Estimation of interaction strength and direction from short and noisy time series

Dmitry A. Smirnov[1,*] and Boris P. Bezruchko[1,2]
[1] *Saratov Branch, Institute of RadioEngineering and Electronics of The Russian Academy of Sciences, 38 Zelyonaya Street, Saratov 410019, Russia*
[2] *Department of Nonlinear Processes, Saratov State University, 155 Moskovskaya Street, Saratov 410026, Russia*
(Received 9 June 2003; published 29 October 2003; publisher error corrected 7 November 2003)

A technique for determination of character and intensity of interaction between the elements of complex systems based on reconstruction of model equations for phase dynamics is extended to the case of short and noisy time series. Corrections, which eliminate systematic errors of the estimates, and expressions for confidence intervals are derived. Analytic results are presented for a particular case of linear uncoupled systems, and their validity for a much wider range of situations is demonstrated with numerical examples. The technique should be useful for the analysis of nonstationary processes in real time, including the situations of significant noise and restrictions on the observation time.

DOI: 10.1103/PhysRevE.68.046209 PACS number(s): 05.45.Tp

I. INTRODUCTION

One of the very important problems, which arises when complex multielement systems (in particular, biological ones) are investigated, is that of determining the presence and direction of interaction (coupling) between two subsystems from an experimental time series of their oscillations [1–14]. Such information allows better understanding of mechanisms of a complex system behavior. Thus, a great deal of attention is paid nowadays to the investigation of interaction between human cardiovascular and respiratory systems [7,11,15–20]. The problem of coupling characterization is also of applied importance for the purposes of medical diagnostics, e.g., for localization of epileptic focus based on the analysis of electroencephalogram and magnetoencephalogram recordings [12,21–28]. Nonstationarity of investigated processes, impossibility to collect sufficient amount of data, and necessity of analysis in real time require estimation of coupling characteristics under the condition of a short observation interval. The task is complicated by the presence of noise, especially if coupling is weak. Here, we develop an approach for estimation of weak coupling with a given degree of belief from short[1] segments of noisy time series.

A very nice and delicate idea for the detection of weak coupling was proposed by Rosenblum and Pikovsky [10,11]. Their technique is based on empirical construction of model maps, describing *phase dynamics* of the two subsystems, and is called *an evolution map approach* (EMA). Having an original time series $\{x_{1,2}(t_i)\}_{i=1}^{N_x}$, where x_1, x_2 are observables, $t_i = i\Delta t$, $i = 1, \ldots, N_x$, Δt is a sampling interval, one calculates time realizations of phases $\{\phi_{1,2}(t_i)\}_{i=1}^{N_\phi}$ and constructs a global model map, which characterizes the dependence of phase increments (over a time interval $\tau\Delta t$) on the phases of subsystems' oscillations, in the form

$$\Delta_{1,2}(t) \equiv \phi_{1,2}(t+\tau\Delta t) - \phi_{1,2}(t) = F_{1,2}(\phi_{1,2}(t), \phi_{2,1}(t), \mathbf{a}_{1,2}), \tag{1}$$

where τ is a positive integer, $F_{1,2}$ are trigonometric polynomials, and $\mathbf{a}_{1,2}$ are vectors of their coefficients. Using the estimates of coefficients $\hat{\mathbf{a}}_{1,2}$, obtained from the time series via the least-squares routine (LSR), one computes intensities of influence of the second subsystem on the first one ($2 \to 1$) $\hat{c}_1$ and of the first subsystem on the second one ($1 \to 2$) $\hat{c}_2$ and directionality index $\hat{d} = (\hat{c}_2 - \hat{c}_1)/(\hat{c}_2 + \hat{c}_1)$. Since $\hat{c}_{1,2} \geqslant 0$, $\hat{d}$ takes the values within the interval $[-1,1]$ only: $\hat{d} = 1$ or $\hat{d} = -1$ corresponds to unidirectional coupling $1 \to 2$ or $2 \to 1$, respectively, and $\hat{d} = 0$ for ideally symmetric coupling.

Numerical experiment showed [10] that a very large amount of data (typically about 10^4–10^5 data points) is necessary for correct and reliable determination of coupling character if noise is considerable. As shown in Secs. II A and II B of the present paper, this is because the estimators $\hat{c}_{1,2}$ and $\hat{d}$ are biased (systematic errors take place generally) while both bias and variance of $\hat{c}_{1,2}$ and $\hat{d}$ decrease with increase in time series length. In Sec. II C, we modify the EMA for correct estimation of coupling from short time series; namely, we do not use characteristics $\hat{c}_{1,2}$ and $\hat{d}$ directly, but propose *unbiased* estimates $\hat{\gamma}_{1,2}$ and $\hat{\delta}$ of different quantities, which are more suitable in the situation considered. Besides, working equations for the confidence intervals are derived (Secs. II C and II D). The suggested *interval statistical estimates* allow the inference about statistical significance of the obtained results. The shorter the accessible time series and the higher the noise level, the more necessary they become. Results of Sec. II are derived *analytically* for a particular case of linear uncoupled processes. In Sec. III their validity is demonstrated *numerically* for more complex and

[1] In practice one must often deal with time series segments of about 1000 data points that appear too short for many analysis techniques. E.g., electroencephalograms are recorded at a typical sampling rate of 250 Hz. The duration of quasistationary segments is about 4 s [29], which corresponds to 1000 data points of recording. Roughly speaking, in this paper we deal with time series lengths of about 1000 data points. The notions of ''short'' and ''long'' time series in the context of coupling characterization are formulated more accurately in Sec. III C.

*Corresponding author. Email address: smirnovda@info.sgu.ru

realistic situations. Conclusions are presented in Sec. IV. Analytic derivations are shown in the appendices.

II. DESCRIPTION OF THE APPROACH TO DETERMINATION OF COUPLING CHARACTER FROM SHORT TIME SERIES

A. Problem setting

For convenience of explanation and notations, let us follow the ideas of Rosenblum and Pikovsky [10,11] to formulate the problem. In Refs. [10,11] the authors consider model (1) and determine coupling characteristics from its coefficients. It is implicitly assumed that if the time series is generated by a mathematical equation, then the model approximates that equation very accurately. This is a plausible hypothesis when a long time series is considered. Here, we focus on short time series (where statistics may be quite poor), hence, we must take model imperfection into account and clearly distinguish between notations for the original mathematical system generating a series and a model constructed from that time series. This is necessary to address the question, to what extent coupling characteristics computed from the model coefficients are close to the corresponding characteristics of the original system. As a rule, we supply quantities belonging to the original system with a superscript "0."

Following the logic of Ref. [10], let us consider sufficiently simple and, simultaneously, universal stochastic differential equations, which reflect adequately the properties of a wide range of oscillatory processes (provided that each of the interacting subsystems exhibits pronounced main rhythm of oscillations [6,9,10]), as an original system:

$$\dot{\phi}_{1,2}(t)=\omega_{1,2}+f_{1,2}(\phi_{1,2}(t),\phi_{2,1}(t))+\xi_{1,2}(t), \qquad (2)$$

where $\phi_{1,2}(t)$ are *unwrapped* phases of subsystem oscillations, $\omega_{1,2}$ are parameters controlling angular frequencies, $f_{1,2}$ are 2π periodic in both argument functions, $\xi_{1,2}$ are random processes normally distributed with zero mean and correlation function $E[\xi_{1,2}(t)\xi_{1,2}(t')]=D_{1,2}\delta(t-t')$ ($E[\cdot]$ stands for mathematical expectation), $\xi_1(t)$ and $\xi_2(t)$ do not depend on each other and on $\phi_1(t)$ and $\phi_2(t)$. Since we aim at dealing with discrete time series, it is more relevant to speak of difference equations instead of differential ones. System (2) can be transformed to such a form if one proceeds from the derivatives $\dot{\phi}_{1,2}$ to finite difference $\Delta_{1,2}$ over a time interval $\tau\Delta t$ and derives

$$\Delta_{1,2}(t)=F^0_{1,2}(\phi_{1,2}(t),\phi_{2,1}(t))+\varepsilon_{1,2}(t), \qquad (3)$$

where $F^0_{1,2}(\phi_1,\phi_2)\equiv E[\Delta_{1,2}(t)|\phi_{1,2}(t),\phi_{2,1}(t)]$ are 2π-periodic functions, $E[\cdot|\cdot]$ stands for conditional expectation, $\varepsilon_{1,2}$ are random processes with zero mean. The form of $F^0_{1,2}$ and characteristics of $\varepsilon_{1,2}$ are determined by the form of $f_{1,2}$, the value of $\tau\Delta t$, and characteristics of $\xi_{1,2}$.

Let functions $F^0_{1,2}$ be approximated accurately with low-order trigonometric polynomials as in Ref. [10]. That is,[2]

$$F^0_1(\phi_1,\phi_2)=\sum_{i=1}^{L^0_1} a^0_{1,i}g_i(\phi_1,\phi_2), \qquad (4)$$

where $g_1=1$, $g_i=\cos(m_i\phi_1+n_i\phi_2)$ for even $i>1$, $g_i=\sin(m_i\phi_1+n_i\phi_2)$ for odd $i>1$, and L^0_1 is the number of terms of the polynomial F^0_1. For $i\geqslant 1$, $m_{2i}=m_{2i+1}$ are nonnegative integers, $n_{2i}=n_{2i+1}$ are arbitrary integers, and $m_1=1$, $n_1=0$ by definition.

The intensity of influence of the second subsystem on the first one, c^0_1, is determined by the steepness of the dependence of F^0_1 on the phase of the second subsystem, ϕ_2, i.e., $\partial F^0_1/\partial\phi_2$. Similarly, c^0_2 is determined by $\partial F^0_2/\partial\phi_1$. More strictly, $(c^0_1)^2=(1/2\pi^2)\int_0^{2\pi}\int_0^{2\pi}(\partial F^0_1/\partial\phi_2)^2 d\phi_1\, d\phi_2$ by definition.[3] By inserting function (4) into this expression and taking the definite integral, one derives

$$c^0_1=\sqrt{\sum_{i=1}^{L^0_1} n_i^2(a^0_{1,i})^2}. \qquad (5)$$

Then, the directionality index is $d^0=(c^0_2-c^0_1)/(c^0_2+c^0_1)$. Thus, the expressions for coupling characteristics $c^0_{1,2},d^0$ in terms of the coefficients $\mathbf{a}^0_{1,2}$ of the original equations (3) are derived.

Since in this context one deals with a short discrete sample and the values $a^0_{1,i}$ are unknown *a priori*, it is impossible to use directly expression (5) in practice for characterization of coupling. So, let us formulate the problem as follows: it is necessary to get the estimates of coupling characteristics (e.g., of $c^0_{1,2},d^0$) from *a single realization* of the random process (2).

B. Properties of Rosenblum's and Pikovsky's estimators

Construction of model maps (1) can be regarded as the first step to obtaining the estimates needed. Let "model" polynomials $F_{1,2}$ be of sufficiently high order to involve all the terms present in the original polynomials $F^0_{1,2}$. According to Ref. [10], the estimates of $a^0_{1,i}$ are obtained via LSR (LS estimates), i.e., from the requirement

$$\sum_{i=1}^{N}[\Delta_1(t_i)-F_1(\phi_1(t_i),\phi_2(t_i),\mathbf{a}_1)]^2\rightarrow\min, \qquad (6)$$

where $N=N_\phi-\tau$. Let us denote the solution to this problem as $\hat{\mathbf{a}}_1$. Then, the estimator of c^0_1 is given by

$$\hat{c}_1=\sqrt{\sum_{i=1}^{L_1} n_i^2\hat{a}^2_{1,i}}, \qquad (7)$$

[2]Here and throughout the paper, we present formulas only for the first subsystem; all the expressions are "symmetric" for the second one.

[3]We introduce the normalizing multiplier $1/2\pi^2$ for convenience.

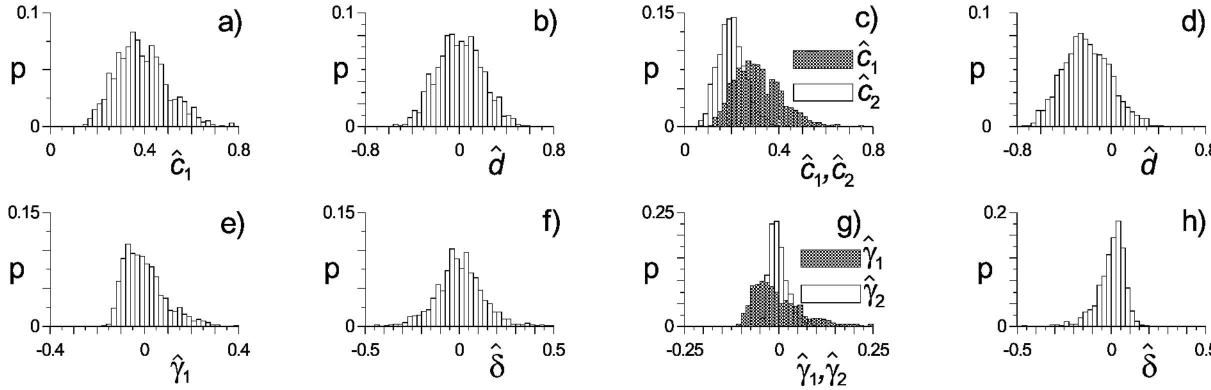

FIG. 1. Histograms for the estimates of coupling, constructed as a result of processing of 1000 time realizations of Eqs. (8) with $\Delta t = 0.2\pi$, $\tau = 10$, $\omega_1 = \omega_2 = 1$ (p is the relative frequency of falling into a bin, bin width is equal to 0.02 for the top figures and to 0.025 for the bottom figures). For identical subsystems with $D_1 = D_2 = 0.4$, (a) $\hat{c}_1$ (a biased estimator), (b) $\hat{d}$, (e) $\hat{\gamma}_{1,2}$, and (f) $\hat{\delta}$ (unbiased estimators). For subsystems with different noise levels $D_1 = 0.4$, $D_2 = 0.1$, (c) $\hat{c}_1$ and $\hat{c}_2$ (exhibit different biases), (d) $\hat{d}$ (exhibits negative bias), (g) $\hat{\gamma}_{1,2}$, and (h) $\hat{\delta}$ (unbiased estimators).

and the estimator of the directionality index is $\hat{d} = (\hat{c}_2 - \hat{c}_1)/(\hat{c}_2 + \hat{c}_1)$. As it is known from the theory of statistical estimation [30], under some conditions LS estimates $\hat{a}_{1,i}$ are consistent. It means that for a very large amount of data ($N \to \infty$) they are unbiased and have practically no scattering; in other words, $\hat{a}_{1,i}$ are almost precisely equal to $a^0_{1,i}$. As a result, $\hat{c}_{1,2}$ and $\hat{d}$ are equal to the true values $c^0_{1,2}$, d^0 as well, which allows correct inference about coupling character.[4] However, if the time series is short (and whether it can be regarded as a long one is not known *a priori*), the following important questions arise.

(1) What is the distribution of the estimates $\hat{c}_{1,2}$ and $\hat{d}$? Are they biased or not?

(2) How can statistical significance of the results be estimated? Or, can one draw a reliable conclusion about coupling presence and direction having computed the numbers $\hat{c}_{1,2}$ and $\hat{d}$?

To illustrate the importance of these questions, let us consider a simple *demonstrative example* when two subsystems are uncoupled and linear, that is, system (2) with $f_{1,2} \equiv 0$. By integrating Eqs. (2) analytically over the interval $\tau\Delta t$, one derives the equations in terms of finite differences

$$\Delta_{1,2}(t) = \omega_{1,2}\tau\Delta t + \varepsilon_{1,2}(t), \qquad (8)$$

where $\varepsilon_{1,2}$ are Gaussian random processes independent of each other with variances $D_{1,2}\tau\Delta t$. Obviously, one has $c^0_1 = c^0_2 = d^0 = 0$ in this case. The estimates $\hat{c}_{1,2}$ and $\hat{d}$ computed from a time series, consisting of 10^3 data points and being simulated numerically (see Sec. III A for details), are misleading. Their distributions are shown in Figs. 1(a–d). Thus, in the case of identical subsystems ($D_1 = D_2$ and $\omega_1 = \omega_2$) $\hat{c}_1$ is always positive and takes sufficiently large values [Fig. 1(a)], i.e., it is a biased estimator for $c^0_1 = 0$. $\hat{d}$ is unbiased, but exhibits quite a large scattering; even the values of $\hat{d} = \pm 0.4$ are encountered quite often [Fig. 1(b)]. Thus, it is very probable to get spurious indication of the presence of interaction from a single realization. The situation becomes even more complicated when subsystems are nonidentical. It is illustrated in Figs. 1(c,d) for the case $D_1 > D_2$, $\omega_1 = \omega_2$. The estimators $\hat{c}_{1,2}$ are biased, bias in $\hat{c}_1$ being greater [Fig. 1(c)]. This leads to biasedness of $\hat{d}$ whose values are *systematically* less than zero [Fig. 1(d)]. Hence, predominant influence $2 \to 1$ is diagnosed, even though coupling is absent in reality. Different biases in $\hat{c}_1$ and $\hat{c}_2$ and indication of coupling direction $2 \to 1$ are observed also in the case of uncoupled subsystems with different angular frequencies: $D_1 = D_2$, $\omega_1 > \omega_2$, and $\tau > 1$. Distributions, qualitatively the same as in Figs. 1(c,d), are obtained, e.g., for $D_1 = D_2 = 0.4$, $\omega_1 = 1.5$, $\omega_2 = 0.5$, $\tau = 10$. Let us consider the cause of the systematic errors and other properties of $\hat{c}_{1,2}$ and $\hat{d}$ in more detail.

Since $\hat{c}_{1,2}$ are functions of $\hat{\mathbf{a}}_{1,2}$ (7), their probabilistic properties can be deduced from the properties of $\hat{\mathbf{a}}_{1,2}$, the latter being determined by the properties of the noise $\varepsilon_{1,2}$. Let the estimates $\hat{\mathbf{a}}_{1,2}$ be unbiased (Appendix A). Then, for each estimate $\hat{a}_{1,i}$ in accordance with the property of variance it holds $E[\hat{a}^2_{1,i}] = (E[\hat{a}_{1,i}])^2 + \sigma^2_{\hat{a}_{1,i}} = (a^0_{1,i})^2 + \sigma^2_{\hat{a}_{1,i}}$. That is, $\hat{a}^2_{1,i}$ is a biased estimator for $(a^0_{1,i})^2$ and its bias equals its variance $\sigma^2_{\hat{a}_{1,i}}$. It follows from this and Eq. (7) that

$$E[\hat{c}^2_1] = \sum_{i=1}^{L_1} n_i^2 E[\hat{a}^2_{1,i}] = (c^0_1)^2 + \sum_{i=1}^{L_1} n_i^2 \sigma^2_{\hat{a}_{1,i}}, \qquad (9)$$

i.e., $\hat{c}^2_1$ is a biased estimator for $(c^0_1)^2$ despite $\hat{a}_{1,i}$ being unbiased. Bias in $\hat{c}^2_1$ equals $\Sigma_{i=1}^{L_1} n_i^2 \sigma^2_{\hat{a}_{1,i}}$. The greater the variance of ε_1 and the shorter the time series, the greater the variances $\sigma^2_{\hat{a}_{1,i}}$ (Appendix A) and, hence, the greater the bias

[4]One additional requirement for obtaining reliable results is that the subsystems should not be in a synchronous regime in order that their phases might serve as independent variables [10].

in $\hat{c}_1^2$. Therefore, $\hat{c}_1^2$ systematically exceeds $(c_1^0)^2$. The same holds for $\hat{c}_1$ being considered as an estimator for c_1^0. This explains the results shown in Fig. 1(a).[5]

Since $\sigma^2_{\hat{a}_{1,i}}$ rises with D_1 and $\sigma^2_{\hat{a}_{2,i}}$ with D_2, bias in $\hat{c}_1$ is greater than bias in $\hat{c}_2$ for $D_1 > D_2$ and other equal conditions. This explains the results of Figs. 1(c,d). Similarly, $\sigma^2_{\hat{a}_{1,i}}$ rises significantly with ω_1 and $\sigma^2_{\hat{a}_{2,i}}$ with ω_2 for $\tau > 1$ (Appendix A), which accounts for different biases in $\hat{c}_1$ and $\hat{c}_2$ for uncoupled systems with different angular frequencies.

C. Unbiased estimators and confidence intervals

To derive expressions for the estimators suitable for analysis of short time series, we use $\hat{c}_{1,2}^2$ as a basis, but remove their biases and estimate confidence intervals for $(c_{1,2}^0)^2$. The latter task appears much more difficult here than the removal of biases. From Eq. (9) one can see that an unbiased estimator of $(c_1^0)^2$ is

$$\hat{\gamma}_1 = \hat{c}_1^2 - \sum_{i=1}^{L_1} n_i^2 \hat{\sigma}^2_{\hat{a}_{1,i}}, \tag{10}$$

where $\hat{\sigma}^2_{\hat{a}_{1,i}}$ are unbiased estimates of variances $\sigma^2_{\hat{a}_{1,i}}$. Derivation of $\hat{\sigma}^2_{\hat{a}_{1,i}}$ is not trivial. Since analytic expressions for $\hat{\sigma}^2_{\hat{a}_{1,i}}$ cannot be derived in general, we confine ourselves to a particular, but sufficiently realistic, case and derive the expressions using simplifying assumptions about the properties of the random processes $\varepsilon_{1,2}(t)$. They are assumed to be Gaussian and statistically independent of each other and of $\phi_{1,2}(t)$ [see conditions (C1)–(C4) and other details in Appendix A]. The derived estimates $\hat{\sigma}^2_{\hat{a}_{1,i}}$ depend on τ, $\hat{a}_{1,1}$, $\hat{a}_{2,1}$ in quite a complicated manner.

Quantities $\hat{\gamma}_{1,2}$ are the estimators for $(c_{1,2}^0)^2$. They allow inference about the presence of influence of one system on another. We do not deal with estimation of $c_{1,2}^0$ since it encounters greater theoretical difficulties: one cannot derive estimates with known distribution law for a sufficiently general case, therefore, derivation of unbiased estimators is possible only under additional strong assumptions. Due to similar reasons in order to characterize coupling direction, we propose just the use of a quantity $\delta^0 = (c_2^0)^2 - (c_1^0)^2$, rather than a normalized quantity d^0. An unbiased estimator for δ^0 is $\hat{\delta} \equiv \hat{\gamma}_2 - \hat{\gamma}_1$.

Now, let us estimate the significance of the numbers $\hat{\gamma}_{1,2}$ and $\hat{\delta}$ obtained from a single realization. The variance of $\hat{\gamma}_1$ for the considered time series lengths $N \sim 10^3$ is equal[6] approximately to the variance of $\hat{c}_1^2$, which is expressed in terms of covariations of $\hat{a}_{1,i}$:

$$\sigma^2_{\hat{\gamma}_1} \approx \sigma^2_{\hat{c}_1^2} = \sum_{i=1}^{L_1} \sum_{j=1}^{L_1} n_i^2 n_j^2 \operatorname{cov}(\hat{a}_{1,i}^2, \hat{a}_{1,j}^2). \tag{11}$$

Since true values of covariations are unknown, they should be estimated from a time series as well. The difficulty of obtaining a "good" estimate $\hat{\sigma}^2_{\hat{\gamma}_1}$ for $\sigma^2_{\hat{\gamma}_1}$ consists in the following. In order to avoid false conclusions about the presence of coupling, it is not allowable to obtain understated estimates of $\sigma^2_{\hat{\gamma}_1}$. In order to detect the presence of weak interaction, it is not allowable to obtain overstated ("pessimistic") estimates of $\sigma^2_{\hat{\gamma}_1}$. Having overcome some technical difficulties, we derive the following semiempirical formula:

$$\hat{\sigma}^2_{\hat{\gamma}_1} = \begin{cases} \displaystyle\sum_{i=1}^{L_1} n_i^4 \hat{\sigma}^2_{\hat{a}^2_{1,i}}, & \hat{\gamma}_1 \geq 5\left(\displaystyle\sum_{i=1}^{L_1} n_i^4 \hat{\sigma}^2_{\hat{a}^2_{1,i}}\right) \\ \displaystyle\frac{1}{2}\sum_{i=1}^{L_1} n_i^4 \hat{\sigma}^2_{\hat{a}^2_{1,i}}, & \text{otherwise,} \end{cases} \tag{12}$$

where $\hat{\sigma}^2_{\hat{a}^2_{1,i}}$ are expressed via the estimates $\hat{a}_{1,i}$ and $\hat{\sigma}^2_{\hat{a}_{1,i}}$ derived earlier (see Appendix B for details).

The estimator $\hat{\gamma}_1$ has asymmetric (right skewed) distribution for low-order trigonometric polynomials typically used (Sec. III A). Therefore, we take "asymmetric" expression $[\hat{\gamma}_1 - \alpha\hat{\sigma}_{\hat{\gamma}_1}, \hat{\gamma}_1 + \beta\hat{\sigma}_{\hat{\gamma}_1}]$ as an estimate of the confidence interval for $(c_1^0)^2$. We choose the values of constants α and β empirically to provide necessary significance level; e.g., a 95% confidence interval is achieved if $\alpha = 1.6$, $\beta = 1.8$ (Appendix B). The conclusion about the presence of influence $2 \to 1$ can be drawn with error probability of 2.5% provided

$$\hat{\gamma}_1 - \alpha\hat{\sigma}_{\hat{\gamma}_1} > 0. \tag{13}$$

[5]Note that in Fig. 1(a), the probability density function (pdf) of the estimator $\hat{c}_1$ is equal to 0 near the true value $c_1^0 = 0$ despite estimates $\hat{a}_{1,i}$ being often equal to their true values $a_{1,i}^0 = 0$. This is explained as follows. In the considered example (see the structure of the model trigonometric polynomial in Sec. III A) there are ten estimates $\hat{a}_{1,i}$ that contribute to the value of $\hat{c}_1$, i.e., with $n_i \neq 0$, see Eq. (7). These ten estimates are independent random quantities. Then, $\hat{c}_1$ is equal to zero only when all these ten estimates are equal to zero *simultaneously*, which is very unlikely. In fact, as one can suggest from expression (7), the shape of the pdf of $\hat{c}_1^2$ is very similar to χ^2 distribution with several degrees of freedom, in our example with ten degrees of freedom. Then, $\hat{c}_1$ is distributed approximately according to χ distribution with ten degrees of freedom, that is, its pdf is approximately $p(x) \propto x^9 e^{-x^2 2\sigma^2_{\hat{a}_{1,i}}}$ for $x > 0$. Hence, the probability that the value of $\hat{c}_1$ would be close to the true value $c_1^0 = 0$ is, indeed, negligibly small.

[6]As numerical experiments show, the variance of $\hat{\gamma}_1$ is greater than the variance of $\hat{c}_1^2$ by approximately 4%. This difference can be neglected within the limits of precision of our deduction that is determined by slight violation of the condition (C1) (Appendix A).

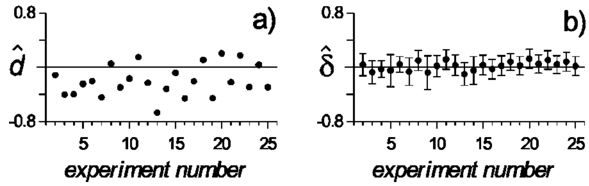

FIG. 2. Estimates of coupling for example (8): the results obtained from the first 25 of 1000 time realizations for the subsystems with different noise levels $D_1=0.4$, $D_2=0.1$. (a) $\hat{d}$ takes predominantly negative values, often large in absolute value. (b) $\hat{\delta}$ (circles) takes negative as well as positive values. For each single time realization the estimated confidence intervals (shown as error bars), as a rule, include zero; the experiment number 20 is the most close to spurious conclusion about the presence of coupling that would correspond to the expected 2.5 % of false conclusions.

The degree of belief can be adjusted by changing α (and, hence, confidence interval width).

Conclusion about predominant direction of interaction can be drawn after estimation of the variance of $\hat{\delta}$. Its reasonable estimator is $\hat{\sigma}^2_{\hat{\delta}}=\hat{\sigma}^2_{\hat{\gamma}_1}+\hat{\sigma}^2_{\hat{\gamma}_2}$. Since distribution law for $\hat{\delta}$ is, as a rule, more or less symmetric, confidence interval for δ^0 is reasonable to be searched for in a "symmetric" form $\hat{\delta}\pm\alpha\hat{\sigma}_{\hat{\delta}}$. Our experiments show that $\alpha=1.6$ again provides $\approx 95\%$ confidence interval. More accurately, if $\alpha=1.6$, then the obtained values

$$\hat{\gamma}_2-\alpha\hat{\sigma}_{\hat{\gamma}_2}>0 \quad \text{and} \quad \hat{\delta}-\alpha\hat{\sigma}_{\hat{\delta}}>0 \tag{14}$$

allow the statement about predominant influence $1\rightarrow 2$ with the error probability of 2.5%. Vice versa, if

$$\hat{\gamma}_1-\alpha\hat{\sigma}_{\hat{\gamma}_1}>0 \quad \text{and} \quad \hat{\delta}+\alpha\hat{\sigma}_{\hat{\delta}}<0, \tag{15}$$

the conclusion that the influence $2\rightarrow 1$ is stronger can be drawn with the same probability of error. If none of the relations (14) and (15) holds, coupling directionality cannot be determined with a given reliability.

Let us consider results of application of the proposed estimators $\hat{\gamma}_{1,2}$ and $\hat{\delta}$ to example (8) (Sec. II B). The absence of systematic errors for $\hat{\gamma}_{1,2}$ and $\hat{\delta}$ is illustrated in Figs. 1(e–h). Figure 2 demonstrates usefulness of the interval estimate $\hat{\delta}\pm\alpha\hat{\sigma}_{\hat{\delta}}$ to ensure reliable conclusions about coupling direction. In accordance with the expected 2.5% error probability, relative frequency of false conclusions about the presence of influence $2\rightarrow 1$ based on Eq. (13) was equal to 0.023 (i.e., false conclusions were drawn for 23 simulated time series of 1000 ones). Approximately the same was true for the frequency of false conclusions about coupling direction based on Eqs. (14) and (15).

D. Conditions for applicability of suggested estimators

The expressions for the estimators derived above are valid under the conditions (C1)–(C4) (Appendix A), which corresponds to $f_{1,2}\equiv 0$. However, they are also applicable when f_1 and f_2 are nonzero, but "small." Let us make the notion "small" more exact using the following rough arguments. Let, for simplicity, angular frequencies of subsystems $\omega_{1,2}$ be approximately equal to each other and equal to ω. Hence, $\omega\propto 1/T_{char}$, where T_{char} is a characteristic time scale. The value of $\tau\Delta t$ optimal for coupling characterization is also equal to T_{char} [11]. "Contribution of nonlinearity $f_{1,2}$" to the Eqs. (3), $\int_t^{t+\tau\Delta t} f_{1,2}(\phi_{1,2}(t'),\phi_{2,1}(t'))dt'$, is, then, of the order of $\|f_{1,2}\|/\omega$ (where $\|\cdot\|$ stands for an appropriately defined norm of a function, e.g., its root-mean-squared value). This contribution should be significantly less than the "contribution of noise $\xi_{1,2}$" to Eqs. (3). The latter is the standard deviations of $\varepsilon_{1,2}$, which are about $\sqrt{D_{1,2}T_{char}}=\sqrt{D_{1,2}/\omega}$. Thus, the smallness of $f_{1,2}$ means $\|f_{1,2}\|\ll\sqrt{D_{1,2}\omega_{1,2}}$.

However, there is also another case where our estimators remain applicable. This is the case of very small noise, i.e., an inverse situation: $\|f_{1,2}\|\gg\sqrt{D_{1,2}\omega_{1,2}}$ (provided that coupling is not so strong as to cause synchronization). The reason is that the variances of all the estimators are very small and, therefore, estimators $\hat{\gamma}_{1,2}$ and $\hat{\delta}$ (as well as Rosenblum's and Pikovsky's estimators) are almost exactly equal to the corresponding true values.

A much more difficult situation is encountered if contributions of nonlinearity and noise to Eqs. (3) are equally strong. Another serious problem is the so-called "error-in-variables" problem, i.e., the presence of significant errors in the observed values of phases $\phi_{1,2}$. This is often the case in practice due to approximate calculation of phases from observed signals. In both cases mentioned, LS estimates of the coefficients can be biased [31], and even more so for all the other considered estimators. To make sure of the applicability of the suggested estimators under moderate violation of the assumptions (C1)–(C4) according to both the scenarios is possible in numerical experiment.

III. NUMERICAL EXAMPLES

A. Procedure of numerical investigation

Properties of the suggested estimators $\hat{\gamma}_{1,2}$ and $\hat{\delta}$ and reliability of conclusions about coupling presence and directionality are investigated using Monte Carlo simulation of time realizations of stochastic differential equations. To obtain the time series, original equations are integrated numerically with the aid of the Euler technique and integration step $h=0.01\pi$. Initial phases of subsystems for each realization are random numbers $\phi_1(0),\phi_2(0)$ distributed uniformly on the interval $[0,2\pi]$. To simulate the influence of Gaussian white noise $\xi_{1,2}$, we use the generator of pseudorandom numbers realized in the subroutine DRNNOR of the library IMSL.

For each of the considered examples we carry out 1000 experiments, i.e., simulate 1000 time realizations (1000 pairs of scalar time series) with the length $N_\phi=10^3$. The estimators $\hat{c}_{1,2}$, $\hat{d}$, $\hat{\gamma}_{1,2}$, and $\hat{\delta}$ are evaluated from each of them. Using the obtained sets of values, we construct histograms, compute mean and variance of each estimator, and count the number of correct and wrong conclusions about coupling presence and direction.

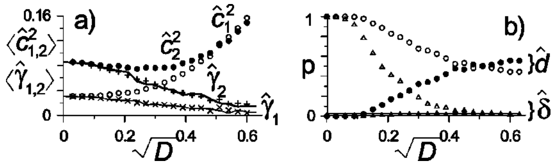

FIG. 3. Results of coupling estimation for system (2) with nonlinear functions $f_{1,2}$, $\Delta t = 20h = 0.2\pi$, and $\tau = 10$ (Sec. III B). (a) Ensemble averages of $\hat{c}_1^2$ and $\hat{c}_2^2$ (open and filled circles, respectively), of $\hat{\gamma}_1$ and $\hat{\gamma}_2$ ($\times$ and $+$, respectively), and values $\hat{c}_1'^2$ and $\hat{c}_2'^2$ calculated from a long time series with $N = 2\times10^5$ (thinner and thicker solid lines, respectively) vs noise level. (b) Relative frequencies of correct and false conclusions about coupling direction (for 1000 simulated time realizations) based on $\hat{d}$ (open and filled circles, respectively) and $\hat{\delta}$ (open and filled triangles, respectively) vs noise level. The level of $p = 0.025$ is shown with a solid line.

The values of τ are chosen so as to provide $\tau\Delta t \approx \min(2\pi/\omega_1, 2\pi/\omega_2)$ [11]. Third-order model polynomial F_1 has the same form for all the examples: $L_1 = 17$, $m_2 = 1, n_2 = 0$, $m_4 = 2, n_4 = 0$, $m_6 = 3, n_6 = 0$, $m_8 = 0, n_8 = 1$, $m_{10} = 0, n_{10} = 2$, $m_{12} = 0, n_{12} = 3$, $m_{14} = 1, n_{14} = -1$, $m_{16} = 1, n_{16} = 1$. The same is true for F_2. Such a form is sufficiently parsimonious and flexible to describe some nontrivial nonlinearities [10].

B. Nonlinearity of original system

An object of investigation is system (2) with $f_1 = 0.03\sin(\phi_2 - \phi_1)$, $f_2 = 0.05\sin(\phi_1 - \phi_2)$, $\omega_1 = 1.1$, $\omega_2 = 0.9$, and $D_1 = D_2 = D$. Equations (3) cannot be derived explicitly for this system. The assumptions (C1)–(C4) are not fulfilled due to the presence of nonlinear functions $f_{1,2}$.

In Fig. 3(a) the mean values $\langle \hat{c}_{1,2} \rangle$ and $\langle \hat{\gamma}_{1,2} \rangle$ (angle brackets denote averaging over the ensemble of 1000 computed values) versus the noise level $\sqrt{D}$ are shown. The values of $(c_1^0)^2$ are unknown here, but "almost true" values $\hat{c}_{1,2}'^2$ (computed from a very long time series with $N_\phi = 2\times10^5$ and, therefore, almost equal to $\hat{\gamma}_{1,2}'$) are shown instead of them. The results of the calculations show that $\hat{c}_{1,2}^2$ exhibit greater bias for stronger noise ($\hat{c}_{1,2}^2$ are 20 times greater than $\hat{c}_{1,2}'^2$ at $\sqrt{D} = 0.6$) while $\hat{\gamma}_{1,2}$ are practically unbiased for any noise level.

Relative numbers of correct (i.e., $1\to2$) and false (i.e., $2\to1$) conclusions about coupling direction are shown in Fig. 3(b). At large noise $\sqrt{D} = 0.6$, one draws false conclusions in more than half of all 10^3 experiments with the aid of the estimator $\hat{d}$. For the same noise level, the relative number of false conclusions drawn with the aid of $\hat{\delta}$, i.e., by checking whether condition (14) or (15) is fulfilled, is equal to 0.02; the relative number of correct conclusions equals 0.024; the former number corresponds well to the expected 2.5% probability of errors. For the rest of the realizations, "cautious" conclusions that it is impossible to state something definite about coupling direction are drawn. In other words, the time series with $N_\phi = 10^3$ is too short for reliable determination of coupling direction at this noise level and the estimator $\hat{\delta}$ diagnoses such a situation.

As the noise level decreases, the number of false conclusions also decreases and the number of correct conclusions rises for both $\hat{d}$ and $\hat{\delta}$. But for the intermediate noise levels $\sqrt{D} = 0.1$–0.3 the number of false conclusions for $\hat{d}$ is still big—about 10–30 %. If $\hat{\delta}$ is used, the frequency of erroneous conclusions is *always* not greater than 0.025.

At weak noise, e.g., $\sqrt{D} = 0.03$, both approaches give a correct conclusion about coupling direction for every experiment. At that, $\hat{\delta}$ is provided with a very narrow confidence interval that diagnoses high reliability of the conclusion.

It follows that at a certain noise level the time series with $N_\phi = 10^3$ becomes sufficiently long for reliable estimation of coupling direction. Thus, the use of $\hat{\delta}$ gives the relative number of correct conclusions greater than 0.95 at $\sqrt{D} \leqslant 0.06$. That is, the time series with $N_\phi = 10^3$ becomes sufficiently long when the value of D decreases ≈ 100 times as compared to $\sqrt{D} = 0.6$. Hence, one can also conclude roughly that for the fixed noise level $\sqrt{D} = 0.6$ the time series becomes sufficiently long if N_ϕ is increased by two orders of magnitude also and becomes $N_\phi = 10^5$. This reasoning makes more precise the terms "long" and "short" time series in the context of our consideration. A concrete value of N_ϕ, separating long and short series, depends on the noise level [and on the difference $(c_2^0)^2 - (c_1^0)^2$ to be resolved].

C. Errors in phases

Finally, let us consider a more complicated and close to reality situation when one observes some signals $x_{1,2}$ rather than observing the phases directly (the latter was implicitly assumed in all the above considerations). Let us take two coupled Van der Pol generators as an object:

$$\ddot{x}_{1,2} = \mu(1 - x_{1,2}^2)\dot{x}_{1,2} - \omega_{1,2}^2 x_{1,2} + b_{1,2}(x_{2,1} - x_{1,2}) + \xi_{1,2}(t), \tag{16}$$

where $\mu = 0.2$, $\omega_1 = 1.02$, $\omega_2 = 0.98$, $b_1 = 0.03$, $b_2 = 0.05$, and $D_1 = D_2 = D$. For this system even Eqs. (2) with Gaussian white noise on the right-hand side can be derived only as asymptotic approximation. But the main difficulty is that the phases should be calculated from the time series of $x_{1,2}$, hence, they are obtained with certain errors.[7] In such a case, LSR may give essentially biased estimates for the coefficients a_i^0 [31]. Therefore, this example represents a more severe test for the suggested approach.

The results of the estimation are shown in Fig. 4. Similar to the previous example of Sec. III B, large biases in $\hat{c}_{1,2}$ are

[7]The length of the analyzed time series is $N_x = 1200$. Phases are calculated using Hilbert transform. A hundred computed values of phases at both edges of the time series are excluded as erroneous. Thus, we obtain the time series of phases with $N_\phi = 10^3$. There are ≈ 20 data points per a characteristic period of oscillations in accordance with the recommendations of Ref. [7].

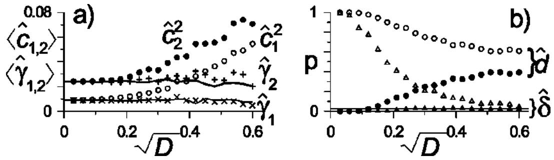

FIG. 4. Results of coupling estimation for system (16) with $\Delta t = 10h = 0.1\pi$ and $\tau = 20$ (Sec. III C). Notations and comments are the same as in Fig. 3.

observed. $\hat{\gamma}_{1,2}$ are again almost unbiased, though a certain difference between $\langle \hat{\gamma}_2 \rangle$ and $\hat{c}'^2_{1,2}$ is still observed at $\sqrt{D} = 0.54, 0.57$. But even this difference is very small compared to the corresponding difference between $\langle \hat{c}_{1,2} \rangle$ and $\hat{c}'^2_{1,2}$. The results concerning numbers of correct and false conclusions about coupling direction [Fig. 4(b)] are the same as in the previous example.

Thus, the estimates $\hat{\gamma}_{1,2}$ and $\hat{\delta}$ turn out to be applicable in both cases (Secs. III B and III C) where the assumptions (C1)–(C4) are violated sufficiently strongly due to different reasons. The main advantages of $\hat{\gamma}_{1,2}$ and $\hat{\delta}$ are as follows.

(1) In the case of very long time series (or very weak noise) they give the same results as Rosenblum's and Pikovsky's estimators and are provided with narrow confidence intervals, which indicates high significance of the results.

(2) In the case of short time series very wide confidence intervals are obtained, which almost excludes false conclusions about coupling presence and directionality.

IV. CONCLUSIONS

In this paper we develop an approach for estimation of intensity and directionality of coupling between two subsystems from short and noisy time series. A crucial requirement is that each of the interacting subsystems should exhibit pronounced main rhythm of oscillations that guarantees the possibility of correct definition of phases and description of their dynamics by an equation of type (2). Besides, the subsystems should not be in a synchronous regime. Under certain additional assumptions (nonlinearity of subsystems and coupling between them are small), *unbiased* estimators for intensity and directionality of interaction *supplied with confidence intervals* are derived. Their applicability in situations when the assumptions are moderately violated is shown in numerical experiments.

Obstacles which limit the applicability of the suggested approach (strong violation of the mentioned assumptions) are the following: (1) contributions of nonlinearity and noise to Eqs. (3) are approximately equal; (2) there are large errors in the values of observables $x_{1,2}$ leading to large errors in the values of their phases.

These situations require, strictly speaking, different approaches, in particular, different techniques of estimating coefficients of model equations [31] and different expressions for the estimators for coupling characteristics. However, it would be difficult to obtain such a universal recipe as that suggested in this paper, since one would need very special assumptions about the properties of noise and form of nonlinearity.

However, our opinion that for a wide range of real-world processes estimators suggested in this work are applicable is quite justified. The reported results should be especially relevant for the analysis of signals of biological origin (electroencephalograms, etc.) where due to nonstationarity it is important to analyze short time series segments and variation of coupling character in real time.

ACKNOWLEDGMENTS

The work was supported by the RFBR (Grant Nos. 02-02-17578 and 03-02-06858), the CRDF (Grant No. REC-006), and the Russian Ministry of Industry, Science and Technology.

APPENDIX A: VARIANCES OF COEFFICIENT ESTIMATES

Bias in, and variance of, $\hat{c}_1^2$ and variance of the suggested estimator $\hat{\gamma}_1$ are related to variances and covariances of the estimates of coefficients $\hat{a}_{1,i}$. Let us formulate simplifying assumptions and derive analytic expressions for variances and covariances of $\hat{a}_{1,i}$.

Let system (2) be the object of investigation. Estimate $\hat{\mathbf{a}}_1$ is obtained via the LSR. Let us rewrite statement (6) in matrix form

$$\|\mathbf{A}\cdot\mathbf{a}_1 - \mathbf{b}\|^2 \to \min, \tag{A1}$$

where $\mathbf{A}$ is a matrix of dimensionality $N \times L_1$ whose elements are $A_{ij} = g_j(\phi_1(t_i), \phi_2(t_i))$, $\mathbf{b}$ is an N-dimensional vector with $b_i = \Delta_1(t_i)$, $\|\cdot\|$ stands for the Euclidean norm. Solution to problem (A1) reduces to the solution to the so-called normal equations [30] and is given by $\hat{\mathbf{a}}_1 = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b}$. Let us assume the following.

(C1) Random matrix $\mathbf{A}^T\mathbf{A}$ of dimensionality $L_1 \times L_1$ can be regarded as constant, that is, independent of random factors: $\mathbf{A}^T\mathbf{A} = E[\mathbf{A}^T\mathbf{A}] = \text{const}$.

(C2) Random quantities $\varepsilon_1(t_i)$ do not depend on phases $\phi_{1,2}(t_i)$ for any time instant t_i.

(C3) Random process $\varepsilon_1(t_i)$, $i = 1,2,\ldots,$ is a sequence of zero mean random quantities distributed identically and normally.

(C4) Random quantities $\varepsilon_1(t_i), \varepsilon_2(t_i)$ are independent of each other.

Validity of (C1) is determined by the number of data points N. For the time series length of $N \approx 10^3$ considered in this work, assumption (C1) is fulfilled within 4% error limit. It can be easily shown that (C2)–(C4) are fulfilled precisely if and only if $f_1 \equiv f_2 \equiv 0$ in Eq. (2). In such a case, one can derive analytically that in Eq. (3) $F_1^0 = \omega_1 \tau \Delta t$ and $\varepsilon_1(t_i) = \int_{t_i}^{t_i+\tau} \xi_1(t)\,dt$ is a random process with variance $\sigma^2_{\varepsilon_1} = D_1 \tau \Delta t$ [see example (8) in Sec. II B]. An expression for the correlation function of ε_1 is derived by the analytic integration:

$$E[\varepsilon_1(t_i)\varepsilon_1(t_j)]=E\left[\int_{t_i}^{t_{i+\tau}}\xi_1(t)dt\int_{t_j}^{t_{j+\tau}}\xi_1(t')dt'\right]$$
$$=\int_{t_i}^{t_{i+\tau}}\int_{t_j}^{t_{j+\tau}}D_1\delta(t-t')dt\,dt'$$
$$=\begin{cases}\sigma_{\varepsilon_1}^2(1-|i-j|/\tau), & |i-j|<\tau\\ 0, & |i-j|\geqslant\tau.\end{cases} \quad \text{(A2)}$$

As is known from the theory of statistical estimation [30], it follows from the conditions (C1)–(C3) that (1) estimates of coefficients are unbiased, i.e., $E[\hat{\mathbf{a}}_1]=\mathbf{a}_1^0$; (2) vector $\hat{\mathbf{a}}_1$ is distributed according to L_1-dimensional Gaussian law; (3) covariation matrix of the components of $\hat{\mathbf{a}}_1$ is given by

$$E[\hat{\mathbf{a}}\cdot\hat{\mathbf{a}}^T]=(\mathbf{A}^T\mathbf{A})^{-1}\cdot E[\mathbf{A}^T\mathbf{e}_1\mathbf{e}_1^T\mathbf{A}]\cdot(\mathbf{A}^T\mathbf{A})^{-1}, \quad \text{(A3)}$$

where $\mathbf{e}_1$ is an N-dimensional vector with components $\varepsilon_1(t_1),\varepsilon_1(t_2),\ldots,\varepsilon_1(t_N)$. Diagonal elements of the matrix $E[\hat{\mathbf{a}}\cdot\hat{\mathbf{a}}^T]$ are variances of $\hat{a}_{1,i}$, other elements are covariances of $\hat{a}_{1,i}$ and $\hat{a}_{1,j}$ for $i\neq j$. Let us derive expressions for them.

Note, first, that under the assumptions (C1)–(C4) the observed values of wrapped phases $(\phi_1(t_i)\mathrm{mod}2\pi,\phi_2(t_i)\mathrm{mod}2\pi)$, $i=1,\ldots,N$, are distributed approximately uniformly in the square $[0,2\pi]\times[0,2\pi]$. Trigonometric monomials $g_i(\phi_1,\phi_2)$ are orthogonal on this set. Hence, the matrix $\mathbf{A}^T\mathbf{A}$ is diagonal [within a certain error limit determined by the violation of (C1)]:

$$\mathbf{A}^T\mathbf{A}=\frac{N}{2}\,\mathrm{diag}(2,1,1,\ldots,1). \quad \text{(A4)}$$

The first case: $\tau=1$. Equation (A2) implies $E[\varepsilon_1(t_i)\varepsilon_1(t_j)]=\sigma_{\varepsilon_1}^2\delta_{ij}$, i.e., subsequent values of ε_1 are uncorrelated. Equations (A3) and (A4) imply $E[\hat{\mathbf{a}}\cdot\hat{\mathbf{a}}^T]=(2\sigma_{\varepsilon_1}^2/N)\mathrm{diag}(\frac{1}{2},1,1,\ldots,1)$, i.e., estimates of coefficients are uncorrelated. Their variances are $\sigma_{\hat{a}_{1,1}}^2=\sigma_{\varepsilon_1}^2/N$, and $\sigma_{\hat{a}_{1,i}}^2=2\sigma_{\varepsilon_1}^2/N$ for $i>1$.

The second case: $\tau>1$. Subsequent values of ε_1 are correlated. By performing transformations similar to those presented in Eq. (A2), remembering about Gaussianity of $\xi_{1,2}$, and taking some definite integrals, one derives that coefficient estimates are again uncorrelated and their variances are given by

$$\sigma_{\hat{a}_{1,i}}^2=\frac{2\sigma_{\varepsilon_1}^2}{N}\left[1+2\sum_{j=1}^{\tau-1}\left(1-\frac{j}{\tau}\right)\right.$$
$$\left.\times\cos[(m_i\omega_1+n_i\omega_2)j\Delta t]e^{-j(m_i^2\sigma_{\varepsilon_1}^2+n_i^2\sigma_{\varepsilon_2}^2)/2\tau}\right] \quad \text{(A5)}$$

for $i>1$. Note that Eq. (A5) is valid for $\tau=1$ as well. To derive expressions for estimates of $\sigma_{\hat{a}_{1,i}}^2$, one may replace *a priori* unknown quantities $\sigma_{\varepsilon_{1,2}}^2$ and $\omega_{1,2}$ in Eq. (A5) by their estimates. Let us insert into Eq. (A5) instead of $\sigma_{\varepsilon_{1,2}}^2$ their estimates $\hat{\sigma}_{\varepsilon_{1,2}}^2$ given by[8]

$$\hat{\sigma}_{\varepsilon_{1,2}}^2=\hat{\sigma}_{\Delta_{1,2}}^2=\frac{1}{N-1}\sum_{i=1}^{N}\left(\Delta_{1,2}(t_i)-\frac{1}{N}\sum_{i=1}^{N}\Delta_{1,2}(t_i)\right)^2. \quad \text{(A6)}$$

The quantities $\hat{a}_{1,1}/(\tau\Delta t)$ and $\hat{a}_{2,1}/(\tau\Delta t)$ can be inserted instead of ω_1 and ω_2. One derives finally (for $i>1$)

$$\hat{\sigma}_{\hat{a}_{1,i}}^2=\frac{2\hat{\sigma}_{\varepsilon_1}^2}{N}\left[1+2\sum_{j=1}^{\tau-1}\left(1-\frac{j}{\tau}\right)\right.$$
$$\left.\times\cos[(m_i\hat{a}_{1,1}+n_i\hat{a}_{2,1})j/\tau]e^{-j(m_i^2\hat{\sigma}_{\varepsilon_1}^2+n_i^2\hat{\sigma}_{\varepsilon_2}^2)/2\tau}\right]. \quad \text{(A7)}$$

Further, we express all other estimates in terms of the derived unbiased estimates $\hat{\sigma}_{\hat{a}_{1,i}}^2$.

APPENDIX B: VARIANCE OF $\hat{\gamma}_1$ AND CONFIDENCE INTERVAL ESTIMATION

In Sec. II C an expression (10) for $\hat{\gamma}_1$ [an unbiased estimator for $(c_1^0)^2$] is presented. From the theoretical point of view, its variance $\sigma_{\hat{\gamma}_1}^2$ is expressed in terms of variances and covariances of the estimates $\hat{a}_{1,i}$ according to expression (11). Let us derive an estimator for $\sigma_{\hat{\gamma}_1}^2$.

First, noncorrelatedness of Gaussian distributed estimates $\hat{a}_{1,i}$ and $\hat{a}_{1,j}$ $(i\neq j)$ implies their statistical independence. Hence, their squares $\hat{a}_{1,i}^2$ and $\hat{a}_{1,j}^2$ are also independent of each other. Then, taking in Eq. (11) $\mathrm{cov}(\hat{a}_{1,i}^2,\hat{a}_{1,j}^2)=0$ for $i\neq j$, one derives

$$\sigma_{\hat{\gamma}_1}^2=\sum_{i=1}^{L_1}n_i^4\sigma_{\hat{a}_{1,i}^2}^2. \quad \text{(B1)}$$

Second, by using the definition of variance, Gaussianity of $\hat{a}_{1,i}$, and taking corresponding definite integral, one derives an expression for the variance of $\hat{a}_{1,i}^2$:

$$\sigma_{\hat{a}_{1,i}^2}^2=2\sigma_{\hat{a}_{1,i}}^4+4(a_{1,i}^0)^2\sigma_{\hat{a}_{1,i}}^2. \quad \text{(B2)}$$

Unbiased estimator for $\sigma_{\hat{a}_{1,i}^2}^2$ would be $2\hat{\sigma}_{\hat{a}_{1,i}}^4+4(\hat{a}_{1,i}^2-\hat{\sigma}_{\hat{a}_{1,i}}^2)\hat{\sigma}_{\hat{a}_{1,i}}^2$. But the use of this expression would often yield

---

[8] $\hat{\sigma}_{\varepsilon_1}^2$ in Eq. (A6) is a somewhat understated estimate of the variance. However, under the usual condition $\tau\ll N$, bias is very small and can be neglected. Therefore, we do not present a more accurate and much more cumbersome expression for a strictly unbiased estimate here.

estimates of $\sigma^2_{\hat{a}^2_{1,i}}$ close to, and even less than, zero. It would lead often to spurious inference of high significance of obtained values $\hat{a}^2_{1,i}$. Therefore, we propose to use a bit overstated estimate

$$\hat{\sigma}^2_{\hat{a}^2_{1,i}}=\begin{cases}2\hat{\sigma}^4_{\hat{a}_{1,i}}+4(\hat{a}^2_{1,i}-\hat{\sigma}^2_{\hat{a}_{1,i}})\hat{\sigma}^2_{\hat{a}_{1,i}}, & \hat{a}^2_{1,i}-\hat{\sigma}^2_{\hat{a}_{1,i}}\geqslant 0,\\ 2\hat{\sigma}^4_{\hat{a}_{1,i}} & \text{otherwise.}\end{cases} \tag{B3}$$

This estimate corresponds to a cautious strategy and excludes frequent spurious conclusions about the presence of coupling. Finally, we propose to use the quantity given by expression (12) as an estimator for $\sigma^2_{\hat{\gamma}_1}$:

$$\hat{\sigma}^2_{\hat{\gamma}_1}=\begin{cases}\sum_{i=1}^{L_1} n_i^4\hat{\sigma}^2_{\hat{a}^2_{1,i}}, & \hat{\gamma}_1\geqslant 5\left(\sum_{i=1}^{L_1} n_i^4\hat{\sigma}^2_{\hat{a}^2_{1,i}}\right)\\ \frac{1}{2}\sum_{i=1}^{L_1} n_i^4\hat{\sigma}^2_{\hat{a}^2_{1,i}} & \text{otherwise.}\end{cases}$$

Such a choice is determined by the following circumstances. For large c_1^0, the top line is a "good" estimate. For $c_1^0\approx 0$, the top line gives an estimate, which is twice as large as the true value on average (this statement is based on our experience with numerical examples). Therefore, it is reasonable to divide the top line by a factor of 2. So, the proposed combination gives a widely acceptable trade-off.

An expression for the confidence interval for $(c_1^0)^2$ depends on the form of the distribution law of $\hat{\gamma}_1$. If $\hat{\gamma}_1$ were distributed normally (this is the case only for very large number of coefficients $a_{1,i}$ with $n_i\neq 0$), then $\hat{\gamma}_1\pm 1.96\sigma_{\hat{\gamma}_1}$ would be a 95% confidence interval and it could be readily estimated as $\hat{\gamma}_1\pm 2\hat{\sigma}_{\hat{\gamma}_1}$. However, the distribution of $\hat{\gamma}_1$ is asymmetric for a typical number of coefficients $a_{1,i}$ with $n_i\neq 0$ (it is about ten, see Sec. III and Ref. [10]). To derive a generally applicable expression for confidence interval analytically is impossible since distributions of $\hat{\gamma}_1$ are different for different numbers of coefficients. But one can expect that in any case it is not essentially different from the expression for Gaussian distribution since the distribution of $\hat{\gamma}_1$ is unimodal, even though a bit skewed. We searched for an estimate of 95% confidence interval in the form $[\hat{\gamma}_1-\alpha\hat{\sigma}_{\hat{\gamma}_1},\hat{\gamma}_1+\beta\hat{\sigma}_{\hat{\gamma}_1}]$. Constants α and β were chosen empirically. As a result of numerical simulations, we obtained the values $\alpha=1.6$, $\beta=1.8$.

As for the quantity $\hat{\delta}$, if $\hat{\gamma}_1$ and $\hat{\gamma}_2$ are uncorrelated, which is guaranteed by the conditions (C1) and (C4), then quite a good estimator for its variance is $\hat{\sigma}^2_{\hat{\delta}}=\hat{\sigma}^2_{\hat{\gamma}_1}+\hat{\sigma}^2_{\hat{\gamma}_2}$. As a rule, the distribution of $\hat{\delta}$ is more symmetric than the distributions of $\hat{\gamma}_{1,2}$. Therefore, we searched for an estimate of 95% confidence interval for δ^0 in a symmetric form $\hat{\delta}\pm\alpha\hat{\sigma}_{\hat{\delta}}$. Again, our numerical experiments resulted in the value of $\alpha=1.6$.

[1] N.F. Rulkov, M.M. Sushchik, L.S. Tsimring, and H.D.I. Abarbanel, Phys. Rev. E **51**, 980 (1995).
[2] L. Kocarev and U. Parlitz, Phys. Rev. Lett. **76**, 1816 (1996).
[3] M.G. Rosenblum, A.S. Pikovsky, and J. Kurths, Phys. Rev. Lett. **76**, 1804 (1996).
[4] U. Parlitz, L. Junge, W. Lauterborn, and L. Kocarev, Phys. Rev. E **54**, 2115 (1996).
[5] L.M. Pecora, T.L. Carroll, G.A. Johnson, D.J. Mar, and J.F. Heagy, Chaos **7**, 520 (1997).
[6] A.S. Pikovsky, M.G. Rosenblum, and J. Kurths, Int. J. Bifurcation Chaos Appl. Sci. Eng. **10**, 2291 (2000).
[7] M.G. Rosenblum, A.S. Pikovsky, J. Kurths, C. Schafer, and P.A. Tass, in *Neuro-informatics*, edited by F. Moss and S. Gielen, Handbook of Biological Physics Vol. 4 (Elsevier Science, New York, 2000), p. 279.
[8] Z. Zheng and G. Hu, Phys. Rev. E **62**, 7882 (2000).
[9] A. Pikovsky, M. Rosenblum, and J. Kurths, *Synchronization, A Universal Concept in Nonlinear Sciences* (Cambridge University Press, Cambridge, 2001).
[10] M.G. Rosenblum and A.S. Pikovsky, Phys. Rev. E **64**, 045202 (2001).
[11] M.G. Rosenblum, L. Cimponeriu, A. Bezerianos, A. Patzak, and R. Mrowka, Phys. Rev. E **65**, 041909 (2002).
[12] R. Quian Quiroga, A. Kraskov, T. Kreuz, and P. Grassberger, Phys. Rev. E **65**, 041903 (2002).
[13] R.Q. Quiroga, T. Kreuz, and P. Grassberger, Phys. Rev. E **66**, 041904 (2002).
[14] B. Bezruchko, V. Ponomarenko, M.G. Rosenblum, and A.S. Pikovsky, Chaos **13**, 179 (2003).
[15] C. Schäfer, M.G. Rosenblum, J. Kurths, and H.-H. Abel, Nature (London) **392**, 239 (1998).
[16] H. Seidel and H. Herzel, IEEE Eng. Med. Biol. Mag. **17**, 54 (1998).
[17] A. Stefanovska and M. Bracic, Contemp. Phys. **40**, 31 (1999).
[18] M. Bracic-Lotric and A. Stefanovska, Physica A **283**, 451 (2000).
[19] N.B. Janson, A.G. Balanov, V.S. Anishchenko, and P.V.E. McClintock, Phys. Rev. Lett. **86**, 1749 (2001).
[20] N.B. Janson, A.G. Balanov, V.S. Anishchenko, and P.V.E. McClintock, Phys. Rev. E **65**, 036212 (2002).
[21] P. Tass, M.G. Rosenblum, J. Weule, J. Kurths, A. Pikovsky, J. Volkmann, A. Schnitzler, and H.-J. Freund, Phys. Rev. Lett. **81**, 3291 (1998).
[22] J. Arnhold, P. Grassberger, K. Lehnertz, and C. E. Elger, in *Chaos in Brain?*, edited by K. Lehnertz, J. Arnhold, P. Grassberger, and C.E. Elger (World Scientific, Singapore, 2000), p. 325.
[23] E. Rodriguez *et al.*, Nature (London) **397**, 430 (1999).
[24] F. Mormann, K. Lehnertz, P. David, and C.E. Elger, Physica D **144**, 358 (2000).
[25] R.G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C.E. Elger, Phys. Rev. E **64**, 061907 (2001).
[26] R.G. Andrzejak, G. Widman, K. Lehnertz, C. Rieke, P. David,

and C.E. Elger, Epilepsy Res. **44**, 129 (2001).
[27] R.G. Andrzejak, F. Mormann, T. Kreuz, C. Rieke, A. Kraskov, C.E. Elger, and K. Lehnertz, Phys. Rev. E **67**, 010901(R) (2003).
[28] F. Mormann, T. Kreuz, R.G. Andrzejak, P. David, K. Lehnertz, and C.E. Elger, Epilepsy Res. **53**, 173 (2003).
[29] A.Ya. Kaplan, Usp. Fiziol. Nauk **29**, 35 (1998).
[30] G.A. Korn and T.M. Korn, *Mathematical Handbook for Scientists and Engineers* (McGraw-Hill, New York, 1968); G. Casella and R.L. Berger, *Statistical Inference* (Wadsworth & Brooks/Cole, 1990).
[31] P.E. McSharry and L.A. Smith, Phys. Rev. Lett. **83**, 4285 (1999).