Estimating causal dependencies in networks of nonlinear stochastic dynamical systems

Linda Sommerlade,[1,2,3,*] Michael Eichler,[4] Michael Jachan,[1,3] Kathrin Henschel,[1,2,3] Jens Timmer,[1,2,3,5] and Björn Schelter[1,2,3]

[1]*Department of Physics, University of Freiburg, Hermann-Herder-Str. 3, 79104 Freiburg, Germany*

[2]*Bernstein Center for Computational Neuroscience, University of Freiburg, Hansastr. 9A, 79104 Freiburg, Germany*

[3]*FDM, Freiburg Center for Data Analysis and Modeling, University of Freiburg, Eckerstr. 1, 79104 Freiburg, Germany*

[4]*Department of Quantitative Economics, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands*

[5]*Freiburg Institute for Advanced Studies (FRIAS), University of Freiburg, Albertstr. 19, 79104 Freiburg, Germany*

(Received 21 July 2009; published 30 November 2009)

The inference of causal interaction structures in multivariate systems enables a deeper understanding of the investigated network. Analyzing nonlinear systems using partial directed coherence requires high model orders of the underlying vector-autoregressive process. We present a method to overcome the drawbacks caused by the high model orders. We calculate the corresponding statistics and provide a significance level. The performance is illustrated by means of model systems and in an application to neurological data.

DOI: 10.1103/PhysRevE.80.051128 PACS number(s): 02.50.Sk, 05.10.−a, 05.45.Xt, 87.19.xp

I. INTRODUCTION

Dynamical networks are ubiquitous in many fields of research. The *direct problem* to infer the behavior of networks consisting of coupled self-sustained possibly stochastic or chaotic oscillators has been thoroughly addressed in the literature, see for instance [1]. The *inverse problem*, i.e., the inference of the network structure from measured signals, including means to statistically evaluate the results, in contrast, is so far mainly underaddressed. This is basically due to the twofold challenge to infer the directionality of the coupling as well as to distinguish direct and indirect interactions in the networks.

For linear systems, several analysis techniques exist that allow differentiating direct and indirect interactions. Partial coherence, for instance, has been suggested to this aim [2]. To a certain extent partial coherence can successfully be applied to signals originating from nonlinear oscillatory networks, too. A nonlinear extension of partial coherence analysis, the so-called partial phase synchronization analysis, enables a more widespread investigation of direct and indirect interactions in oscillatory networks [3].

Concerning the direction of coupling in nonlinear dynamical networks, bivariate analysis techniques dominate the literature. Albeit some techniques are mainly able to infer the pronounced direction of coupling [4,5], others also provide information about mutually coupled oscillators [6–8]. A multivariate extension of these techniques is however missing with the exception of [9]. A linear Granger-causality measure is used therein to cope with nonlinearities based on a delay embedding reconstruction of the phase-space attractors. For linear systems methods based on the common sense idea that causes precede their effects in time, formalized by the Granger causality [10], have been suggested, which also enable multivariate analyses [11–16]. Granger causality is used to determine causal influences mainly utilizing linear vector-autoregressive models. Partial directed coherence has been introduced for inference of linear Granger causality in the frequency domain [17].

The applicability of partial directed coherence to nonlinear systems is hampered by the fact that high model orders are required to capture the second-order characteristics of the underlying dynamics sufficiently well [18]. These high model orders introduce strong fluctuations to partial coherence spectra. These strong fluctuations hinder the ability to reveal the true underlying network structure since the coupling has to be rather high to become clearly significant as the significance level for high model orders is rather high. *False negative* conclusions about the network structure are, thus, rather likely when applying partial directed coherence. Additionally, the fluctuations eventually lead to *false positive* conclusions about the network structure as demonstrated below.

To overcome these limitations of partial directed coherence analysis, we introduce the concept of smoothing as a regularization method. Especially for statistical considerations, smoothing partial directed coherence is not straightforward. An analytic significance level as it was derived for ordinary partial directed coherence [18] is calculated here. The smoothed partial directed coherence together with this significance level allows disentangling the underlying network structure of coupled nonlinear possibly chaotic oscillators. Compared to the extended Granger-causality index [9], which uses a delay embedding reconstruction of the phase-space attractors to estimate Granger causality in the nonlinear case, our approach does not require a delay embedding, which is particularly challenging when analyzing noise data.

The paper is structured as follows. In Sec. II, partial directed coherence and its significance level are summarized. The smoothed partial directed coherence is introduced and the corresponding significance level is derived. The performance of this smoothed partial directed coherence is demonstrated by linear and nonlinear stochastic model systems with different dynamic behavior in Sec. III. An exemplary application to electroencephalographic (EEG) and electromyographic (EMG) recordings of a patient suffering from essential tremor is presented in Sec. IV.

II. SMOOTHED PARTIAL DIRECTED COHERENCE

To motivate smoothed partial directed coherence, first partial directed coherence and its pointwise significance level

*linda.sommerlade@fdm.uni-freiburg.de

are summarized. Then, we introduce the smoothing of the partial directed coherence and derive the corresponding statistics.

A. Partial directed coherence

Using the concept of Granger causality [10], which is based on the common sense conception that causes precede their effects in time, the partial directed coherence $|\pi_{i\leftarrow j}(\omega)|$ for an n-dimensional vector-autoregressive process of order p (VAR[p] process),

$$\mathbf{x}(t) = \sum_{r=1}^{p} \mathbf{a}(r)\mathbf{x}(t-r) + \boldsymbol{\varepsilon}(t), \tag{1}$$

where $\boldsymbol{\varepsilon}(t)$ is a multivariate Gaussian white-noise process with covariance matrix $\boldsymbol{\Sigma}$, is defined as [17]

$$|\pi_{i\leftarrow j}(\omega)| = \frac{|\mathbf{A}_{ij}(\omega)|}{\sqrt{\sum_m |\mathbf{A}_{mj}(\omega)|^2}} \tag{2}$$

with

$$\mathbf{A}(\omega) = I - \sum_{r=1}^{p} \mathbf{a}(r) e^{-\mathrm{i}\omega r}. \tag{3}$$

The $\mathbf{a}(r)$ are the $n\times n$ coefficient matrices of the vector-autoregressive model of order p.

The partial directed coherence $|\pi_{i\leftarrow j}(\omega)|$ provides a measure for the directed—strictly speaking—linear influence of $x_j(t)$ on $x_i(t)$ at frequency ω. It is estimated by fitting an n-dimensional VAR[p] model to the data and using Eqs. (2) and (3) with the parameter estimates $\hat{\mathbf{a}}_{ij}(r)$ substituted for the true coefficients $\mathbf{a}_{ij}(r)$.

The pointwise α-significance level for the partial directed coherence $|\pi_{i\leftarrow j}(\omega)|$ is given by [18]

$$\left(\frac{\hat{C}_{ij}(\omega)\chi^2_{1,1-\alpha}}{N\sum_m |\hat{\mathbf{A}}_{mj}(\omega)|^2}\right)^{1/2}, \tag{4}$$

where $\chi^2_{1,1-\alpha}$ is the $1-\alpha$ quantile of the χ^2 distribution with 1 degree of freedom. $\hat{C}_{ij}(\omega)$ is an estimate of

$$C_{ij}(\omega) = \boldsymbol{\Sigma}_{ii}\left\{\sum_{l,m=1}^{p} \mathbf{H}_{jj}(l,m)[\cos(l\omega)\cos(m\omega) + \sin(l\omega)\sin(m\omega)]\right\}. \tag{5}$$

Here, $\mathbf{H}_{jj}(l,m)$ are entries of the inverse $\mathbf{H}=\mathbf{R}^{-1}$ of the covariance matrix $\mathbf{R}$ of the VAR[p]-process $\mathbf{x}(t)$, which is composed of the entries

$$\mathbf{R}_{ij}(l,m) = \mathrm{cov}[x_i(t-l), x_j(t-m)] \tag{6}$$

for $i,j=1,\ldots,n$ and $l,m=1,\ldots,p$ [19]. The $\mathbf{H}_{jj}(l,m)$ are related to the estimates $\hat{\mathbf{a}}_{ij}(r)$ of the VAR[p] coefficients $\mathbf{a}_{ij}(r)$ by

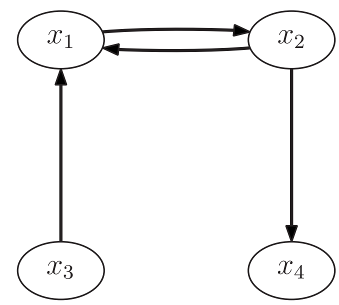

FIG. 1. Graph showing the interaction of the system of four coupled stochastic van der Pol oscillators [Eq. (8)].

$$\lim_{N\to\infty} N\,\mathrm{cov}(\hat{\mathbf{a}}_{ij}(l), \hat{\mathbf{a}}_{ij}(m)) = \boldsymbol{\Sigma}_{ii}\mathbf{H}_{jj}(l,m) \tag{7}$$

with the covariance matrix $\boldsymbol{\Sigma}$ of the white-noise input $\boldsymbol{\varepsilon}(t)$ [Eq. (1)].

We like to emphasize here that the significance level depends on the order of the vector-autoregressive process as the number of addends for $C_{ij}(\omega)$ changes. Higher model orders lead to higher significance levels which in turn indicate that for higher model orders the ability to detect weak coupling decreases.

B. System of four coupled stochastic van der Pol oscillators

To illustrate another limitation of partial directed coherence when applied to a network of stochastic nonlinear oscillators, four coupled stochastic van der Pol oscillators

$$\ddot{x}_i = \mu(1-x_i^2)\dot{x}_i - \omega_i^2 x_i + \sum_{j\neq i} \gamma_{ij}(x_j - x_i) + \sigma\varepsilon_i \tag{8}$$

are investigated with $i,j=1,2,3,4$, $N=10\,000$ data points, $\omega_1=1.5$, $\omega_2=1.48$, $\omega_3=1.53$, $\omega_4=1.44$, $\mu=5$, $\sigma=1.5$, and Gaussian distributed white noise ε_i with unit variance. The data were simulated using the Euler method with an integration time step of 1 ms and a samplingrate of 2 Hz. The coupling parameters are set to zero except for $\gamma_{12}=\gamma_{21}=\gamma_{13}=\gamma_{42}=0.4$ for which the oscillators synchronize.

The causal influences are summarized in the graph in Fig. 1. For partial directed coherence analysis, an autoregressive process of order $p=100$ is chosen in order to model the second-order properties of the nonlinear system sufficiently well. The estimated partial directed coherence as well as the ordinary autospectra are shown in Fig. 2. The corresponding 5%-significance levels are indicated by the almost horizontal red (gray) lines. All direct causal influences are detected as $|\pi_{2\leftarrow1}|$, $|\pi_{1\leftarrow2}|$, $|\pi_{1\leftarrow3}|$, and $|\pi_{4\leftarrow2}|$ are significant at the corresponding oscillation frequencies. But $|\pi_{4\leftarrow3}|$ is also significant, Fig. 2; a false detection. The expected number of false positive crossings of the significance level for partial directed coherence is $p\cdot\alpha$, i.e., the order of the process times the significance level. At first glance, this might appear to be in contrast to ordinary coherence, where the expected number of false positives is $m\cdot\alpha$, where m is the number of independent entries in the coherence spectrum—typically fixed by the spectral estimator. Here, the order p of the VAR process determines the number of independent entries in the

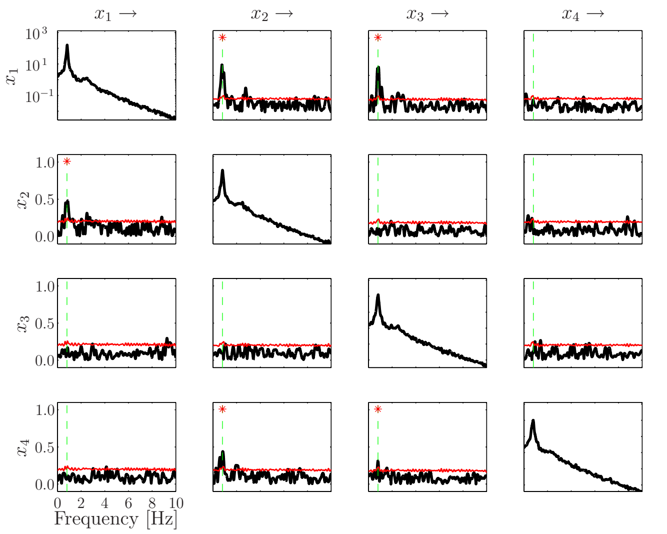

FIG. 2. (Color online) Partial directed coherence for the system of four coupled stochastic van der Pol oscillators [Eq. (8)]. 5%-significance levels are depicted by the almost horizontal red (gray) lines. Autospectra are shown on the diagonal. The peak frequency is marked by a dashed green line, significant values are highlighted by a red asterisk.

partial directed coherence estimate. For example for a VAR[1]-process there is only a single off-diagonal parameter that essentially determines the entire partial directed coherence spectrum. The order determines the covariance structure of the partial directed coherence in a nonintuitive and difficult to derive way. This indeed also hampers the interpretability of partial directed coherence once it is supposed to be evaluated not just at one single predefined frequency.

The dependence of false positives on the order of the process, however, has severe consequences for partial directed coherence estimates. By increasing the order of the process, it becomes more likely to obtain false positives. For nonlinear systems high model orders are required to capture the properties of the nonlinear system sufficiently well to enable reasonable conclusions.

Consequently, it would be desirable to effectively decrease the model order without loosing the ability to reveal the network structure in synchronizing possibly chaotic systems. A certain regularization strategy is mandatory to achieve both, avoiding false positive conclusions about the coupling as well as increasing the ability to detect couplings that are actually present. A higher model order decreases the correlation between the estimated partial directed coherence at different frequencies whereas smoothing increases this correlation. Thus, smoothing changes the correlation between the estimated partial directed coherence at different frequencies in a way similar to a reduction in the model order. Motivated by this, smoothing partial directed coherence by convolution with an appropriate kernel presents an approach to regularize the partial directed coherence.

C. Smoothing

Without loss of generality, for the kernel used to smooth the partial directed coherence, we choose a normalized triangular window

$$K(d) = \begin{cases} \dfrac{1}{h} - \dfrac{|d|}{h^2} & \text{for } |d| \le h \\ 0 & \text{else} \end{cases} \tag{9}$$

of width $(2h+1)$ frequency bins corresponding to a frequency width of $h/N \cdot \Delta f$ where Δf is the sampling frequency.

A convolution of the $\mathbf{A}_{ij}(\omega)$ with this kernel leads to the smoothed $\mathbf{A}_{ij}^{\text{smooth}}(\omega)$, which are used to estimate the smoothed partial directed coherence

$$|\pi_{i\leftarrow j}^{\text{smooth}}(\omega)| = \frac{|\mathbf{A}_{ij}^{\text{smooth}}(\omega)|}{\sqrt{\sum_m |\mathbf{A}_{mj}^{\text{smooth}}(\omega)|^2}}. \tag{10}$$

This naïve approach to smooth the partial directed coherence by convolution with an appropriate kernel cannot be transferred to the significance level due to the covariance structure of $\mathbf{A}(\omega)$, which was actually the reason why the interpretation of significance level crossings is extremely difficult for partial directed coherence once investigating not just a single predefined frequency as described in the previous section. The statistical properties need to be revisited for the case of smoothed $\mathbf{A}_{ij}^{\text{smooth}}(\omega)$.

The convolution in the frequency domain is represented by a multiplication with the inverse Fourier transformation of the kernel in the time domain

$$\mathbf{A}^{\text{smooth}}(\omega) = \mathbf{A}(\omega) * K(d) = I - \mathcal{FT}[\mathbf{a}(r) \cdot k(r)], \tag{11}$$

with

$$k(r) = i\mathcal{FT}(K(d)) = \frac{\sin^2\left(\dfrac{\pi r}{N} h\right)}{h^2 \sin^2\left(\dfrac{\pi r}{N}\right)}. \tag{12}$$

Using Eq. (7) we find

$$\mathbf{H}_{jj}^{\text{smooth}}(l,m) = k(l)k(m)\mathbf{H}_{jj}(l,m), \tag{13}$$

and thus

$$C_{ij}^{\text{smooth}}(\omega) = \mathbf{\Sigma}_{ii}\left\{ \sum_{l,m=1}^{p} k(l)k(m)\mathbf{H}_{jj}(l,m)[\cos(l\omega)\cos(m\omega) + \sin(l\omega)\sin(m\omega)] \right\}, \tag{14}$$

leading to the α-significance level

$$\left(\frac{\hat{C}_{ij}^{\text{smooth}}(\omega)\chi_{1,1-\alpha}^2}{N\sum_m |\hat{\mathbf{A}}_{mj}^{\text{smooth}}(\omega)|^2} \right)^{1/2} \tag{15}$$

for the smoothed partial directed coherence of Eq. (10).

Analog to smoothing, this approach could be interpreted as a different estimator for the VAR coefficients which uses a well chosen weighting for the different orders [Eq. (11)]. In this case $\mathbf{b}(u)=\mathbf{a}(u)\cdot k(u)$ are the weighted coefficients. The weights represent the chosen smoothing. The fitted VAR co-

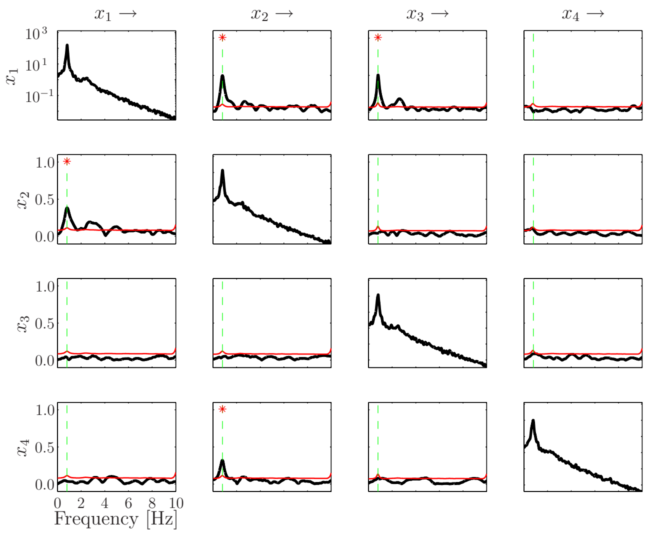

FIG. 3. (Color online) Smoothed partial directed coherence for the system of four coupled stochastic van der Pol oscillators [Eq. (8)]. 5%-significance levels are depicted by red (gray) lines. Autospectra are shown on the diagonal. The peak frequency is marked by a dashed green line, significant values are highlighted by a red asterisk.

efficients are correlated by the weighting. This again might appear counterproductive at a first glance, but the introduced correlation effectively decreases the order of the VAR process. More precisely, an increase in the correlation leads to a decreased number of independent frequency bins. Thus, the Fourier transformation of fewer but different VAR coefficients would be sufficient to achieve the same partial directed coherence spectrum. Hence, an increase in the correlation can be interpreted as an effective decrease of the model order to p_{eff}. This in turn allows a more rigorous interpretation of the results evaluated at predefined frequencies. The number of expected false positive crossings of the significance level for a single spectrum, i.e., in the case that a single realization is analyzed, is $p \cdot \alpha$. Since by smoothing the effective model order is reduced to p_{eff}, this leads to a lower number of expected false positive crossings of the significance level, $p_{\text{eff}} \cdot \alpha < p \cdot \alpha$. As shown in the following simulations, the ability to detect the true interaction structure is increased.

D. System of four coupled stochastic van der Pol oscillators revisited

For the system of four coupled stochastic van der Pol oscillators [Eq. (8)], which was analyzed previously, smoothed partial directed coherence was estimated using the same estimated coefficients as for the partial directed coherence ($p=100$) and a smoothing window width of $h=0.41$ Hz. In Fig. 3 the estimated smoothed partial directed coherence as well as the spectra are shown. The corresponding 5%-significance levels are indicated by the almost horizontal red (gray) lines. Smoothed partial directed coherence correctly detects the causal influences in the van der Pol system as only $|\pi_{2\leftarrow 1}^{\text{smooth}}|$, $|\pi_{1\leftarrow 2}^{\text{smooth}}|$, $|\pi_{1\leftarrow 3}^{\text{smooth}}|$, and $|\pi_{4\leftarrow 2}^{\text{smooth}}|$ are significant at the corresponding oscillation frequencies. As motivated by the heuristic arguments in the previous sections, smoothing improves the correct detection of interaction in nonlinear systems. The former erroneously detected interaction $|\pi_{4\leftarrow 3}|$ between oscillators 3 and 4 is correctly revealed as not being present.

III. SIMULATIONS

In order to analyze the performance of the smoothed partial directed coherence in more detail with particular emphasis on the ability to detect interactions that are actually present, we investigated its statistical performance in terms of power and coverage based on simulated data. Coverage describes the ability of a method to keep to $\alpha \cdot 100\%$ false detections when using an α-significance level. Here, it is of crucial importance that coverage for smoothed partial directed coherence as for partial directed coherence can only be evaluated by generating a number of independent realizations. It is impossible to assess it on the basis of few realizations by just looking at various frequencies as described above.

The term power is used to describe the ability of a method to detect present interactions. For a power analysis a so-called power curve is calculated. To this aim the sensitivity of the method is tested by increasing the coupling strength. For each coupling strength the percentage of detections is measured. A strong power is characterized by a steep increase to 100% already at low coupling strengths.

First, we choose a vector-autoregressive process for which the partial directed coherence was originally developed. Second, a system of two coupled stochastic van der Pol oscillators is analyzed. Third, to demonstrate the performance of smoothed partial directed coherence when applied to a system consisting of chaotic oscillators, it is applied to a system of coupled Rössler oscillators. For each of the systems 100 realizations were simulated with a length of $N=5000$ data points. Transients have been removed. For all simulations a 5%-significance level is used.

A. Vector-autoregressive process

The following system of a two-dimensional vector-autoregressive process of order one

$$x_1(t) = 0.7 \cdot x_1(t-1) + \varepsilon_1(t)$$

$$x_2(t) = \gamma \cdot x_1(t-1) + \varepsilon_2(t) \tag{16}$$

with γ the variable interdependence parameter is simulated with Gaussian distributed white noise $\varepsilon_i(t)$, $i=1,2$ of unit variance. The value of γ is varied between 0 and 0.5 in steps of 0.05. For each γ, the partial directed coherence as well as the smoothed partial directed coherence are estimated, using an order of $p=150$ and a smoothing window width of $h=0.05$ Hz. Partial directed coherence and smoothed partial directed coherence are tested for significance at 0.1 Hz. The percentage of connections detected within the 100 realizations is displayed in dependence on γ, Fig. 4. Figure 4(a) shows the results for the connection from $x_2(t)$ to $x_1(t)$ which

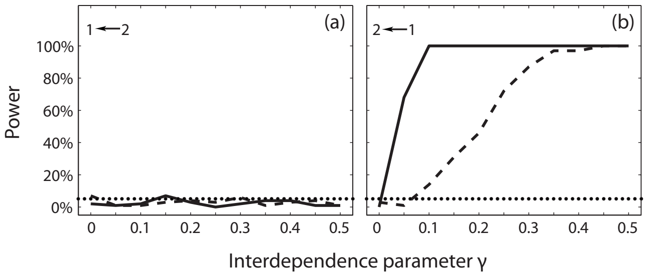

FIG. 4. Power curve for vector-autoregressive processes with increasing interdependence parameter $\gamma=0,\ldots,0.5$ ($p=150$ and $h=0.05$ Hz). Solid line: smoothed partial directed coherence, dashed line: partial directed coherence, dotted line: 5%-significance level.

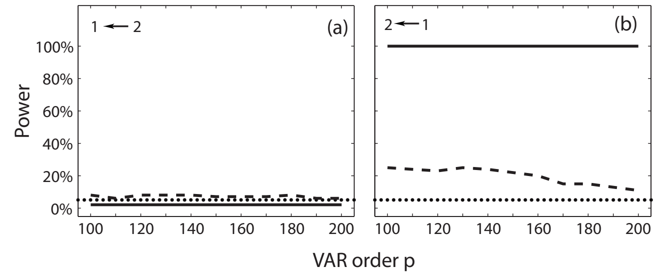

FIG. 6. Power curve for vector-autoregressive processes with increasing order of the autoregressive process $p=100,\ldots,200$ ($\gamma=0.1$ and $h=0.05$ Hz). Solid line: smoothed partial directed coherence, dashed line: partial directed coherence, dotted line: 5%-significance level.

is absent in the simulations. Both methods detect this connection in less than 5% of the realizations for a given 5%-significance level. Also for the direction from $x_1(t)$ to $x_2(t)$, both methods show a good coverage, meaning that if no connection is present ($\gamma=0$), it is pretended in less than 5% of the realizations. In Fig. 4(b) the results of the partial directed coherence (dashed line) increase slower with increasing γ than those of the smoothed partial directed coherence (solid line). In other words the power of the partial directed coherence is increased by smoothing. These results indicate that the smoothed partial directed coherence is superior to the ordinary partial directed coherence as the former is characterized by a higher power while the correct coverage of the partial directed coherence is not changed.

In a second scenario the interdependence parameter is fixed at $\gamma=0.1$. Using an order of $p=150$, smoothed partial directed coherence is estimated for 11 different smoothing window widths ranging from $h=0$ Hz to $h=0.1$ Hz and compared to ordinary partial directed coherence. A power curve is achieved according to the same rules as above and again both method show good coverage Fig. 5(a). Smoothing with increasing window width improves the detection of the true connection from $x_1(t)$ to $x_2(t)$ and thus the smoothed partial directed coherence is found to be superior to the ordinary partial directed coherence [Fig. 5(b)].

Using a smoothing window of $h=0.05$ Hz and an interdependence parameter of $\gamma=0.1$, the performance for different orders $p=100$ to $p=200$ in steps of 10 is analyzed. Results are shown in Fig. 6. For increasing model orders, the performance of the partial directed coherence becomes worse while for the chosen interdependence parameter the smoothed partial directed coherence detects the true connections reliably, Fig. 6(b). As both methods show good coverage, smoothed partial directed coherence again outperforms the ordinary partial directed coherence due to its increased power.

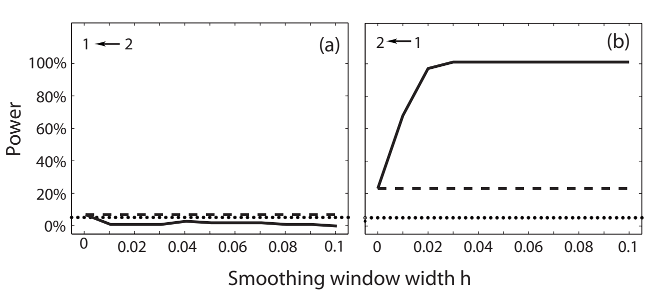

FIG. 5. Power curve for vector-autoregressive processes with increasing smoothing window width $h=0,\ldots,0.1$ Hz ($\gamma=0.1$ and $p=150$). Solid line: smoothed partial directed coherence, dashed line: partial directed coherence, dotted line: 5%-significance level.

B. Two coupled stochastic van der Pol oscillators

To demonstrate the performance of smoothed partial directed coherence when dealing with the situation of nonlinear oscillators and bidirectional coupling, we simulated two symmetrically coupled stochastic van der Pol oscillators

$$\ddot{x}_1 = \mu(1-x_1^2)\dot{x}_1 - \omega_1^2 x_1 + \sigma\varepsilon_1 + \gamma_{12}(x_2 - x_1)$$

$$\ddot{x}_2 = \mu(1-x_2^2)\dot{x}_2 - \omega_2^2 x_2 + \sigma\varepsilon_2 + \gamma_{21}(x_1 - x_2) \qquad (17)$$

with $\omega_1=1.5$, $\omega_2=1.48$, nonlinearity parameter $\mu=5$, $\sigma=1.5$, and Gaussian distributed white noise ε_i, $i=1,2$ of unit variance. The data were simulated using the Euler method with an integration time step of 1 ms and a samplingrate of 2 Hz. The coupling strength $\gamma=\gamma_{12}=\gamma_{21}$ is varied between 0 and 0.3 in steps of 0.05. The connections were analyzed at the peak frequency of the oscillators.

The performance is analyzed for different coupling strengths $\gamma=\gamma_{12}=\gamma_{21}=0,\ldots,0.3$ ($h=0.8$ Hz, $p=150$, Fig. 7). The coverage for both methods is correct. This can be

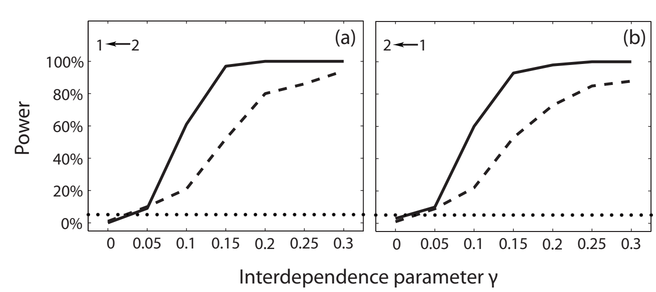

FIG. 7. Power curve for symmetrically coupled stochastic van der Pol oscillators with increasing coupling strength $\gamma=\gamma_{12}=\gamma_{21}=0,\ldots,0.3$ ($h=0.8$ Hz and $p=150$). Solid line: smoothed partial directed coherence, dashed line: partial directed coherence, dotted line: 5%-significance level.

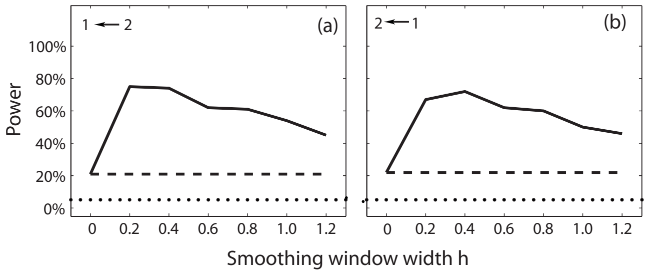

FIG. 8. Power curve for symmetrically coupled stochastic van der Pol oscillators with increasing smoothing window width $h=0,\ldots,1.2$ Hz ($\gamma_{12}=\gamma_{21}=0.1$ and $p=150$). Solid line: smoothed partial directed coherence, dashed line: partial directed coherence, dotted line: 5%-significance level.

observed in Fig. 7, in which for both directions in less than 5% of the realizations a connection is pretended for a 5%-significance level for coupling strengths $\gamma=0$. The power is increased by smoothing as the smoothed partial directed coherence (solid line) reaches 100% for a smaller coupling strength.

The performance for different smoothing window widths increasing from 0 to 1.2 Hz in steps of 0.2 ($\gamma=0.1$, $p=150$), is shown in Fig. 8. Smoothed partial directed coherence is characterized by a higher power than ordinary partial directed coherence. The maximum of the curve indicates that an optimal smoothing window width exists. As the stochastic van der Pol oscillator has a pronounced peak in its spectrum it can be oversmoothed. Thereby, the peak that is related to the interaction may be reduced such that it becomes nonsignificant. Analyzing data it is possible to estimate the smoothed partial directed coherence for different smoothing window width and base the final choice of the window width on these estimations. To show this, we used the very same realizations for all smoothing window width in this study. This is precisely the situation one has to face in applications once it comes to decide about the smoothing width. Using the same realizations also explains the absence of fluctuations for partial directed coherence in these analyses. For increasing model orders p ranging from 100 to 220 ($\gamma=0.1$ and $h=0.8$ Hz, Fig. 9) the results resemble those for the vector-autoregressive processes, partial directed coherence looses power (dashed line) while smoothed partial directed

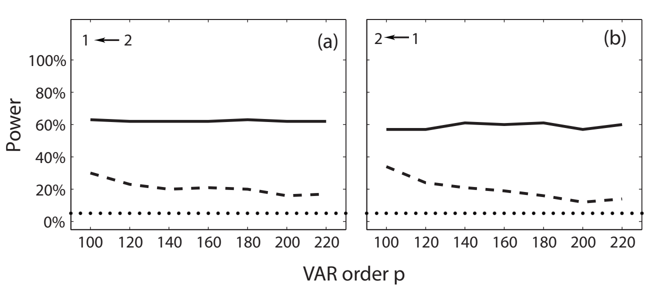

FIG. 9. Power curve for symmetrically coupled stochastic van der Pol oscillators with increasing order of the autoregressive process $p=100,\ldots,220$ ($\gamma_{12}=\gamma_{21}=0.1$ and $h=0.8$ Hz). Solid line: smoothed partial directed coherence, dashed line: partial directed coherence, dotted line: 5%-significance level.

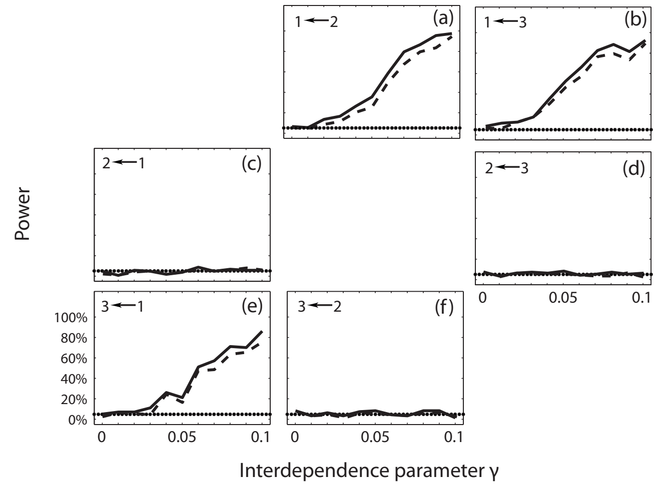

FIG. 10. Power curves for coupled stochastic Rössler oscillators [Eq. (18)] with increasing coupling strength $\gamma_{13}=\gamma_{31}=\gamma_{12}=0,\ldots,0.1$ ($h=0.02$ Hz and $p=150$). Solid line: smoothed partial directed coherence, dashed line: partial directed coherence, dotted line: 5%-significance level.

coherence detects the connection equally well for the different model orders (solid line). All the resulting power curves, Figs. 7–9, show that also for this nonlinear system the performance of the partial directed coherence is clearly improved by smoothing.

C. Coupled stochastic Rössler oscillators

To demonstrate the performance of smoothed partial directed coherence when dealing with the smallest indeed multivariate network which is in its chaotic regime, we simulated three coupled stochastic Rössler oscillators

$$\dot{x}_i = -\omega_i y_i - x_i + \sum_{j\neq i} \gamma_{ij}(x_j - x_i) + \sigma\varepsilon_i$$

$$\dot{y}_i = \omega_i x_i + a y_i$$

$$\dot{z}_i = b + (x_i - c) z_i, \qquad (18)$$

with $i,j=1,2,3$, $\omega_1=1.03$, $\omega_2=1.01$, $\omega_3=0.99$, $a=0.15$, $b=0.2$, $c=10$, $\sigma=1.5$, and Gaussian distributed white noise ε_i of unit variance. The data were simulated using the Euler method with an integration time step of 10 ms and a sampling rate of 10 Hz. The coupling strengths $\gamma_{13}=\gamma_{31}=\gamma_{12}$ were varied between 0 and 0.1 in steps of 0.01 while all other coupling strengths were set to zero. The connections were analyzed only for the x_i at the peak frequency of the driving oscillators.

In Fig. 10 the results for the Rössler oscillators are shown. The partial directed coherence (dashed line) as well as the smoothed partial directed coherence (solid line) show good coverage as they pretend the absent connections $1\to2$, $3\to2$, and $2\to3$ in less than 5% of the realizations given a 5%-significance level, Figs. 10(c), 10(d), and 10(f). Smoothing improves the power of the detection of a present connection as the power curves for the smoothed partial directed coherence rise for slightly lower coupling strengths than those of partial directed coherence [Figs. 10(a), 10(b), and 10(e)]. Both methods show power above chance level for coupling strength above 0.01. For the deterministic case phase synchronization sets in at a bidirectional coupling of 0.02. Thus, we are able to reveal the presence of coupling for a rather large range of coupling strengths including those which represent the regime of phase synchrony. In addition both methods are capable of detecting only direct connections since the connection $2\rightarrow 3$ is correctly identified as indirect [Fig. 10(f)]. We like to emphasize that for this simulation study, the difference between smoothed and nonsmoothed partial directed coherence is rather small. This is due to the fact that the smoothing width has to be chosen rather small because of the narrow peak exhibited by the Rössler system.

IV. APPLICATION TO ESSENTIAL TREMOR

Essential tremor, a neurological disorder with a prevalence of 0.4–4 % [21], manifests itself mainly in the upper limbs, when the hands are in a postural outstretched position. Usually the trembling frequency of the hands is 4–8 Hz. The pathophysiological basis of essential tremor is not precisely known and therefore relationships between the brain and trembling muscles are of particular interest to elucidate the tremor generating mechanisms of essential tremor. Tremor correlated cortical activity has been observed by coherence analysis of simultaneously recorded EEG and EMG [22]. Within that study it was not possible to differentiate whether the cortex imposes its oscillatory activity on the muscles via the corticospinal tract or whether the muscle activity is just reflected to the cortex via proprioceptive afferences. Therefore, to get deeper insights into tremor generating mechanisms, partial directed coherence is applied to data recorded from patients suffering from essential tremor.

For one representative patient with essential tremor, the EMG from the left wrist flexor as well as the EEG recorded over the left and right sensorimotor cortex are analyzed. Bilateral postural tremor was recorded for 300 s using a sampling rate of 1 000 Hz. EEG data were band pass filtered between 0.5 and 200 Hz. To avoid movement artifacts, EMG data were band pass filtered between 30 and 200 Hz and rectified afterwards. Before the analysis, downsampling to 100 Hz was performed.

We estimated partial directed coherence using a VAR order of $p=200$. The results are shown in Fig. 11, where the marked frequency is the dominant tremor frequency, 5.3 Hz, estimated from the power spectrum of the EMG data. Figure 12 shows the results for the smoothed partial directed coherence using a smoothing window width of 0.5 Hz. The analysis at the tremor frequency reveals the graphs shown in Fig. 13. The partial directed coherence corresponding to the directed connection from the right EEG to the muscle, Fig. 11 bottom left, shows significant values around the tremor frequency but not at this specific frequency. Therefore partial directed coherence declares this edge as missing. Smoothed partial directed coherence has increased power as demonstrated above. It reveals, besides the connection from the

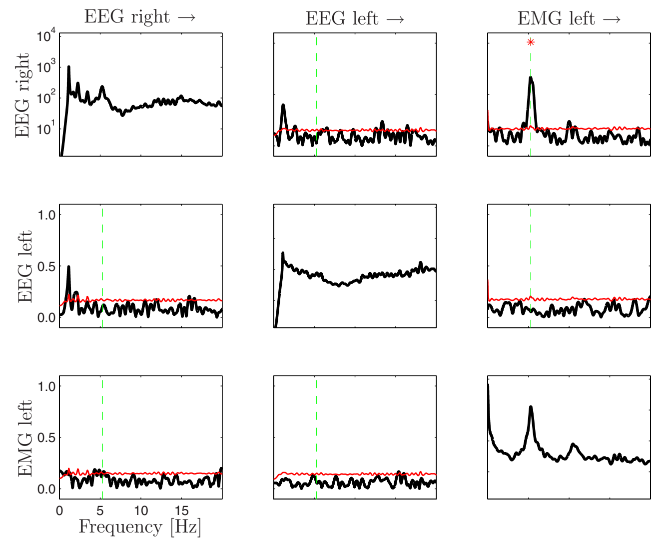

FIG. 11. (Color online) Partial directed coherence for the application to essential tremor. 5%-significance levels are depicted by red (gray) lines. Autospectra are shown on the diagonal. The tremor frequency is marked by a dashed green line, significant values are highlighted by a red asterisk.

muscle to the right EEG, the influence from EEG to the muscle. The graph deduced from partial directed coherence, Fig. 13(a) alone leads to the conclusion that the muscle only reports its tremor activity to the cortex, which would indicate that the cortex just receives feedback from the muscles. The brain would not be assigned a particular role in tremor generation or maintenance, indicating a direct connection from the subcortical basal ganglia as suggested by [23]. The results of smoothed partial directed coherence, Fig. 13(b), which have turned out to be more reliable from our simulation studies, suggest an involvement of the brain in tremor generation.

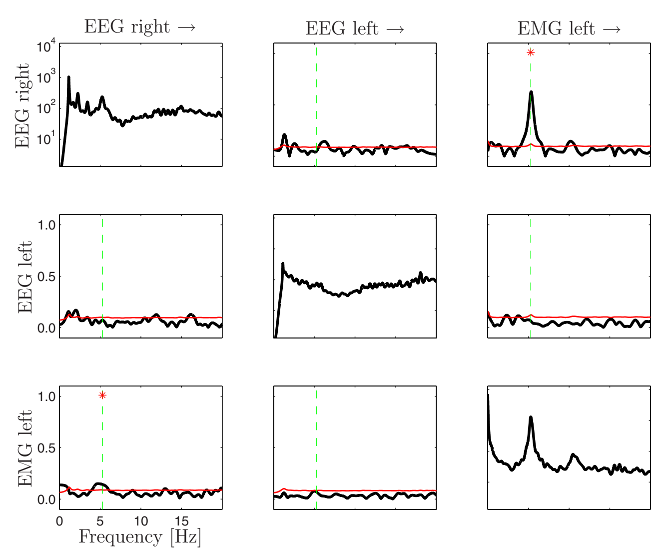

FIG. 12. (Color online) Smoothed partial directed coherence for the application to essential tremor. 5%-significance levels are depicted by red (gray) lines. Autospectra are shown on the diagonal. The tremor frequency is marked by a dashed green line, significant values are highlighted by a red asterisk.

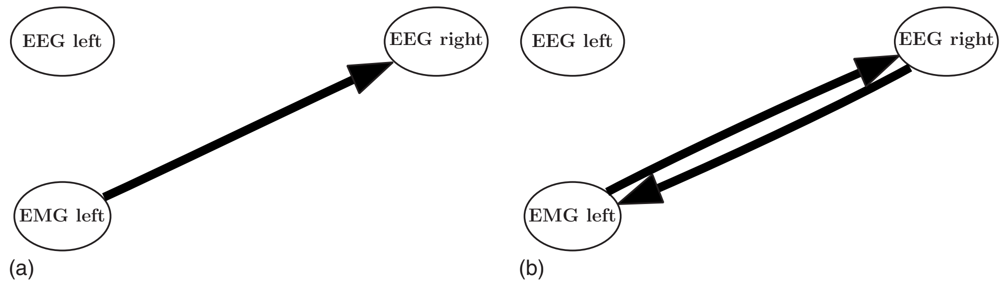

FIG. 13. Graphs for the application to essential tremor: (a) Graph revealed by estimation of partial directed coherence, (b) Graph revealed by estimation of smoothed partial directed coherence.

For partial coherence it was shown, that observational noise might hamper the detection of the underlying interaction structure [20]. However, approaches on denoising signals related to Granger causality have been suggested in the literature, see for instance [24–26].

V. CONCLUSION

Partial directed coherence was suggested as a means for inferring the network structure in linear systems. It was shown to be capable of inferring the underlying network structure of nonlinear systems for certain parameter ranges. To this aim high model orders for the applied vector-autoregressive models were inevitable. It stands to reason that the high model orders are required to capture the second-order properties of the nonlinear systems sufficiently well to reveal the underlying network structure.

High model orders cause highly fluctuating partial directed coherence spectra. Thus, for nonlinear systems it is difficult to gather the true underlying interaction structure as false positive crossings of the significance level might occur with increased probability within each partial directed coherence spectrum. Due to the increased significance level for high model orders, partial directed coherence values are nonsignificant at the frequency of interest, although the connection is actually present. These false negative conclusions that manifest in a lack of power are the second important drawback of ordinary partial directed coherence analysis.

In this paper, we presented an extension of partial directed coherence, the smoothed partial directed coherence, which is capable to reliably reveal the interaction structure of stochastic nonlinear oscillators, including networks of chaotic oscillators which are phase synchronized. The concept of smoothing to regularize the highly fluctuating estimates of partial directed coherence achieves a considerably improved power to reveal the underlying network structure of nonlinear synchronizing oscillators only in combination with the significance level. The significance level could be derived analytically. For many other techniques an analytic expression of the significance level is not known. Surrogate or bootstrap techniques are often used to obtain the statistics in those cases. Surrogate or bootstrap techniques however do base on certain assumptions which may or may not be fulfilled. This can be scarcely controlled in applications. Thus, surrogate or bootstrap techniques could eventually lead to false positive conclusions. The analytic derivation of the statistics presented here, depends only on the Gaussianity of the vector-autoregressive process parameter estimates, which is justified by the central limit theorem.

We have presented an exemplary application to EEG and EMG data from a patient suffering from essential tremor. Estimation of smoothed partial directed coherence allows detection of causal influences between EEG and EMG recordings in essential tremor and provides thus closer insights into the tremor generating mechanisms.

In summary, smoothed partial directed coherence enables to correctly reveal the network structure of nonlinear dynamical systems. It provides both, information on the direction as well as the directness of an interaction.

ACKNOWLEDGMENTS

This work was supported by the German Science Foundation (Grant No. Ti315/4-2), the German Federal Ministry of Education and Research (BMBF Grant No. 01GQ0420), and the Excellence Initiative of the German Federal and State Governments. B.S. and L.S. are indebted to the Landesstiftung Baden-Württemberg for the financial support of this research project by the Eliteprogramme for Postdocs. The authors thank B. Hellwig, C.H. Lücking, and F. Amtage for their support regarding the tremor data.

[1] A. S. Pikovsky and M. G. Rosenblum, *Universal Concept in Nonlinear Sciences* (Cambridge University Press, Cambridge, 2001).
[2] R. Dahlhaus, Metrika **51**, 157 (2000).
[3] B. Schelter, M. Winterhalder, R. Dahlhaus, J. Kurths, and J. Timmer, Phys. Rev. Lett. **96**, 208103 (2006).
[4] G. Nolte, A. Ziehe, V. V. Nikulin, A. Schlögl, N. Krämer, T. Brismar, and K.-R. Müller, Phys. Rev. Lett. **100**, 234101 (2008).
[5] M. C. Romano, M. Thiel, J. Kurths, and C. Grebogi, Phys. Rev. E **76**, 036211 (2007).
[6] L. Cimponeriu, M. G. Rosenblum, T. Fieseler, J. Dammers, M.

Schiek, M. Majtanik, P. Morosan, A. Bezerianos, and P. A. Tass, Prog. Theor. Phys. **150**, (Suppl.), 22 (2003).
[7] M. G. Rosenblum and A. S. Pikovsky, Phys. Rev. E **64**, 045202(R) (2001).
[8] D. A. Smirnov and B. P. Bezruchko, Phys. Rev. E **68**, 046209 (2003).
[9] Y. Chen, G. Rangarajan, J. Feng, and M. Ding, Phys. Lett. A **324**, 26 (2004).
[10] J. Granger, Econometrica **37**, 424 (1969).
[11] M. Dhamala, G. Rangarajan, and M. Ding, Phys. Rev. Lett. **100**, 018701 (2008).
[12] M. Dhamala, G. Rangarajan, and M. Ding, Neuroimage **41**, 354 (2008).
[13] W. Hesse, E. Möller, M. Arnold, and B. Schack, J. Neurosci. Methods **124**, 27 (2003).
[14] M. J. Kaminski and K. J. Blinowska, Biol. Cybern. **65**, 203 (1991).
[15] M. J. Kamiński, M. Ding, W. A. Truccolo, and S. L. Bressler, Biol. Cybern. **85**, 145 (2001).
[16] K. Sameshima and L. A. Baccalá, J. Neurosci. Methods **94**, 93 (1999).
[17] L. A. Baccalá and K. Sameshima, Biol. Cybern. **84**, 463 (2001).
[18] B. Schelter, M. Winterhalder, M. Eichler, M. Peifer, B. Hellwig, B. Guschlbauer, C. H. Lücking, R. Dahlhaus, and J. Timmer, J. Neurosci. Methods **152**, 210 (2006).
[19] H. Lütkepohl, *Introduction to Multiple Time Series Analysis* (Springer, New York, 1993).
[20] Z. Albo, G. Viana Di Prisco, Y. Chen, G. Rangarajan, W. Truccolo, J. Feng, R. P. Vertes, and M. Ding, Biol. Cybern. **90**, 318 (2004).
[21] E. Louis, B. Ford, K. Wendt, and G. Cameron, Mov. Disord. **13**, 803 (1998).
[22] B. Hellwig, S. Häußler, B. Schelter, M. Lauk, B. Guschlbauer, J. Timmer, and C. H. Lücking, Lancet **357**, 519 (2001).
[23] D. M. Halliday, B. A. Conway, S. F. Farmer, U. Shahani, A. J. Russell, and J. R. Rosenberg, Lancet **355**, 1149 (2000).
[24] H. Nalatore, M. Ding, and G. Rangarejan, J. Neurosci. Methods **179**, 131 (2009).
[25] M. Winterhalder, B. Schelter, W. Hesse, K. Schwab, L. Leistritz, D. Klan, R. Bauer, J. Timmer, and H. Witte, Signal Process. **85**, 2137 (2005).
[26] M. Winterhalder, B. Schelter, W. Hesse, K. Schwab, L. Leistritz, J. Timmer, and H. Witte, Biomed. Tech. **51**, 281 (2006).