) = u_i(t) + a_i\{b_i v_i(t) - u_i(t)\}, \quad (26)$$

where  $dt$  is a small time step(0.5[msec]),  $a_i, b_i, c_i, d_i$  are parameters specifying neuron behavior (excitatory or inhibitory).

In the numerical simulation, the total number of neurons is set as  $N' = 50$ , among which 40 neurons are excitatory and 10 neurons are inhibitory based on the biological evidence. In addition, the number of connections from one neuron is limited to at most 10. The strengths of connections are randomly generated from uniform distribution on  $(0, 1]$  for excitatory connections and  $[-1, 0)$  for inhibitory connections.

To investigate the performance for partial observations, we assume only  $N = 32$  neurons (all excitatory) among 50 neurons are observable and used for estimation.

The resting membrane potential  $v_i(t)$  is set to be between  $-70[\text{mV}]$  and  $-60[\text{mV}]$ , and the neuron fires when  $v_i(t)$  exceeds  $30[\text{mV}]$  and it is reset to  $c_i$ . In this simulation, we do not take into account the signal delay and the synaptic plasticity.

For the time window parameters  $\epsilon$  and  $\epsilon'$ , we fixed  $\epsilon$  by  $3.0[\text{msec}]$  and we examined 4 different length of  $\epsilon'$ ,  $3.0, 4.0, 5.0, 6.0[\text{msec}]$ .

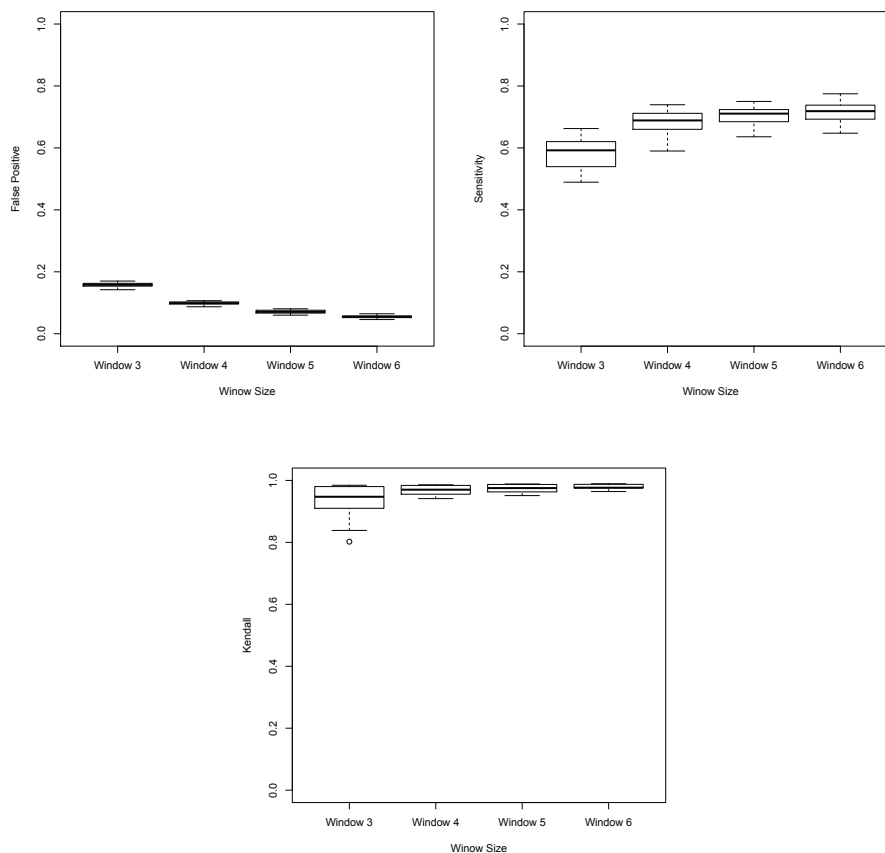
The number of iterations of the algorithm is set to 5.

## 5.2. Results

Since the resulting connection matrix represents a weighted graph structure, we evaluate the performance by two kinds of consistency: (1) topology of graph structure, and (2) strength of connection weights.

*5.2.1. Topology of graph structure* Suppose  $\hat{\Lambda}$  and  $\Lambda^*$  be the estimation and the ground truth of connection weight matrices respectively. Here we define the TP (true positive) by the number of connections correctly estimated that satisfy both  $\hat{\lambda}_{ij} \neq 0$  and  $\lambda_{ij}^* \neq 0$ , and the FN (false negative) by the number of connections that satisfy  $\hat{\lambda}_{ij} = 0$  and  $\lambda_{ij}^* \neq 0$ . The sensitivity is calculated by  $\text{Sensitivity} = \text{TP} / (\text{TP} + \text{FN})$ , which represents the ratio of correctly extracted connections among all existing connections.

It has been demonstrated that the majority of cortical excitatory connections is weak [7] and therefore long term continuous recordings (e.g., 25 hours or longer) would be necessary to detect weak connections from extracellular spike data [10]. The sensitivity using a small amount of spiking data (eg., 1 hour or less) like this investigation would be unavoidably low because of these weak connections. However, in terms of appropriate scientific inference of connections, it would be more important to avoid large FP (false positive) that is the number of estimated connections that do not exist, because FP suggests an effect that is not actually there. In the following, we calculate FP rate, which is defined by  $\text{FP} / (\text{FP} + \text{TN})$ , where TN denotes true negative.



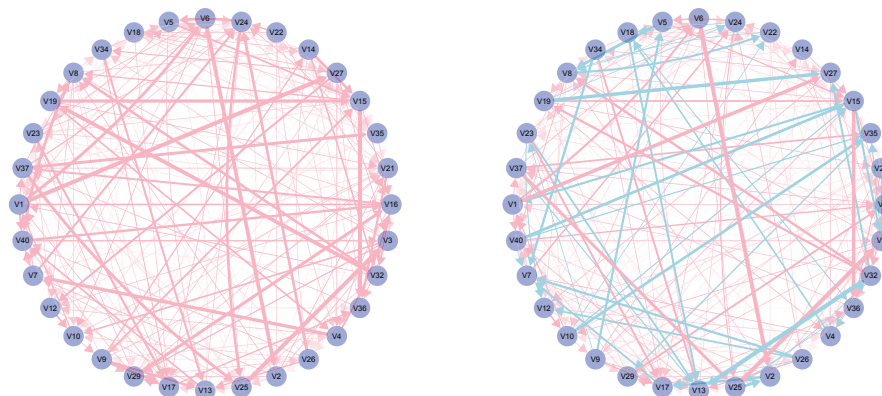
**Figure 4.** Performance of the proposed algorithm. Upper-left: False positive rate, upper-right: Sensitivity, and lower: Kendall's  $\tau$

*5.2.2. Evaluation of connection weights* In order to evaluate how the proposed method extract strengths of connections correctly, we use Kendall rank correlation coefficient (Kendall's  $\tau$ ) that measures the degree of coincidence of rank statistics between  $\hat{\Lambda}$  and  $\Lambda^*$ . To focus on nonzero connections, first we exclude the pairs with no connections ( $\lambda_{ij} = 0$ ). However, since the nonzero components are not always the same in the two matrices, we calculate Kendall's  $\tau$  by choosing the connections included both in  $\hat{\Lambda}$  and in  $\Lambda^*$ .

### 5.3. Evaluation

We performed the simulation for 20 different networks generated by different random seeds. Boxplots of FP rate, Sensitivity and Kendall's  $\tau$  are plotted in figure 4. Horizontal axis represents the length of window size  $\epsilon'$ . For  $\epsilon' = 5$  and 6, FP rate is quite small, which means that the proposed method correctly estimates the absence of connections in the true network. We see relatively lower values of sensitivity mainly caused by the sparsity of the graph. Kendall's  $\tau$  is quite high for  $\epsilon' = 5$  and 6, so the strong connections are extracted correctly. In figure 5, examples of true and estimated network structures are shown.





**Figure 5.** Examples of graph structure. Left: True, Right: Estimated. Red edges represent excitatory connections, and blue edges represent inhibitory connections.

## 6. Conclusion

In this paper, we proposed a method to extract directed connectivity between neurons based on a simple generative model. Using the synthetic data that were generated by the Izhikevich neuron model, we demonstrated that the proposed method successfully detected the direction and strength of both excitatory and inhibitory connections. In neuroscience, cross-correlation has been widely used to assess synaptic connections [10, 11]. Our method would complement the existing methods and can be easily applied to real data because it does not involve many tuning parameters. By designing appropriate statistics, the proposed method would be also applicable to other application areas.

## Acknowledgments

The present study was financially supported by JSPS KAKENHI 25120009, 25120011, 26120504, 15H01678, 25870811.

## References

- [1] Hebb D O 1949 *The Organization of Behavior: A Neuropsychological Theory* (Wiley)
- [2] Wilson M A and McNaughton B L 1994 *Science* **265** 676–8
- [3] Ganmor E, Segev R and Schneidman E 2011 *PNAS* **108** 9679–84
- [4] Ohiorhenuan I E, Mechler F, Purpura K P, Schmid A M and Hu Q 2010 *Nature* **466** 617–21
- [5] Shimazaki H, Amari S, Brown E N and Grün S 2012 *PLoS Comp. Biol.* **8** e1002385
- [6] Noda A, Hino H, Tatsuno M, Akaho S and Murata N 2014 *Neural Computation* **26** 1455–83
- [7] Song S, Sjöström P J, Reigl M, Nelson S and Chklovskii D B 2005 *PLoS Biol.* **3** e68
- [8] Roeder K, Liu H and Wasserman L 2010 *Advances in Neural Information Processing Systems* **23** 1432–40
- [9] Izhikevich E M 2003 *IEEE Trans. on Neural Networks* **14** 1569–72
- [10] Schwindel C D, Ali K, McNaughton B L and Tatsuno M 2014 *Journal of Neuroscience* **34** 5454–67
- [11] Ostojic S, Brunel N and Hakim V 2009 *Journal of Neuroscience* **29** 10234–53