Chaos, Solitons and Fractals 82 (2016) 11–21

Contents lists available at ScienceDirect

Chaos, Solitons and Fractals

Nonlinear Science, and Nonequilibrium and Complex Phenomena

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

Frontiers

Choosing the optimal model parameters for Granger causality in application to time series with main timescale

Maksim V. Kornilov [a,b,*], Tatiana M. Medvedeva [a], Boris P. Bezruchko [a,b], Ilya V. Sysoev [a,b]

[a] Saratov State University, 83, Astrakhanskaya street, 410012 Saratov, Russia
[b] Saratov Branch of Kotel'nikov Institute of Radioengineering and Electronics of Russian Academy of Sciences, 38, Zelyonaya street, 410019 Saratov, Russia

ARTICLE INFO

Article history:
Received 2 September 2015
Accepted 20 October 2015

Keywords:
Coupling identification
Granger causality
Time series

ABSTRACT

The problem of determining the presence and direction of coupling between experimentally observed time series is of immediate interest in many relevant areas of knowledge. One of the approaches to its solution is the method of nonlinear Granger causality. The algorithm is based on the construction of predictive models and its effectiveness depends on the proper selection of model parameters.

The most important of them for signals with a characteristic time scale fluctuations are the time lag used in the reconstruction of the state vector, and the range forecast. In this paper, we propose two criteria for evaluating performance of the method of nonlinear Granger causality, which allows one to select the lag and range forecast and achieves the best sensitivity and specificity. The sensitivity is determined by range of weakness the method can detect and specificity means the ability to avoid false positive results. Because of the proposed criteria on the example of several unidirectionally coupled reference systems were received practical advice on the selection of the following model parameters: lag and range forecast.

1. Introduction

Detection of the presence and direction of interactions between subsystems of complex systems from their experimentally observed time series is an up-to-date problem, which has applications in various fields of knowledge. A variety of methods were developed in order to solve this task, including the cross-correlation function, coherence, phase synchronisation index, information based measures [1,2], a partial directed coherence [3] and approaches based on building predictive models, which include the Granger causality approach [4], and phase dynamics approach [5,6]. The main idea of approaches Granger causality method, transfer entropy and partial directed coherence is similar, and in some cases they can be shown to be completely equal [7]. However in general this is not the case [8].

The author developed Granger causality in relation to economical studies, but now it is successfully applied, for example, in neuroscience to identify the coupling between different brain regions [9–14], in climatology—to predict the behaviour of the monsoon [15]. There are a number of studies, where the Granger causality is used to indicate the evolution of coupling in time [16–18].

Despite the transparency of the idea of the method, its efficiency depends critically on the details of the implementation. For example, the special attention has to be paid to effects of measurement noise [19–21].

* Corresponding author at: Saratov State University, 83, Astrakhanskaya street, 410012 Saratov, Russia. Tel.: +7 9173085778.
E-mail addresses: kornilovmv@gmail.com (M.V. Kornilov), golovatanya@rambler.ru (T.M. Medvedeva), bezruchkobp@gmail.com (B.P. Bezruchko), ivssci@gmail.com (I.V. Sysoev).

http://dx.doi.org/10.1016/j.chaos.2015.10.027
0960-0779/

In this paper, we propose a new approach to choose the method parameters for the case, when the measured time series have a single main time scale which can be detected, for example, by using auto-correlation function. Such systems are quite common in nature, e.g. rhythmic changes in solar activity (sunspot number), oscillatory changes of brain potentials (for example, spike-wave discharges at the absence epilepsy, or θ-rhythm during the sleep), the signals of the heart (the main heart rate at the cardiogram) etc. The idea of the approach is that the parameters of the model, with the meaning of time: the time lag l, used for the reconstruction of the state vector [22–24], and the prediction time τ should be associated with the characteristic scale of the observed oscillations. Some research in this direction has already been held: in [25] it was shown that the lack of sample rate and, therefore, too large values of τ and l lead to systematic, fatal errors in determining the direction of coupling, primarily to the appearance of false positive conclusions. Similarly, in the paper [26] it was shown that mistakes can be caused by too small values of the prediction time, therefore it was proposed to use the value of τ equal to a quarter of the characteristic period of the observed oscillations.

In general, however, the question of the optimal choice of τ and l is still open, even for individual narrow classes of signals, including the class considered in this paper, since the previously used numerical criteria took into account only one parameter of the Granger causality method: either prediction time, as in [26], or sampling interval, as in [25], or a kind of approximation functions, as in [27]. Therefore, in this paper we propose two new numerical criteria characterising the performance of the method according to τ and l and test them in a number of nonlinear coupled reference systems. The application of these criteria helped to make general conclusions about the best and worst values of τ and l for the models used in the method of Granger causality to determine the coupling of signals having a distinct time scale (oscillation period). The degree of efficiency of the method is shown according to the level of nonlinearity of the original data, which is determined by the highest Lyapunov exponent and the effective coefficient of phase diffusion.

In the original study [4] linear approximating functions were used, but now nonlinear functions of different type: polynomials [27,28], radial basis functions [29] and kernel Granger causality [30] can be used instead. The other way to solve the problem of approximation function choice is to use local linear models, as it was proposed in [31]. We examined the model with polynomial nonlinearity of the general form and the local linear models, since they are most frequently used in practice due to simplicity and generality. In this regard, we chose the small values of both the model dimension and polynomial order. So, the models do not require many coefficients and can be reliably estimated from the short time series. This helps to make the results useful from a practical point of view, when the amount of data is very limited due to the features of the experiment, nonstationarity, or a desire for acceptable temporal resolution, when constructing models in a time window, as is done, for example, in [14,17,32].

2. Granger causality

Let us remind the key point of Granger causality. Supposing that we have time series of two systems—a series $\{x_n\}_{n=1}^N$ from the system X and a series $\{y_n\}_{n=1}^N$ from the system Y, where $n = 1, 2, \ldots, N$ is discrete time, N is the length of the series. It requires to determine whether the system Y drives the system X or not by analysing realisations $\{x_n\}_{n=1}^N$ and $\{y_n\}_{n=1}^N$. To solve this problem an individual model (dynamical system) is constructed on the first step:

$$x'_{n+\tau} = f(x_n, x_{n-l}, \ldots, x_{n-(D_s-1)l}, \mathbf{c}^s), \tag{1}$$

where x'_n is a predicted value at the time moment n and it may differ from the measured value x_n, f is an approximating function (if it is nonlinear, method is called a nonlinear Granger causality), l—lag of the model, i.e. the number of discrete time points between the two subsequent values from $\{x_n\}_{n=1}^N$, forming D_s-dimensional state vector of the model $\mathbf{x}_n(x_n, x_{n-1}, \ldots, x_{n-(D_s-1)l}$, τ is the prediction time—the distance in time between the predicted point and the closest point of the state vector, D_s—dimension of the individual model (the number of points of the time series which form the state vector, which is being reconstructed by the time delay method [23,24]), $\mathbf{c}^s$—unknown vector of coefficients which is chosen using least squares fit to minimise the standard error of approximation (1):

$$\varepsilon_s^2 = \frac{1}{N} \sum_{n=\tau+(D_s-1)l+1}^{N} (x'_n - x_n)^2 \tag{2}$$

The next step is to build the joint model, in which D_a members from the series $\{y_n\}_{n=1}^N$ are used besides the data of the series $\{x_n\}_{n=1}^N$:

$$x''_{n+\tau} = g(x_n, x_{n-l}, \ldots, x_{n-(D_s-1)l}, y_n, y_{n-l}, \ldots, y_{n-(D_a-1)l}, \mathbf{c}^j), \tag{3}$$

where x''_n is a model predicted value, $\mathbf{c}^j$—joint model coefficients. The standard prediction error of the joint model similarly to (2) has the form:

$$\varepsilon_j^2 = \frac{1}{N} \sum_{n=\tau+(\max(D_s,D_a)-1)l+1}^{N} (x''_n - x_n)^2. \tag{4}$$

If $\varepsilon_j^2 < \varepsilon_S^2$, the system Y is considered to drive the system X (systems are coupled). *Prediction improvement* index is typically used as a measure of coupling:

$$PI = 1 - \frac{\varepsilon_j^2}{\varepsilon_S^2}. \tag{5}$$

If $PI = 0$ (considering the signal $\{y_n\}_{n=1}^N$ did not help in predicting $\{x_n\}_{n=1}^N$), it is considered that Y has no effect on X. If the $PI \to 1$ (considering the signal $\{y_n\}_{n=1}^N$ has significantly improved the prediction of $\{x_n\}_{n=1}^N$), it should be regarded as Y drives X.

Practice shows that the choice of the parameters of the described procedure (lag l, prediction time τ, dimensions D_s and D_a, type of nonlinear functions f and g) significantly determines the efficiency of the method. For example, the use of too small or too large τ may cause a large number of errors: positive conclusions about the coupling that in fact does not exist [25,26]. Neglecting the nonlinearity in the modelling often leads to a situation, when really existing links are not detected [27,33]. The problem seems to be major since most coupling analysis techniques are very model-dependent [8].

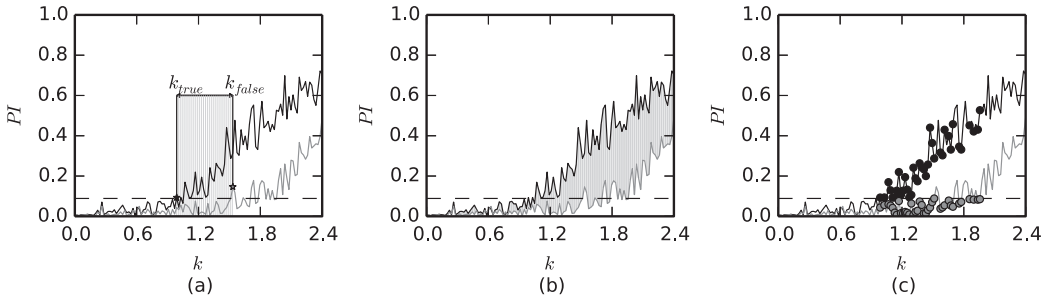

Fig. 1. The dependency of the prediction improvement *PI* on the coupling coefficient *k* for unidirectionally coupled oscillators. Fig. 1(a) is the typical plot of *PI* on *k*, where k_{true} is the minimum value at which coupling is found significantly in the right direction, and k_{false} is the minimum value of *k* at which the coupling is found significantly in the wrong direction (grey line). Fig. 1(b) illustrates the criterion (6). Fig. 1(c) illustrates the criterion (7), dots indicate the values at which $\delta = 1$.

Efficiency of the method can be achieved by developing specialised technologies of its implementation for the selected fairly narrow class of systems. This can be done due to additional considerations, such as a priori information about the properties of the system, or based on the results of analytic and numerical analysis of the etalon oscillators. Here, the systems having main characteristic time scales are considered.

3. Method efficiency investigation technique

The efficiency of methods for detecting the coupling of the two systems is commonly characterised by their sensitivity and specificity. The sensitivity is determined by the weakest coupling that method can detect. Various factors: finite length of observed time series, measurement noises, the use of improper basis functions, insufficient accuracy of calculations, etc. do not allow detecting the coupling at the arbitrarily low level. Specificity means the ability of the method to avoid false positive results. Good specificity corresponds to a relatively small number of false couplings (found where in fact they do not exist), and bad—to the large number of them. Typically, the method of Granger causality (Fig. 1) gives a non-zero *PI* both ways, even for unidirectionally coupled systems [34]. Such errors are often caused by the finite volume of experimental data or the unfortunate choice of the method parameters. Therefore, when calculating the *PI* one should test the *significance* of these values—the probability that non-zero values of the prediction improvement index were received not accidentally. The significance can be measured in different ways, for example, based on consideration of the theoretical properties of the residuals of the model [25] or through the generation of various surrogate time series [26,32,34].

In [26] an approach for evaluating the effectiveness of Granger causality in the coupled equations systems was proposed. This approach consists in determining the critical values of the coupling coefficient $\bar{k}$, for which the value of *PI* becomes significant. Let $\bar{k} = k_{true}$ be the minimum value of *k* at which method significantly detects the coupling in the right direction, and $\bar{k} = k_{false}$—minimum value of *k*, for which the coupling is significantly detected in a false direction. Since the dependency $PI(k)$ is usually not smooth (see Fig. 1(a)) due to the presence of noise, different initial conditions and finite length of considered time series (non-uniform bypass of an attractor), it is often necessary averaging over realisations.

However, for our goal—studying dependency of the results on two parameters: l and τ—obtained curves $PI(k)$ require further processing, because in fact one has to consider the 3-dimensional dependency $PI(k, l, \tau)$. This dependency is impossible to submit in a chart using an intuitive manner. Therefore it is necessary to develop additional, integral criteria. In the practical application the value of the coupling strength is unknown, so the widest possible range of values *k* should be considered for research purposes. The lowest boundary of this range corresponds to the lack of coupling—$k = 0$. The upper one—to the synchronisation of the driving and the driven systems, since in this case to detect the coupling direction is no longer possible due to the fact that the two measured time series become equal or functionally dependent. The phase synchronisation index was estimated to evaluate the upper boundary value of the coupling coefficient $k = k_{\max}$ by the method proposed in [35], since the phase synchronisation is typical for the systems with the main time scale and usually precedes other types of synchronisation [36].

Practice shows that, when choosing the model parameters, one have to comply with the optimal balance between high sensitivity and high specificity, since an improvement in one of these model properties usually leads to a deterioration of another. To select the optimal value of l and τ two numerical criteria are proposed here. In some way they are similar to well known Schwartz criterion [37], since they are based on the number of successful coupling findings in the actual direction, but introduce the penalty for findings in the opposite (wrong) direction.

1. The first criterion is given by the formula (6):

$$S_1 = \langle PI_{true}(k) - PI_{false}(k) \rangle_k. \tag{6}$$

It takes into account the average difference between the values of $PI_{true}(k)$ and $PI_{false}(k)$ for all possible values of *k*. Visually, it matches the average width of the shaded area in Fig. 1(b). Since, according to (5), *PI* is a standardised value which ranges from 0 to 1, the value of S_1 can vary from 1 to −1. If for all values of *k* $PI_{true} = 1$ (coupling in the actual direction was detected ideally for any strength of coupling) and $PI_{false} = 0$, (coupling in wrong direction is always absent), than according to (6) $S_1 = 1$. If it turns

out that $\langle PI_{true}\rangle_k = \langle PI_{false}\rangle_k$—the average prediction improvement in the actual and the false directions are the same, i.e. the method is not able to determine the direction of the coupling—$S_1 = 0$. If $PI_{false} = 1$ and $PI_{true} = 0$, coupling is always found in the false direction and never in the actual one, $S_1 = -1$.

2. The second criterion (see Fig. 1(c)) takes into account the significance of the *PI*, but it does not consider its absolute value. For this the binary function $\delta(k)$ is introduced. If the coupling in the actual side is defined as a significant, and in the false one—as insignificant $\delta(k) = 1$, otherwise $\delta(k) = 0$. The lowest possible value of S_2 is 0, that corresponds to the case, when for all k at least either the coupling in the actual direction detected as insignificant, or coupling in the coupling in the wrong direction is detected as significant. The highest $S_2 = 1$, that means coupling to be always significantly detected in the actual direction and there is no significant results in the wrong one.

$$S_2 = \frac{1}{Z_k}\sum \delta(k),$$
$$\delta(k) = \begin{cases} 1, PI_{true} > PI_{sign} \wedge PI_{false} < PI_{sign}, \\ 0, PI_{true} \le PI_{sign} \vee PI_{false} \ge PI_{sign}, \end{cases} \quad (7)$$

where Z_k is the number of considered values of the coupling coefficient; PI_{sign} is the 95% significance level obtained by using surrogate time series, derived from the same systems as studied, but without coupling (similar to [34]).

Criteria (6) and (7) complement each other. If the criterion S_1 reaches a value close to 1, it is arguable that the method works well with the given parameters, and if it is close to 0 or negative—works poorly. In this case test of significance may be waived. If the criterion S_1 takes some intermediate values between 0 and 1, testing the significance is important, and efficiency is determined by the criterion S_2.

4. Numerical experiments on different systems

4.1. Coupled Kiyashko–Pikovsky–Rabinovich generator equations

Let us consider the equations of Kiyashko–Pikovsky–Rabinovich radio-engineering generator [39]:

$$\begin{cases} \dot{x}_1 = 0.15x_1 + x_2 - 0.93x_3 + ky_1 \\ \dot{x}_2 = -x_1 \\ 0.2\dot{x}_3 = x_1 - f(x_3), \end{cases}$$
$$\begin{cases} \dot{y}_1 = 0.23y_1 + y_2 - 0.755y_3 \\ s\dot{y}_2 = -y_1 \\ 0.2\dot{y}_3 = y_1 - f(y_3), \end{cases}$$
$$f(p) = 8.592p - 22p^2 + 14.408p^3 \quad (8)$$

These equations are very interesting since they demonstrate the chaotic regimes which are very regular, so the only main time scale in the spectrum is present (see Fig. 2 (a)), the autocorrelation function decreases very slow (see Fig. 2 (b)) and phase can be easily introduced.

Equations of coupled systems were solved numerically with Runge–Kutta 4th order method with sampling rate $\Delta t = 0.01$. Series of $N = 20480$ points were considered, the transient process of $T_{trans} = 1000$ unit of dimensionless time was cut[1]. The main timescale T is equal approximately to 611 time points (6.11 unit of dimensionless time). Coupling coefficient ranged from 0 to 0.038. The positive Lyapunov exponent of both driving and driven systems was equal $\Lambda_1 \approx 0.059$, while their parameters were slightly different. The effective coefficient of phase diffusion of the driven system $D < 10^{-5}$.

From Fig. 2(d)–(g) one can see that the values of l approximately equal to one or one half of the driven system main time scale are absolutely pessimal for both types if models. This is not a big surprise, because in such a case the components of the reconstructed state vector are close to be linearly dependent, since the signal is very regular. Also one can see that the value of τ equal to one time point is not a good choice, at least for the polynomial model, as it was shown in [26].

Actually the local-linear model performs badly as it can be seen from Fig. 2(g). Even the best choice of l and τ correspond only to 20% of successful coupling detections (when the coupling in the actual direction is found and the coupling in the wrong one is considered as insignificant). For the polynomial model the choice of l and τ occurs to be very important, since the number of successful coupling recognitions may vary from 18% in the worst considered case to 72% in the best one (4 times).

The dependencies plotted on Fig. 2(d)–(g) were obtained for $D_s = 2$, but for $D_s = 3$ they are similar. We also used combination of $D_s = 3$ and polynomial order $\nu = 2$ for the polynomial model with the same effect.

4.2. Coupled Rössler oscillators

The Rössler oscillator [38] is one of the fundamental models of nonlinear dynamics. Two Rössler oscillators coupled as shown in (9) were considered:

$$\begin{cases} \dot{x}_1 = -x_2 - x_3 \\ \dot{x}_2 = x_1 + a_1x_2 + ky_2 \\ \dot{x}_3 = b_1 - (c_1 - x_1)x_3 \end{cases}$$
$$\begin{cases} \dot{y}_1 = -y_2 - y_3 \\ \dot{y}_2 = y_1 + a_2y_2 \\ \dot{y}_3 = b_2 - (c_2 - y_1)y_3 \end{cases} \quad (9)$$

The time series of coordinates x_2 and y_2 were observed. The first Lyapunov exponent for both driving and driven subsystems was approximately equal to 0.060. The effective coefficient of diffusion equal to ≈ 0.003. There is a single main peak in the spectrum (see Fig. 9(a)) corresponding to one main timescale, that is 620 time points (the same value can be obtained from auto-correlation function, see Fig. 9(b)). However the correlation function decreases faster and the spectral peak is wider than for Kiyashko–Pikovsky–Rabinovich generator.

For the polynomial model the (Fig. 9(d) and (e)) pronounced minima present in both dependencies of S_1 and S_2 on l for all considered values of τ. These results match the results for equation of coupled Kiyashko–Pikovsky–Rabinovich

[1] The same Δt, N and T_{trans} were used for all considered further systems.

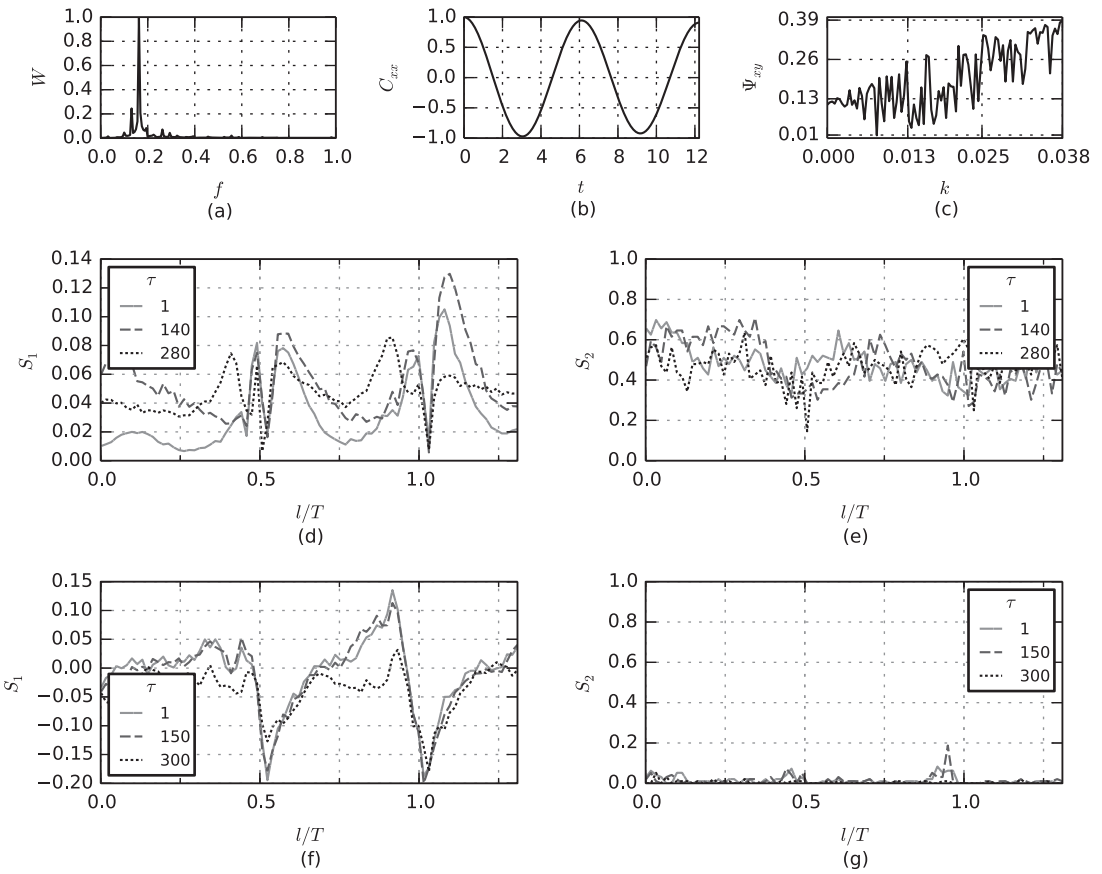

Fig. 2. Spectrum (a), auto-correlation function (b) and dependency of phase synchronisation index Ψ_{xy} on coupling coefficient k for driven Kiyashko–Pikovsky–Rabinovich oscillator. Dependencies of criteria (6) and (7) for unidirectionally coupled Kiyashko–Pikovsky–Rabinovich generator equations on model lag l for different values of prediction time: (d) and (f) for the criterion S_1, (e) and (g) for the criterion S_2. Dependencies (d, e) correspond to the polynomial approximating functions with polynomial order $\nu = 3$, while dependencies (f, g)—to the local linear ones. $D_s = 2$, $D_a = 1$ for all cases.

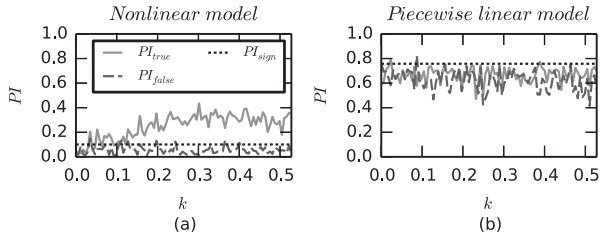

Fig. 3. The typical dependency of the prediction improvement *PI* on the coupling coefficient k for unidirectionally coupled Rössler oscillators for different types of predicting model: (a) polynomial; (b) piecewise linear.

generators. At the same time, the local-linear model gave completely opposite results. Mainly it is inappropriate due to very low specificity and sensitivity, but for $l = T/2$ and $l = T$ the method demonstrates significant improvement.

The detailed consideration of results achieved using local-linear models showed that for most values of l the main problem is insensitivity of the method to growth of coupling, i. e. *PI* does not increase with rise of k, as it is illustrated in Fig. 3. The local-linear model does not benefit from introducing the coupling term in most cases. Since using $l = T/2$ or $l = T$ makes information accumulated in the state vector insufficient, using the coupling term largely improves the forecasting capabilities of the model. But for these values of l not only PI_{true} but also PI_{false} increase in all considered examples except coupled Rössler oscillator in combination with local-linear models. The atypical dependency of PI_{false} on l for local-linear models caused by individual properties of Rössler attractor leads to the fact that values pessimal in most cases become optimal in this one.

4.3. Coupled Anishchenko–Astakhov generator equations

Let us consider the coupled equations of another popular radio-engineering chaotic generator [40]:

$$\begin{cases} \dot{x}_1 = 1.106x_1 + x_2 - x_1x_3 + ky_1 \\ \dot{x}_2 = -x_1 \\ \epsilon\dot{x}_3 = -0.68x_3 + \theta(x_1)x_1^2 \end{cases}$$

$$\begin{cases} \dot{y}_1 = 1.111y_1 + y_2 - y_1y_3 \\ \dot{y}_2 = -y_1 \\ \dot{y}_3 = -0.7y_3 + \theta(y_1)y_1^2 \end{cases}$$

$$\theta(x) = \begin{cases} 0, & \text{at } x <= 0 \\ 1, & \text{at } x > 1 \end{cases} \qquad (10)$$

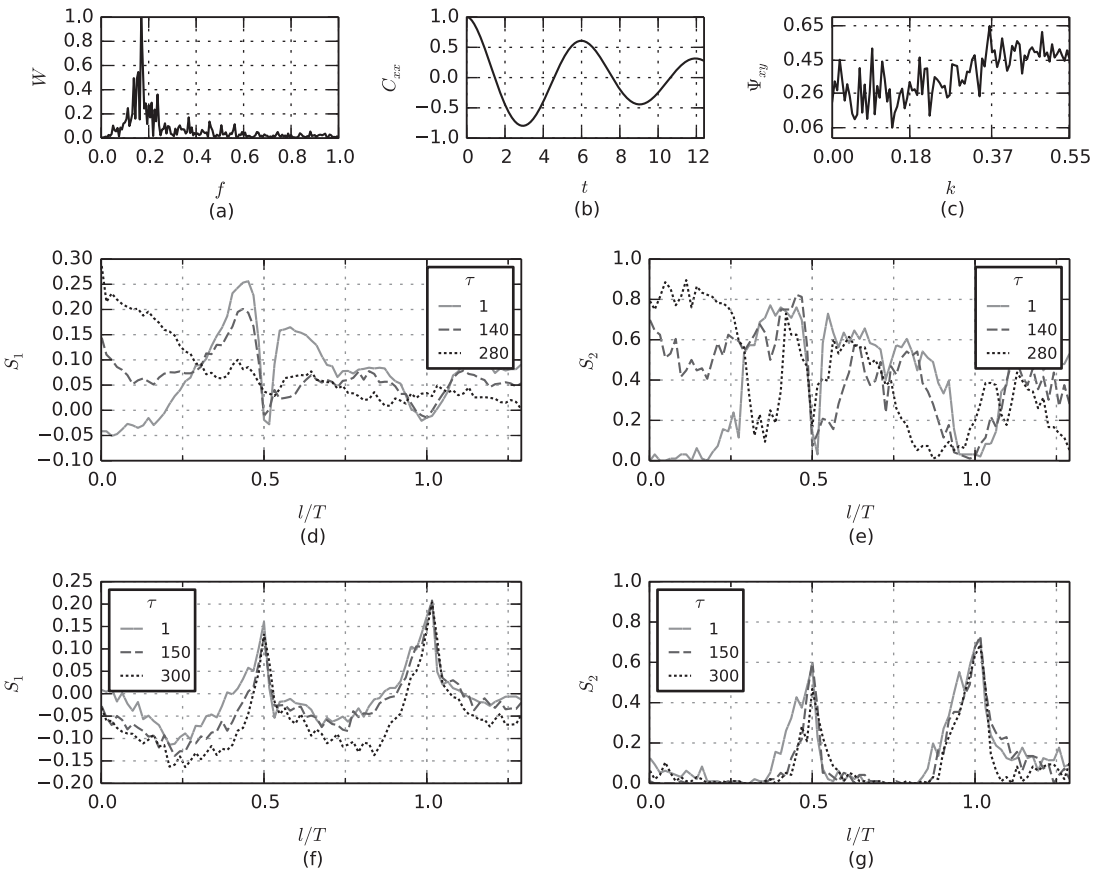

Fig. 4. Spectrum (a), auto-correlation function (b) and dependency of phase synchronisation index Ψ_{xy} on coupling coefficient k for driven Rössler oscillator. Dependencies of criteria (6) and (7) for unidirectionally coupled Rössler systems on model lag l for different values of prediction time: (d) and (f) for the criterion S_1, (e) and (g) for the criterion S_2. Dependencies (d, e) are correspond to the polynomial approximating functions with polynomial order $\nu = 3$, while dependencies (f, g)—to the local linear ones. $D_s = 2$, $D_a = 1$ for all cases.

The values of coupling coefficient were limited to 0.15. The positive Lyapunov exponent of the driven subsystem was $\Lambda_1 = 0.136$.

The auto-correlation function (see Fig. 5(b)) shows one main timescale approximately equal to 683 time points (6.83 units of dimensionless time), while the power spectrum indicates two main peaks (Fig. 5(a)). Also auto-correlation function is explicitly asymmetric, that was not the case for Kiyashko–Pikovsky–Rabinovich and Rössler systems (compare to Figs. 4(b) and 2(b)).

In Fig. 5 (d), the results of S_1 criterion for the polynomial model are plotted. One can see the same two minima as for previously considered systems, however they are shifted to the lower values of l than $T/2$ and T. Also they are not so sharp and their position differs for different τ. This can be explained by the presence of the second peak in the spectrum. The same picture can be seen for the second proposed criterion (Fig. 5(e)), while minima being better localised at the values $T/2$ and T. The best results are in the range $T/4 < l < T/3$ that is mainly the same range as for Kiyashko–Pikovsky–Rabinovich and Rössler systems.

The dependencies calculated for local-linear models (Fig. 5(f) and (g)) do not demonstrate any pronounced minimum in most curves. A minimum is still present at both curves $S_1(l)$ and $S_2(l)$ at $l = T$ for $\tau = 1$. The value $\tau = 1$ seems to be pessimal. The detailed analysis shows that a lot of false positives occur for $\tau = 1$, that is in a good correspondence with the results reported in [26].

4.4. Coupled 3D generator equations

Let us consider other systems which demonstrate complex non-linear dynamics—equations of coupled 3D generator [41] (also known as generator with 1.5 degrees of freedom).

$$\begin{cases} \dot{x}_1 = (F_1(x_3) - x_1)/3 + k y_3 \\ \dot{x}_2 = x_1 - x_3 \\ \dot{x}_3 = x_2 - 0.21 x_3 \end{cases}$$
$$\begin{cases} \dot{y}_1 = (F_2(y_3) - y_1)/3 \\ \dot{y}_2 = y_1 - y_3 \\ \dot{y}_3 = y_2 - 0.22 y_3 \end{cases} \quad (11)$$

$$F_1(z) = 26 z \exp(-z^2)$$
$$F_2(z) = 28 z \exp(-z^2) \quad (12)$$

Values of k were limited to 0.2. The first Lyapunov exponent was $\Lambda_1 = 0.148$. Coordinates $x_1 y_1$ were considered as observables. The sampling interval was chosen to be equal to 0.03 in order to have approximately the same number of

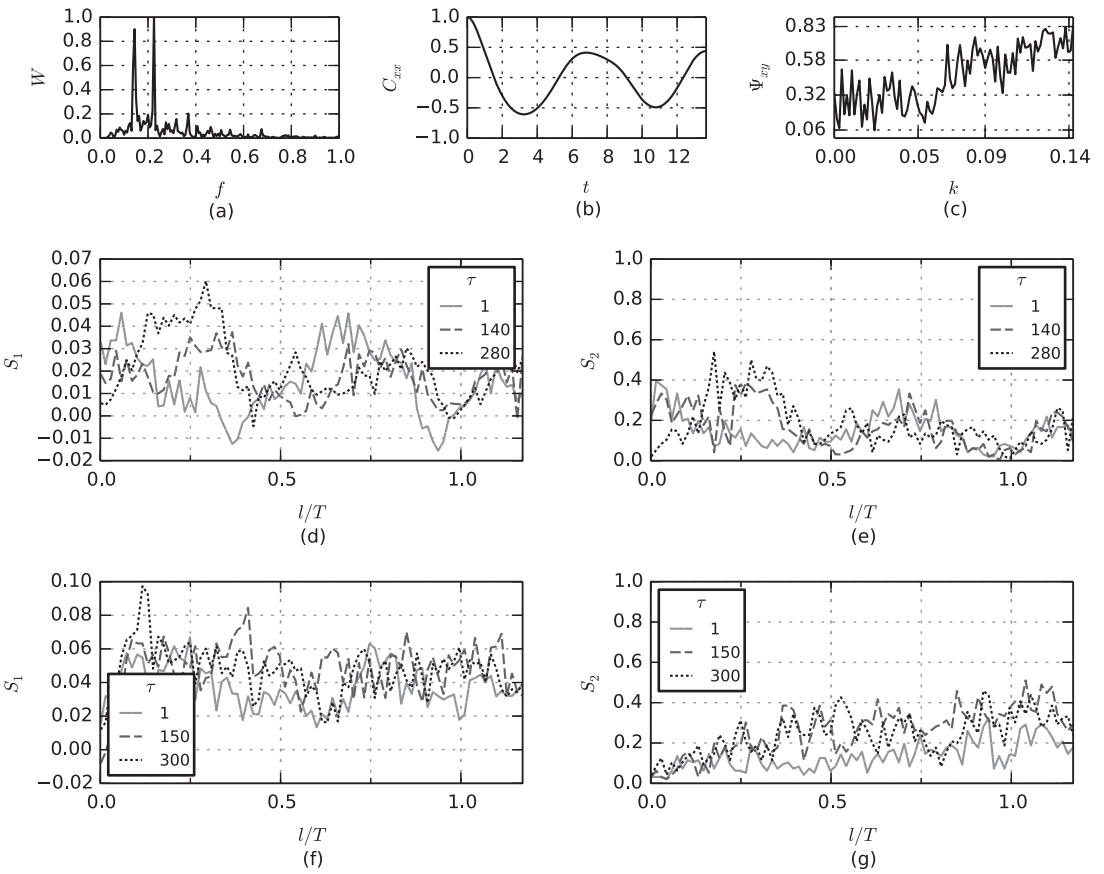

Fig. 5. Spectrum (a), auto-correlation function (b) and dependency of phase synchronisation index Ψ_{xy} on coupling coefficient k for driven Anishchenko–Astakhov oscillator. Dependencies of criteria (6) and (7) unidirectionally coupled Anishchenko–Astakhov generator equations on model lag l for different values of prediction time: (d) and (f) for the criterion S_1, (e) and (g) for the criterion S_2. Dependencies (d, e) are correspond to the polynomial approximating functions with polynomial order $\nu = 3$, while dependencies (f, g)—to the local linear ones. $D_s = 2$, $D_a = 1$ for all cases.

data points for each main timescale as in previously considered examples. The main timescale cannot be estimated from power spectrum (Fig. 6(a)) since there were many peaks. However its value $T = 20.2$ (i.e. 673 time points) can be extracted from auto-correlation function (Fig. 6(b)).

From Fig. 6(d) and (e) one can see that there are minima at both dependencies $S_1(l)$ and $S_2(l)$ for the values of l close to $T/2$ and T, though they are less pronounced due to more complex, in comparison with the previously considered systems, dynamics of coupled 3D generators. Also the value $\tau = 140$ that is close to $T/4$ seems to be much preferable in comparison to other considered values based on S_1, that confirms the outcomes of [26].

The local-linear model seems to be less efficient and mainly insensitive to the choice of l, though there are local minima for $S_1(l)$ (Fig. 6(f)) and $S_2(l)$ (Fig. 6(g)) dependencies for values of l not far from $T/2$ and T, and these minima are common for all three considered values of τ.

4.5. Dependency of optimal choice of l and τ on first Lyapunov exponent

Comparing the results obtained for different systems one can see that the dependencies $S_1(l)$ and $S_2(l)$ are affected by the value of the largest Lyapunov exponent: with increase of Λ minima and maxima become less pronounced. To test this fact again and to exclude possible influences of individual properties of considered system, the coupled Rössler oscillators (9) in different chaotic regimes correspondent to the different values of first Lyapunov exponent are considered: $\Lambda = 0.06$[2], $\Lambda = 0.08$[3], $\Lambda = 0.1$[4], and $\Lambda = 0.12$[5] (these values were calculated for $k = 0$). Power spectra of oscillators for different values of Λ are shown on Fig. 7(e)–(h), with main peak locating at ~ 0.18. the main timescale can be calculated using auto-correlation function (Fig. 7(a)–(d)) as a position of first local maximum. For $\Lambda_1 = 0.06$, $\Lambda_1 = 0.08$, and $\Lambda_1 = 0.10$ $T \approx 6.2$ (620 time points with sampling interval $\Delta t = 0.01$), for $\Lambda_1 = 0.12$ $T = 6.55$. One can notice that signal regularities decrease with the increase of Λ that can be seen analysing auto-correlation function which first local maximum is lower for larger values of Λ.

Since the local-linear models demonstrated atypical dependencies $S_1(l)$ and $S_2(l)$ for coupled Rössler oscillators we

[2] $a_1 = 0.146$, $b_1 = 0.3$, $c_1 = 10$, $a_2 = 0.16$, $b_2 = 0.2$, $c_2 = 16$
[3] $a_1 = 0.176$, $b_1 = 0.19$, $c_1 = 11$, $a_2 = 0.172$, $b_2 = 0.2$, $c_2 = 10$
[4] $a_1 = 0.2$, $b_1 = 0.3$, $c_1 = 10$, $a_2 = 0.2$, $b_2 = 0.15$, $c_2 = 7.5$
[5] $a_1 = 0.205$, $b_1 = 0.05$, $c_1 = 20$, $a_2 = 0.20$, $b_2 = 0.05$, $c_2 = 11$

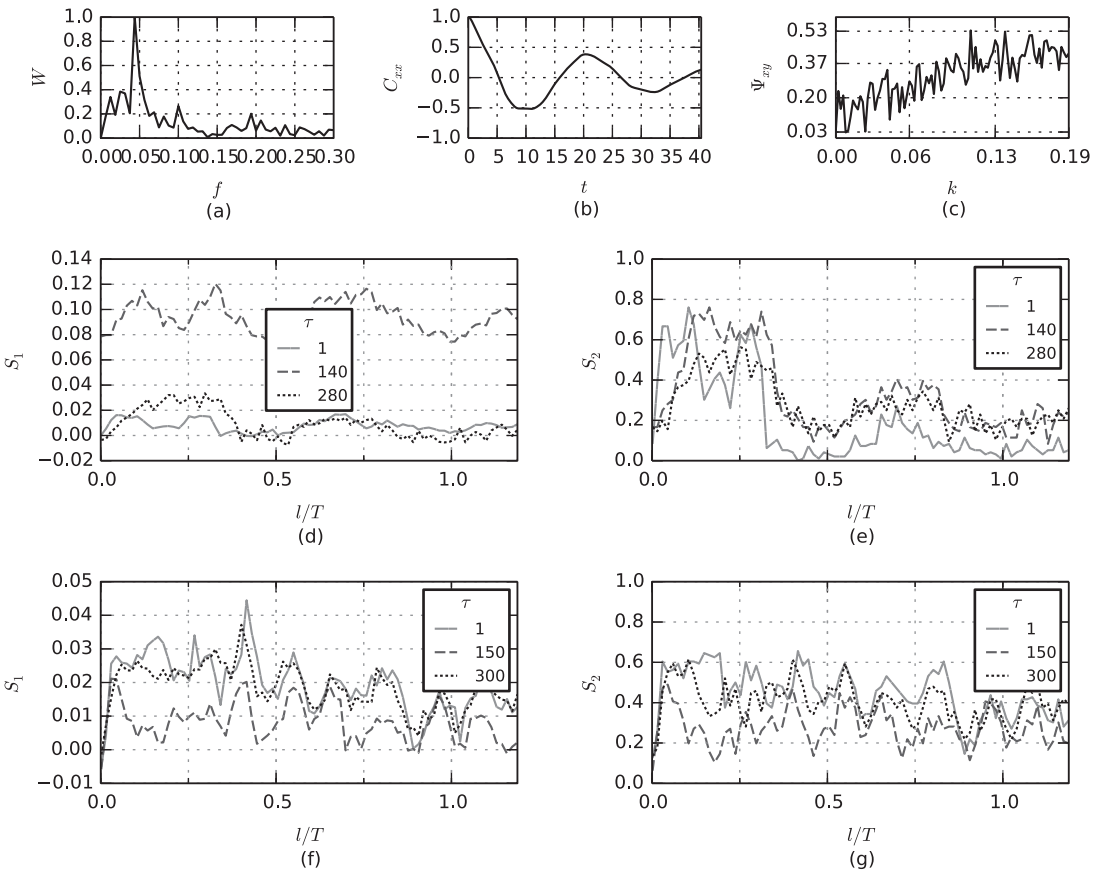

Fig. 6. Spectrum (a), auto-correlation function (b) and dependency of phase synchronisation index Ψ_{xy} on coupling coefficient k for driven 3D generator oscillator. Dependencies of criteria (6) and (7) for unidirectionally coupled 3D generator equations on model lag l for different values of prediction time: (d) and (f) for the criterion S_1, (e) and (g) for the criterion S_2. Dependencies (d, e) are correspond to the polynomial approximating functions with polynomial order $\nu = 3$, while dependencies (f, g)—to the local linear ones. $D_s = 2$, $D_a = 1$ for all cases.

focus here only on results for polynomial models. Comparing different curves on Fig. 8 one can see that the four curves plotted for S_1 for all free values of τ behave similar, but extrema are better pronounced for lower values of Λ_1.

Following criterion (6) the best results of method application correspond to l near but less $T/2$ for all values of τ. The criterion (7) approves this idea, though low (but not equal to one) values of l and large τ give the best results: $> 80\%$ of correct findings in the actual direction for which the coupling in the opposite one is considered to be insignificant. The 3D plots of $S_1(\tau, l)$ and $S_2(\tau, l)$ (see Fig. 9), constructed for $\Lambda_1 = 0.1$ demonstrate the same outcomes. One has to notice that both criteria show the method to be ineffective for $\tau = 1$ in combination with low values of lag ($l < T/4$), that proves the conclusions made in [26].

It can be noticed that the dependency $S_2(l)$ at $\Lambda_1 = 0.06$ differs from dependencies for other values of Λ_1 and mainly lies lower. The detailed analysis of curves $PI(k)$ in this case indicates the role of relatively (in comparison to other values of Λ_1) high phase synchronisation ($\Psi_{xy} = 0.7$ for most values of k). The relatively high synchronisation prevents the coupling direction from being established since signals of the driving and the driven systems become too similar. Therefore the specificity test usually fails (the coupling is detected as bidirectional), while PI in the actual direction is higher than in the wrong one.

The results similar to plotted on the Fig. 8 were obtained for other values of dimension D_s and polynomial order ν, namely for $D_s = 3$, $\nu = 3$ and $D_s = 2$, $\nu = 2$.

5. Conclusion and discussion

The active use of coupling estimation methods directly raises the questions of their reliability. Among other problems, the dependency of achieved results on method parameters is one of the most important. Here, some solution is provided for the Granger causality approach in application to the signals with one main timescale.

The general idea is to link the model parameters which have the meaning of time with the characteristic timescale of the measured signals. These parameters are: lag l, used to reconstruct the state vector from the scalar observable with the method of delays, and prediction time τ.

To implement this idea, we have developed two integral criteria of method efficiency, allowing to determine the best and the worst value of τ and l for a wide range of coupling strengths for unidirectionally coupled systems. These criteria were applied to different etalon systems. The first criterion is

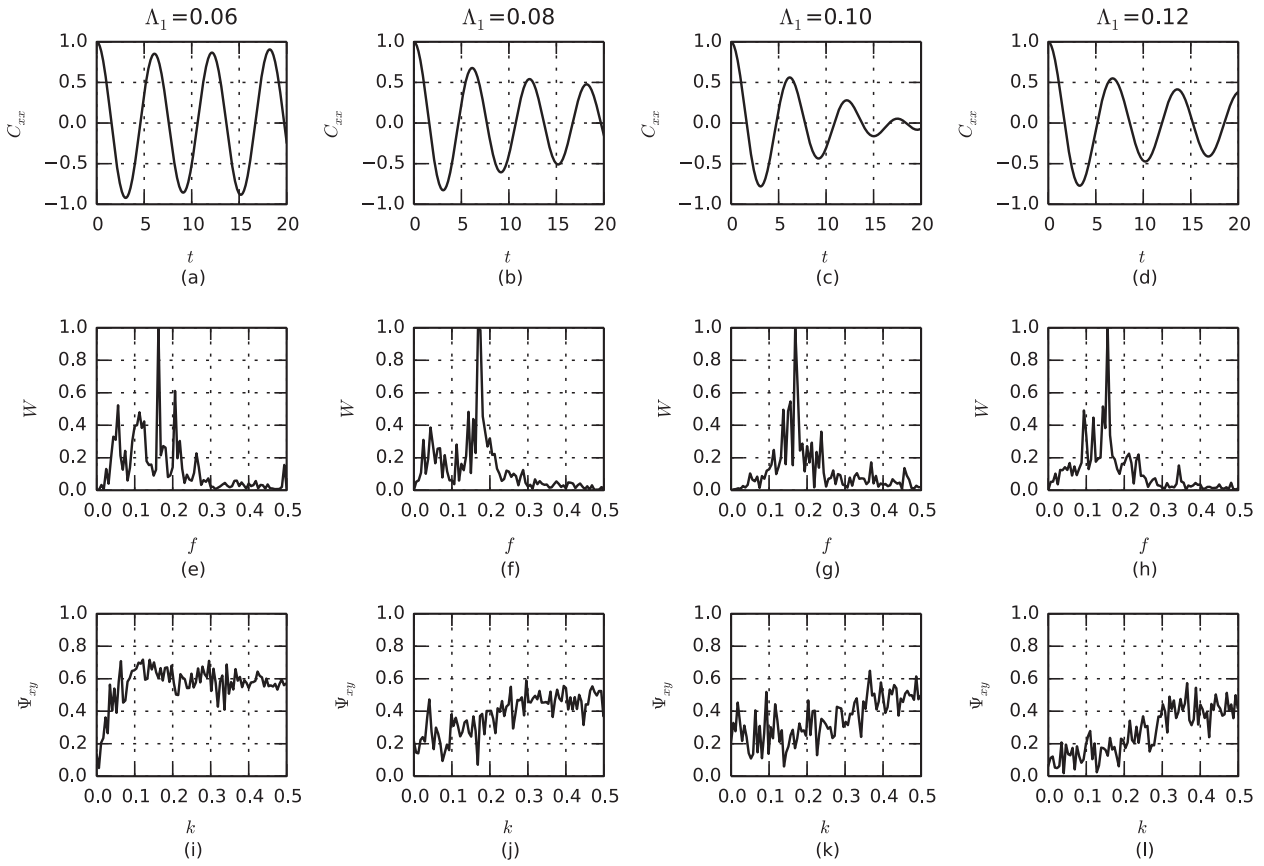

Fig. 7. Auto-correlation functions (a–d) and power spectra (e–h) of driven Rössler system for different values of Λ_1. Dependencies of phase synchronisation index on coupling coefficient $\Psi_{xy}(k)$ (i–l). Plots (a, e, i) correspond to $\Lambda = 0.06$, plots (b, f, j)—to $\Lambda = 0.08$, plots (c, g, k)—to $\Lambda = 0.1$, and plots (d, h, l)—to $\Lambda = 0.12$.

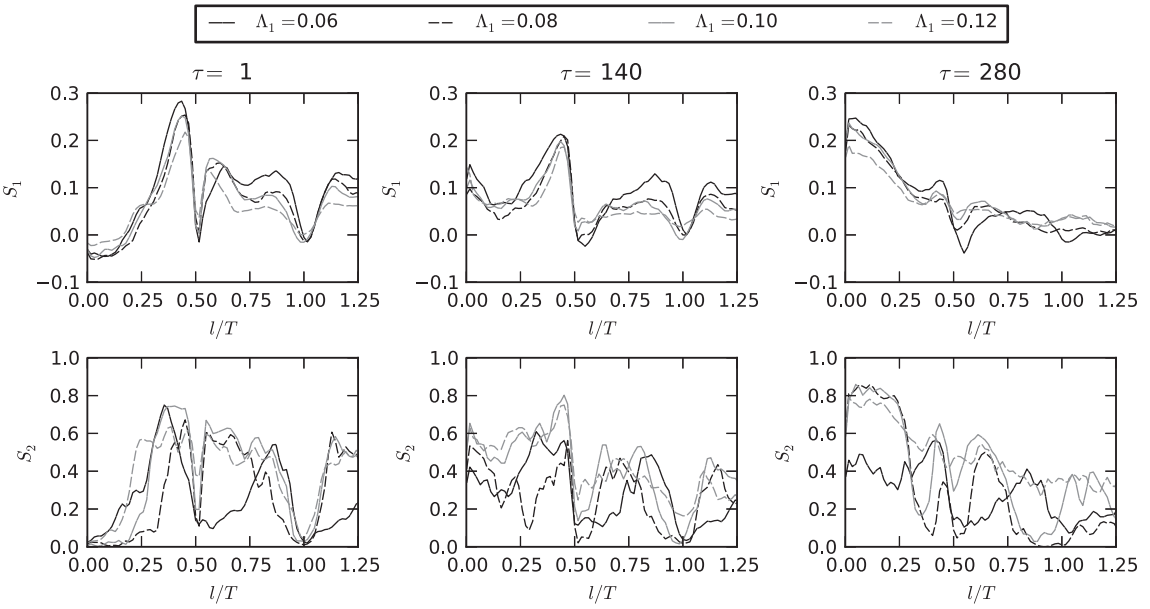

Fig. 8. Dependencies of criteria (6) and (7) on model lag l for different prediction times τ calculated for polynomial models.

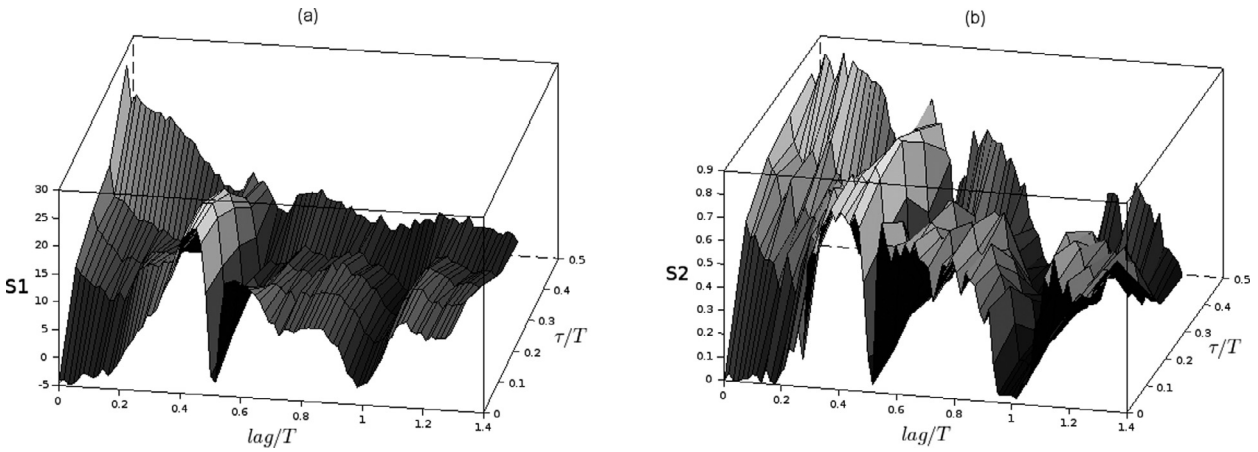

Fig. 9. 3D dependencies of criteria (6)(a) and (7)(b) on τ and l for unidirectionally coupled Rössler oscillators (9), constructed using polynomial approximation with $D_s = 2$, $D_a = 1$ and $\nu = 3$.

based on the average difference between the prediction improvement in the actual direction and in the false one. The second criterion mainly takes into account the level of significance of prediction improvement values. Here, to calculate significance level surrogate time series were used, while any other way to obtain significance level can be implemented.

To make decisions more general, the entire range of coupling strength from zero to the level, at which synchronisation occurs, was considered, while calculating proposed criteria. Consideration of larger coupling is usually senseless since synchronisation makes the measured series indistinguishable, preventing any attempt to detect coupling directionality.

We limited enumeration of other parameters of Granger causality in the research, otherwise it would be too extensional. Also one has to keep in the mind that arbitrary small models (i. e. models with lesser amount of coefficients) are more suitable for evaluation in moving window, since their coefficients can be estimated more reliably. Biological data are of primary interest nowadays, and since criticality was multiple times mentioned as a major problem of the biological signal analysis [42,43], using compact models can solve the problem of non-stationary series, at least partly, as it was illustrated for fast transient processes in [18]. However, in a number of examples we performed the similar calculations for systems with larger dimensions and polynomial order to improve the reliability. The results were very similar.

Since the polynomial models usually demonstrate the better specificity and sensitivity, it is preferable to focus on main issues for this class of models:

1. Selecting l equal or close to a half (or any integer or semi-integer number) of characteristic timescale of oscillations is unsuccessful. This is due to the high correlation (large linear dependence) between the selected components of the state vector, since the resulting vector carries too little (insufficient) information on the studied time series. This conclusion is reliable, as demonstrated in most of the considered systems.
2. The dependency of the efficiency of Granger causality method on l and τ is expressed more strongly (its minima are deeper and its maxima are higher) at lower values of first Lyapunov exponent of driven subsystem. This conclusion is based on the consideration of both the same systems with different Lyapunov exponents (with different parameter values), and by comparing the results obtained for different etalon systems.
3. One can suggest two strategies for choosing the optimal values of l and τ. The first one is to choose the value of lag less than half the characteristic period, for example $T/4$ or $T/3$, because for most systems there is a global maximum of values of the proposed criteria in this range. In the first strategy the value of prediction improvement is not very important, however choosing $\tau \approx T/4$ is still preferable for number of cases. To realise the second strategy it is necessary to focus on a combination of long-range prediction (around $T/2$) and small lag, for example, $l = T/10$. Which strategy is preferable—largely depends on the specifics of the object.

These recommendations have a certain range of application related to how well the regarded time series fit in the selected class. They are more applicable for weakly nonlinear systems (largest Lyapunov exponent is small, say, less than 0.15) with an obviously expressed peak in the spectrum in which the auto-correlation function decays slowly (no more than twice during the period).

Comparing results achieved for polynomial and local-linear models, we have to say that the approximating functions are of great importance. While for systems close to regular with very low effective coefficient of phase diffusion, the conclusions are the same for both considered functional bases, for more irregular systems the local-linear models occur less effective or demonstrate untypical results originating from the individual properties of considered objects. For local-linear model the dependency of method efficiency on l and τ disappears for lesser values of Lyapunov exponent and effective coefficient of phase diffusion than for polynomial one. This shows that only the different approaches to coupling analysis cannot be completely reduced one to another as it was shown in [8], but even for the same method the results depend on all details of its realisation. In particular, for the Ganger causality method, the choice of approximating functions, l and τ cannot be done separately.

Acknowledgment

This research was funded by Russian Science Foundation, grant no. 14-12-00291.

References

[1] Schreiber T. Measuring information transfer. Phys Rev Lett 2000;85:461.
[2] Faes L, Nollo G, Porta A. Information-based detection of nonlinear Granger causality in multivariate processes via a nonuniform embedding technique. Phys Rev E 2011;83:051112.
[3] Schelter B, Timmer J, Eichler M. Assessing the strength of directed influences among neural signals using renormalized partial directed coherence. J Neurosci Methods 2009;179:121–30.
[4] Granger CWJ. Investigating causal relations by econometric models and cross-spectral methods. Econometrica 1969;37(3):424–38.
[5] Rosenblum M, Pikovsky A. Detecting direction of coupling in interacting oscillators. Phys Rev E 2001;64:045202(R).
[6] Smirnov D, Bezruchko B. Estimation of interaction strength and direction from short and noisy time series. Phys Rev E 2003;68:046209.
[7] Barnett L, Barrett AB, Seth AK. Granger causality and transfer entropy are equivalent for Gaussian variables. Phys Rev Lett 2009;103:238701.
[8] Smirnov DA. Quantifying causal couplings via dynamical effects: a unifying perspective. Phys Rev E 2014;90:062921.
[9] Baccala LA, Sameshima K, Ballester G, Valle ACD, Timo-Laria C. Studying the interactions between brain structures via directed coherence and Granger causality. Appl Signal Process 1998;5:40–8.
[10] Tass P, Smirnov D, Karavaev A, Barnikol U, Barnikol T, Adamchic I, et al. The causal relationship between subcortical local field potential oscillations and parkinsonian resting tremor. J Neural Eng 2010;7:016009.
[11] Gourevitch B, Bouquin-Jeannes RL, Faucon G. Linear and nonlinear causality between signals: methods, examples and neurophysiological applications. Biol Cybern. 2006;95:349–69.
[12] Vakorin VA, Mišić B, Krakovska O, Bezgin G, McIntosh AR. Confounding effects of phase delays on causality estimation. PLoS One 2013;8(1):e53588.10.1371
[13] Sitnikova E, Dikanev T, Smirnov D, Bezruchko B, van Luijtelaar G. Granger causality: cortico-thalamic interdependencies during absence seizures in WAG/Rij rats. J Neurosci Methods 2008a;170:245–54.
[14] Sysoeva MV, Sitnikova E, Sysoev IV, Bezruchko BP, van Luijtelaar G. Application of adaptive nonlinear Granger causality: disclosing network changes before and after absence seizure onset in a genetic rat model. J Neurosci Methods 2014;226:33–41.
[15] Mokhov II, Smirnov DA. Empirical estimates of the influence of natural and anthropogenic factors on the global surface temperature. Docklady Earth Sciences 2009;427(1):798–803.
[16] Sommerlade L, Thiel M, Platt B, Plano A, Riedel G, Grebogi C, et al. Inference of Granger causal time-dependent influences in noisy multivariate time series. J Neurosci Methods 2012;203:173–85.
[17] Hesse W, Möller E, Arnold M, Schack B. The use of time-variant EEG Granger causality for inspecting directed interdependencies of neural assemblies. J Neurosci Methods 2003;124:27–44.
[18] Sysoev IV, Sysoeva MV. Detecting changes in coupling with Granger causality method from time series with fast transient processes. Phys D 2015;309:9–19.
[19] Winterhalder M, Schelter B, Hesse W, Schwab K, Leistritz L, Klan D, et al. Comparison of linear signal processing techniques to infer directed interactions in multivariate neural systems. Signal Process 2005;85:2137–60.
[20] Nalatore H, Ding M, Rangarajan G. Mitigating the effects of measurement noise on Granger causality. Phys Rev E 2007;75:031123.
[21] Nalatore H, N S, Rangarajan G. Effect of measurement noise on Granger causality. Phys Rev E 2014;90:062127.
[22] Yakhno YV, Molkov YI, Mukhin DN, Loskutov EM, Feigin AM. Reconstruction of an evolution operator as a technique of analysis of epileptiform electric brain activity. Izv VUZ Appl Nonlinear Dyn 2011;19(6):156–72.
[23] Packard N, Crutchfield J, Farmer J, Shaw R. Geometry from a time series. Phys Rev Lett 1980;45:712–16.
[24] Kougioumtzis D. State space reconstruction parameters in the analysis of chaotic time series—the role of the time window length. Phys D 1996;95(1):13–28.
[25] Smirnov D, Bezruchko B. Spurious causalities due to low temporal resolution: towards detection of bidirectional coupling from time series. Europhys Lett 2012;100:10005.
[26] Sysoeva MV, Dikanev TV, Sysoev IV. Selecting time scales for empirical model construction. Izv VUZ Appl Nonlinear Dyn 2012;20(2):54–62.
[27] 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.
[28] 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. Physics-Uspekhi 2008;51:304–10.
[29] Marinazzo D, Pellicoro M, Stramaglia S. Nonlinear parametric model for Granger causality of time series. Phys Rev E 2006;73:066216.
[30] Marinazzo D, Pellicoro M, Stramaglia S. Kernel method for nonlinear Granger causality. Phys Rev Lett 2008;100:144103.
[31] Casdagli M. Nonlinear prediction of chaotic time series. Physica D 1985;35:335–56.
[32] Sitnikova E, Dikanev T, Smirnov D, Bezruchko B, van Luijtelaar G. Granger causality: cortico-thalamic interdependencies during absence seizures in WAG/Rij rats. J Neurosci Methods 2008;170(2):245–54.
[33] Sysoev IV, Karavaev AS, Nakonechny PI. Role of model nonlinearity for Granger causality based coupling estimation for pathological tremor. Izv VUZ Appl Nonlinear Dyn 2010;18(2):81–90.
[34] Kornilov MV, Sysoev IV. Influence of the choice of the model structure for working capacity of nonlinear Granger causality approach. Izv VUZ Appl Nonlinear Dyn 2013;21(2):74–87.
[35] Allefeld C, Kurths J. Testing for phase synchronization. Int J Bifurc Chaos 2004;14(2):405–16.
[36] Pikovsky A, Rosenblum M, Kurths J. Synchronization: A Universal Concept in Nonlinear Sciences. Cambridge, UK: Cambridge University Press; 2001.
[37] Schwarz G. Estimating the dimension of a model. Ann Stat 1978;6(2):461–4.
[38] Rössler OE. An equation for continuous chaos. Phys Lett 1976;A57(5):397–8.
[39] Kijashko SV, Pikovsky AS, Rabinovich MI. A radio-frequency generator with stochastic behavior. Radio Eng Electron Phys 1980;25(2).
[40] Astakhov V, Anishchenko VS, Pecora LM. Bifurcations and chaos in oscillator with inertial nonlinearity. In: Nonlinear Dynamics in Circuits. Singapore: World Scientific; 1995. p. 55–88.
[41] Dmitriev AS, Kislov VY. Stochastic oscillations in a self-excited oscillator with a first-order inertial delay. Radiotekhnika i elektronika 1984;29:2389.
[42] Lovecchio E, Allegrini P, Geneston E, West BJ, Grigolini P. From self-organized to extended criticality. FrontPhysiol 2012;3. doi:10.3389/fphys.2012.00098.
[43] Luković M, Vanni F, Svenkeson A, Grigolini P. Transmission of information at criticality. Physica A 2014;416:430–8.