Commun Nonlinear Sci Numer Simulat 57 (2018) 342–351

Contents lists available at ScienceDirect

Commun Nonlinear Sci Numer Simulat

journal homepage: www.elsevier.com/locate/cnsns

Research paper

Reconstruction of coupling architecture of neural field networks from vector time series

Ilya V. Sysoev[a,b,*], Vladimir I. Ponomarenko[a,b], Arkady Pikovsky[c,d]

[a] Saratov State University, 83, Astrakhanskaya str., Saratov, Russia
[b] Saratov Branch of Institute of Radioengineering and Electronics of RAS, 38, Zelyonaya str., Saratov, Russia
[c] University of Potsdam, Institute for Physics and Astronomy, Karl-Liebknecht-Str. 24–25, Potsdam 14476, Germany
[d] Institute for Supercomputing, Lobachevsky State University of Nizhni Novgorod, 23 Prospekt Gagarina, Nizhni Novgorod, Russia

ARTICLE INFO

Article history:
Received 11 July 2017
Revised 5 October 2017
Accepted 11 October 2017
Available online 13 October 2017

Keywords:
Network reconstruction
Time series
Neurooscillators
Time delay

ABSTRACT

We propose a method of reconstruction of the network coupling matrix for a basic voltage-model of the neural field dynamics. Assuming that the multivariate time series of observations from all nodes are available, we describe a technique to find coupling constants which is unbiased in the limit of long observations. Furthermore, the method is generalized for reconstruction of networks with time-delayed coupling, including the reconstruction of unknown time delays. The approach is compared with other recently proposed techniques.

1. Introduction

Interaction mechanisms play a crucial role in the organization of complex networks. One of the main empirical approaches here, widely used in neuroscience, climatology, cardio-physiology, economics, social sciences etc [1–3]., is infering the properties of the network by virtue of the analysis of the available time series from the dynamics. Here, one particularly aims at reconstruction of the properties of coupling and, possibly, of some properties of individual oscillators, from the observations of all involved subsystems. For networks of oscillatory elements, these time series are typically represented by records from all the nodes of a network. Special approaches of model reconstruction, suggested for cases when the data are insufficient due to hidden variables [4,5], or observations of some subsystems are not available[6], or there is additional unknown external driving [7], are potentially interesting and practically important, but they lie beyond the scope of this paper.

Most approaches to the coupling analysis could be split in two groups:

1. Pairwise approaches such as Granger causality [8]. We mention here various nonlinear extensions [9–11], transfer entropy [12], mutual information [13], different types of nonlinear correlation [14,15].
2. Full ensemble approaches, including partial directed coherence [16], some linear autoregressive frameworks [17], generalizations of nonlinear correlation [18], approaches based on minimization of length (or some other functional) of nonlinear functions, developed for time-delayed equations [19].

* Corresponding author at: Saratov State University, 83, Astrakhanskaya str., Saratov, 410012, Russia.
E-mail addresses: ivssci@gmail.com (I.V. Sysoev), ponomarenkovi@gmail.com (V.I. Ponomarenko), pikovsky@uni-potsdam.de (A. Pikovsky).

https://doi.org/10.1016/j.cnsns.2017.10.006
1007-5704/

The pairwise approaches consider each pair of elements of the network under study as independent of the rest. This means that a model (e.g. for Granger causality) or a statistical distribution function (for transfer entropy or mutual information function) are constructed only for individual nodes and pairs, but not for the whole ensemble. This simplification allows one to reduce the demands for the amount of data (time series could be shorter), to use simpler models and to increase robustness of estimates of coupling constants. On the other hand, in such setups indirect interactions or presence of a common external driving could be falsely interpreted as a valid coupling, because information about other participating nodes is not taken into account.

To avoid this artifact, one modifies the approach to include information about possible third party. Conditional Granger causality [21] and interaction information [22] are examples of such modifications. These techniques could be efficient, if one deals with a small number of nodes, usually three or four. For larger ensembles, the model structure occurs to be too complex and the number of unknown coupling coefficients becomes too large for a reliable estimation.

The full ensemble approaches are usually protected from false positive results due to an indirect coupling or a common force, since all subsystems that are important for the dynamics of the network are considered simultaneously. However, the complexity of the problem appears to be high, and in most cases one needs to know model equations to apply these techniques, otherwise the situation becomes too uncertain. It is therefore highly desirable to develop ensemble approaches requiring minimal knowledge on the specific equations and/or assumptions, at least for practical important classes of oscillatory networks.

Sompolinsky et al. [20] have formulated and analyzed the basic network model of an ensemble of oscillators, describing complex dynamics of neural bursting activity. The possibility of transition to chaos was theoretically proved in case of $D \to \infty$, where D is the number of units. Here we propose a technique for reconstruction of coupling constants for this model, given the time series of all elements. Furthermore, we consider a slightly modified system, which includes generally different delays for all coupling links, and extend the technique for this generalized system, being able to reconstruct not only all coupling constants, but also all the time delays. The asymptotic behavior of the estimates is also studied theoretically and numerically.

2. Models and methods

2.1. Voltage-based neural field network model

Our basic model, described in [20], belongs to a class of voltage-based neural field networks. Each node is described by the membrane potential of a neuron x_i, so that its activity (firing rate) $h(x_i)$ affects through j the connections the potentials of other neurons:

$$\begin{aligned} \dot{x}_i &= -x_i + \sum_{j=1, j\neq i}^{D} k_{i,j} h(x_j), \\ & i = 1, 2, \ldots, D, \\ & h(x) = \tanh(gx). \end{aligned} \tag{1}$$

Here coupling constants $k_{i,j}$ — elements of the coupling matrix $\hat{K}$ — characterize the synaptic efficacy and are generated from a normal distribution with zero mean and standard deviation $J/\sqrt{D}$, where J is a parameter defining the coupling strength. It has been proved in [20], that the ensemble (1) demonstrates chaotic behavior in the thermodynamic limit $D \to \infty$ and $(Jg) > 2$, for an arbitrary typical matrix $\hat{K}$.

While it was not proved theoretically, a chaotic behavior could be found even for finite D, although here stable regimes appear to be dominant for small D. For example, even for small networks with $D = 16$, the numerical solution with 4th order Runge–Kutta method revealed that 3 ensembles of 120 trial ones (ensembles differed one from another by the coupling matrix $\hat{K}$) demonstrate irregular behavior, the example of which (time series and spectra) is plotted in Fig. 1. To determine the chaotic nature of oscillations, the largest Lyapunov exponent was calculated for all ensembles. However, with this approach it is not possible to distinguish between a strange attractor and very a long chaotic transient process. This difference is not important for the reconstruction below, provided the chaotic part of the time series is used. The portion of matrices $\hat{K}$ yielding a chaotic regime increases with the network size D.

2.2. Technique of reconstruction

Given a time series of all nodes $\{x_i(t_n)\}_{n=1}^{N}$ of length N (a vector or multivariate time series), obtained with a sampling time Δt, let us denote $x_i(n) = x_i(t_n)$ and $\dot{x}_i(n) = \dot{x}_i(t_n)$. The time series of the derivatives are calculated numerically from $x_i(n)$. In these notations, Eq. (1) can be rewritten as:

$$x_i(n) = -\dot{x}_i(n) + \sum_{j=1, j\neq i}^{D} k_{i,j} h(x_j(n)) \tag{2}$$

As the next step, we, for a given i, sort all the values of $\{x_i(n)\}_{n=1}^{N}$ (e.g. in the ascending order, this in fact does not matter). Let us introduce the map Q_i which maps the position of the point n in the original series to its position $Q_i(n)$ in

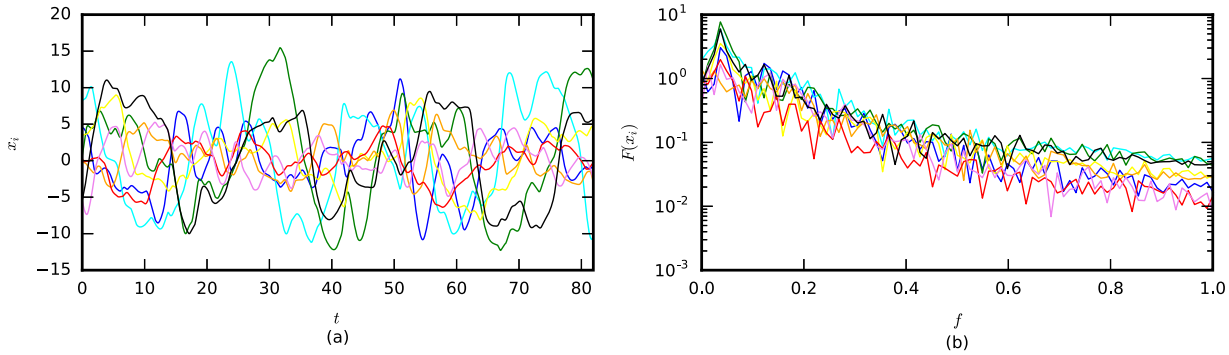

Fig. 1. Panel (a): time series of first 8 oscillations of ensemble (1) with $D = 16$ nodes; Panel (b): power spectrum of oscillations of first 8 oscillators out of 16 in total.

the sorted one. Also, let us consider the inverse map Q_i^{-1} which maps the position $Q_i(n)$ in the sorted series to the position n in the original one. So, $Q_i^{-1}(Q_i(n)) = n$.

Let us consider two neighbor values in the sorted series, which in the original time series have the positions n and $p_n = Q_i^{-1}(Q_i(n) - 1)$: in the sorted series $x_i(p_n)$ just precedes $x_i(n)$. We denote the difference between these values as:

$$\delta_i(n) = x_i(n) - x_i(p_n) \tag{3}$$

Let us also consider the simultaneous values of all other series $\{x_j\}$: $x_j(n)$ and $x_j(p_n)$. Thus we can express $\delta_i(n)$ in (3), using (2), as follows:

$$\delta_i(n) = \sum_{j=1, j\neq i}^{D} k_{i,j}\Delta h_{i,j}(n) - \Delta\dot{x}_i(n), \tag{4}$$

$$\Delta\dot{x}_i(n) = \dot{x}_i(n) - \dot{x}_i(p_n), \tag{5}$$

$$\Delta h_{i,j}(n) = h(x_j(n)) - h(x_j(p_n)). \tag{6}$$

Since the amplitude of oscillations is finite, the values $x_i(n)$ and $x_i(p_n)$, being neighbors in the sorted time series, are very close to each other, if N is large enough. To characterize this smallness, we introduce the sum of squares of $\delta_i(n)$:

$$S_i^2(\mathbf{k}_i) = \sum_{n=1}^{N}\delta_i^2(n) = \sum_{n=1}^{N}\left(\sum_{j=1, j\neq i}^{D} k_{i,j}\Delta h_{i,j}(n) - \Delta\dot{x}_i(n)\right)^2, \tag{7}$$

where $\mathbf{k}_i = \left(k_{i,1}, \ldots, k_{i,D}\right)$, and additionally $Q_i(n) \neq 1$.

Let us now examine the r.h.s. of Eq. (7). Due to chaotic and asynchronous behavior, closeness of values $x_i(n)$ and $x_i(p_n)$ does not lead to closeness of simultaneous values of other coordinates: $x_j(n)$ and $x_j(p_n)$, as well as derivatives $\dot{x}_i(n)$ and $\dot{x}_i(p_n)$ are generally not close. Therefore, $\Delta h_{i,j}(n)$ and $\Delta\dot{x}_i(n)$ are not small. Thus, the only way the r.h.s. of Eq. (7) could be as small as the l.h.s., is the case, when the correct values of $k_{i,j}$ are employed and the terms $k_{i,j}\Delta h_{i,j}(n)$ and $\Delta\dot{x}_i(n)$ cancel each other.

If we know the coupling functions h, the formula (7) can be considered as a linear least square problem, where S_i^2 is a target function, which has to be minimized. Here $\Delta\dot{x}_i(n)$ and $\Delta h_{i,j}(n)$ are known from the time series (the former quantities approximately as a result of numerical differentiation). Thus, the coupling matrix $\hat{K}$ can be estimated from the time series as a solution of the linear least square problem (7).

To check that this method indeed works for long time series, let us consider behavior of the target function (7) in the limit $N \to \infty$. In such a case, if the oscillation amplitude is limited, all $\delta_i(n)$ decay as $\sim N^{-1}$ on average, so the squares $\delta_i^2(n)$ decay as $\sim N^{-2}$. Therefore, the sum (7) has to decay as $\sim N \cdot N^{-2} = N^{-1}$. This means that $S_i^2(\mathbf{k})$ tends to zero in the limit $N \to \infty$, and thus the estimates $\tilde{k}_{i,j}$ of the coupling constants $k_{i,j}$ obtained with the proposed algorithm are asymptotically unbiased.

2.3. Interpolation approach

There is a possibility to improve the proposed technique if high order derivatives can be calculated with enough precision. As one can see from the analysis above, the main error for finite N comes from the finiteness of the differences $\delta_i(n) = x_i(n) - x_i(p_n)$. The idea is to use an interpolation between $x_i(n)$ and $x_i(p_n)$, to construct the nearly equal points.

Let us consider the value $X_i(n)$, located somewhere between $x_i(n)$ and $x_i(p_n)$. For simplicity, let it be in the middle:

$$X_i(n) = \frac{1}{2}(x_i(n) + x_i(p_n)). \tag{8}$$

Considering x as a function of time, one can replace $X_i(n)$ by its Taylor's series, calculated near the point $x_i(n)$ (Eq. (9)), and a similar Taylor's series but calculated near the point $x_i(p_n)$ (Eq. (10)):

$$X_i(n) = x_i(n) + \dot{x}_i(n)T_i(n) + \frac{\ddot{x}_i}{2}T_i^2(n) + \ldots, \tag{9}$$

$$X_i(n) = x_i(p_n) + \dot{x}_i(p_n)T_i'(n) + \frac{\ddot{x}_i}{2}T_i'^2(n) + \ldots. \tag{10}$$

Here $T_i(n)$ is a small time step, needed by oscillator with index i to move to the state $X_i(t)$ from the state $x_i(t)$, while continuing to follow the same trajectory on which the measured value $x_i(n)$ was located (the time $T_i(n)$ may be negative, i. e. one can move back along the trajectory). Similarly, $T_i'(n)$ is another small time step, needed to move to the point $X_i(t)$ from the point $x_i(p_n)$, following the trajectory on which $x_i(p_n)$ is located.

Since $T_i(n)$ and $T_i'(n)$ are rather small time intervals (in typical case they are less or much less than one sampling time unit), all the terms of the second and higher orders in Eqs. (9) and (10) can be neglected. Thus, $T_i(n)$ and $T_i'(n)$ can be easily calculated from formulae (8,9) and (8,10) respectively:

$$T_i(n) = \frac{x_i(p_n) - x_i(n)}{\dot{x}_i(n)}, \tag{11}$$

$$T_i'(n) = \frac{x_i(n) - x_i(p_n)}{\dot{x}_i(p_n)}. \tag{12}$$

Next, we substitute the interpolated value $X_i(n)$ in Eq. (2) :

$$X_i(n) = \sum_{j=1, j\neq i}^{D} k_{i,j}h(X_j(n)) - \dot{X}_i(n). \tag{13}$$

Let us substitute (9) and (10) in (13), keeping only the first-order terms:

$$X_i(n) \approx \sum_{j=1, j\neq i}^{D} k_{i,j}h\big(x_j(n) + \dot{x}_j(n)T_i(n)\big) - \big(\dot{x}_i(n) + \ddot{x}_i(n)T_i(n)\big),$$

$$X_i(n) \approx \sum_{j=1, j\neq i}^{D} k_{i,j}h\big(x_j(p_n) + \dot{x}_j(p_n)T_i'(n)\big) - \big(\dot{x}_i(p_n) + \ddot{x}_i(p_n)T_i'(n)\big).$$

In this way we get equations for $\delta_i(n)$ similar to Eqs. (4–6)

$$\delta_i(n) = \sum_{j=1, j\neq i}^{D} k_{i,j}\Delta h_{i,j}(n) - \Delta\dot{x}_i(n) - \Delta\ddot{x}_i(n) \approx 0, \tag{14}$$

$$\Delta h_{i,j}(n) = h\big(x_j(n) + \dot{x}_j(n)T_i(n)\big) - h\big(x_j(p_n) + \dot{x}_j(p_n)T_i'(n)\big), \tag{15}$$

$$\Delta\dot{x}_i(n) = \dot{x}_i(n) - \dot{x}_i(p_n), \tag{16}$$

$$\Delta\ddot{x}_i(n) = \ddot{x}_i(n)T_i(n) - \ddot{x}_i(p_n)T_i'(n). \tag{17}$$

The advantage of this model are much smaller values of $\delta_i(n)$. In fact, because we used the first-order approximation, we can estimate $\delta \sim N^{-2}$. Actually, one can generalize this interpolation by using derivative of higher order, but numerical calculation of high order derivatives leads to large errors and is hardly feasible. Moreover, we expect this interpolation method to be practical only for data with small amount of noise.

2.4. Delay coupled model

The model (1) can be slightly modified and generalized by introducing delays in coupling terms:

$$\dot{x}_i(t) = -x_i(t) + \sum_{j=1, j\neq i}^{D} k_{i,j}h(x_j(t - \tau_{i,j})), \tag{18}$$

where $\tau_{i,j}$ are delay times, individual for each coupling.

Introduction of coupling delays increases model complexity and forces it to generate chaotic behavior for smaller number of nodes, as it is illustrated in Fig. 2. The main frequency moves to significantly higher values for most nodes (see Fig. 2(b)), and the spectrum demonstrates a plateau at the frequency range $f \in [0.3; 0.7]$, which means that more degrees of freedom become involved into the dynamics.

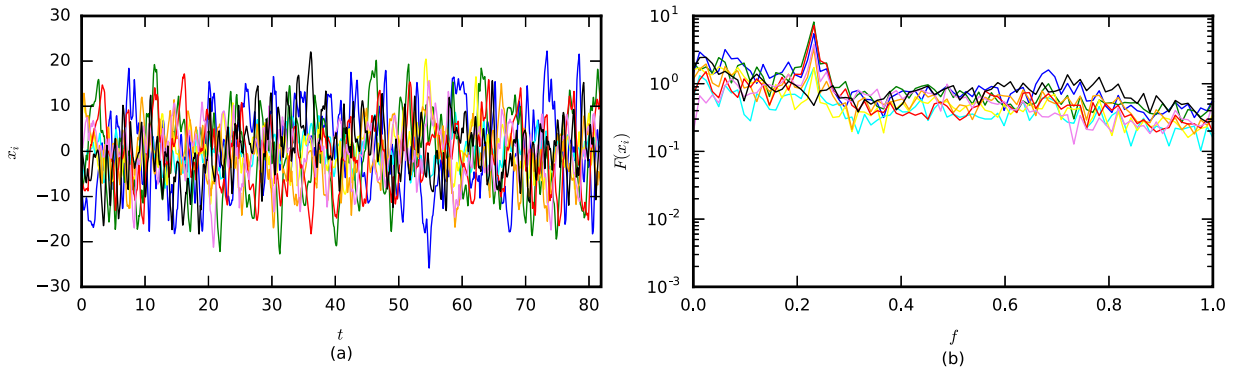

Fig. 2. Panel (a): time series of first 8 oscillations of ensemble (18) with $D = 16$ nodes; panel (b): spectrum of oscillations of first 8 oscillators of 16 in total. Delays were chosen to be random integer numbers, uniformly distributed in the range $\theta_{i,j} \in [50; 70]$ with $\Delta t = 0.05$.

2.5. Technique for reconstruction of the delay coupled model

Reconstruction of the modified model (18) can be performed mainly in the same way as for the original one, if delay times are known. Let us denote $\theta_{i,j} = \tau_{i,j}/\Delta t$ — integer time of delay in terms of the discrete time series. Then, Eq. (2) can be rewritten as (19):

$$x_i(n) = -\dot{x}_i(n) + \sum_{j=1, j\neq i}^{D} k_{i,j} h(x_j(n - \theta_{i,j})). \tag{19}$$

Eq. (3) does not change, but the position of the nearest previous sorted point in the original series $r_{i,j,n}$ (instead of p_n) becomes:

$$r_{i,j,n} = Q_i^{-1}(Q_i(n - \theta_{i,j}) - 1).$$

Eq. (6) is to be rewritten as

$$\Delta h_{i,j}(n) = h(x_i(n - \theta_{i,j})) - h(x_i(r_{i,j,n}))\ , \tag{20}$$

while (4) and (5) do not change, except for replacement $p_n \to r_{i,j,n}$. In formula (7), the indexing changes from $n = 1$ to $n = \theta_i + 1$, where $\theta_i = \max\limits_{j=1,\ldots,D, j\neq i}(\theta_{i,j})$:

$$S_i^2(\mathbf{k}_i) = \sum_{n=\theta_i+1}^{N} \delta_i(n) = \sum_{n=\theta_i+1}^{N} \left(\sum_{j=1, j\neq i}^{D} k_{i,j} \Delta h_{i,j}(n) - \Delta\dot{x}_i(n) \right)^2. \tag{21}$$

Thus, the same approach as above can be applied to the reconstruction of a delayed-coupled network.

2.6. Reconstruction of time delays

To reconstruct the delay times, if these are unknown, one can scan over all possible combinations of $\theta_{i,j}$ for each i separately, like it was proposed for time-delayed oscillators in [23]. One can perform reconstruction with the proposed method for all trial delay times in the range, and the true values of delay times should minimize the target function (21). Such a search is feasible, since delay times $\tilde{\theta}_{i,j}$ can be only nonnegative integers and values larger than the length of time series do not have sense. However, for large networks it is very CPU-time consuming, since it is necessary to scan in the $(D-1)$-dimensional space.

We propose to use an iterative algorithm, similar to the gradient descent method. First, some initial trial values for all $\theta_{i,j}$ are set for fixed i. Then, the reconstruction is performed for this set. Then trial guesses are changed to explore the closest neighbors: each component of the vector $\tilde{\theta}_i$ is shifted by ± 1 with respect to the original trial — this is an elementary perturbation (only one component is changed for each neighbor). For each of the $2(D-1)$ perturbed values of the time delay $\tilde{\theta}_i$, the reconstruction is performed, and the neighbor corresponding to the smallest value of S_i^2 is chosen. If this S_i^2 is lower than that for the original trial $\tilde{\theta}_{i,j}$, the new vector is adopted and all the steps are repeated until a set of delays $\tilde{\theta}_{i,j}$ is arrived which yields the value of the target function S_i^2 smaller than those of all the neighbors.

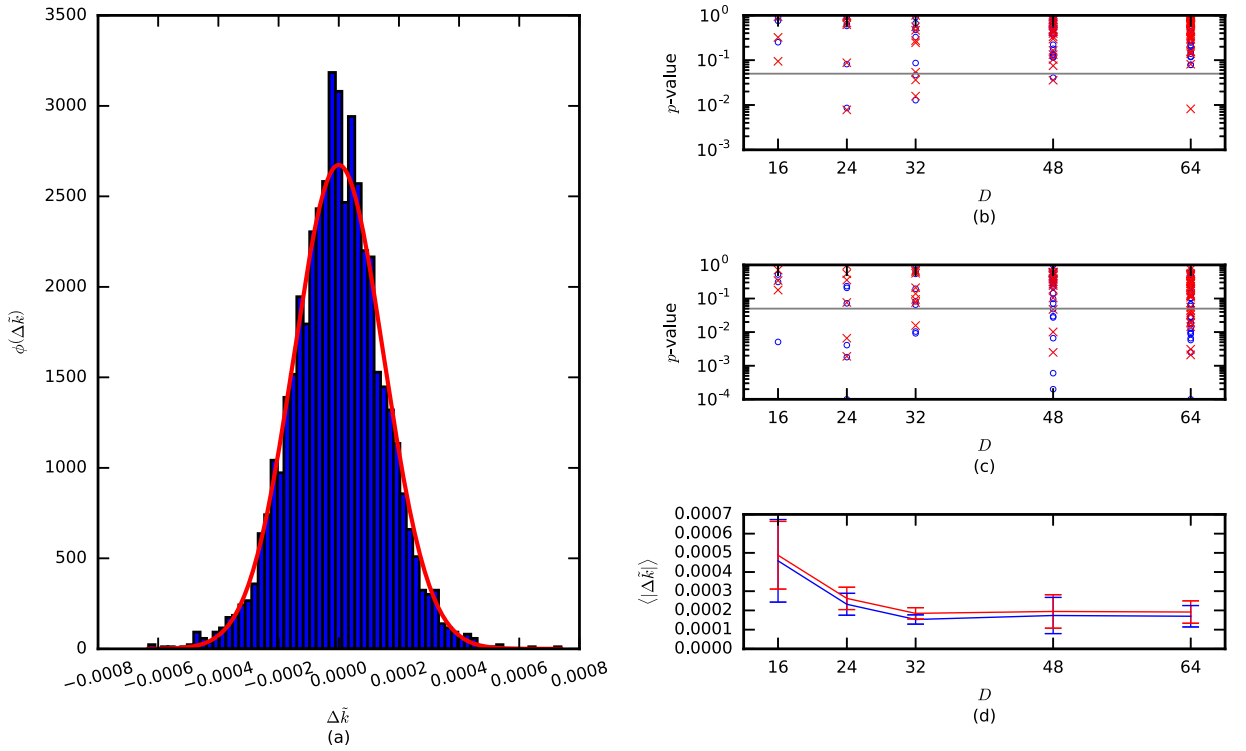

Fig. 3. Panel (a): the histogram of relative deviations $\Delta\tilde{k}_{i,j}$ of coefficient estimates $\tilde{k}_{i,j}$, normed to have a unity integral (the estimation of probability density function) for an ensemble of $D = 64$ nodes and its approximation with Gaussian (red line). Panel (b): p-value for Student's t-test for zero mean for $\Delta\tilde{k}$. Panel (c): p-value for Kolmogorov–Smirnov test for normality. Panel (d): the dependency of the mean $\Delta\tilde{k}$ on the dimension of ensemble D with errorbars, indicating standard deviation of $\Delta\tilde{k}$ from $\langle\Delta\tilde{k}\rangle$. On panels (b–d), the blue circles correspond to the original method (see Section 2.2), while red crosses — to the interpolation approach (see Section 2.3). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3. Results

3.1. Reconstruction of networks in the Sompolinsky et al. model

We generated 100 networks (1) with $D = 16$, $D = 24$, $D = 32$, $D = 48$ and $D = 64$ oscillator (500 samples in total). Equations of motion were solved numerically using adaptive Runge–Kutta algorithm of 4th order with sampling time of time series $\Delta t = 0.01$. The transient process of 2^{17} points was omitted. To determine the dynamic regime, the largest Lyapunov exponent was calculated using the classical Benettin's algorithm [24]. Only ensembles with chaotic behavior (for the considered setups there were 3, 6, 8, 20 and 38 such ensembles, respectively) were considered for reconstruction.

Then, the multivariate time series of $N = 2^{14}$ data points from all D oscillators were stored from each network, and the reconstruction of coupling constants $\hat{K}$ was performed with the method described in Section 2.2.

To characterize the accuracy of coupling matrix reconstruction, the histogram of relative deviations from the original values calculated as (22) is plotted in Fig. 3. Here the values $\tilde{k}_{i,j}$ indicate how much the reconstructed value deviates from the original one:

$$\Delta\tilde{k}_{i,j} = (\tilde{k}_{i,j} - k_{i,j})/(J\sqrt{D}), \tag{22}$$

where $k_{i,j}$ are the original coupling constants, and $\tilde{k}_{i,j}$ are the reconstructed ones. Since for different D the standard deviation of the original coupling constants was different and equal to $J/\sqrt{D}$, we included the corresponding normalization to formula (22).

Visually, the distribution of $\Delta\tilde{k}$ (Fig. 3a) seems to be very similar to a Gaussian distribution. The statistical tests were performed to characterize, whether this distribution has a zero mean (Student's t-test) — see Fig. 3(b), and whether it is Gaussian (classical Kolmogorov–Smirnov test with empirically estimated variance and zero mean) — see Fig. 3(c) for all considered ensembles for different number of nodes. One can see that the hypothesis of a zero mean cannot be rejected in most cases on level 0.05 — see gray line in Fig. 3(b). The hypothesis of zero mean has to be adopted for all cases for $D = 64$ and for all cases except one if $D = 48$. The hypothesis of Gaussian distribution is rejected in 40% of cases — see Fig. 3(c). This means that estimates of coefficients are unbiased (that is good), but the proposed method does not provide a maximal likelihood estimation, otherwise errors $\Delta\tilde{k}_{i,j}$ would follow the Gaussian law.

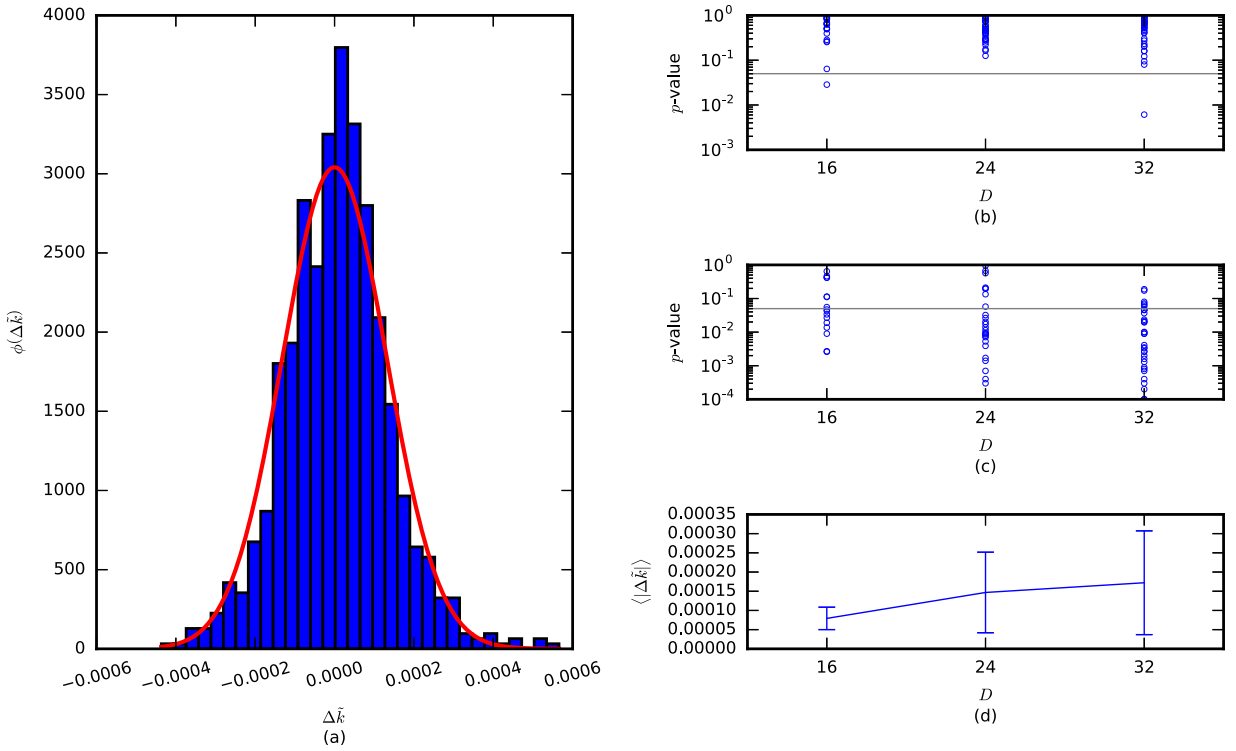

Fig. 4. Panel (a): the histogram of relative deviations $\Delta\tilde{k}_{i,j}$ of coefficient estimates $\tilde{k}_{i,j}$, normed to have a unity integral (the estimation of probability density function) for an ensemble of $D = 32$ nodes and its approximation with Gaussian (red line). Panel (b): p-value for Student's t-test for zero mean for $\Delta\tilde{k}$. Panel (c): p-value for Kolmogorov–Smirnov test for normality. Panel (d): the dependency of the mean $\Delta\tilde{k}$ on the dimension of ensemble D with errorbars, indicating standard deviation of $\Delta\tilde{k}$ from $\langle\Delta\tilde{k}\rangle$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

For generality, the same approach was applied for networks of $D = 32$ nodes, but with coefficients $k_{i,j}$ distributed uniformly with the same variance as considered previously and zero mean. Such a coupling was not considered in the original paper [20]. Seven of 100 randomly generated ensembles demonstrated well pronounced chaotic behavior. The errors of resulting estimates $\Delta\tilde{k}_{i,j}$ occurred to be distributed similar to the errors in the case when the original coefficients were normally distributed.

Generally, adopting the interpolation procedure did not give an advantage, possibly due to errors in the calculation of the derivative — compare blue and red points in Fig. 3(b–d). In all the cases the results of the straightforward approach (Section 2.2) and of its modification (2.3) are very similar. However, for small networks (e. g. $D = 16$) the results of the modified algorithm appear to be slightly better than of the original one. The advantage of the modification could be possibly found if the second or higher derivative could be directly measured or calculated with higher precision.

3.2. Reconstruction of delayed coupled networks

We generated 40 networks of $D = 16$, $D = 24$, and $D = 32$ oscillators (120 in total). All delays were chosen to be uniformly distributed in the range [2.5; 3.5]. From these networks 19, 34, and 38 respectively demonstrated chaotic behavior. Equations were solved numerically using the Euler algorithm with sampling time $\Delta t = 0.01$ and step of integration equal to $\Delta t/100$. Such an approach was chosen due to delay in the equations, and was rather CPU-time-consuming. Therefore, larger networks were not studied. The transient process of 2^{17} points was omitted. Only ensembles with chaotic behavior were considered for reconstruction. Multivariate time series of $N = 2^{14}$ data points from all D oscillators were recorded from an each ensemble, and the reconstruction of coupling constants $\hat{K}$ was performed using the method described in Section 2.5.

Results of the coupling matrix reconstruction in the case of known delay times are shown in Fig. 4 in the same manner as in Fig. 3. For small values of D, the mean error of the coupling constant reconstruction is much less for delayed ensembles (18) than for ones without delay (1). This could be explained by the fact that only very week chaos with small largest Lyapunov exponent can be found in the system (1) for small D, while with delay, chaos is stronger. In stronger chaotic regime the same errors in coupling coefficients lead to larger errors in the target function. Therefore, the method becomes more sensitive to errors and estimates of coupling coefficients can be achieved with better precision.

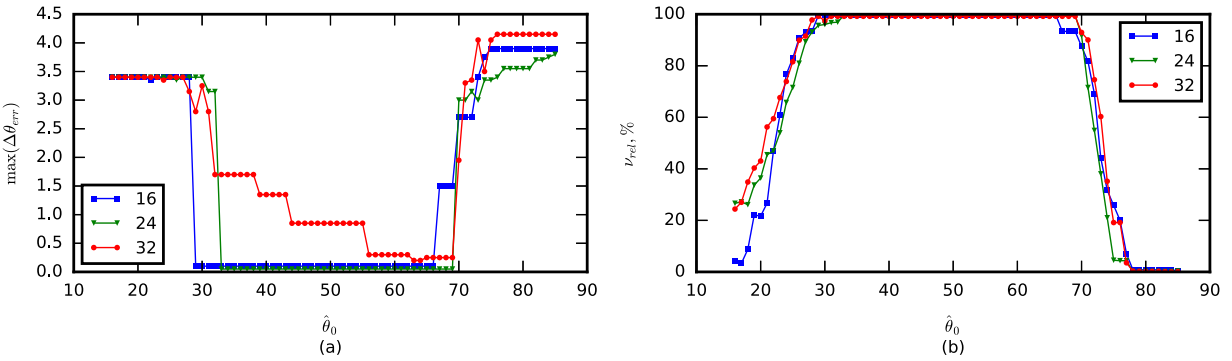

Fig. 5. Panel (a): the maximal absolute error in delay over all considered delays in the ensemble — $\max(\Delta\theta_{err})$. Panel (b): the percentage ν_{rel} of correctly detected delays.

3.3. Reconstruction of delay times

To reconstruct delayed time networks as described above, one has to know all the delay times $\tau_{i,j}$, or their discrete analogs $\theta_{i,j}$. One can, however, hardly assume that all of them can be directly measured or calculated from first principles. Therefore, in Section 2.6 we proposed an iterative technique for reconstructing delay times, to be tested below. To illustrate its efficiency, the reconstruction was performed for different initial trial values of the time delays, to check convergence to the true values. In the considered case of $D \geq 16$ oscillators in the ensemble (1), there are 240 or more delays in the links, so it is unlikely to test all possible variations of starting guesses for them, even assuming that starting guesses can attain only discrete values with the step equal to the sampling interval, rather than continuous values.

We restricted ourselves to a most realistic starting guesses that all the delays are equal; let us denote the value of such an initial guess as $\hat{\theta}_0$. To test quality of reconstruction, two measures were evaluated for different $\hat{\theta}_0$:

1. The number ν_{abs} (and percentage ν_{rel}) of correctly detected delays.
2. The maximal absolute error in delay times, calculated from all considered delays in the ensemble: $\max(\Delta\theta_{err})$.

The results for these two quantities are plotted in Fig. 5. The original values of delays $\tau_{i,j}$ were generated uniformly distributed in the range [2.5; 3.5] (i. e. $\theta_{i,j} \in [50; 70]$). One can see that there is a large domain of $\hat{\theta}_0$ values around the center of this range ($\hat{\theta}_0 = 60$), from which the coupling delays are reconstructed very precisely: 2 errors of value ± 1 for $D = 16$ of 240 couplings in total, one error equal to 2 for $d = 24$ of 554 couplings in total, and six errors not exceeding ± 6 for $D = 32$ of 992 couplings in total. In contradistinction, reconstruction started from very small or very large initial guesses typically failed. Reconstruction of larger networks tends to be more erroneous at the same time series length. For a network with $D = 32$ nodes, the maximal error $\max(\Delta\theta_{err})$ is achieved for the estimation of delay of one particular coupling, for all starting guesses in a wide range of from $\hat{\theta}_0 = 32$ to $\hat{\theta}_0 = 69$, while delays for other couplings are reconstructed equally fine in the whole range. This particular coupling corresponds to a relatively small coupling constant, with absolute value 2.5 times less than the mean value.

The results of reconstruction of coupling delays severely depend on the dynamical regime. For plots in Fig. 5, only chaotic regimes with the largest Lyapunov exponent $\Lambda > 0.12$ were chosen. For smaller Λ, the number of errors could be significantly larger, e. g. for an ensemble of $D = 24$ oscillators with $\Lambda \approx 0.04$, all coupling coefficients can be reconstructed correctly with all delays only for 6 oscillators from 24, and total percentage of errors $\nu_{rel} > 60\%$ even for the best initial guesses $\hat{\tau}_0 = 3$.

4. Conclusion and discussion

4.1. Comparison with Granger causality and other approaches

To compare the proposed technique with already existing ones, let us consider Granger causality first. Usual pairwise approach [8] is not applicable for the problem to be solved since each node has too many links with others, so it will cause very large errors. However, one can consider the so called “conditional Granger causality” [25]. The general multivariate forecasting models to be constructed in order to estimate coupling coefficients can be written in the form (23).

$$x_i(n+1) = \sum_{j=1}^{j=D} c_{i,j} x_j(n) + \xi_i(n), \tag{23}$$

where $\xi_i(n)$ are model residuals and $c_{i,j}$ are unknown coefficients of coupling, but different from the original coefficients $k_{i,j}$. If we take into account that we exactly know the model structure (1) and the coupling functions h, the model (23) can

be rewritten as (24) by approximating the derivative with a simple difference. Let us denote such an approach as "specified Granger technique".

$$x_i(n+1) \approx c_{i,i}x_i(n) + \sum_{j=1, j\neq i}^{D} c_{i,j}h(x_j(n)), \tag{24}$$
$$c_{i,i} = 1 - \Delta t,$$
$$c_{i,j} = \Delta t \tilde{k}_{i,j}, \quad j \neq i.$$

The model (24) can be fitted to data, so the estimates of coupling coefficients $\tilde{k}_{i,j}$ can be easily found.

We tested specified Granger technique for all ensembles of 32 nodes, exactly the same as for the proposed method in the Section 2.2, and estimates of coupling coefficients $\tilde{k}_{i,j}$ were found to be better for the proposed approach. The reason for this is the fact that the main source of errors for the proposed approach is a difference between nearby values of $x_i(n)$ in the sorted series, which is going to zero for $N \to \infty$, while the main source of errors for the specified Granger technique is the finite sampling rate, which is usually constant. But for short series the Granger causality can give even better estimates.

The other techniques which in principle are able to give some coupling estimation for the considered system are partial directed coherence [16] and transfer entropy [12]. But they cannot give the direct estimates of coupling coefficients since there is not possible to take into account the information about the exact structure of the system considered. A new approach based on residual analysis was proposed in [28] but it also aims to estimate only the fact of coupling presence. Additionally, the transfer entropy should be applied pairwise, or too many data must be used for estimating the probability densities, even if the advanced algorithms based on nearest neighbors (like one proposed in [13]). There are many approaches specialized for some particular system type (many of them one can find in the review [26]), which also cannot be applied directly. For example, for stochastic oscillators see [27], for hidden variables see [4], for low dimensional nonautonomous ODEs see [7].

4.2. General conclusion

In this paper we propose a technique for reconstruction of neural networks described in [20]. The network should operate in a chaotic regime, and we demonstrate that all the coupling constants in the ensemble can be estimated with a high precision if time series of all elements are provided and nonlinear function of coupling is known. The numerical consideration was performed for different distributions of coupling coefficients. In addition, we consider the generalized form of the original equations, for which delay times (generally different for all the links) are present in all the couplings. The proposed approach is applicable for such a generalized system with minor changes, if all delay times are given.

To find delay times if these are unknown, we proposed to use an iterative search approach similar to gradient decent, but adopted to the discrete case, since delay time may be set only as integer multipliers of sampling intervals. Initial guess values for all delays are important parameters, here we assigned them to be equal. This technique showed a rather extended domain of convergence in the space of starting guesses, for networks of different sizes. This domain is possibly dependent on the main time scales and regimes of the network. Remarkably, for all three different choices for number of nodes: $D = 16$, $D = 24$ and $D = 32$, the domains of convergence are very similar; they differ in some minor details only (actually, they are slightly different for different $\hat{K}$ and $\tau_{i,j}$ even for the same D). This could indicate that this domain originates from the fundamental properties of the system (18) such as structure of equations and the way how the coupling is introduced, rather than from individual parameters of the network.

The proposed technique can be easily extended to the case where a smooth unknown nonlinear function of the variable x_i is present in the right side of equations (1) or (18) instead of simple term $-x_i$, using the same approach as in [23], even if this function is different for different oscillators.

The technique proposed here differs from the approach described in [29], while having the same basic idea. In Ref. [29], a single coupling function from all inputs is considered (firing-rate neural network), while here we consider separate coupling functions for each input, so direct implementation of the approach from [29] is not possible. Also, here a delay time in coupling is considered, and this delay is reconstructed for each coupling link, while there is no delay in coupling in the system considered in Ref. [29]. Another difference is that here the standard least squares minimization routine is used, with estimates of coupling coefficients being asymptotically unbiased, while in Ref. [29] homogeneous linear equations are solved, using singular value decomposition approach.

Acknowledgments

This research was funded by Russian Foundation for Basic Research, grant No 16-02-00091 and Stipendium of President of Russian Federation for support of young scientists СП-1510.2015.4. Study of the interpolation approach was supported by Russian Science Foundation, grant 17-12-01534.

References

[1] Schelter B, Winterhalder M, Timmer J. Handbook of time series analysis: recent theoretical developments and applications. Wiley-VCH; 2006.
[2] Box G, Jenkins GM, Reinsel GC, Ljung GM. Time series analysis: forecasting and control. — fifth edition. Willey; 2015.

[3] Pollock D. A handbook of time series analysis, signal processing and dynamics. Academic Press; 1999.
[4] Baake E, Baake M, Bock HG, Briggs KM. Fitting ordinary differential equations to chaotic data. Phys Rev A 1992;45(8):5524–9.
[5] Quach M, Brunel N, d'Alché Buc F. Estimating parameters and hidden variables in non-linear state-space models based on ODEs for biological networks inference. Bioinformatics 2007;23:3209–16.
[6] Kim J-W, Pachepsky YA. Reconstructing missing daily precipitation data using regression trees and artificial neural networks for SWAT streamflow simulation. J Hydrol (Amst) 2010;394(3–4):305–14.
[7] Bezruchko BP, Smirnov DA. Constructing nonautonomous differential equations from experimental time series. Phys Rev E 2000;63:016207.
[8] Granger C. Investigating causal relations by econometric models and cross-spectral methods. Econometrica 1969;37(3):424–38.
[9] Chen Y, Rangarajan G, Feng J, Ding M. Analyzing multiple nonlinear time series with extended granger causality. Phys Lett A 2004;324(1):26–35.
[10] Marinazzo D, Pellicoro M, Stramaglia S. Nonlinear parametric model for granger causality of time series. Phys Rev E 2006;73:066216.
[11] Bezruchko BP, Ponomarenko VI, Prokhorov MD, Smirnov DA, Tass PA. Modeling nonlinear oscillatory systems and diagnostics of coupling between them using chaotic time series analysis: applications in neurophysiology. Phys Usp 2008;51:304–10.
[12] Schreiber T. Measuring information transfer. Phys Rev Lett 2000;85:461.
[13] Kraskov A, Stögbauer H, Grassberger P. Estimating mutual information. Phys Rev E 2004;69:066138.
[14] Pijn J, Vijn P, Silva FHLd, van Em W, Blanes W. The use of signal-analysis for the location of an epileptogenic focus: a new approach. Adv Epilptol 1989;17:272–6.
[15] Aguirre LA. A nonlinear correlation function for selecting the delay time in dynamical reconstructions. Phys Lett A 1995;203(2–3):88–94.
[16] Baccala LA, Sameshima K. Partial directed coherence: a new concept in neural structure determination. Biol Cybern 2001;84:463–74.
[17] Schelter B, Mader M, Mader W, Sommerlade L, Platt B, Lai Y-C, et al. Overarching framework for data-based modelling. EPL 2014;105:30004.
[18] Wang Q, Shen Y, Zhang JQ. A nonlinear correlation measure for multivariable data set. Physica D 2005;200(3–4):287–95.
[19] Sysoev IV, Prokhorov MD, Ponomarenko VI, Bezruchko BP. Reconstruction of ensembles of coupled time-delay system from time series. Phys Rev E 2014;89:062911.
[20] Sompolinsky H, Crisanti A, Sommers HE. Chaos in random neural networks. Phys Rev Lett 1988;61(3):259–62.
[21] Liao W, Mantini D, Zhang Z, Pan Z, Ding J, Gong Q, et al. Evaluating the effective connectivity of resting state networks using conditional granger causality. Biol Cybern 2010;102(1):57–69.
[22] McGill W. Multivariate information transmission. Psychometrika 1954;19:97–116.
[23] Sysoev IV, Ponomarenko VI, Kulminskiy DD, Prokhorov MD. Recovery of couplings and parameters of elements in networks of time-delay systems from time series. Phys Rev E 2016;94:052207.
[24] Benettin G, Galgani L, Strelcyn J-M. Kolmogorov entropy and numerical experiments. Phys Rev A 1976;14:2338.
[25] Ding M, Chen Y, Bressler SL. Granger causality: Basic theory and application to neuroscience, Handbook of time series analysis: recent theoretical developments and applications ch. 17. 2006. 437–460.
[26] Bezruchko B, Smirnov D. Extracting knowledge from time series. Berlin: Springer; 2010.
[27] Smirnov DA, Bezruchko BP. Detection of coupling in ensembles of stochastic oscillators. Phys Rev E 2009;79:046204.
[28] Marques VM, Munaro CJ, Shah S. Detection of causal relationships based on residual analysis. IEEE Trans Autom Sci Eng 2015;12(4):1525–34.
[29] Pikovsky A. Reconstruction of a neural network from a time series of firing rates. Phys Rev E 2016;93:062313.