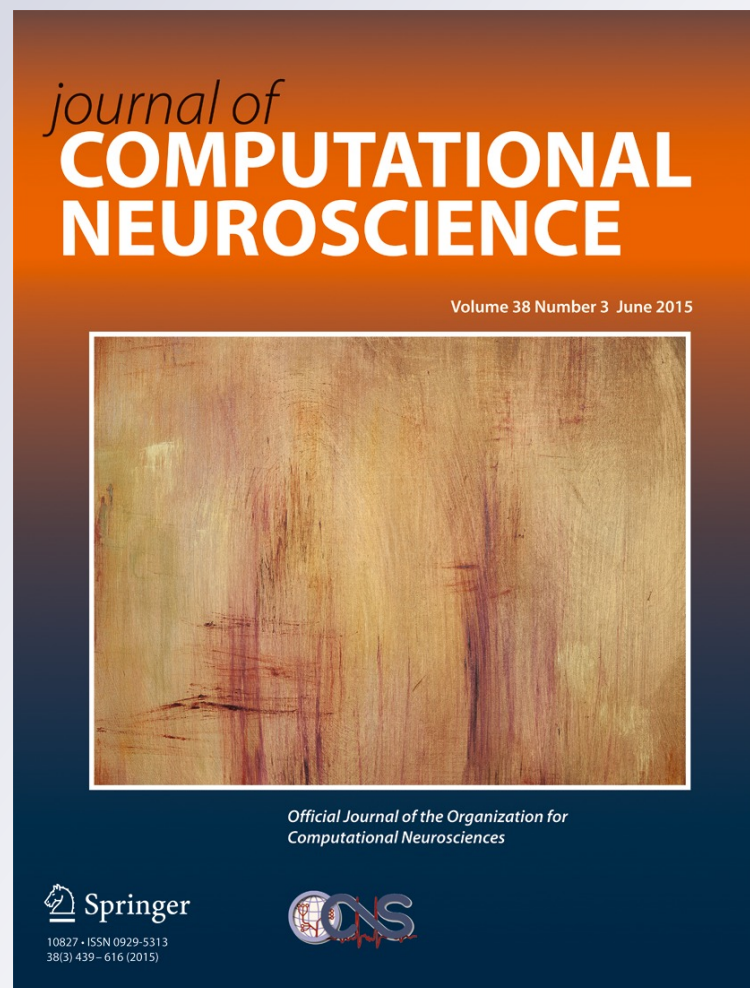# Granger causality-based synaptic weights estimation for analyzing neuronal networks

Pei-Chiang Shao, Jian-Jia Huang, Wei-Chang Shann, Chen-Tung Yen, Meng-Li Tsai & Chien-Chang Yen

journal of
COMPUTATIONAL NEUROSCIENCE

Volume 38 Number 3 June 2015

Official Journal of the Organization for Computational Neurosciences

Springer

10827 · ISSN 0929-5313
38(3) 439 – 616 (2015)

Springer

Springer

Granger causality-based synaptic weights estimation for analyzing neuronal networks

Pei-Chiang Shao · Jian-Jia Huang · Wei-Chang Shann · Chen-Tung Yen · Meng-Li Tsai · Chien-Chang Yen

Abstract Granger causality (GC) analysis has emerged as a powerful analytical method for estimating the causal relationship among various types of neural activity data. However, two problems remain not very clear and further researches are needed: (1) The GC measure is designed to be nonnegative in its original form, lacking of the trait for differentiating the effects of excitations and inhibitions between neurons. (2) How is the estimated causality related to the underlying synaptic weights? Based on the GC, we propose a computational algorithm under a best linear predictor assumption for analyzing neuronal networks by estimating the synaptic weights among them. Under this assumption, the GC analysis can be extended to measure both excitatory and inhibitory effects between neurons. The method was examined by three sorts of simulated networks: those with linear, almost linear, and nonlinear network structures. The method was also illustrated to analyze real spike train data from the anterior cingulate cortex (ACC) and the striatum (STR). The results showed, under the quinpirole administration, the significant existence of excitatory effects inside the ACC, excitatory effects from the ACC to the STR, and inhibitory effects inside the STR.

Keywords Granger causality analysis · Synaptic weights estimation · Vector autoregressive model · Neuronal networks

P.-C. Shao · W.-C. Shann
Department of Mathematics, National Central University,
Jhongli 32001, Taiwan

J.-J. Huang
Graduate Institute of Electronics Engineering, National Taiwan
University, Taipei 10617, Taiwan

C.-T. Yen
Department of Life Sciences, National Taiwan University,
Taipei 10617, Taiwan

M.-L. Tsai (✉)
Department of Biomechatronic Engineering, National Ilan
University, Ilan 26047, Taiwan
e-mail: mltsai@niu.edu.tw

C.-C. Yen (✉)
Department of Mathematics, Fu-Jen Catholic University,
Xinzhuang 24205, Taiwan
e-mail: yen@math.fju.edu.tw

1 Introduction

Granger causality (GC) (Granger 1969, 1980) has been shown to be an effective method for analyzing the causal relationship between continuous-valued neural activity data (Baccala and Sameshima 2001; Dhamala et al. 2008; Cadotte et al. 2010; Bressler and Seth 2011) and has been widely deployed in recent neuroscience research. To further understand how neurons cooperate to generate specific brain functions, several extended GC methods were also proposed for identifying directional interactions between neurons through multiple spike trains (Sameshima and Baccala 1999; Nedungadi et al. 2009; Krumin and Shoham 2010; Kim et al. 2011; Zhou et al. 2014). Being the fundamental knowledge used in this paper, the time domain GC analysis will be briefly introduced in the next section and the readers are referred to an article by Barnett and Seth Barnett and Seth (2014) for more details.

The term *synaptic weight* is widely used in neural network research and typically refer to the coupling strength of a connection between two nodes in the network. A large

synaptic weight usually means that a large signal (i.e., high-frequency spikes) from the pre-synaptic neuron can result in a large signal of the post-synaptic one. Therefore, in neuroscience and biology, it can also be interpreted as the amount of influence of one neuron has on the firing activity of another.

The spikes of a pre-synaptic neuron are carried by the axon, which will release excitatory or inhibitory neurotransmitter into the synapse. When the post-synaptic neuron receives the neurotransmitter, an excitatory post-synaptic potential (EPSP) or an inhibitory post-synaptic potential (IPSP) is then induced to temporarily depolarize or hyperpolarize the membrane potential. An EPSP makes the neuron more likely to generate an action potential (AP), while an IPSP makes the neuron to do the opposite. However, a single EPSP is not sufficient for the membrane to generate an AP, temporal or spatial summations are required. This means that the firing pattern of the post-synaptic neuron is generally not a direct consequence of the influence of a single pre-synaptic neuron; but a weighted result of the effects of several pre-synaptic neurons with possibly different synaptic weights. Furthermore, IPSPs will diminish EPSPs, playing a much more crucial role of determining whether or not an AP generation will occur at the post-synaptic membrane.

The GC analysis has emerged as a powerful analytical method for estimating the causal strength of complex networks (Seth 2005; Seth and Edelman 2007). However, the effects of excitations and inhibitions could not be differentiated in its original form. Based on the GC, we propose a computational algorithm (presented in Section 2.3) under the assumption of best linear predictor (BLP) for analyzing neuronal networks by estimating the synaptic weights among them. The idea of the mathematical assumption BLP is that the weighted voltage-fluctuation of the pre-synaptic neurons should be the best linear explanation for the voltage-fluctuation of the post-synaptic neuron among the network. Using this interpretation, the GC analysis can be extended to measure both excitatory and inhibitory effects between neurons without too much extra computational complexity. The appropriateness of the BLP assumption was examined by three sorts of simulated networks: those with linear, almost linear, and nonlinear network structures. To illustrate the application of the proposed method, real spike trains from the anterior cingulate cortex (ACC) and the striatum (STR) were analyzed.

It is worth noting that spike trains are non-equally spaced data and are regarded as being from a point process. Filtering is usually required for converting them to equally spaced time series for further GC analyses (Sameshima and Baccala 1999). This study adopted the Gaussian kernel filtering or binning (depending on the situation) to convert spike trains into time series data for the following three main reasons:

(1) it reduces the complexity of analysis, and considers also the effect of temporal summation of action potentials, (2) under suitable preprocessing, even short, sparse spike trains can be converted, so that the standard autoregression modeling can be applied (Zhu et al. 2003), (3) most important of all, spike trains can be filtered to form close approximations to the firing rates or the voltage-fluctuations of the underlying neurons (Lehky 2010).

The rest of this paper is organized as follows. In Section 2, we briefly introduce the so-called Granger causality index, and then extend it to measure both excitatory and inhibitory effects between network nodes by using the BLP assumption. Section 3 presents three network models to ensure the appropriateness of the proposed algorithm. In Section 4, we apply the algorithm to analyze real spike train data. Section 5 provides some discussion about the results obtained from Section 3–4, shortcomings of the method, and related future works.

2 Modeling and analysis

Based on the framework of Granger causality analysis and a BLP interpretation of synaptic weights, we propose a procedure for weights estimation and define a synaptic measure between neuronal time series using the estimated weights.

2.1 An introduction to the GC

Let $X = (x^1, x^2, \cdots, x^n)$ be a stationary n-dimensional time series process with zero mean. The p-th order linear autoregressive model for X is given by

$$
\begin{cases}
x_t^1 = \sum_{r=1}^{p} a_r^{1,1} x_{t-r}^1 + \cdots + \sum_{r=1}^{p} a_r^{1,n} x_{t-r}^n + \epsilon_t^1 \\
x_t^2 = \sum_{r=1}^{p} a_r^{2,1} x_{t-r}^1 + \cdots + \sum_{r=1}^{p} a_r^{2,n} x_{t-r}^n + \epsilon_t^2 \\
\quad \vdots \\
x_t^n = \sum_{r=1}^{p} a_r^{n,1} x_{t-r}^1 + \cdots + \sum_{r=1}^{p} a_r^{n,n} x_{t-r}^n + \epsilon_t^n,
\end{cases}
\tag{1}
$$

where $a_r^{i,j}$ is the projection coefficient from the i-th time series onto the j-th time series at time lag r, representing the coupling strength from node j to node i in the network. The residuals $\epsilon^1, \epsilon^2, \cdots, \epsilon^n$ are zero-mean uncorrelated white noises with covariance matrix Σ. The diagonal entries $\{\Sigma_{ii} = Var(\epsilon^i), i = 1, \cdots, n\}$ measure the accuracy of the autoregressive prediction to each node based on the information from time stamps $t - 1$ to $t - p$.

To see whether the information contained in time series x^j is useful in explaining the state of time series x^i, namely, the importance of node j to node i, we can exclude the

time series variable x^j from Eq. (1) to obtain a reduced[1] $(n-1)$-dimensional autoregressive model with residual series $\eta^{i,j}$ of x^i and corresponding prediction error $\Gamma^j_{ii} = Var(\eta^{i,j})$. Here Γ^j_{ii} measures the accuracy of the prediction of x^i based only on the previous values in time series $\{x^1, \cdots, x^{j-1}, x^{j+1}, \cdots, x^n\}$. If Σ_{ii} in Eq. (1) is significantly less than Γ^j_{ii} in the reduced model in some suitable statistical sense, then we say that x^j Granger-cause x^i. This causality can be quantified by the GC index from x^j to x^i formulated as:

$$F_{j \to i} = \ln \frac{\Gamma^j_{ii}}{\Sigma_{ii}}. \tag{2}$$

It is clear that $F_{j \to i} = 0$ when $\Gamma^j_{ii} = \Sigma_{ii}$, i.e., x^j has no causal influence on x^i, and $F_{j \to i} > 0$ when x^j Granger-cause x^i. Notice that $F_{j \to i}$ is defined only for $j \neq i$ and is always nonnegative, i.e., Σ_{ii} is bounded above by Γ^j_{ii}, since the full model defined in Eq. (1) should fit the data better than the reduced model. Finally, we note that the GC index Eq. (2) is significant if the corresponding coefficients $a^{i,j}_r$ are jointly significantly different from zero. This can be assessed via an F-test on the null hypothesis that $a^{i,j}_r$ are zero (Greene 2002; Seth 2010). The projection coefficients $a^{i,j}_r$ and prediction errors Σ_{ii} can be obtained by solving the Yule-Walker equation (Kitagawa 2010) and an efficient model order can be determined using the Akaike Information Criterion (AIC):

$$AIC(p) = 2\log(\det(\Sigma)) + \frac{2n^2 p}{T}, \tag{3}$$

where T is the total length of the time series. More information about the GC can be found in the author's previous work (Shao et al. 2013) and also Ding et al. (2006), Cadotte et al. (2010), and Barnett and Seth (2014).

2.2 Synaptic weights estimation

Now, consider the multivariate zero-mean time series, $X = (x^1, x^2, \cdots, x^n)$, consisting of the trajectories of membrane voltage of n distinct neurons. Suppose that the i-th neuron is triggered by some other k neurons in the network, say $\{i_1, i_2, \cdots, i_k\}$-th neurons, with synaptic weights $\{\alpha^i_1, \alpha^i_2, \cdots, \alpha^i_k\}$. For convenience, we assume that $1 \leq k := k(i) \leq n-1$. The case $k = 0$ means that the i-th neuron is not triggered by others, thus is relatively easy to deal with. The weights are assumed to be nonzero, if some α^i_s is zero, then we can just remove the corresponding index

---

[1]The i−th equation of the reduced model reads $x^i_t = \sum_{r=1}^p b^{i,1}_r x^1_{t-r} + \cdots + \sum_{r=1}^p b^{i,j-1}_r x^{j-1}_{t-r} + \sum_{r=1}^p b^{i,j+1}_r x^{j+1}_{t-r} + \cdots + \sum_{r=1}^p b^{i,n}_r x^n_{t-r} + \eta^{i,j}$, where the b's are the corresponding projection coefficients.

i_s from the trigger set. Positive and negative weights represent excitatory and inhibitory influences on the i-th neuron, respectively.

In general, the trigger set $I_i := \{i_1, i_2, \cdots, i_k\}$ and the corresponding weights $\alpha^i := \{\alpha^i_1, \alpha^i_2, \cdots, \alpha^i_k\}$ in the network can not be identified and estimated easily due to the underlying complex dynamics. However, under the assumption of the best linear predictor (BLP) (Definition 1 below), I_i and α^i can be approximated effectively. The results are described in the following proposition.

Definition 1 Let x and y be two stationary time series with zero-means. Then we say that y forms the best linear predictor (BLP) of x among a variable set $\mathcal{G}$ if $\sigma^2(x|\bar{x}, \bar{y}) < \sigma^2(x|\bar{x}, \bar{z}), \forall z \in \mathcal{G}$, where $\sigma^2(x|\bar{x}, \bar{y}) := \min_{p,\{f_r\},\{d_r\}} E\{x_t - \sum_{r=1}^p [f_r y_{t-r} + d_r x_{t-r}]\}^2$.

Proposition 1 *In the situation described above, if further the weighted trajectory $u^i := \alpha^i_1 x^{i_1} + \alpha^i_2 x^{i_2} + \cdots + \alpha^i_k x^{i_k}$, made by the trigger set and the corresponding weights, forms the BLP to the trajectory of the i-th neuron (namely, x^i) among the whole network; then based on the GC framework, I_i can be completely identified and the estimate of α^i can be obtained as $\hat{\alpha}^i := \left\{\sum_{r=1}^p a^{i,i_1}_r, \sum_{r=1}^p a^{i,i_2}_r, \cdots, \sum_{r=1}^p a^{i,i_k}_r\right\}$ up to a scale factor.*

Proof Let $u^i := \alpha^i_1 x^{i_1} + \cdots + \alpha^i_k x^{i_k}$ form the BLP of x^i, then there exist a positive integer p and projection coefficients $\{f^i_r, r = 1, 2, \cdots, p\}$, $\{d^i_r, r = 1, 2, \cdots, p\}$ such that $x^i_t = \sum_{r=1}^p [f^i_r u^i_{t-r} + d^i_r x^i_{t-r}] + \epsilon^i_t$, where ϵ^i is a stationary white noise possessing the smallest variance among the whole network. Replacing u^i with the weighted trajectory, we obtain

$$\begin{aligned}
x^i_t &= \sum_{r=1}^p \left[f^i_r u^i_{t-r} + d^i_r x^i_{t-r}\right] + \epsilon^i_t \\
&= \sum_{r=1}^p \left[f^i_r \left(\alpha^i_1 x^{i_1}_{t-r} + \cdots + \alpha^i_k x^{i_k}_{t-r}\right) + d^i_r x^i_{t-r}\right] + \epsilon^i_t \\
&= \sum_{r=1}^p \left[\alpha^i_1 f^i_r x^{i_1}_{t-r} + \cdots + \alpha^i_k f^i_r x^{i_k}_{t-r} + d^i_r x^i_{t-r}\right] + \epsilon^i_t, \tag{4}
\end{aligned}$$

which represents the underlying but unknown network structure of $\{x^i, x^{i_1}, x^{i_2}, \cdots, x^{i_k}\}$. On the other hand, fitting to data the same equation as the i-th equation in Eq. (1), we have the following empirical regression (compared to the theoretical regression Eq. (4))

$$x^i_t = \sum_{r=1}^p \left[a^{i,1}_r x^1_{t-r} + \cdots + a^{i,n}_r x^n_{t-r}\right] + \tilde{\epsilon}^i_t. \tag{5}$$

Let $\bar{I}_i := \{1, 2, \cdots, i-1, i+1, \cdots, n\} - I_i$ be the complement of the trigger set I_i. We note that $\tilde{\epsilon}_t^i \equiv \epsilon_t^i$ if $a_r^{i,s} = 0, \forall r = 1, \cdots, p$ and $s \in \bar{I}_i$; otherwise $\tilde{\epsilon}^i$ and ϵ^i are totally different but with $Var(\tilde{\epsilon}_t^i) = Var(\epsilon_t^i)$ since Eq. (4) has the smallest residual variance among the whole network and Eq. (5) has more degree of freedom (coefficients) than Eq. (4).

If the trajectories of the $\bar{I}_i$-th neurons are stochastically independent of both the i-th and I_i-th neurons, then we have $a_r^{i,s} = 0$, for $r = 1, \cdots, p$ and $s \in \bar{I}_i$. Comparing Eq. (5) with Eq. (4), we then have

$$\sum_{r=1}^{p} a_r^{i,i_1} = \alpha_1^i \sum_{r=1}^{p} f_r^i \quad, \cdots, \quad \sum_{r=1}^{p} a_r^{i,i_k} = \alpha_k^i \sum_{r=1}^{p} f_r^i. \quad (6)$$

Since the projection coefficients in Eq. (5) can be obtained by solving the Yule-Walker equation or simply by the least-squares method, Eq. (6) immediately leads to

$$\frac{\alpha_1^i}{\sum\limits_{r=1}^{p} a_r^{i,i_1}} = \frac{\alpha_2^i}{\sum\limits_{r=1}^{p} a_r^{i,i_2}} = \cdots = \frac{\alpha_k^i}{\sum\limits_{r=1}^{p} a_r^{i,i_k}}, \quad (7)$$

provided $\sum_{r=1}^{p} f_r^i \neq 0$. For α_s^i and $\sum_{r=1}^{p} a_r^{i,i_s}$ to have the same sign for $s = 1, 2, \cdots, k$, we can, without loss of generality, assume that $\sum_{r=1}^{p} f_r^i > 0$. If $\sum_{r=1}^{p} f_r^i < 0$, then $-u^i$ is used to replace u^i.

If some of the trajectories of the $\bar{I}_i$-th neurons are linearly dependent of the i-th or I_i-th neurons, then $a_r^{i,s} \neq 0$, for some $r \in \{1, \cdots, p\}, s \in \bar{I}_i$ and the projection coefficients in Eq. (5) are thus affected, resulting in a biased estimation of Eq. (7). However, this predicament can be solved by virtue of the assumption of BLP and the concept of GC. Since the ϵ^i in Eq. (4) has the smallest variance among the whole network, taking out any element of $\{x^s : s \in \bar{I}_i\}$ from the regression Eq. (5) does not increase the variance of $\tilde{\epsilon}^i$. According to the concept in Eq. (2), we can correct the model coefficients by ruling out all the useless information of the $\bar{I}_i$-th neurons. $\square$

We end this subsection by the following remarks.

Remark 1 The idea behind the BLP mathematical assumption is that the weighted voltage-trajectory of the trigger neurons should be the best linear explanation for the voltage-trajectory of the target neuron among the whole network. Based on this interpretation, the GC index can be extended to measure both excitatory and inhibitory effects in virtue of the estimated synaptic weights.

Remark 2 The synaptic weight α_s^i and the summation of the projection coefficients $\sum_{r=1}^{p} a_r^{i,i_s}$ are forced to have the same sign for $s = 1, 2, \cdots, k$, because positive and negative $\sum_{r=1}^{p} a_r^{i,i_s}$ refer to positive and negative correlations between x^i and x^{i_s}, respectively.

Remark 3 The case $k = 0$ means that the i-th neuron is not triggered by other neurons in the network, therefore $F_{j \to i} = 0, \forall j$ and there is no synaptic weights to be estimated.

Remark 4 For readers dealing with sparse networks, L1 regularization (or LASSO) would be an useful technique for fitting to data a sparse regression to avoid overfitting and the problem of multiple testing (Arnold et al. 2007; Michailidis and d'Alche-Buc 2013). In this scenario, a computational much more efficient approach would be first running LASSO to learn the network structure and then using GC to get the causal strength.

Remark 5 Some arguments in the proof such as the stochastic independence and the estimations of projection coefficients assume that the law of large numbers (LLN) holds. In the case of small samples or limited data, estimation errors come into exsitence thus some statistical inferences in the proof may fail. However, the proof holds for most large-sample cases.

2.3 The algorithm

Here, we present a step by step algorithm for computing the proposed index (named neuron synaptic index, NSI) from multiple spike train data.

Step 1   Properly smooth the spike train data by kernel filtering (Gaussian kernels are commonly used) to acquire the approximate membrane voltage trajectories to the underlying voltage evolution of the neurons in the network.

Step 2   Subtract the mean value from each trajectory to form zero-mean time series and then fit the vector autoregressive model in Eq. (1). Appropriate model order can be obtained beforehand by using AIC in Eq. (3).

Step 3   Compute all the GC indices by Eq. (2) for all pairs of neurons, i.e., $i, j = 1, 2, \cdots, n$ with $i \neq j$ and also perform F-tests to ensure statistical significance.

Step 4   For each node $i = 1, 2, \cdots, n$, refine the autoregressive model by ruling out the information about the $\bar{I}_i$-th neurons, i.e., the neurons with

insignificant $F_{j\to i}$, to correct the projection coefficients.

Step 5 For each node, compute the synaptic weights of the trigger neurons by simply summing the projection coefficients up to the model order p as shown in Eq. (7).

Step 6 For each node, take the weighted trajectory as a new explanatory time series and then compute the GC index from this weighted time series to that of the node.

Step 7 Finally, the NSI is then defined to be the l_1-normalized[2] estimated weights obtained in Step 5 multiplying the GC index obtained in Step 6 (see Eq. (A.5) in Appendix).

3 Simulation study

Here, three sorts of network models are simulated to investigate the proposed algorithm: the linear, almost-linear, and nonlinear networks. The linear network, derived directly from autoregressive framework, examines the validity of the synaptic weights estimation in Proposition 1. Additionally, the almost-linear and nonlinear networks, derived both from integrate-and-fire (IF) neuron models, examine the appropriateness of the BLP assumption in Remark 1 through the reflected subthreshold dynamics of the models. The Matlab code used for this study is available to interested readers upon request.

Significance tests on the GC indices should be corrected for multiple testing and we adopted the approach of False Discovery Rate (FDR) (Benjamini and Hochberg 1995) which has greater statistical power than the conservative Bonferroni correction. In this and next sections, significant Granger causality interaction between input neurons and output neurons are calculated using an F-test corrected by FDR for multiple comparison with confidence threshold at $P-value = 0.05$ (i.e., 95 % significance level).

3.1 Linear network

A linear network, depicted in Fig. 1, is presented here. The time series variable w serves as the trajectory of post-synaptic neuron, while x, y, and z serve as the trajectories of pre-synaptic neurons with synaptic weights α, β, and γ, respectively. For readers not familiar with the multivariate settings in Section 2, a much clear-cut derivation of the NSI using this simple network is given in Appendix.
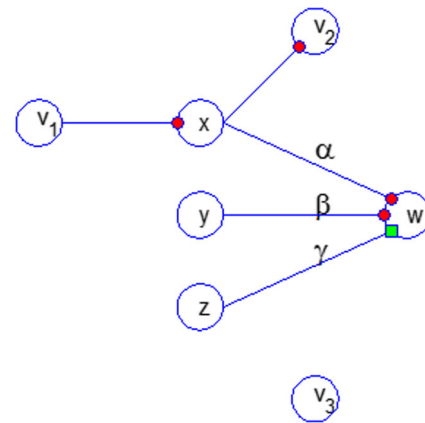
Fig. 1 A simple linear network. Red circles represent excitation and green squares represent inhibition. Variable w serves as the trajectory of post-synaptic neuron, variables x, y, and z serve as the trajectories of pre-synaptic neurons with synaptic weights $\alpha > 0$, $\beta > 0$, and $\gamma < 0$, respectively. v_1, v_2, and v_3 are collateral variables, consisting of source, target, and independent nodes. A much clear-cut derivation of the NSI using this simple network is given in Appendix

Trajectories are generated by the following equations:

$$\begin{cases} v_{1,t} = \epsilon_{1,t} \\ x_t = g_1 v_{1,t-1} + g_2 v_{1,t-2} + g_3 v_{1,t-3} + \epsilon_{2,t} \\ v_{2,t} = h_1 x_{t-1} + h_2 x_{t-2} + h_3 x_{t-3} + \epsilon_{3,t} \\ y_t = \epsilon_{4,t} \\ z_t = \epsilon_{5,t} \\ u_t = \alpha x_t + \beta y_t + \gamma z_t \\ w_t = f_1 u_{1,t-1} + f_2 u_{1,t-2} + f_3 u_{1,t-3} + d_1 w_{1,t-1} \\ \qquad + d_2 w_{1,t-2} + d_3 w_{1,t-3} + \epsilon_{6,t} \\ v_{3,t} = \epsilon_{7,t} \end{cases} \quad (8)$$

where $\epsilon_k \sim^{iid} \mathcal{N}(0,1), k = 1, \cdots, 7$ are zero-mean uncorrelated Guassian white noises.

Settings of the simulation are: $[g_1, g_2, g_3] = [0.4, 0.2, 0.1]$, $[h_1, h_2, h_3] = [0.1, 0.2, 0.4]$, $[f_1, f_2, f_3] = [0.5, 0.3, 0.1]$, and $[d_1, d_2, d_3] = [0.1, 0.3, 0.5]$. Variable x is triggered by v_1 and variable v_2 is triggered by x. Variable v_3 serves as an independent node in the network. The length of each trajectory is 1000. We note that the above coefficients can be replaced by any stable coefficients and the simulation results shown below will remain unaffected. The stability condition (Lutkepohl 2005) ensures Eq. (8) to generate stationary processes. There have been many convincing examples published (Baccala and Sameshima 2001; Zou et al. 2010) showing that the GC framework is a well-established method for identifying the causal relationship among stationary time series. If some unstable coefficients are used instead, the generated time series will be nonstationary, then some signal preprocessing technique (e.g, differencing) will be needed to convert nonstationary processes to stationary processes. More details can be found in Barnett and Seth (2014).

[2]This means that the estimated weight vector $\hat{\alpha}^i = (\hat{\alpha}_{i_1}, \hat{\alpha}_{i_2}, \cdots, \hat{\alpha}_{i_k})$ is normalized by its l_1 norm $\|\hat{\alpha}^i\|_1 := |\hat{\alpha}_{i_1}| + |\hat{\alpha}_{i_2}| + \cdots + |\hat{\alpha}_{i_k}|$. Then the normalized weight vector $\hat{\alpha}^i / \|\hat{\alpha}^i\|_1$ will have unit l_1 vector norm.

For $[\alpha, \beta, \gamma] = [1.0, 0.5, -0.5]$, averaging from 100 repeated simulations gives: the synaptic weights estimates $[\hat{\alpha}, \hat{\beta}, \hat{\gamma}] = [0.9012, 0.4549, -0.4539]$, the weighted GC index $F_{\hat{\alpha}x+\hat{\beta}y+\hat{\gamma}z\to w} = 0.4515$, the normalized weights $[1.0000, 0.5064, -0.5053]$ (divided by $|\hat{\alpha}|$ instead of $|\hat{\alpha}| + |\hat{\beta}| + |\hat{\gamma}|$ defined in Eq. (A.5) for easy comparison with the underlying weights), and the NSIs $[N_{x\to w}, N_{y\to w}, N_{z\to w}] = [0.4515, 0.2286, -0.2281]$ with standard deviations $[0.0359, 0.0396, 0.0347]$. Notice that no knowledge of the parameters for generating the data was used in the estimation procedure, only the generated trajectories were used.

The results show that the normalized weight estimates $[1.0000, 0.5064, -0.5053]$ are consistent with the underlying weights $[1.0, 0.5, -0.5]$, and the collateral variables v_1, v_2, and v_3 do not affect the estimation since $F_{v_1\to w}$, $F_{v_2\to w}$, and $F_{v_3\to w}$ are all zero, thus these information can be ruled out directly by the GC analysis. The NSI Eq. (A.5) unites the estimated synaptic weights and the weighted GC index. A positive (negative) value represents excitatory (inhibitory) synaptic connection and it's magnitude at the same time reflects the degree of synaptic influence.

3.2 Almost-linear network

To illustrate the synaptic weights estimation in a neural spiking context, a simple feedforward IF neuron network was simulated (depicted in Fig. 2). Briefly, Neurons #2 − 5 were modeled by independent Poisson processes with
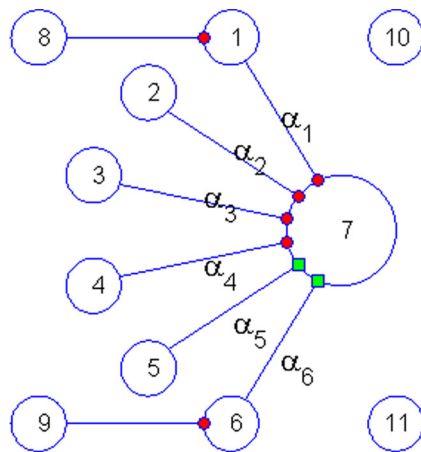
Fig. 2 A simple feedforward integrate-and-fire neuron network. Red circles represent excitation and green squares represent inhibition. Neurons #2 − 5 are modeled by independent Poisson processes with firing rate λ. Neurons #8, 9 were modeled as single strong inputs by independent Poisson processes with firing rate 1.5λ. Neuron #7 is implemented by a direct discrete time summation of the synaptic inputs $\alpha_1, \cdots, \alpha_6 (mV)$. Neurons #1, 6 are implemented in the same way that Neuron #7 is done with synaptic inputs $30 (mV)$ from Neurons #8, 9 respectively. Neurons #10, 11 are modeled as independent nodes by independent Poisson processes with firing rate λ. Neurons #5, 6 are inhibitory, i.e., $\alpha_5, \alpha_6 < 0$

firing rate λ. Neurons #8, 9 were modeled as single strong inputs by independent Poisson processes with firing rate 1.5λ. Neuron #7 was implemented by a direct discrete time summation of the synaptic inputs $\alpha_i (mV), i = 1, \cdots, 6$ (i.e., the weighted outputs of Neurons #1 − 6 after some propagation delay), leading to its internal potential that was reset to $V_{reset} = -80 (mV)$ and produced a spike when the threshold value $V_{th} = -55 (mV)$ was reached. During the refractory period, the potential will linearly recover from V_{reset} to the resting potential $V_{rest} = -70 (mV)$. Time resolution was set to be 1 ms and there was a 2 mV decrease/increase of the potential to the V_{rest} every unit time depending on the status of de/hyper -polarizations, respectively to model the diffusion of ions. The internal potential was forced to lie in the range $[E_{K^+}, E_{Na^+}] = [-90, 60]$, the equilibrium potential of K^+ and Na^+, respectively and action potentials were normalized to 30 mV for display. Neurons #1, 6 were implemented in the same way that Neuron #7 was done with synaptic inputs 30 (mV) from Neurons #8, 9 respectively. Neurons #10, 11 were modeled as independent nodes by independent Poisson processes with firing rate λ.

We begin with Simulation 1 in which the synaptic weights were fixed at $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 5 (mV)$ and $\alpha_5 = \alpha_6 = -2.5 (mV)$, and the propagation delay of each source neuron was set to be 10 ms. 60 sec. voltage-trajectories of Neurons #1, 6, 7 were then simulated according to the way described above. The first 1 sec. of the trajectory of Neuron #7 with input rate λ = 40 Hz is shown in Fig. 3 and the corresponding simulated spike train data is shown in Fig. 4. The subthreshold trajectory of
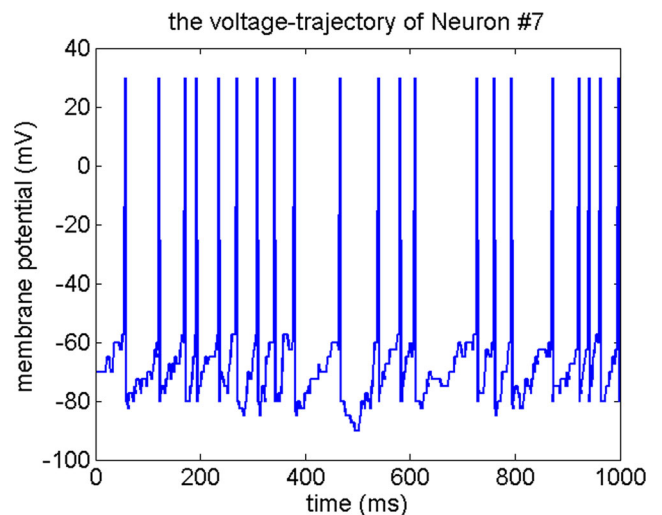
Fig. 3 The first 1 sec. of a simulated voltage-trajectory of Neuron #7 with input rate λ = 40 Hz. The simulation was done according to the way described in the context with $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 5 (mV)$, $\alpha_5 = \alpha_6 = -2.5(mV)$, and 10ms propagation delay. The subthreshold trajectory is not very regular due to the lack of self dynamics, in other words, Neuron #7 is completely triggered by Neurons #1 − 6

simulated spike train data
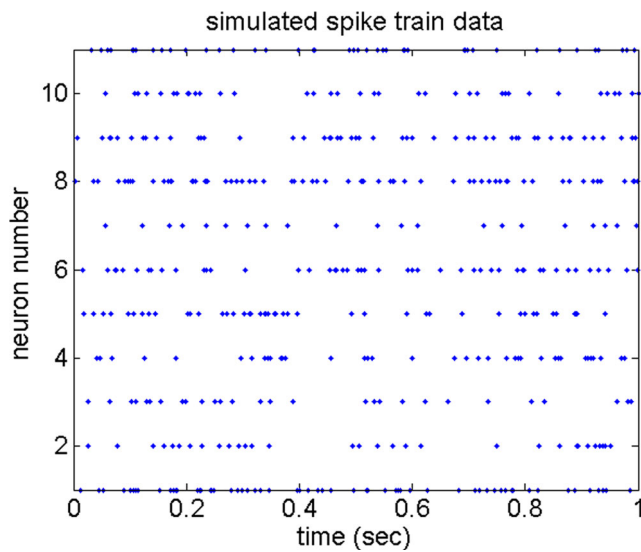
Fig. 4 The first 1 *sec.* of a simulated spike train data of the simple feedforward network with $\lambda = 40\,Hz$, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 5\,(mV)$, $\alpha_5 = \alpha_6 = -2.5\,(mV)$, and 10 *ms* propagation delay

Neuron #7 is not very regular due to the lack of self dynamics, compared to the nonlinear network Eq. (9a) introduced next. However, in this case it faithfully reflects the effects of the input neurons, that is, the actual degree of effects of the input neurons are to be proportional to the corresponding synaptic weights. To analyze the network directly through the simulated spike train data, a Gaussian kernel filtering with bandwidth 5 *ms* was performed to obtain an approximation of the subthreshold dynamics of each neuron in the network, the result is depicted in Fig. 5. Based on the filtered

normalized amplitude
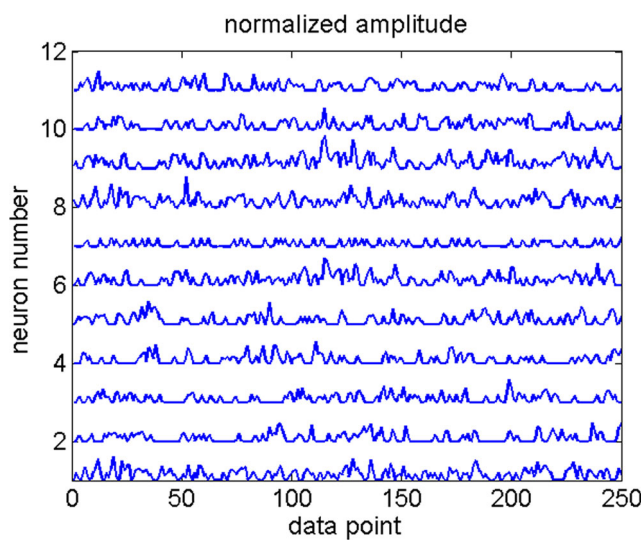
Fig. 5 A Gaussian kernel filtering with bandwidth 5 *ms* was performed on the spike train data depicted in Fig. 4 to obtain an approximation of the subthreshold dynamics of each neuron in the network. The computations of the GCI and NSI were based on the filtered results and this figure shows the first 250 data points

data, both GC and NS indices were computed for different input rates $\lambda = 40, 60$, and $80\,Hz$. In each case, the indices were both obtained from the average of 100 simulations and the results are summarized in Table 1.

We can find from Table 1 that although the GC indices correctly identify the direction of information flow among the network, the effects of excitations and inhibitions could not be differentiated directly by the sign of the indices since they are by definition to be nonnegative. From the GC indices, we can only tell that Neurons #2 − 4 have more influences than Neurons #1, 5 have on Neuron #7. Information about the underlying synaptic weights was not provided. As can be found in the lower part of Table 1, the synaptic weights were successfully reconstructed from the spike train data by the NS indices in the sense that the ratio between excitatory and inhibitory sources was close to $5.0 : -2.5 = 1 : -0.5$ for all different input rates. We note that the GC and NS indices from Neurons #8 − 12 to Neuron #7 are all zero (i.e., insignificant). As the results show, a large NSI does not necessarily imply a large GCI. That is, a strong synaptic transmission can not always guarantee a strong causal relationship; it depends also on the firing pattern/timing of the source and the coordination with other neurons. So, from this perspective, NSI can be treated as a better proxy for synaptic weights rather than a new causality measure. GCI provides information on causal structure while NSI provides complementary information on synaptic transmission.

In Simulation 2, the input rate λ was fixed at $60\,Hz$ while the synaptic weights varied. Let $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 5\,(mV)$ and $\alpha_5 = \alpha_6 = -k \times 5\,(mV)$. Three different weight-ratio $k = 0.5, 1.0, 1.5$ were considered, and the computed NS indices are presented in Table 2. We can find that the ratio between excitatory and inhibitory sources was still close to $1 : -k$ as weight-ratio changes.

3.3 Nonlinear network

Following the same network topology (Fig. 2) presented in the previous subsection, here the dynamics of Neuron #7 was modeled instead by the Izhikevich's simple spiking neuron model (Izhikevich 2003; Nageswaran et al. 2009; Nedungadi et al. 2009) as it can provide more neural responses compared to classical IF neurons. Briefly, Izhikevich neurons are modeled by the following 2-D differential equation with an after-spiking resetting:

$$v' = 0.04v^2 + 5v + 140 - u + I \tag{9a}$$

$$u' = a(bv - u) \tag{9b}$$

$$\text{if } (v \geq 30mV) \text{ then } v \leftarrow c \text{ and } u \leftarrow u + d. \tag{9c}$$

The variable v represents the membrane potential of the neuron, u represents a membrane recovery variable, and I

Table 1 The numerical results of Simulation 1 in Section 3.2

Input rate	$1 \rightarrow 7$	$2 \rightarrow 7$	$3 \rightarrow 7$	$4 \rightarrow 7$	$5 \rightarrow 7$	$6 \rightarrow 7$	$8 \rightarrow 1$	$9 \rightarrow 6$
Granger Causality Index (GCI)								
$\lambda = 40$	0.0223	0.1972	0.1906	0.1919	0.0497	0.0066	2.9976	3.0085
	(0.0035)	(0.0125)	(0.0155)	(0.0129)	(0.0059)	(0.0031)	(0.0868)	(0.1016)
$\lambda = 60$	0.0411	0.2770	0.2768	0.2769	0.0791	0.0115	2.5959	2.6063
	(0.0038)	(0.0138)	(0.0141)	(0.0156)	(0.0076)	(0.0020)	(0.0573)	(0.0829)
$\lambda = 80$	0.0591	0.3466	0.3425	0.3444	0.1020	0.0156	2.3026	2.3235
	(0.0080)	(0.0188)	(0.0155)	(0.0168)	(0.0083)	(0.0037)	(0.0531)	(0.0575)
Neuron Synaptic Index (NSI)								
$\lambda = 40$	0.1155	0.1145	0.1140	0.1137	-0.0571	-0.0555	2.9972	3.0083
	(0.0043)	(0.0040)	(0.0052)	(0.0046)	(0.0034)	(0.0069)	(0.0868)	(0.1018)
$\lambda = 60$	0.1530	0.1502	0.1505	0.1503	-0.0756	-0.0754	2.5959	2.6054
	(0.0052)	(0.0048)	(0.0045)	(0.0054)	(0.0038)	(0.0027)	(0.0570)	(0.0819)
$\lambda = 80$	0.1799	0.1795	0.1793	0.1783	-0.0893	-0.0898	2.3030	2.3237
	(0.0044)	(0.0049)	(0.0059)	(0.0053)	(0.0036)	(0.0046)	(0.0532)	(0.0571)

The effects of excitations and inhibitions can be differentiated directly by the sign of the NS indices and the ratio of effects between them was close to $5.0 : -2.5 = 1 : -0.5$. Numbers in parentheses are corresponding standard errors

represents the total input synaptic current. The parameter a describes the time scale of u, characterizing the recovery rate. The parameter b describes the sensitivity of u to the subthreshold fluctuations of v. Parameters c and d are spike reset values of v and u, respectively. Two sets of parameter values were considered in this study: the fast spiking (FS) neurons ($a = 0.1$, $b = 0.25$, $c = -65$, $d = 2$) and the low-threshold spiking (LTS) neurons ($a = 0.02$, $b = 0.25$, $c = -65$, $d = 2$).

Here, the parameter settings are: $\lambda = 60\ Hz$, and $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 5\ (mV)$, $\alpha_5 = \alpha_6 = -k \times 5\ (mV)$, all of which with $10\ ms$ propagation delay. The numerical results for the FS and LTS neuron models with weight-ratio $k = 1.0, 1, 5, 2.0$ are summarized in Table 3. The NS indices were obtained from the average of 100 simulations, and the first $0.5\ sec.$ of the voltage fluctuations of the Neuron #7 under FS and LTS models are depicted in Figs. 6 and 7, respectively.

The voltage-trajectories are much regular than that of the almost-linear network (Fig. 3) due to the self-dynamics of the term $0.04v^2 + 5v + 140$ in Eq. (9a). It can be considered that the nature (or the type) of Neuron #7, to some degree, affects its own behavior; therefore, the effects of the inputs (Neurons #1 − 6) will not be equivalent to the underlying mechanism, meaning that the actual degree of effects of the input neurons will not to be proportional to the corresponding synaptic weights.

Now, it can be found from Table 3 that under both the FS and LTS neuron models the negative NSIs from Neurons #5, 6 to Neuron #7 grow in magnitude, relative to the positive ones from Neurons #1−4 to Neuron #7, as weight-ratio k increase. Although the NS indices are not proportional to the underlying weights as what mentioned above, the trend is correctly captured for the increase of the negative NSIs with increasing inhibition strength. Finally, we note that the reason for the absence of the case $k = 0.5$ was

Table 2 The numerical results of Simulation 2 in Section 3.2

weight ratio	$1 \rightarrow 7$	$2 \rightarrow 7$	$3 \rightarrow 7$	$4 \rightarrow 7$	$5 \rightarrow 7$	$6 \rightarrow 7$	$8 \rightarrow 1$	$9 \rightarrow 6$
Neuron Synaptic Index (NSI)								
$k = 0.5$	0.1530	0.1502	0.1505	0.1503	-0.0756	-0.0754	2.5959	2.6054
	(0.0052)	(0.0048)	(0.0045)	(0.0054)	(0.0038)	(0.0027)	(0.0570)	(0.0819)
$k = 1.0$	0.1077	0.1072	0.1063	0.1067	-0.1025	-0.1031	2.5928	2.6151
	(0.0041)	(0.0031)	(0.0044)	(0.0043)	(0.0034)	(0.0037)	(0.0857)	(0.0671)
$k = 1.5$	0.0615	0.0621	0.0612	0.0625	-0.0857	-0.0866	2.5849	2.5869
	(0.0030)	(0.0031)	(0.0036)	(0.0037)	(0.0029)	(0.0039)	(0.0764)	(0.0735)

The input rate λ was fixed at $60\ Hz$, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 5\ (mV)$, and $\alpha_5 = \alpha_6 = -k \times 5\ (mV)$. The ratio of effects between excitatory and inhibitory sources was still close to $1 : -k$ as weight ratio changes. Numbers in parentheses are corresponding standard errors

Table 3 The numerical results of the simulations in Section 3.3

weight ratio	$1 \to 7$	$2 \to 7$	$3 \to 7$	$4 \to 7$	$5 \to 7$	$6 \to 7$	$8 \to 1$	$9 \to 6$
NSI of the FS neuron model								
$k = 1.0$	0.0634	0.0688	0.0691	0.0695	-0.0444	-0.0425	2.3089	2.3414
	(0.0092)	(0.0058)	(0.0070)	(0.0065)	(0.0049)	(0.0094)	(0.0514)	(0.0531)
$k = 1.5$	0.0514	0.0565	0.0572	0.0548	-0.0522	-0.0516	2.3269	2.3450
	(0.0095)	(0.0061)	(0.0062)	(0.0059)	(0.0074)	(0.0096)	(0.0551)	(0.0513)
$k = 2.0$	0.0428	0.0467	0.0460	0.0483	-0.0502	-0.0572	2.3309	2.3379
	(0.0097)	(0.0081)	(0.0074)	(0.0082)	(0.0066)	(0.0099)	(0.0423)	(0.0513)
NSI of the LTS neuron model								
$k = 1.0$	0.0512	0.0536	0.0539	0.0538	-0.0256	-0.0238	2.0241	2.0213
	(0.0095)	(0.0092)	(0.0085)	(0.0071)	(0.0051)	(0.0084)	(0.0351)	(0.0399)
$k = 1.5$	0.0404	0.0438	0.0443	0.0436	-0.0373	-0.0387	1.9861	2.0182
	(0.0093)	(0.0071)	(0.0067)	(0.0059)	(0.0051)	(0.0075)	(0.0434)	(0.0569)
$k = 2.0$	0.0352	0.0394	0.0398	0.0390	-0.0401	-0.0443	2.0244	2.0124
	(0.0093)	(0.0061)	(0.0068)	(0.0042)	(0.0068)	(0.0094)	(0.0554)	(0.0530)

The parameter settings are: $\lambda = 60\ Hz$, and $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 5\ (mV)$, $\alpha_5 = \alpha_6 = -k \times 5\ (mV)$, all of which with $10\ ms$ propagation delay. Numbers in parentheses are corresponding standard errors

that the inhibitory input was so weak to the dynamical system that the resulting NSIs were not significantly different from zero; that is, Neurons #5 − 6 actually did not have any influence on Neuron #7 even though the underlying synaptic weights were not zero.

4 Real data analysis

In this section, we illustrate the application of the proposed method to real spike train data and a simulation is also

given to examine the validity of the results. Note that significant Granger causality interaction are again calculated using an F-test corrected by FDR for multiple comparison with confidence threshold at $P - value = 0.05$.

4.1 Setup and results

Multichannel electrophysiological recording was used for tracking neuronal activity in the anterior cingulate cortex (ACC) and the striatum (STR) and are the same data set used in our previous study (Huang et al. 2013). Briefly,
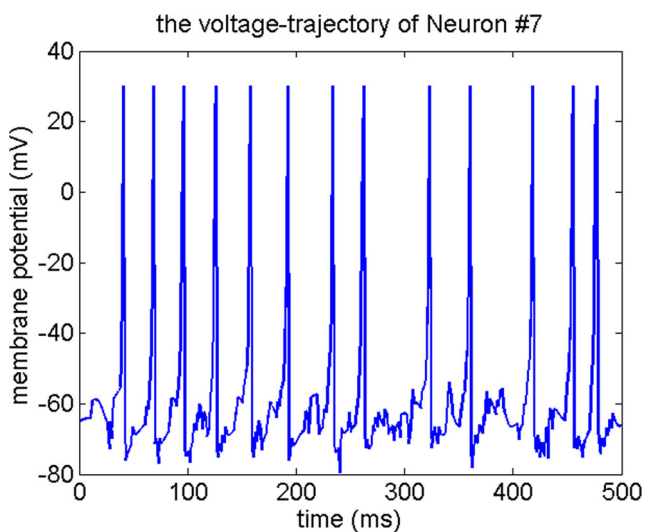
Fig. 6 The first $0.5\ sec.$ of a simulated voltage-trajectory of Neuron #7 with $k = 1$ under the fast spiking (FS) neuron model ($a = 0.1, b = 0.25, c = -65, d = 2$)
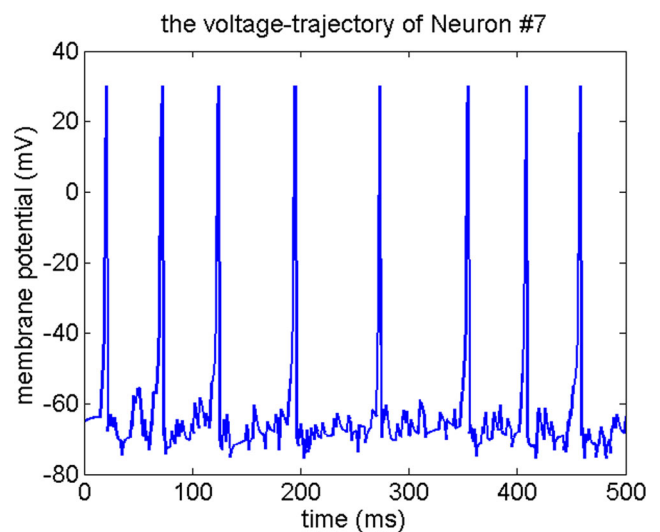
Fig. 7 The first $0.5\ sec.$ of a simulated voltage-trajectory of Neuron #7 with $k = 1$ under the low-threshold spiking (LTS) neuron model ($a = 0.02, b = 0.25, c = -65, d = 2$)

neuronal spikes were recorded from the ACC and the STR of urethane-anesthetized rats after administration of saline or 0.05 or 0.5 mg/kg quinpirole. A multichannel neuronal acquisition processor system (Plexon, Dallas, TX, USA) was used for unit recording, with a filter range of 400 Hz to 8.8 kHz and a sampling rate of 40 kHz. Spikes were further sorted using Offline Sorter (Plexon), based on principle-component clustering with a user-defined template. All animal procedures were approved by the Institutional Animal Care and Use Committee of National Ilan University and adhered to the guidelines established by the Codes for Experimental Use of Animals from the Council of Agriculture, Taiwan.

For this study, data from two independent rats were considered and the numbers of neurons recorded were: 8 (ACC, Rat#1), 9 (STR, Rat#1), 16 (ACC, Rat#2), and 15 (STR, Rat#2). The neurons were randomly put into 2 groups (ACC, Rat#1), 3 groups (STR, Rat#1), 4 groups (ACC, Rat#2), and 3 groups (STR, Rat#2), and each group had 4 (8/2) neurons (ACC, Rat#1), 3 9/3) neurons (STR, Rat#1), 4 (16/4) neurons (ACC, Rat#2), and 5 (15/3) neurons (STR, Rat#2). After the random grouping described above, the single unit spike trains in each group were pooled as a whole for investigating the brain network. Hence there were 5 (2 + 3) pools in Rat#1 and 7 (4 + 3) pools in Rat#2.

After binning with bin width 2 *sec*., the GCIs between these random pools can then be computed (Section 2.3). Twenty minutes after quinpirole injection, 400 *sec*. data from both Rat#1 and Rat#2 were used to compute the GCIs. Significant GCIs were found only when certain random group appear, meaning that certain neurons should be pooled together to perform causality. These specific combinations are summarized in Table 4, and the corresponding GCIs and NSIs are summarized in Table 5. The results show that, under the quinpirole administration, there were excitatory effects inside the ACC (Fig. 8a), excitatory effects from the ACC to the STR (Fig. 8b and c), and inhibitory effects inside the STR (Fig 8d).

Notice that, for single-input case, e.g., in rat 1 from Pool #2 (ACC) to Pool #3 (STR), both GCI (0.1232) and NSI (0.1116) reflect the degree of causal effect. However, the NSI will be more appropriate than the GCI since the NSI is obtained by fitting a more refined autoregressive model (Step 4 in Section 2.3). For multiple-input case, e.g., in rat 2 from Pool #2 (ACC) to Pool #5 (STR) and from Pool #7 (STR) to Pool #5 (STR), the GCI (0.0879 and 0.1598) still reflects the degree of causal effects while the NSI (0.1445 and -0.1058) reflects the degree and the type of synaptic transmission. From this perspective, NSI can be treated as a new complement to provide information on synaptic weights that original GCI does not provide.

Finally, we have to note that, under the saline administration, the same combinations performed no significant GCIs

Table 4 Groups found performing significant NSIs in Section 4.1

Rat no.	Pool no.	location	elements
1	1	ACC	2,4,5,7
1	2	ACC	1,3,6,8
1	3	STR	3,5,7
1	4	STR	1,4,6
1	5	STR	2,8,9
2	1	ACC	4,8,10,16
2	2	ACC	1,5,6,11
2	3	ACC	3,7,9,12
2	4	ACC	2,13,14,15
2	5	STR	6,7,8,9,15
2	6	STR	1,3,10,11,14
2	7	STR	2,4,5,12,13

The numbers of neurons recorded were: 8 (ACC, Rat#1), 9 (STR, Rat#1), 16 (ACC, Rat#2), and 15 (STR, Rat#2). Each group had $8/2 = 4$, $9/3 = 3$, $16/4 = 4$, and $15/3 = 5$ neurons. Hence there were $2 + 3 = 5$ pools in Rat#1 and $4 + 3 = 7$ pools in Rat#2

(i.e., GCIs = 0). Furthermore, the GCIs between the ACC of Rat#1 and the STR of Rat#2, and the GCIs between the STR of Rat#1 and the ACC of Rat#2 were all computed, and they were all zero.

4.2 Implications of the pooled data

A spike train obtained by superimposing individual spike trains and disregarding where each spike came from is called a *pooled spike train* (Gomez et al. 2005). Adjacent neurons usually work together with each other to generate suitable pooled spike trains to perform specific tasks. An illustration is provided in Fig. 9. On the cause side (left brown), the pooled train (pool 1) can be considered as a collective input with respect to the effects of temporal or spatial summation of one of the following two types: (i) the additive effect produced by many PSPs that have been generated from several very close synapses on the same post-synaptic neuron at the same time. (ii) the additive effect produced by many PSPs that have been generated from

Table 5 The NSIs between the pooled data from the combinations summarized in Table 4

| Rat no.	From	To	GCI	NSI
1	Pool #2 (ACC)	Pool #1 (ACC)	0.1966	0.1035
1	Pool #2 (ACC)	Pool #3 (STR)	0.1232	0.1116
1	Pool #2 (ACC)	Pool #4 (STR)	0.1163	0.1995
2	Pool #2 (ACC)	Pool #5 (STR)	0.0879	0.1445
2	Pool #2 (ACC)	Pool #6 (STR)	0.1098	0.1121
2	Pool #7 (STR)	Pool #5 (STR)	0.1598	-0.1058
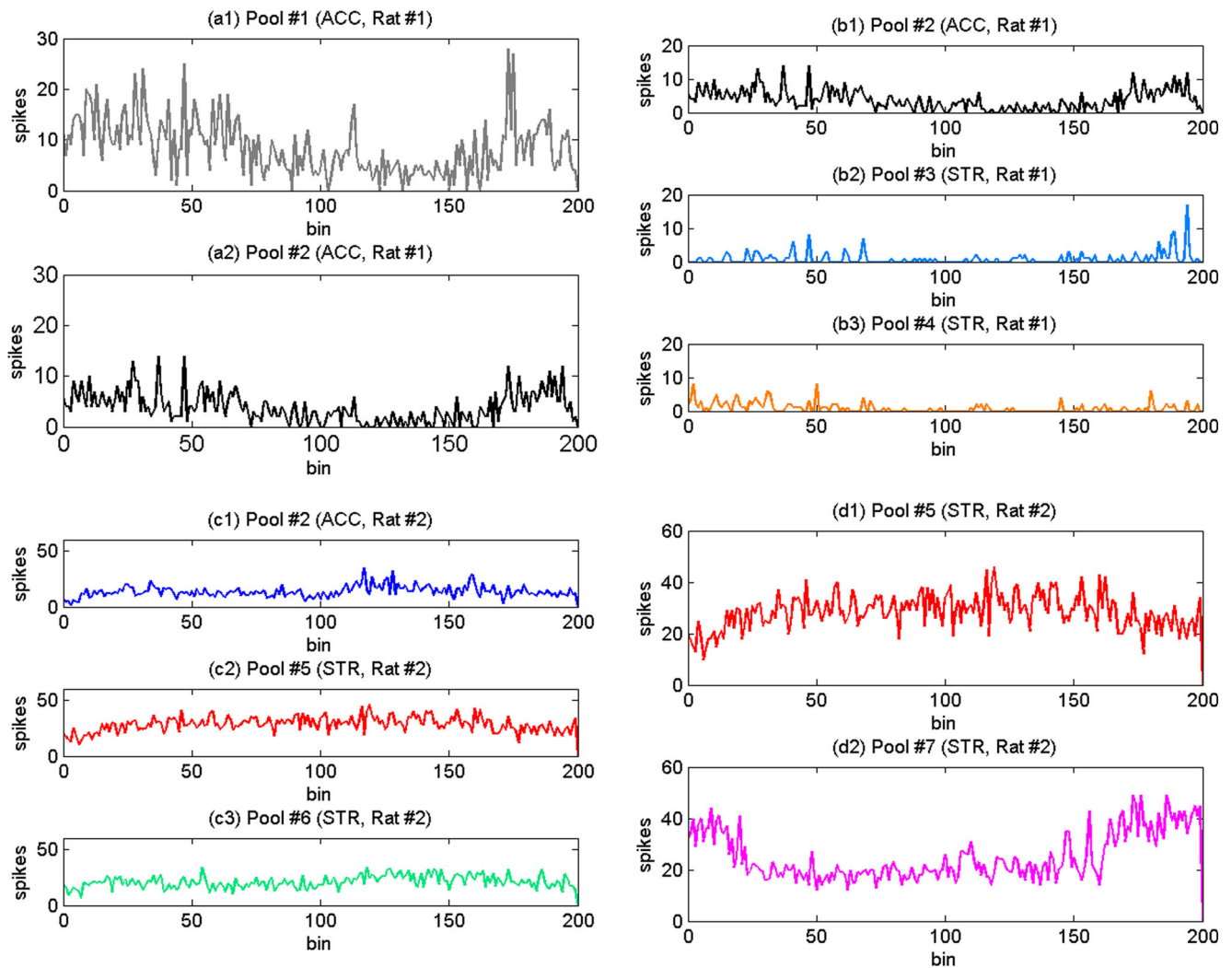
Fig. 8 The firing trajectories (400 *sec.* with 2 *sec.* bin) under the quinpirole administration. **a** excitatory effects inside the ACC (NSI = 0.1035). **b** excitatory effects from the ACC to the STR (NSIs = 0.1116, 0.1995). **c** excitatory effects from the ACC to the STR (NSIs = 0.1445, 0.1121). **d** inhibitory effects inside the STR (NSI = −0.1058). It can be found that positive (negative) NSIs exhibit positive (negative) correlations in the fluctuations of the signals

several synapses which have similar effects on the axon hillock of the same post-synaptic neuron. On the effect side (middle blue), the pooled train (pool 2) can be considered as a collective output, which represents the total discharge of a group of cooperative neurons in function. Again, the collective output (pool 2) can then be treated as a collective input to trigger others (right purple).

Cross correlations can be dramatically amplified by pooling, that is, weak correlations between pairs of neurons in two populations can lead to strong correlations between the summed activity of these two populations (Rosenbaum et al. 2011). Similar results should hold for the GC analyses. To check this, a simulation is designed as follows: Let $P = \{p_1, p_2, ..., p_n\}$ be a Poisson spike train of time length T with firing rate λ. Let Q be the output spike train of the almost linear system in Section 3.2 with input P, synaptic weight w, and time delay d. Since P will be treated as a pooled spike train, we uniformly decompose it into k sub-trains $\{P_i, i = 1, \cdots, k\}$, that is, each $p_j \in P$ has the same probability to be distributed into the sub-train P_i, for $j = 1, \cdots, n$ and $i = 1, \cdots, k$. As a result, $\cup_i P_i = P$, $\cap_i P_i = \emptyset$, and the firing rate of each P_i is λ/k. Further, let $\{U_i, i = 1, \cdots, m\}$ be m uncorrelated spike trains with $\{P_i, i = 1, \cdots, k\}$ to serve as the role of environment neurons. The distribution of each U_i is also Poisson with rate λ/k. Now, for $T = 10 (sec.)$, $\lambda = 20 (spikes/sec.)$, $w = 8 (mV)$, $d = 10 (ms)$, $k = 5$, $m = 2$, and bin width $0.1(sec.)$, the GCI from P to Q is 0.2655, a quite large value; while the GCIs from P_i to Q is about 0.005, a very low causality. The results are obtained from the average of 100 such simulations, and the first $2sec.$ of one of the realizations is shown in Fig. 10; where P is labeled neuron
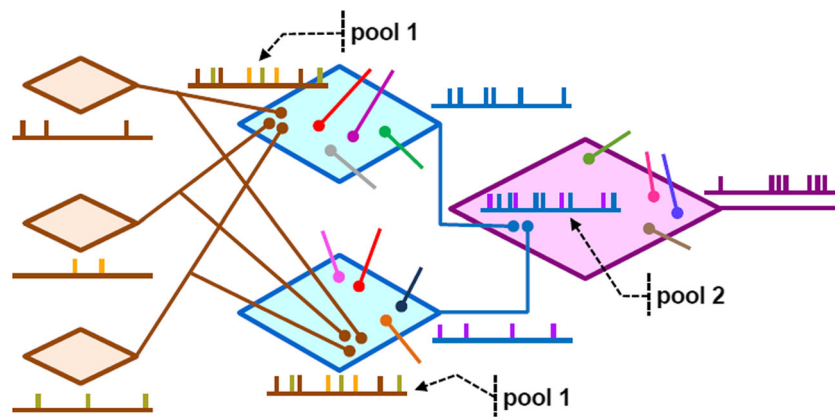
Fig. 9 An illustration for pooled data. On the cause side (left brown), the pooled train (pool 1) can be considered as a collective input with respect to the effects of temporal or spatial summation of one of the two types ((i) and (ii) in the context.) On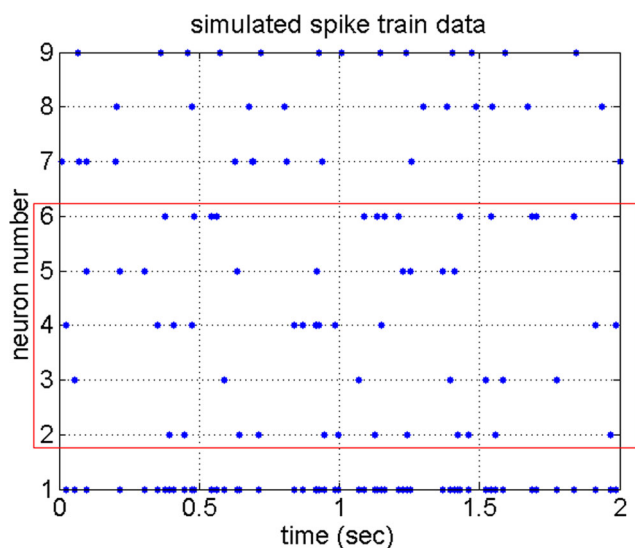 the effect side (middle blue), the pooled train (pool 2) can be considered as a collective output, which represents the total discharge of a group of cooperative neurons in function. Again, the collective output (pool 2) can then be treated as a collective input to trigger others (right purple)

#1, P_i, neuron $\#2 - 6$, U_i, neuron $\#7 - 8$, and Q is labeled neuron #9.

To check the appropriateness of random grouping used in the previous subsection, 5 spike trains are randomly chosen from P_i and U_i and then be pooled together to compute the GCI from such pooled data to the target Q. The averaged results are: when these 5 spike trains are all chosen from P_i then the GCI attends the maximum 0.2655. When 4 is chosen from P_i and 1 from U_i the GCI is destroyed and is 0.0641. Finally, when 3 is chosen from P_i and the other 2 are from U_i, the GCI is 0.0209. The results show that if the pooled data contains the spikes of other irrelevant neurons then the GCIs will be small and destroyed.

For real-world spike train data, individual neurons usually perform weak contributions to each other while groups of neurons perform very significant contributions. In the former case, causal influences are difficult to be detected via most statistical methods, grouping and pooling are usually needed to enhance the causation. Since our data set is small, random grouping approach is both reasonable and sufficient to explore the network structure. Significant NSIs are also found within an acceptable period of time. How to efficiently group neurons is absent in our current analysis, but has been being studied via numerical simulations. Efficient grouping strategy is an interesting research topic and will be included in a separate article in the future.

5 Discussion

The original Granger causality index is by definition non-negative, thereby lacking of the trait for differentiating the effects of excitations and inhibitions between neurons. Inspired by the concept that the firing pattern of the post-synaptic neuron is generally a weighted result of the effects of several pre-synaptic neurons with possibly different synaptic weights; a computational algorithm was proposed (Section 2.3) under a BLP assumption for analyzing neuronal networks by estimating the synaptic weights among them. The extended Granger causality index, the NSI, was shown to be able to measure both excitatory and inhibitory effects between neurons by several numerical simulations. The method was also illustrated to analyze real spike train data from the ACC and the STR. Significant

Fig. 10 The first 2 *sec*. simulated spike trains of one of the realizations. The parameters are: $T = 10$ (*sec*.), $\lambda = 20$ (*spikes/sec*.), $w = 8$ (*mV*), $d = 10$ (*ms*), $k = 5$, $m = 2$, and bin width 0.1(*sec*.), The source spike train P is labeled neuron #1, the decomposed trains P_i are labeled neuron $\#2 - 6$, the environment trains U_i are labeled neuron $\#7 - 8$, and the target Q is labeled neuron #9

NSIs were found only when certain random groups appear (Tables 4 and 5); and the results showed, under the quinpirole administration, the significant existence of excitatory effects inside the ACC, excitatory effects from the ACC to the STR, and inhibitory effects inside the STR.

Mathematically, the proposed NSI is just a weighted version of the original GCI, that is, the NSI is obtained by multiplying the normalized weights with the GCI from the weighted trajectory to the target trajectory. Physiologically, under the prerequisite conditions that (i) all of the connectivity relations between neurons are correctly identified and (ii) the behavior of influences follow the vector autoregressive model with some finite order; the original GCI can reveal the degree of causal influences between neurons while the extended NSI can reveal the degree and the type of synaptic transmission. Theoretically, the GCI and NSI can only be approximations or even be spurious if either of these two conditions fail to hold. It seems to be a very strong constraint for researchers to apply these approaches in practice. However, things are not really that bad. The simulated IF spiking networks in Section 3.2–3.3 are tested to be correctly captured by vector autoregressive model and show a consistency between network structures and the proposed indices. Finally and also importantly, note that significant GCI or NSI do not necessarily signify a anatomical connectivity since they are fundamentally statistical concepts. Treated with care, the GCI and NSI could be useful for researchers to infer possible relationships between network structure or to construct a description of network dynamics in neuroscience.

It should be pointed out that the significance of the NSI can be checked directly via the significance test on the GCI, since the proposed method is GC-based. In addition, if voltage-trajectories are available then the method can be applied directly without any ambiguity. For spike train data, however, the role of voltage-trajectory can be replaced by the trajectory of firing rate estimated using binning or Gaussian kernel filtering. The high firing rate in the simulations are to ensure significant inhibitory effects among the simple networks. In practice, the high firing rate can be considered as a consequence of pooling spike trains. Low firing rate leads to sparse spike trains, for which the Gaussian kernel filtering will introduce highly artificial signals which will hinder the autoregression modeling for computing the GCI. The solution is applying some suitable preprocessing (Zhu et al. 2003) or applying binning (Shimazaki and Shinomoto 2007), which is more stable for converting sparse spike trains.

In closing, it is worth noting some related articles for possible future works. (i) In Section 3.3, the NSI was shown to successfully capture the behavior of synaptic transmissions. Although the NSI values are correlated with the underlying synaptic weights, it is not clear what the actual relationship

might be. Cadotte et al. have found that under certain settings, the GCI and the synaptic weight has the following nonlinear relationship (Cadotte et al. 2008):

$$F_{Y \to X} = \frac{1}{1 + 384e^{-02124*sw}}, \tag{10}$$

where sw denotes the synaptic weight. Similar results for the NSI may be derived in the future using Eq. (10). (ii) When faced with most scientific computing problem, a linear model is generally a first, basic, and winning strategy to try. Neural spiking networks are highly complicated and nonlinear, a linear model could be considered to be satisfactory if it can well approximates the behavior of a nonlinear dynamics to some degree. In the current study, the behavior of the simulated nonlinear network is well captured by our approach; obtaining the exact ratio of the synaptic weights can be set as an important objective for us to strive. Our proposed method and the BLP assumption could be reformulated and generalized to handle nonlinear dynamics by means of the nonparametric kernel modelling (Marinazzo et al. 2011) in the future. (iii) This work did not consider any hidden network structure in both theoretical and numerical parts. The partial Granger causality, proposed by Guo et al. (2008), may be used to extend the method to deal with the effects of exogenous inputs or hidden neurons. (iv) Granger causality is originally designed to measure effect, not mechanism (Barrett and Barnett 2013). Numerical evidences showed that Granger causality can, to some degree, be used to infer the underlying mechanism. The GCI only use the information of residual noises, there are still some useful information which could be extracted from the regression coefficients (Hu et al. 2011). The NSI uses the summation of regression coefficients Eq. (7), other forms of information may be developed in the future. (v) The Gaussian filtering and binning techniques link the spike train data and the statistical models for continuous signals, leading to both mathematically easy derivations and computationally efficient algorithms. However, distortions may arise after the filtering is applied. A generalized linear model (GLM)-based point process framework was proposed for directly applying the GC on neural spike trains without any filtering (Kim et al. 2011). A conceptually similar but more robust measure, called directed information, was also proposed (Quinn et al. 2011). The modality-independent nature allows the measure to characterize statistically causal relationship between arbitrary stochastic processes. A more sophisticated coupling model was also proposed (Pillow et al. 2008). It's parameters consist of a bank of stimulus filters, spike-history filters, and couping filters. Splines can also be used to fit nonlinearity in the stimulus filter. The mathematics used in these frameworks are more involved than that used in this paper, but we surmise that the proposed algorithm and assumption can somehow be translated

into them to obtain similar or more powerful results in the future. (vi) Recently, a new framework of spatio-temporal Granger causality has been proposed to reliably estimate the Granger causality from experimental datasets possessing time-varying properties (Luo et al. 2013). The NSI may be extended to its dynamic version for automatically analyzing experimental datasets without laborious jobs on windowing.

Conflict of interests The authors declare that they have no conflict of interest.

Appendix: Derivation of the NSI using simple network

Here, we re-formulate the NSI using the simple network (Fig. 1). Let $u = \alpha x + \beta y + \gamma z$ form the BLP of w, then there exist p, $\{f_r, r = 1, 2, \cdots, p\}$, and $\{d_r, r = 1, 2, \cdots, p\}$ such that $w_t = \sum_{r=1}^{p}[f_r u_{t-r} + d_r w_{t-r}] + \epsilon_t$, where ϵ is a stationary white noise possessing the smallest variance among $\mathcal{G} = span(\{x, y, z, v_1, v_2, v_3\})$. Replacing u with the weighted trajectory, we obtain

$$
\begin{aligned}
w_t &= \sum_{r=1}^{p}[f_r u_{t-r} + d_r w_{t-r}] + \epsilon_t \\
&= \sum_{r=1}^{p}[f_r(\alpha x_{t-r} + \beta y_{t-r} + \gamma z_{t-r}) + d_r w_{t-r}] + \epsilon_t \\
&= \sum_{r=1}^{p}[\alpha f_r x_{t-r} + \beta f_r y_{t-r} + \gamma f_r z_{t-r} + d_r w_{t-r}] + \epsilon_t.
\end{aligned}
$$
(A.1)

On the other hand, fitting to data the following empirical regression

$$
w_t = \sum_{r=1}^{p}[a_r x_{t-r} + b_r y_{t-r} + c_r z_{t-r} + g_r v_{t-r} + d_r w_{t-r}] + \tilde{\epsilon}_t,
$$
(A.2)

where $g_r v_{t-r} := \sum_{k=1}^{3} g_{k,r} v_{k,t-r}$ for convenience.

- If v is stochastically independent of x, y, z, w, then we have $g_r \equiv 0$. Since $\{a_r\}, \{b_r\}, \{c_r\}$ can be obtained

through Least-Squares method, comparing Eq. (A.2) with Eq. (A.1), we have

$$
\sum_{r=1}^{p} a_r = \alpha \sum_{r=1}^{p} f_r, \ \sum_{r=1}^{p} b_r = \beta \sum_{r=1}^{p} f_r, \ \sum_{r=1}^{p} c_r = \gamma \sum_{r=1}^{p} f_r,
$$
(A.3)

and get

$$
\alpha : \beta : \gamma = \sum_{r=1}^{p} a_r : \sum_{r=1}^{p} b_r : \sum_{r=1}^{p} c_r, \text{ provided } \sum_{r=1}^{p} f_r > 0,
$$
(A.4)

where $sgn(\alpha) = sgn\left(\sum_{r=1}^{p} a_r\right)$, $sgn(\beta) = sgn\left(\sum_{r=1}^{p} b_r\right)$, and $sgn(\gamma) = sgn\left(\sum_{r=1}^{p} c_r\right)$.

- If v is linear dependent of x, y, z, w, then $g_r \gg 0$ and $\{a_r\}, \{b_r\}, \{c_r\}$ will be affected. However, since ϵ in Eq. (A.1) possessing the smallest variance among $\mathcal{G}$, taking out v does not increase the variance of $\tilde{\epsilon}$, therefore we still can correct the model coefficients by ruling out the useless information v.

Finally, the neuron synaptic index from x, y, z to w are defined respectively as

$$
\begin{aligned}
N_{x \to w} &:= \frac{\alpha}{|\alpha| + |\beta| + |\gamma|} F_{u \to w}, \\
N_{y \to w} &:= \frac{\beta}{|\alpha| + |\beta| + |\gamma|} F_{u \to w}, \\
N_{z \to w} &:= \frac{\gamma}{|\alpha| + |\beta| + |\gamma|} F_{u \to w},
\end{aligned}
$$
(A.5)

where $|N_{x \to w}| + |N_{y \to w}| + |N_{z \to w}| = F_{u \to w}$ is the GC index from the weighted trajectory $u = \alpha x + \beta y + \gamma z$ to the target trajectory w.

References

Arnold, A., Liu, Y., & Abe, N. (2007). Temporal causal modeling with graphical Granger methods. *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*, 7, 66–75.

Baccala, L.A., & Sameshima, K. (2001). Partial directed coherence : a new concept in neural stucture determination. *Biological Cybernetics*, 84, 463–474.

Barnett, L., & Seth, A.K. (2014). The MVGC multivariate Granger causality toolbox: A new approach to Granger-causal inference. *Journal of Neuroscience Methods*, 223, 50–68.

Barrett, A.B., & Barnett, L. (2013). Granger causality is designed to measure effect, not mechanism. *Frontiers in Neuroinformatics*, 7, 1–2.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society*, 57, 289–300.

Bressler, S.L., & Seth, A.K. (2011). Wiener-Granger Causality: A well established methodology. *NeuroImage*, 58, 323–329.

Cadotte, A.J., DeMarse, T.B., He, P., & Ding, M. (2008). Causal measures of structure and plasticity in simulated and living neural networks. *PLoS Computational Biology*, 3, 1–14.

Cadotte, A.J., DeMarse, T.B., Mareci, T.H., Parekh, M.B., Talathi, S.S., Hwang, D.U., Ditto, W.L., Ding, M., & Carney, P.R. (2010). Granger causality relationships between local field potentials in an animal model of temporal lobe epilepsy. *Journal of Neuroscience Methods*, *189*, 121–129.

Dhamala, M., Rangarajan, G., & Ding, M. (2008). Analyzing information flow in brain networks with nonparametric Granger causality. *NeuroImage*, *41*, 354–362.

Ding, M., Chen, Y., & Bressler, S.L. (2006). Granger Causality: Basic Theory and Application to Neuroscience. *Handbook of Time Series Analysis: Recent Theoretical Developments and Applications*, (pp. 437–460). Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA.

Gomez, L., Budelli, R., Saa, R., Stiber, M., & Segundo, J.P. (2005). Pooled spike trains of correlated presynaptic inputs as realizations of cluster point processes. *Biological Cybernetics*, *92*, 110–127.

Granger, C. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, *37*, 424–438.

Granger, C. (1980). Testing for causality: A personal viewpoint. *Journal of Economic Dynamics and Control*, *2*, 329–352.

Greene, W.H. (2002). *Econometric Analysis*, fifth ed. Upper Saddle River, NJ: Prentice-Hall.

Guo, S., Seth, A.K., Kendrick, K.M., Zhou, C., & Feng, J. (2008). Partial Granger causality – Eliminating exogenous inputs and latent variables. *Journal of Neuroscience Methods*, *172*, 79–93.

Hu, S., Dai, G., Worrel, G.A., Dai, Q., & Liang, H. (2011). Causality analysis of neural connectivity: critical examination of existing methods and advances of new methods. *IEEE Transactions on Neural Networks*, *22*, 829–844.

Huang, J.J., Yen, C.T., Liu, T.L., Tsao, H.W., Hsu, J.W., & Tsai, M.L. (2013). Effects of dopamine D2 agonist quinpirole on neuronal activity of anterior cingulate cortex and striatum in rats. *Psychopharmacology*, *227*, 459–466.

Izhikevich, E.M. (2003). Simple models of spiking neurons. *IEEE Transactions on Neural Networks*, *14*, 1569–1572.

Kim, S., Putrino, D., Ghosh, S., & Brown, E.N. (2011). A Granger causality measure for point process models of ensemble neural spiking activity. *PLoS Computational Biology*, *7*, 3.

Kitagawa, G. (2010). *Introduction to Time Series Modeling*: Chapman & Hall/CRC Monographs on Statistics & Applied Probability.

Krumin, M., & Shoham, S. (2010). Multivariate autoregressive modeling and Granger causality analysis of multiple spike trains. *Computational Intelligence and Neuroscience*, 752428.

Lehky, S.R. (2010). Decoding poisson spike trains by gaussian filtering. *Neural Computation*, *22*, 1245–1271.

Luo, Q., Lu, W., Cheng, W., Valdes-Sosa, P.A., Wen, X., Ding, M., & Feng, J. (2013). Spatio-temporal Granger causality: A new framework. *NeuroImage*, *79*, 241–263.

Lutkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*: Springer.

Marinazzo, D., Liao, W., Chen, H., & Stramaglia, S. (2011). Nonlinear connectivity by Granger causality. *NeuroImage*, *58*, 330–338.

Michailidis, G., & d'Alche-Buc, F. (2013). Autoregressive models for gene regulatory network inference: Sparsity, stability and causality issues. *Mathematical Biosciences*, *246*, 326–334.

Nageswaran, J.M., Dutt, N., Krichmar, J.L., Nicolau, A., & Veidenbaum, A.V. (2009). A configurable simulation environment for the efficient simulation of large-scale spiking neural networks on graphics processors. *Neural Networks*, *22*, 791–800.

Nedungadi, A.G., Rangarajan, G., Jain, N., & Ding, M. (2009). Analyzing multiple spike trains with nonparametric granger causality. *Journal of Computational Neuroscience*, *27*, 55–64.

Pillow, J.W., Shlens, J., Paninski, L., Sher, A., Litke, A.M., Chichilnisky, E.J., & Simoncelli, E.P. (2008). Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature*, *454*, 995–999.

Quinn, C.J., Coleman, T.P., Kiyavash, N., & Hatsopoulos, N.G. (2011). Estimating the directed information to infer causal relationships in ensemble neural spike train recordings. *Journal of Computational Neuroscience*, *30*, 17–44.

Rosenbaum, R., Trousdale, J., & Josic, K. (2011). The effects of pooling on spike train correlations. *Frontiers in Neuroscience*, *5*, 1–10.

Sameshima, K., & Baccala, L.A. (1999). Using partial directed coherence to describe neuronal ensemble interactions. *Journal of Neuroscience Methods*, *94*, 93–103.

Seth, A.K. (2005). Causal connectivity of evolved neural networks during behavior. *Network*, *16*, 35–54.

Seth, A.K., & Edelman, G.M. (2007). Distinguishing causal interactions in neural populations. *Neural Computation*, *19*, 910–933.

Seth, A.K. (2010). A MATLAB toolbox for Granger causal connectivity analysis. *Journal of Neuroscience Methods*, *186*, 262–273.

Shao, P.C., Tseng, W.T., Kuo, C.C., Shann, W.C., Tsai, M.L., & Yen, C.C. (2013). Effects of spike sorting error on the Granger causality index. *Neural Networks*, *46*, 249–259.

Shimazaki, H., & Shinomoto, S. (2007). A method for selecting the bin size of a time histogram. *Neural Computation*, *19*, 1503–1527.

Zhou, D., Xiao, Y., Zhang, Y., Xu, Z., & Cai, D. (2014). Granger causality network reconstruction of conductance-based integrate-and-fire neuronal systems. *PLoS ONE*, *9*, 2.

Zhu, L., Lai, Y.C., Hoppensteadt, F.C., & He, J. (2003). Probing changes in neural interaction during adaptation. *Neural Computation*, *15*, 2359–2377.

Zou, C., Ladroue, C., Guo, S., & Feng, J. (2010). Identifying interactions in the time and frequency domains in local and global networks - A Granger causality approach. *BMC Bioinformatics*, *11*, 337.