ELSEVIER

Journal of Neuroscience Methods 152 (2005) 210–219

NEUROSCIENCE METHODS

www.elsevier.com/locate/jneumeth

Testing for directed influences among neural signals using partial directed coherence

Björn Schelter[a,b,*], Matthias Winterhalder[a,b], Michael Eichler[c], Martin Peifer[a], Bernhard Hellwig[d], Brigitte Guschlbauer[d], Carl Hermann Lücking[d], Rainer Dahlhaus[c], Jens Timmer[a,b]

[a] *FDM, Freiburg Center for Data Analysis and Modeling, University of Freiburg, Eckerstr. 1, 79104 Freiburg, Germany*
[b] *Bernstein Center for Computational Neuroscience, University of Freiburg, Freiburg, Germany*
[c] *Institute of Applied Mathematics, University of Heidelberg, Im Neuenheimer Feld 294, 69120 Heidelberg, Germany*
[d] *Department of Neurology, University Hospital of Freiburg, Breisacher Str. 64, 79106 Freiburg, Germany*

Received 13 May 2005; received in revised form 2 September 2005; accepted 5 September 2005

Abstract

One major challenge in neuroscience is the identification of interrelations between signals reflecting neural activity. When applying multivariate time series analysis techniques to neural signals, detection of directed relationships, which can be described in terms of Granger-causality, is of particular interest. Partial directed coherence has been introduced for a frequency domain analysis of linear Granger-causality based on modeling the underlying dynamics by vector autoregressive processes. We discuss the statistical properties of estimates for partial directed coherence and propose a significance level for testing for nonzero partial directed coherence at a given frequency. The performance of this test is illustrated by means of linear and non-linear model systems and in an application to electroencephalography and electromyography data recorded from a patient suffering from essential tremor.

Keywords: Partial directed coherence; Granger-causality; Multivariate time series; Graphical models; Essential tremor

1. Introduction

Examinations of interrelations and especially causal influences between different brain areas are of particular interest in neuroscience. These investigations are based on considering the brain as a dynamic system and analyzing signals reflecting neural activity, for example, electroencephalographic (EEG) or magnetoencephalographic (MEG) recordings. This approach has been used, for instance, in application to data sets recorded from patients suffering from neurological diseases, in order to increase the understanding of underlying mechanisms generating these dysfunctions (Grosse et al., 2002; Hellwig et al., 2000, 2001, 2003; Hesse et al., 2003; Tass et al., 1998; Volkmann et al., 1996).

Various time series analysis techniques have been proposed for the description of interdependencies between dynamic processes and for the detection of causal influences in multivariate systems (Boccaletti et al., 2002; Dahlhaus et al., 1997; Dahlhaus, 2000; Eichler et al., 2003; Pikovsky et al., 2001; Rosenblum and Pikovsky, 2001; Smirnov and Bezruchko, 2003; Timmer et al., 1998). In the frequency domain the interdependencies between two dynamic processes are investigated by the cross-spectrum and the coherence. For multivariate systems with more than two components, the distinction between indirect and direct relationships becomes important, the latter can be described in terms of the partial coherence (Brillinger, 1981). The partial coherence has been used to define graphical interaction models that describe the dependence structure of multivariate time series by undirected graphs (Brillinger, 1996; Dahlhaus, 2000).

The concept of Granger-causality (Granger, 1969) is usually utilized for determination of causal influences. This probabilistic concept of causality is based on the common sense idea that causes precede their effects in time and is formulated in terms of predictability. Empirically, Granger-causality is commonly evaluated by fitting vector autoregressive models. A

* Corresponding author. Tel.: +49 761 203 7710; fax: +49 761 203 7700.
E-mail address: schelter@fdm.uni-freiburg.de (B. Schelter).

0165-0270/$ – see front matter
doi:10.1016/j.jneumeth.2005.09.001

graphical approach for modeling Granger-causal relationships in multivariate processes has been discussed (Eichler, Preprint 2001); an overview is provided in Dahlhaus and Eichler (2003). More generally, graphs provide a convenient framework for causal inference and allow, for example, the discussion of so-called spurious causalities due to confounding by latent variables (Eichler, 2005).

The partial directed coherence has been introduced for inference of linear Granger-causality in the frequency domain based on vector autoregressive models of appropriate order p (VAR[p]) (Baccala and Sameshima, 2001; Sameshima and Baccala, 1999). Unlike coherence and partial coherence analysis, the statistical properties of the partial directed coherence have not yet been investigated. In particular, significance levels for testing for nonzero partial directed coherences at fixed frequencies are not available and are usually determined by simulations (Baccala and Sameshima, 2001; Schnider et al., 1989). On the one hand, without a significance level, detection of causal influences becomes more hazardous for increasing model order as the variability of estimated partial directed coherences increases leading to false detections. On the other hand, a high model order is often required to describe the dependencies of a multivariate process examined sufficiently well.

In this paper, the statistical properties of partial directed coherence estimates are discussed and a significance level for testing nonzero directed influences between frequency components is derived. In Section 2, partial directed coherence is introduced and simulations illustrating the problem of a missing significance level are presented. Furthermore, statistical properties of partial directed coherence are calculated. Since not only linear but also non-linear stochastic systems may be present in applications, the performance of this significance level is demonstrated by one linear and one non-linear stochastic model system with different dynamic behavior in Section 3. An exemplary application to electroencephalographic and electromyographic (EMG) recordings of a patient suffering from essential tremor is presented in Section 4.

2. Partial directed coherence

In the following, the concepts of Granger-causality and partial directed coherence are presented and illustrated in examples of vector autoregressive processes. Furthermore, we discuss the statistical properties of estimators for partial directed coherence and derive a pointwise significance level for testing for a nonzero partial directed coherence.

2.1. Definition

The concept of Granger-causality, which originates from econometrics (Granger, 1969), is a fundamental tool for the description of directed dynamic relationships among the components of a multivariate process and has been applied recently to problems in neuroscience (e.g. Eichler, 2005; Goebel et al., 2003; Hesse et al., 2003). Based on the common sense conception that causes precede their effects in time, this probabilistic concept of causality exploits the temporal structure of signals and defines causal relationships in terms of predictability. In a linear framework, the notion of Granger-causality is closely related to vector autoregressions.

More precisely, let $\mathbf{x} = (\mathbf{x}(t))_{t\in\mathbb{Z}}$ with $\mathbf{x}(t) = (x_1(t), \ldots, x_n(t))'$ be a stationary n-dimensional time series with mean zero. Then a vector autoregressive model of order p, abbreviated VAR[p], for $\mathbf{x}$ is given by

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

where $\mathbf{a}(r)$ are the $n \times n$ coefficient matrices of the model and $\varepsilon(t)$ is a multivariate Gaussian white noise process. The covariance matrix of the noise process is denoted by $\mathbf{\Sigma}$. To guarantee stationarity of the model we assume that

$$\det(I - \mathbf{a}(1)\,z - \ldots - \mathbf{a}(p)\,z^p) \neq 0 \tag{2}$$

for all $z \in \mathbb{C}$ such that $|z| \leq 1$ (e.g. Lütkepohl, 1993).

In this model, the coefficients $\mathbf{a}_{ij}(r)$ describe how the present values of x_i depend linearly on the past values of the components x_j. Thus, we say that x_j does not Granger-cause another process x_i with respect to the full process $\mathbf{x}$ if in the autoregressive representation (1) all entries $\mathbf{a}_{ij}(r)$ are zero for $r = 1, \ldots, p$; or in other words, if linear prediction of $x_i(t+1)$ based on the past and present values of all variables but x_j cannot be improved by adding the past and present values of x_j. We note that the vector autoregressive modeling approach allows only the description of linear relationships among the variables and hence, strictly speaking, relates to linear Granger-causality. In the sequel, we will use “Granger-causality” in this restricted meaning.

In order to provide a frequency domain description of Granger-causality, Baccala and Sameshima (2001) introduced the concept of partial directed coherence. Let

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

denote the difference between the n-dimensional identity matrix I and the Fourier transform of the coefficient series. Then the partial directed coherence $|\pi_{i\leftarrow j}(\omega)|$ for a VAR[p]-process is defined as

$$|\pi_{i\leftarrow j}(\omega)| = \frac{|\mathbf{A}_{ij}(\omega)|}{\sqrt{\sum_k |\mathbf{A}_{kj}(\omega)|^2}}. \tag{4}$$

We note that condition (2) guarantees that the denominator is strictly positive and hence that the partial directed coherence is well defined. From the definition, it follows that $|\pi_{i\leftarrow j}(\omega)|$ vanishes for all frequencies ω if and only if all coefficients $\mathbf{a}_{ij}(r)$ are zero and hence x_j does not Granger-cause x_i given the other variables. This suggests that the partial directed coherence $|\pi_{i\leftarrow j}(\omega)|$ provides a measure for the directed linear influence of x_j on x_i at frequency ω. Furthermore, because of the normalization in Eq. (4), the partial directed coherence takes values in the interval [0,1]. It compares the effect of the past of x_j on the presence of x_i with the effect of the past of x_j on the other variables. Thus, partial directed coherence ranks the interaction strengths with respect to a given signal source.

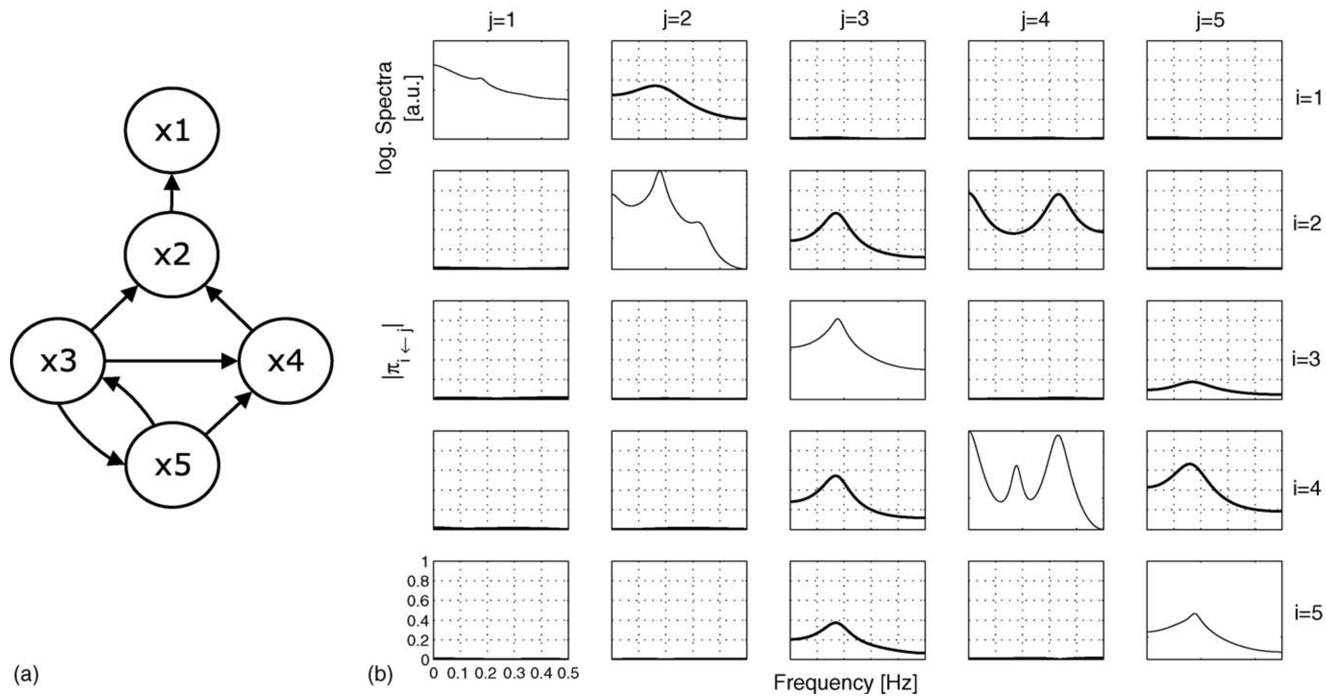

Fig. 1. Graph summarizing the causal influences for the example of a five-dimensional VAR[4] (a). Corresponding spectra (diagonal) and partial directed coherence (off-diagonal) (b). The simulated causal influences are reproduced correctly by the estimated partial directed coherence.

2.2. Background

To illustrate the performance of partial directed coherence in detecting causal influences, the following five-dimensional vector autoregressive process of order $p = 4$

$$\begin{aligned}
x_1(t) &= 0.6x_1(t-1) + 0.65x_2(t-2) + \varepsilon_1(t)\\
x_2(t) &= 0.5x_2(t-1) - 0.3x_2(t-2) - 0.3x_3(t-4)\\
&\quad + 0.6x_4(t-1) + \varepsilon_2(t)\\
x_3(t) &= 0.8x_3(t-1) - 0.7x_3(t-2) - 0.1x_5(t-3) + \varepsilon_3(t)\\
x_4(t) &= 0.5x_4(t-1) + 0.9x_3(t-2) + 0.4x_5(t-2) + \varepsilon_4(t)\\
x_5(t) &= 0.7x_5(t-1) - 0.5x_5(t-2) - 0.2x_3(t-1) + \varepsilon_5(t)
\end{aligned} \tag{5}$$

is investigated. The covariance matrix of the noise $\mathbf{\Sigma}$ is set to the identity, the number of simulated data points is $N = 50.000$ for each component of the VAR-process. The structure of the autoregressive model can be summarized by the graph in Fig. 1(a). In this graph, a directed edge from x_j to x_i is drawn if x_j Granger-causes x_i. The properties of such graphs have been investigated in detail in Eichler (2001, 2002). In particular, they allow to differentiate between direct and indirect influences. For instance, process x_4 has an indirect influence on x_1 that is mediated by x_2 since, for all times t, the present value of $x_1(t)$ is influenced by $x_2(t-2)$, which in turn is influenced by $x_4(t-3)$. If x_2 would be blocked somehow, changes in x_4 would no longer affect x_1.

The corresponding partial directed coherence is given in Fig. 1(b). Partial directed coherence is estimated by fitting a VAR[4]-model to the data and direct usage of Eq. (4). The spectra of the processes x_i are displayed on the diagonal. Partial directed coherences $|\pi_{i\leftarrow j}|$ are shown on the off-diagonal elements. The results show that the autoregressive structure of the simulated process is detected correctly by partial directed coherence, because $|\pi_{1\leftarrow 2}|$, $|\pi_{2\leftarrow 3}|$, $|\pi_{4\leftarrow 3}|$, $|\pi_{5\leftarrow 3}|$, $|\pi_{2\leftarrow 4}|$, $|\pi_{3\leftarrow 5}|$, and $|\pi_{4\leftarrow 5}|$ do not vanish. The example demonstrates that partial directed coherence provides a frequency domain approach for the identification of causal influences in multivariate systems.

In this example, the order of the fitted vector autoregressive process has been chosen to be the correct order of the simulated process $p = 4$. But in applications to empirical time series, the order is unknown that is necessary to describe the processes by a VAR-model sufficiently well, especially for non-linear stochastic processes.

To demonstrate the effects of different values for the order of the fitted model, a two-dimensional vector autoregressive process VAR[2]

$$\begin{pmatrix} x_1(t)\\ x_2(t)\end{pmatrix} = \begin{pmatrix} a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}\begin{pmatrix} x_1(t-1)\\ x_2(t-1)\end{pmatrix} + \begin{pmatrix} b_{11} & b_{12}\\ b_{21} & b_{22}\end{pmatrix}\begin{pmatrix} x_1(t-2)\\ x_2(t-2)\end{pmatrix} + \begin{pmatrix} \varepsilon_1(t)\\ \varepsilon_2(t)\end{pmatrix} \tag{6}$$

is examined, that can be interpreted as a system of relaxators and damped oscillators. The parameters

$$a_{11} = a_{22} = 2\cos\left(\frac{2\pi}{T}\right)\exp\left(-\frac{1}{\tau}\right) \tag{7}$$

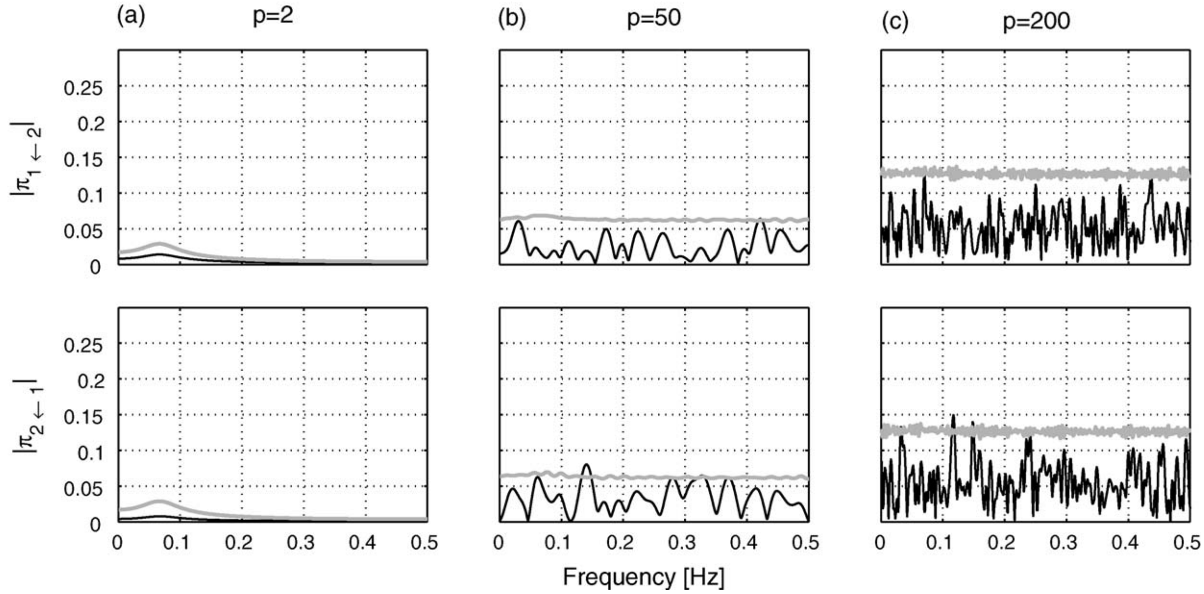

Fig. 2. Partial directed coherence for three different values of the order of the fitted vector autoregressive process $p = 2$ (a), $p = 50$ (b), and $p = 200$ (c). Variability increases with increasing model order p. As in almost all applications the correct value of p is unknown and especially for non-linear stochastic processes a high order of p is required to describe the processes sufficiently well, a significance level for partial directed coherence is strongly required. The almost horizontal, gray lines indicate the 5%-significance level that will be derived in the following.

$$b_{11} = b_{22} = -\exp\left(-\frac{2}{\tau}\right) \tag{8}$$

are chosen using $T = 15$ for the characteristic period and $\tau = 5$ for the characteristic relaxation time of the oscillators (Timmer, 1998). The problem of interpreting partial directed coherence for increasing orders of the fitted VAR-model is illustrated by two independent AR[2]-processes ($a_{12} = a_{21} = b_{12} = b_{21} = 0$), simulated each with $N = 10.000$ data points.

In Fig. 2 the partial directed coherence is shown for different values of the order p of the fitted VAR[p]-model to the true VAR[2]-process (cf. Eq. (6)). In (a), the correct order of $p = 2$ is chosen and partial directed coherence is mostly vanishing as there is no influence between both AR[2]-processes. Increasing the order of the fitted vector autoregressive process to $p = 50$ (b) and $p = 200$ (c), respectively, yields a higher variability in partial directed coherence. This increasing variability complicates the interpretation of partial directed coherence. Partial directed coherence obtains values up to 0.15 for $p = 200$ (c). Thus, it is hardly possible to decide without a significance level whether or not there are causal influences between the processes. As in several applications a high value of p is required to describe especially non-linear processes sufficiently well by vector autoregressive processes, a significance level for partial directed coherence is required. The gray lines represent the corresponding significance level that will be derived in the next section.

2.3. Statistical properties

The partial directed coherence $|\pi_{i\leftarrow j}(\omega)|$ is estimated by fitting a n-dimensional VAR[p] model to the data and using Eqs. (3) and (4) with the parameter estimates $\hat{\mathbf{a}}_{ij}(k)$ substituted for the true coefficients $\mathbf{a}_{ij}(k)$. Thus, the statistical properties of the estimates of partial directed coherence $|\hat{\pi}_{i\leftarrow j}(\omega)|$ can be derived from the properties of the parameter estimates $\hat{\mathbf{a}}_{ij}(k)$. Commonly used estimates $\hat{\mathbf{a}}_{ij}(k)$ of the coefficients $\mathbf{a}_{ij}(k)$ such as least squares estimates or Yule-Walker estimates are asymptotically normally distributed with mean $\mathbf{a}_{ij}(k)$ and covariances

$$\lim_{N\to\infty} N \operatorname{cov}(\hat{\mathbf{a}}_{ij}(k), \hat{\mathbf{a}}_{ij}(l)) = \mathbf{\Sigma}_{ii}\, \mathbf{H}_{jj}(k, l). \tag{9}$$

Here, $\mathbf{H}_{jj}(k, l)$ are entries of the inverse $\mathbf{H} = \mathbf{R}^{-1}$ of the covariance matrix $\mathbf{R}$ of the VAR-process $\mathbf{x}$. The covariance matrix $\mathbf{R}$ is composed of $n \times n$ sub-matrices

$$\mathbf{R}(k, l) = \begin{pmatrix} \mathbf{R}_{11}(k, l) & \cdots & \mathbf{R}_{1n}(k, l) \\ \vdots & \ddots & \vdots \\ \mathbf{R}_{n1}(k, l) & \cdots & \mathbf{R}_{nn}(k, l) \end{pmatrix} \tag{10}$$

with entries

$$\mathbf{R}_{ij}(k, l) = \operatorname{cov}(x_i(t - k), x_j(t - l)) \tag{11}$$

for $i, j = 1, \ldots, n$ and $k, l = 1, \ldots, p$ (Lütkepohl, 1993). Consequently, the real and imaginary part of $\hat{\mathbf{A}}_{ij}(\omega)$ are also asymptotically jointly normally distributed. In the appendix, it is shown that under the null hypothesis of $|\mathbf{A}_{ij}(\omega)|^2 = 0$ the asymptotic distribution of

$$\frac{N}{C_{ij}(\omega)}\, |\hat{\mathbf{A}}_{ij}(\omega)|^2 \tag{12}$$

is a weighted average of two independent χ^2-distributions with one degree of freedom, where

$$C_{ij}(\omega) = \boldsymbol{\Sigma}_{ii}\left[\sum_{k,l=1}^{p} \mathbf{H}_{jj}(k,l)(\cos(k\omega)\cos(l\omega)+\sin(k\omega)\sin(l\omega))\right]. \quad (13)$$

Thus, the critical value of the distribution is bounded by the critical value of a χ^2-distribution with one degree of freedom.

Next, we consider the complex valued function

$$\hat{\pi}_{i\leftarrow j}(\omega) = \frac{\hat{\mathbf{A}}_{ij}(\omega)}{\sqrt{\sum_k |\hat{\mathbf{A}}_{kj}(\omega)|^2}} \quad (14)$$

from which we obtain the partial directed coherence by taking the absolute value. This function is non-linear in the parameter estimates $\hat{\mathbf{a}}_{ij}(k)$. Taylor expansion of $\hat{\pi}_{i\leftarrow j}(\omega)$ about $\mathbf{a}_{ij}(k)$ yields

$$\hat{\pi}_{i\leftarrow j}(\omega) = \frac{\hat{\mathbf{A}}_{ij}(\omega)}{\sqrt{\sum_k |\mathbf{A}_{kj}(\omega)|^2}} + \mathbf{A}_{ij}(\omega)\,R_1 + R_2 \quad (15)$$

with $|R_1| \leq C\,\|\hat{\mathbf{a}}-\mathbf{a}\|$ and $|R_2| \leq C\,\|\hat{\mathbf{a}}-\mathbf{a}\|^2$. The estimate of $\mathbf{a} = (\mathbf{a}(1),\ldots,\mathbf{a}(p))$ is denoted by $\hat{\mathbf{a}}$. Under the hypothesis of $|\mathbf{A}_{ij}(\omega)|^2 = 0$ the second term $\mathbf{A}_{ij}(\omega)\,R_1$ vanishes and we have

$$|\hat{\pi}_{i\leftarrow j}(\omega)|^2 = \frac{|\hat{\mathbf{A}}_{ij}(\omega)|^2}{\sum_k |\mathbf{A}_{kj}(\omega)|^2} + R, \quad (16)$$

where the remainder R is of order $\|\hat{\mathbf{a}}-\mathbf{a}\|^3$ and is negligible compared with the main term. Therefore, the α-significance level for the partial directed coherence $|\pi_{i\leftarrow j}(\omega)|$ can be approximated by

$$\left(\frac{\hat{C}_{ij}(\omega)\,\chi^2_{1,1-\alpha}}{N\sum_k |\hat{\mathbf{A}}_{kj}(\omega)|^2}\right)^{1/2}, \quad (17)$$

where $\chi^2_{1,1-\alpha}$ is the $1-\alpha$ quantile of the χ^2-distribution with one degree of freedom and $\hat{C}_{ij}(\omega)$ is an estimate of $C_{ij}(\omega)$ in Eq. (13).

An interesting feature of the significance level in Eq. (17) is that it does depend on the frequency unlike e.g. the significance levels of ordinary coherence (Bloomfield, 1976). In particular, it compensates for the effects of normalization by $\sqrt{\sum_k |\hat{\mathbf{A}}_{kj}(\omega)|^2}$, that is, significance depends not on the relative but the absolute strength of the relationship at that frequency. As a consequence, it could happen that the largest value of $|\hat{\pi}_{i\leftarrow j}(\omega)|$ is found non-significant while at other frequencies smaller values with a larger normalization factor are significant. This demonstrates again the importance of a significance level that can be used to detect local deviations from the null hypothesis of $|\pi_{i\leftarrow j}(\omega)| = 0$.

A possible drawback of the proposed significance level is that it is only a pointwise level, which is typically exceeded for a certain number of frequencies even if there is no causal influence. We expect on average $p\alpha$ crossings of the pointwise α-significance level, where p is the order of the VAR-process. However, since the values of $|\pi_{i\leftarrow j}(\omega)|$ at different frequencies ω are in general highly correlated, it seems hardly possible to derive a uniform non-constant significance level for the partial directed coherence. We mention though that it is rather common in other applications to use pointwise levels.

Calculation of this significance level for the example of two independent AR[2]-processes leads to the gray lines in Fig. 2 ($\alpha = 5\%$). Since the significance level is only crossed at a small number of frequencies, a causal influence between both processes is rejected. We note that for $p = 2$ the significance level exhibits a strong dependence on the frequency illustrating the fact that the significance level in Eq. (17) varies with ω.

3. Simulations

In the first part of this section, we demonstrate the performance of the proposed significance level for a vector autoregressive process. We shall see that the method even works well when the true model order is strongly overfitted.

Partial directed coherence has been developed in the frame of linear stochastic processes. In several applications, however, non-linear stochastic processes are expected to generate the time series. Nevertheless, in most of these cases the dependence structure is reflected in the linear second order structure and for this reason the partial directed coherence also works well for many non-linear processes. In the second part of this section we therefore demonstrate the performance of the method for a non-linear stochastic oscillatory system. Again we show that the proposed significance level works well for rather high model orders. High model orders are required for a sufficient description of non-linear systems by linear vector autoregressive processes.

3.1. Vector autoregressive process

The five-dimensional VAR[4]-process from Eq. (5) serves as a first example for the performance of the significance level. The interdependence structure is given in Fig. 1(a). As shown in Fig. 1(b), the interdependence structure can be intuitively reproduced correctly without knowledge of a significance level. However, this holds if the order of the fitted VAR[p]-process is chosen correctly. Akaike's information criterion (Akaike, 1973) or different model selection strategies (Hastie et al., 1996; Rissanen, 1978, 1983; Schwartz, 1979) might give a hint for the appropriate selection of the model order.

In applications, if a huge time lag is present between time series or if combinations between linear and non-linear processes have to be investigated, rather high model orders are required. In these circumstances the significance level allows to decide whether a partial directed coherence is significantly different from zero or not.

To investigate the effects of high model orders, a VAR[200]-model is fitted to the time series generated by the VAR[4]-process. The results are shown in Fig. 3. The corresponding 5%-significance levels are shown by the gray lines. The higher order increases the variability of the partial directed coherence but, correctly, only direct causal influences are revealed by sig-

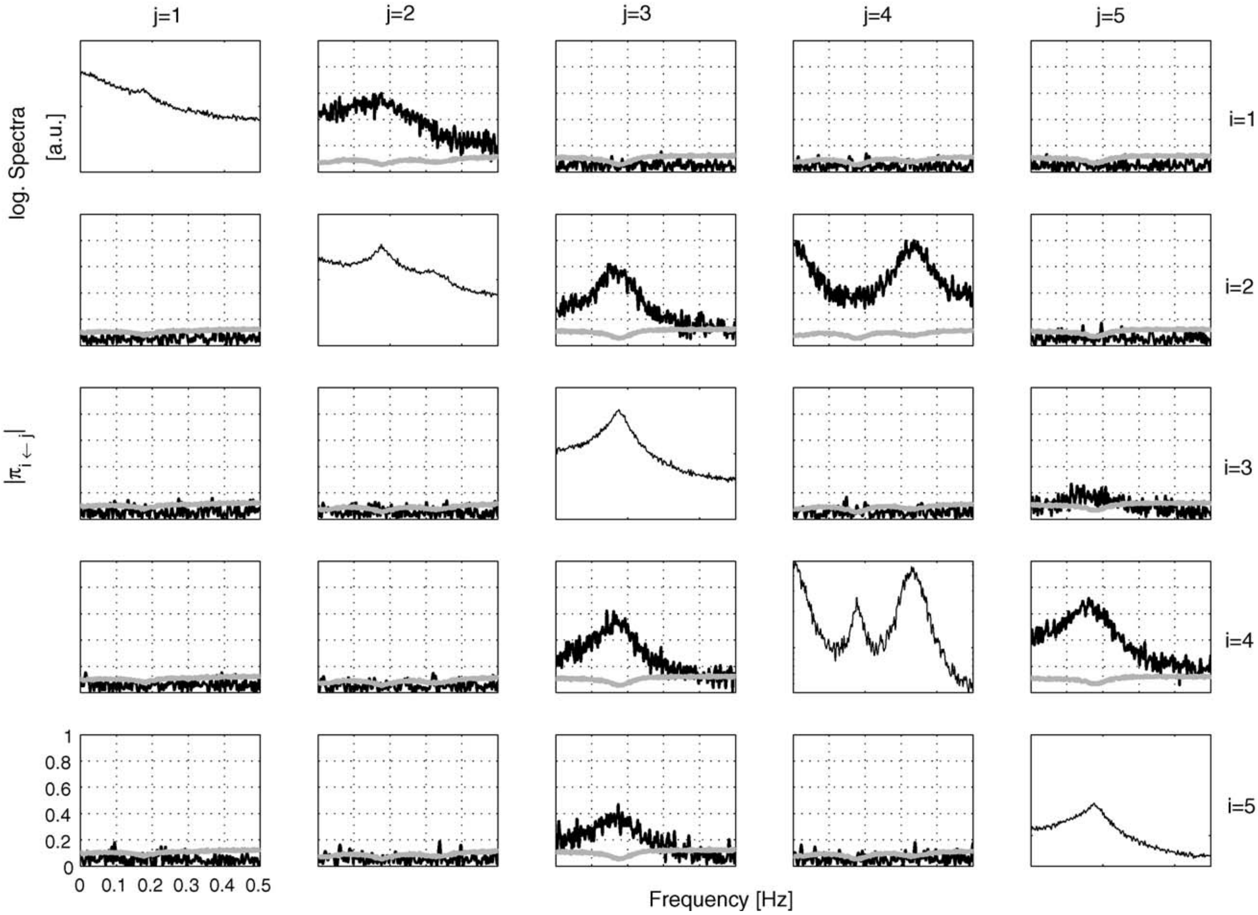

Fig. 3. Spectra (diagonal) and partial directed coherence (off-diagonal) for the example of the VAR[4]-process. The corresponding 5%-significance levels are indicated by the gray lines. The simulated causal influences are reproduced correctly since only partial directed coherences $|\pi_{1\leftarrow 2}|$, $|\pi_{2\leftarrow 3}|$, $|\pi_{4\leftarrow 3}|$, $|\pi_{5\leftarrow 3}|$, $|\pi_{2\leftarrow 4}|$, $|\pi_{3\leftarrow 5}|$, and $|\pi_{4\leftarrow 5}|$ are significant.

nificant partial directed coherences. These are the partial directed coherences $|\pi_{1\leftarrow 2}|$, $|\pi_{2\leftarrow 3}|$, $|\pi_{4\leftarrow 3}|$, $|\pi_{5\leftarrow 3}$, $|\pi_{2\leftarrow 4}|$, $|\pi_{3\leftarrow 5}|$, and $|\pi_{4\leftarrow 5}|$.

3.2. Four coupled stochastic van der Pol oscillators

To test the validity of the significance level for a non-linear system, an example of coupled stochastic van der Pol oscillators (van der Pol, 1922)

$$\ddot{x}_1 = \mu(1 - x_1^2)\dot{x}_1 - \omega_1^2 x_1 + \sigma\eta_1 + \varepsilon_{12}(x_2 - x_1) + \varepsilon_{13}(x_3 - x_1) + \varepsilon_{14}(x_4 - x_1) \quad (18)$$

$$\ddot{x}_2 = \mu(1 - x_2^2)\dot{x}_2 - \omega_2^2 x_2 + \sigma\eta_2 + \varepsilon_{23}(x_3 - x_2) + \varepsilon_{24}(x_4 - x_2) + \varepsilon_{21}(x_1 - x_2) \quad (19)$$

$$\ddot{x}_3 = \mu(1 - x_3^2)\dot{x}_3 - \omega_3^2 x_3 + \sigma\eta_3 + \varepsilon_{34}(x_4 - x_3) + \varepsilon_{31}(x_1 - x_3) + \varepsilon_{32}(x_2 - x_3) \quad (20)$$

$$\ddot{x}_4 = \mu(1 - x_4^2)\dot{x}_4 - \omega_4^2 x_4 + \sigma\eta_4 + \varepsilon_{41}(x_1 - x_4) + \varepsilon_{42}(x_2 - x_4) + \varepsilon_{43}(x_3 - x_4) \quad (21)$$

is investigated with $N = 50.000$ data points for each process, $\omega_1 = 1.5, \omega_2 = 1.48, \omega_3 = 1.53, \omega_4 = 1.44, \sigma = 1.5$, and Gaussian distributed white noise η_i. Although the interaction between the four oscillators is still linear, the system is non-linear due to terms weighted by the parameter μ. The parameter μ is fixed to 5, leading to a highly non-linear behavior of the van der Pol oscillators. The unidirectional and bidirectional coupling between these four non-identical oscillators are set to $\varepsilon_{12} = \varepsilon_{21} = 0.2$, $\varepsilon_{24} = \varepsilon_{42} = 0.2$, $\varepsilon_{31} = 0.2$, and $\varepsilon_{34} = 0.2$.

The causal influences are summarized in the graph in Fig. 4(a). Estimated partial directed coherence as well as the spectra are given in Fig. 4(b). The order of the vector autoregressive process is chosen to be $p = 200$. This high model order is required to reproduce the spectra sufficiently well, compared to non-parametric spectral estimates.

The corresponding 5%-significance levels are indicated by gray lines. Partial directed coherence correctly detects the causal influences in the van der Pol system, as only $|\pi_{2\leftarrow 1}|$, $|\pi_{3\leftarrow 1}|$, $|\pi_{1\leftarrow 2}|$, $|\pi_{4\leftarrow 2}|$, $|\pi_{2\leftarrow 4}|$, and $|\pi_{3\leftarrow 4}|$ are significant at the corresponding oscillation frequencies. The significance level depends on the investigated frequency. At the peaks in the spectra of the van der Pol oscillators, the significance level is slightly higher than at the remaining frequencies. This allows only those partial

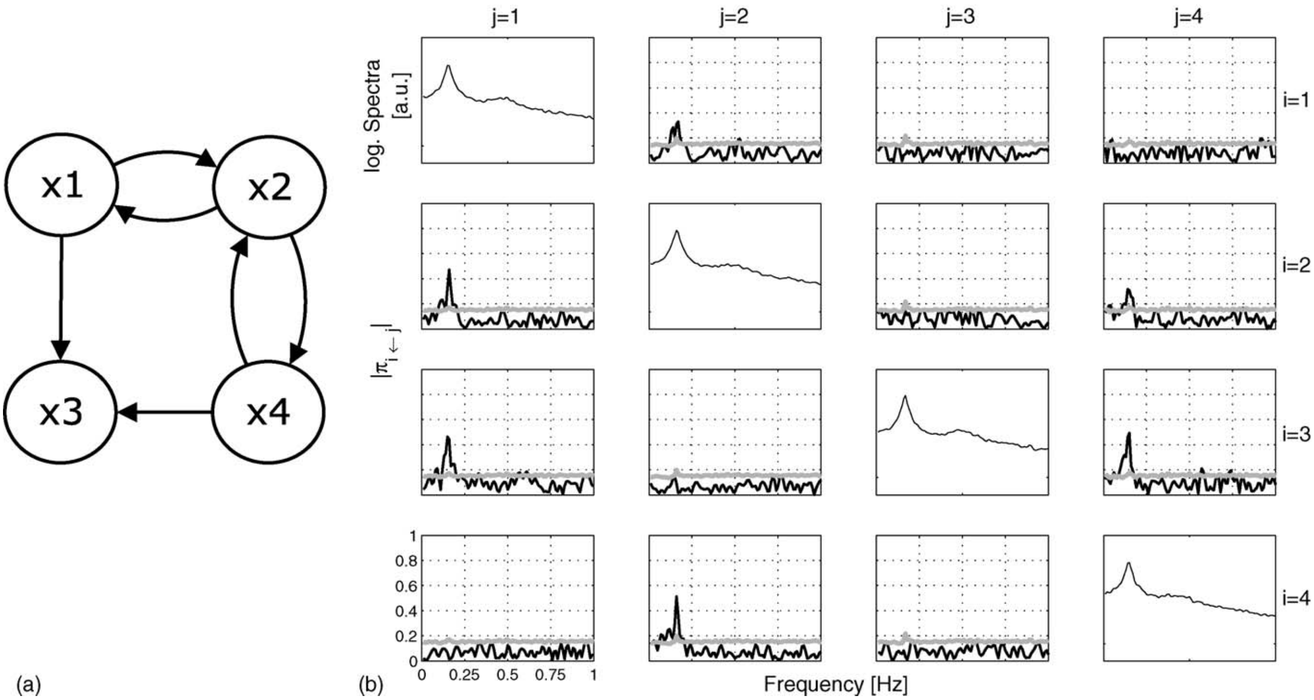

Fig. 4. Causal influences for the example of four coupled stochastic van der Pol oscillators (a). Corresponding spectra (diagonal) and partial directed coherence (off-diagonal) (b). The corresponding 5%-significance levels are indicated by gray lines. The simulated causal influences are reproduced correctly since only partial directed coherences $|\pi_{2\leftarrow 1}|$, $|\pi_{3\leftarrow 1}|$, $|\pi_{1\leftarrow 2}|$, $|\pi_{4\leftarrow 2}|$, $|\pi_{2\leftarrow 4}|$, and $|\pi_{3\leftarrow 4}|$ are significant at the oscillation frequencies.

directed coherencies to be significant that are related to direct causal influences between the oscillators.

4. Application to essential tremor

The pathophysiological basis of essential tremor, a common neurological disease with a prevalence of 0.4–4% (Louis et al., 1998), is not precisely known. Essential tremor manifests itself mainly in the upper limbs, when the hands are in a postural outstretched position. Usually the trembling frequency of the hands is 4–10 Hz. To elucidate the tremor generating mechanisms in essential tremor, relationships between the brain and trembling muscles are of particular interest. Tremor correlated cortical activity has been observed by coherence analysis of simultaneously recorded electroencephalography and electromyography (Hellwig et al., 2001). 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 in the cortex via proprioceptive afferences. Therefore, to get deeper insights into tremor generation, partial directed coherence is applied to data recorded from patients suffering from essential tremor.

For one patient with essential tremor, the EMG from the left and right wrist extensor as well as the EEG recorded over the left and right sensorimotor cortex are analyzed. Bilateral postural tremor was recorded for 250 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.

In Fig. 5, results of the partial directed coherence analysis for the two EMG and the two EEG channels are shown. On the diagonal the spectra of the processes are given. The tremor frequency indicated by the sharp peak in the left EMG-spectrum is almost 5 Hz. The tremor amplitude is lower for the right wrist extensor and its frequency is slightly higher compared to the left wrist extensor. Significant partial directed coherences at the corresponding tremor frequencies are detected for the direction from the left EMG to the right, contralateral EEG, from the right EEG to the left EMG, and from the left EEG to the right EMG. All remaining partial directed coherences between EEG and EMG are non-significant at the corresponding frequencies of the tremor. Caused by the property of the significance level to be valid only pointwise, partial directed coherence should only be evaluated at the tremor frequency.

However, since the EEG recordings are mutually influencing each other over a range of frequencies, we consider this result as significant even if we do not apply a correction for multiple testing. Actually, this weak interaction indicating an interhemispheric coupling is not unexpected, as shown for bilaterally activated essential tremor by means of bivariate coherence analysis (Hellwig et al., 2003). Moreover, there are no significant partial directed coherences between both EMG recordings. The partial directed coherence indicating a causal influence from the left EMG to the right EEG is much higher than the partial directed coherence indicating a causal influence from the right EEG to the left EMG (Fig. 5). This might be caused by the signal-to-noise ratios of the EEG and the EMG. While the EMG has hardly any noise influence, the EEG is characterized by a poor signal-to-noise ratio. The scalp EEG contains information

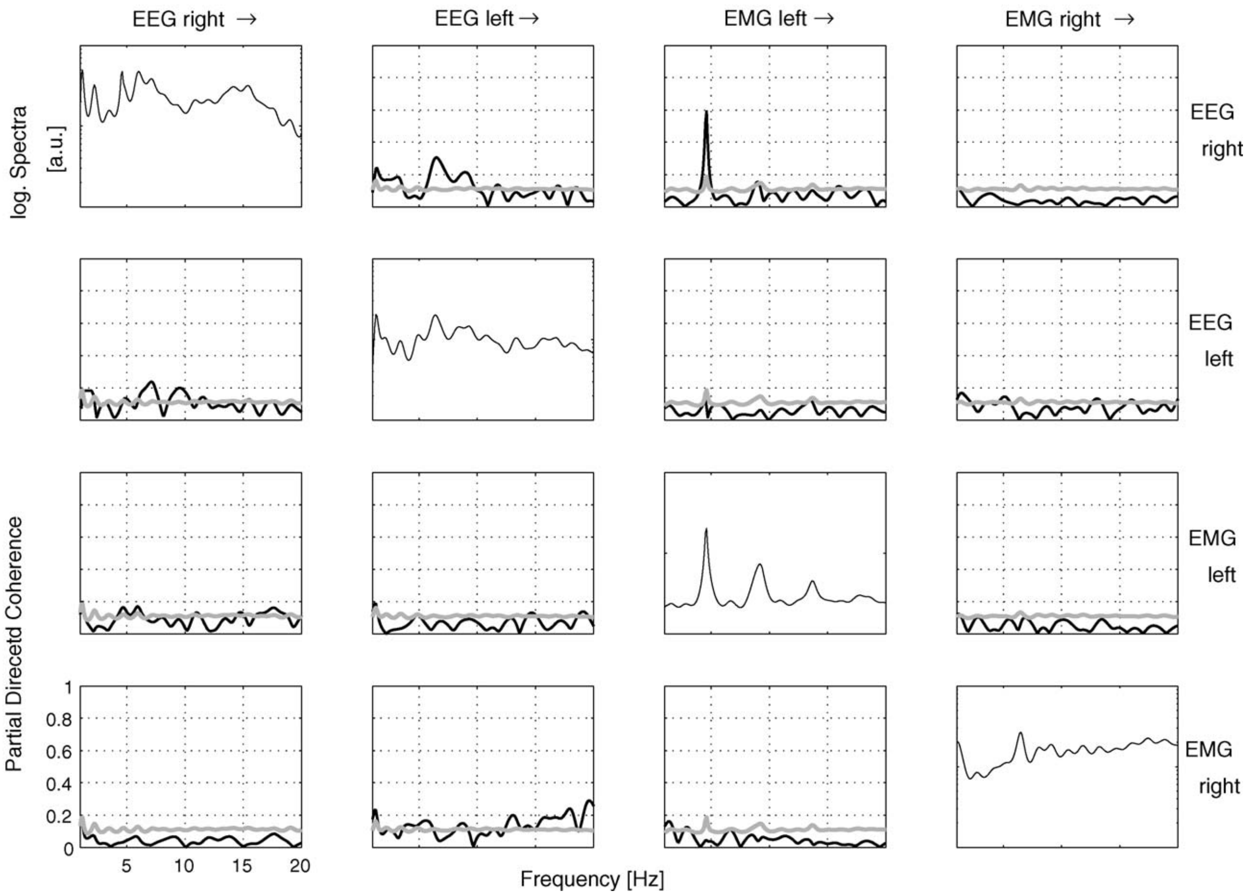

Fig. 5. Partial directed coherences for EEG- and EMG-recordings of one patient suffering from essential tremor. The corresponding 5%-significance levels are indicated by gray lines. Significant partial directed coherences at the corresponding frequencies of the tremor are found between EEG left → EMG right, EEG right → EMG left, and EMG left → EEG right indicating a directed influence between these recordings. Both EEG recordings are mutually influencing each other over a range of frequencies around the tremor frequencies. Neither ipsilateral interdependencies nor an interaction between both muscles are significant at the tremor frequencies.

from rather extended cortical areas and is not selective about tremor-related information from the motor cortex. Therefore, a high partial directed coherence from the EEG to the EMG is not likely to be detected. Similarly, a causal influence from the right EMG to the left EEG might be absent due to the weak tremor amplitude.

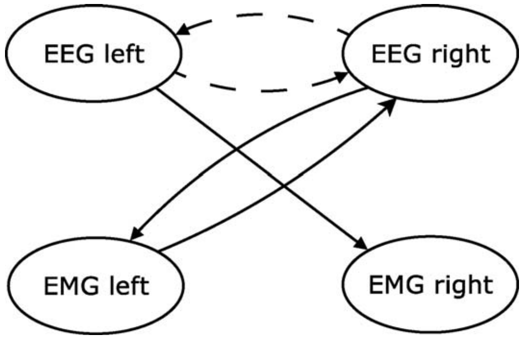

Fig. 6. Graph for partial directed coherence analysis of the tremor application (cf. Fig. 5). The arrows indicate a direct and directed interrelation at the tremor frequency. The dashed arrows indicate partial directed coherence between the EEGs, which are significant over a range of frequencies close to the tremor frequency.

The graph summarizing the results of partial directed coherence analysis is presented in Fig. 6. The influences between both EEGs are marked by dashed arrows. This is due to the fact, that partial directed coherence is not significant exactly at the tremor frequency but over a range of frequencies close to the tremor frequency. Since causal influences from both EEGs to the corresponding contralateral EMGs are present, participation of the motor cortex in tremor generation is strongly indicated. Moreover, there is also a significant partial directed coherence from the EMG to the contralateral EEG at the tremor frequency. This corresponds to a feedback from the muscles to the somatosensory cortex. For this patient, unexpected ipsilateral interrelations are not detected by partial directed coherence analysis.

5. Conclusion

Partial directed coherence is a powerful analysis technique to detect causal influences in multivariate stochastic systems with respect to Granger-causality. We illustrated the problem of inference based on partial directed coherence without a significance level, especially if a high model order is required. In order to draw reliable conclusions in applications, we analyzed the statistical properties and analytically derived a significance level.

The performance of the proposed significance level has been shown by means of linear stochastic as well as non-linear stochastic model systems. Even in application to the non-linear stochastic van der Pol system, partial directed coherence in combination with the significance level is able to detect causal influences and to distinguish direct from indirect ones.

The significance level allows for more freedom in the choice of the fitted VAR-process order. This has been demonstrated by the analysis of a VAR[4]-process, in which the order of the fitted process has been chosen to be $p = 200$. In spite of highly variable estimates for partial directed coherence, the correct interrelation structure could be inferred using the significance level. The order of the fitted process should not be smaller than the true order, but a higher order is unproblematic in combination with the significance level.

Choosing a high model order requires a large sample size for a reliable estimation of the parameters and the significance level. In general, the sample size required strongly depends on the system investigated and the corresponding model order. Especially if only a small sample size is available in applications, we recommend a tailored study on simulated data sets in advance.

Furthermore, we have presented an exemplary application to EEG and EMG data from a patient suffering from essential tremor. Using partial directed coherence in combination with the significance level allows to detect causal influences between EEG and EMG recordings in essential tremor and provides thus closer insights into the tremor generating mechanisms. The significance level introduced for partial directed coherence provides a rigorous framework for statistical inference of causal influences in general enabling a more widespread use of partial directed coherence in neuroscience applications.

Acknowledgements

This work was supported by the German Science Foundation (Ti315/2-1) and by the German Federal Ministry of Education and Research (BMBF grant 01GQ0420).

Appendix A

In this appendix, we derive the asymptotic distribution of $|\hat{\mathbf{A}}_{ij}(\omega)|^2$ under the hypothesis of $|\mathbf{A}_{ij}(\omega)|^2 = 0$. We first note that for $p = 1$ we have $|\hat{\mathbf{A}}_{ij}(\omega)|^2 = \hat{\mathbf{a}}_{ij}(1)^2$ and $C_{ij}(\omega) = \mathbf{\Sigma}_{ii} \boldsymbol{H}_{jj}(1,1) = \mathrm{var}(\hat{\mathbf{a}}_{ij}(1))$, which implies that the ratio $N|\hat{\mathbf{A}}_{ij}(\omega)|^2/C_{ij}(\omega)$ is χ^2-distributed with one degree of freedom.

For $p \geq 2$, we can write $N\,|\hat{\mathbf{A}}_{ij}(\omega)|^2$ as $\mathbf{X}'_{ij}(\omega)\mathbf{X}_{ij}(\omega)$, where

$$\mathbf{X}_{ij}(\omega) = \begin{pmatrix} \sqrt{N}\mathrm{Re}\hat{\mathbf{A}}_{ij}(\omega) \\ \sqrt{N}\mathrm{Im}\hat{\mathbf{A}}_{ij}(\omega) \end{pmatrix}. \tag{A.1}$$

Since the real and imaginary part of $\hat{\mathbf{A}}_{ij}(\omega)$ are linear functions in the parameter estimates $\hat{\mathbf{a}}_{ij}(1), \ldots, \hat{\mathbf{a}}_{ij}(p)$, $\mathbf{X}_{ij}(\omega)$ is asymptotically normally distributed with mean zero and covariance matrix

$$\mathbf{V}_{ij}(\omega) = \sum_{k,l=1}^{p} \mathbf{H}_{jj}(k,l)\,\mathbf{\Sigma}_{ii} \times \begin{pmatrix} \cos(k\omega)\cos(l\omega) & \cos(k\omega)\sin(l\omega) \\ \sin(k\omega)\cos(l\omega) & \sin(k\omega)\sin(l\omega) \end{pmatrix}. \tag{A.2}$$

For $\omega \neq 0 \bmod \pi$, the matrix $\mathbf{V}_{ij}(\omega)$ is positive definite. To see this, let

$$\mathbf{M}(\omega) = \begin{pmatrix} \cos(\omega) & \sin(\omega) \\ \vdots & \vdots \\ \cos(p\omega) & \sin(p\omega) \end{pmatrix}. \tag{A.3}$$

Then for $\boldsymbol{x} \in \mathbb{R}^2$ we have

$$\boldsymbol{x}'\mathbf{V}_{ij}(\omega)\boldsymbol{x} = \mathbf{\Sigma}_{ii}(\mathbf{M}(\omega)\boldsymbol{x})'\mathbf{H}^{(j)}\mathbf{M}(\omega)\boldsymbol{x} = 0 \tag{A.4}$$

where $\mathbf{H}^{(j)}$ is the $p \times p$ matrix with entries $\mathbf{H}^{(j)}_{k,l} = \mathbf{H}_{jj}(k,l)$ for $k, l = 1, \ldots, p$. Since $\mathbf{H}^{(j)}$ is positive definite and $\mathbf{M}(\omega)$ has rank 2 if and only if $p \geq 2$ and $\omega \neq 0 \bmod \pi$, we have $\boldsymbol{x}'\mathbf{V}_{ij}(\omega)\boldsymbol{x} = 0$ if and only if $\boldsymbol{x} = 0$, which proves the positive definiteness of $\mathbf{V}_{ij}(\omega)$. Consequently, $\mathbf{V}_{ij}(\omega)$ can be factorized as $\mathbf{V}_{ij}(\omega) = \mathbf{Q}_{ij}(\omega)\mathbf{D}_{ij}(\omega)\mathbf{Q}'_{ij}(\omega)$ where $\mathbf{Q}_{ij}(\omega)$ is some orthogonal matrix and $\mathbf{D}_{ij}(\omega)$ is the diagonal matrix of the eigenvalues of $\mathbf{V}_{ij}(\omega)$. Then $\mathbf{Y}_{ij}(\omega) = \mathbf{D}_{ij}(\omega)^{-1/2}\mathbf{Q}'(\omega)\mathbf{X}_{ij}(\omega)$ is asymptotically bivariately standard normally distributed, from which it follows that

$$N\,|\mathbf{A}_{ij}(\omega)|^2 = \mathbf{X}'_{ij}(\omega)\mathbf{X}_{ij}(\omega) = \mathbf{Y}'_{ij}(\omega)\,\mathbf{D}_{ij}(\omega)\,\mathbf{Y}_{ij}(\omega) \tag{A.5}$$

has asymptotically the same distribution as

$$\mathbf{Z}'\mathbf{D}_{ij}(\omega)\mathbf{Z} = \mathbf{D}_{11,ij}(\omega)\mathbf{Z}_1^2 + \mathbf{D}_{22,ij}(\omega)\mathbf{Z}_2^2, \tag{A.6}$$

where $\mathbf{Z} = (\mathbf{Z}_1, \mathbf{Z}_2)'$ has a bivariate standard normal distribution. The two eigenvalues $\mathbf{D}_{11,ij}(\omega)$ and $\mathbf{D}_{22,ij}(\omega)$ of $\mathbf{V}_{ij}(\omega)$ are in general not equal. Calculation of the eigenvalues of $\mathbf{V}_{ij}(\omega)$

$$\det \begin{pmatrix} \mathbf{V}_{11,ij}(\omega) - \mathbf{D}_{kk,ij}(\omega) & \mathbf{V}_{12,ij}(\omega) \\ \mathbf{V}_{21,ij}(\omega) & \mathbf{V}_{22,ij}(\omega) - \mathbf{D}_{kk,ij}(\omega) \end{pmatrix} \stackrel{!}{=} 0, \quad k = 1, 2 \tag{A.7}$$

leads to

$$\mathbf{D}^2_{kk,ij}(\omega) - \mathbf{D}_{kk,ij}(\omega)(\mathbf{V}_{11,ij}(\omega) + \mathbf{V}_{22,ij}(\omega)) + \det \mathbf{V}_{ij}(\omega) \stackrel{!}{=} 0. \tag{A.8}$$

The eigenvalues $\mathbf{D}_{kk,ij}(\omega)$ satisfy

$$\frac{\mathbf{D}_{kk,ij}(\omega)}{C_{ij}(\omega)} = \frac{1}{2} \pm \sqrt{\frac{1}{4} - \frac{\det \mathbf{V}_{ij}(\omega)}{C^2_{ij}(\omega)}} > 0 \tag{A.9}$$

where $C_{ij}(\omega) = \mathbf{V}_{11,ij}(\omega) + \mathbf{V}_{22,ij}(\omega)$ is given by the constant in Eq. (13). It follows that $N\,|\mathbf{A}_{ij}(\omega)|^2/C_{ij}(\omega)$ is the weighted average of two independent χ^2-distributed random variables with one degree of freedom because $\mathbf{Z}_1$ and $\mathbf{Z}_2$ are standard normally distributed. Only if the two eigenvalues of $\mathbf{V}_{ij}(\omega)$ were identical, we would obtain a χ^2-distribution with two degrees of freedom scaled by a factor 1/2.

Finally, we note that for $p \geq 2$ and $\omega = 0 \bmod \pi$, we have $\mathbf{X}_{ij,2} = \mathrm{Im}\,\hat{\mathbf{A}}_{ij}(\omega) = 0$, while $\mathbf{X}_{ij,1}$ is asymptotically normally distributed with mean zero and variance $\mathbf{V}_{11,ij}(\omega) = C_{ij}(\omega)$. It follows that $N|\hat{\mathbf{A}}_{ij}(\omega)|^2/C_{ij}(\omega)$ is χ^2-distributed with one degree of freedom.

References

Akaike H. Information theory and an extension of the maximum likelihood principle. In: Petrov B, Csaki F, editors. Second International Symposium on Information Theory. Budapest: Akademiai Kiado; 1973.

Baccala L, Sameshima K. Partial directed coherence: a new concept in neural structure determination. Biol Cybern 2001;84:463–74.

Bloomfield P. Fourier Analysis of Time Series: An Introduction. New York: John Wiley & Sons; 1976.

Boccaletti S, Kurths J, Osipov G, Valladares D, Zhou C. The synchronization of chaotic systems. Phys Rep 2002;366:1–101.

Brillinger D. Time Series: Data Analysis and Theory. San Francisco: Holden-Day, Inc; 1981.

Brillinger D. Remarks concerning graphical models for time series and point processes. Revista de Econometrica 1996;16:1–23.

Dahlhaus R, Eichler M. Causality and Graphical Models for time series. In: Green P, Hjort N, Richardson S, editors. Highly Structured Stochastic Systems. Oxford: Oxford University Press; 2003.

Dahlhaus R. Graphical interaction models for multivariate time series. Metrika 2000;51:157–72.

Dahlhaus R, Eichler M, Sandkühler J. Identification of synaptic connections in neural ensembles by graphical models. J Neurosci Methods 1997;77:93–107.

Eichler M. Granger causality and path diagrams for multivariate time series, J. Econometrics, in press.

Eichler M. Graphical modelling of multivariate time series. University of Heidelberg: Preprint; 2001.

Eichler M. A graphical approach for evaluating effective connectivity in neural systems. Philos Transact R Soc B 2005;360:953–67.

Eichler M, Dahlhaus R, Sandkühler J. Partial correlation analysis for the identification of synaptic connections. Biol Cybern 2003;89:289–302.

Goebel R, Roebroeck A, Kim DS, Formisano E. Investigating directed cortical interactions in time-resolved fMRI data using vector autoregressive modeling and granger causality mapping. Magn Reson Imaging 2003;21:1251–61.

Granger J. Investigating causal relations by econometric models and cross-spectral methods. Econometrica 1969;37:424–38.

Grosse P, Cassidy MJ, Brown P. EEG-EMG, MEG-EMG and EMG-EMG frequency analysis: physiological principals and clinical applications. Clin Neurophysiol 2002;113:1523–31.

Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning. New York: Springer; 1996.

Hellwig B, Häußler S, Lauk M, Köster B, Guschlbauer B, Kristeva-Feige R, Timmer J, Lücking CH. Tremor-correlated cortical activity detected by electroencephalography. Electroencephalogr Clin Neurophysiol 2000;111:806–9.

Hellwig B, Häußler S, Schelter B, Lauk M, Guschlbauer B, Timmer J, Lücking CH. Tremor correlated cortical activity in essential tremor. Lancet 2001;357:519–23.

Hellwig B, Schelter B, Guschlbauer B, Timmer J, Lücking CH. Dynamic synchronisation of central oscillators in essential tremor. Clin Neurophysiol 2003;114:1462–67.

Hesse W, Möller E, Arnold M, Schack B. The use of time-variant EEG Granger causality for inspecting directed interdependencies of neural assemblies. J Neurosci Methods 2003;124:27–44.

Louis E, Ford B, Wendt K, Cameron G. Clinical characteristics of essential tremor: data from a community-based study. Mov Disord 1998;13:803–8.

Lütkepohl H. Introduction to Multiple Time Series Analysis. Springer; 1993.

Pikovsky A, Rosenblum M, Kurths J. Synchronization—a universal concept in Non-linear Sciences. Cambridge: Cambridge University Press; 2001.

Rissanen J. Modeling by shortest data description. Automatica 1978;14:465–71.

Rissanen J. A universal prior for integers and estimation by minimum description length. Ann Stat 1983;11:416–31.

Rosenblum M, Pikovsky A. Detecting direction of coupling in interacting oscillators. Phys Rev E 2001;64:045202.

Sameshima K, Baccala L. Using partial directed coherence to describe neuronal ensemble interactions. J Neurosci Methods 1999;94:93–103.

Schnider S, Kwong R, Lenz F, Kwan H. Detection of feedback in the central nervous system using system identification techniques. Biol Cybern 1989;60:203–12.

Schwartz G. Estimating the dimension of a model. Ann Stat 1979;6:461–64.

Smirnov D, Bezruchko B. Estimation of interaction strength and direction from short and noisy time series. Phys Rev E 2003;68:046209.

Tass P, Rosenblum MG, Weule J, Kurths J, Pikovsky A, Volkmann J, Schnitzler A, Freund HJ. Detection of n:m phase locking from noisy data: application to magnetoencephalography. Phys Rev Lett 1998;81:3291–95.

Timmer J. Modeling noisy time series: physiological tremor. Int J Bif Chaos 1998;8:1505–16.

Timmer J, Lauk M, Pfleger W, Deuschl G. Cross-spectral analysis of physiological tremor and muscle activity I Theory and application to unsynchronized EMG. Biol Cybern 1998;78:349–57.

van der Pol B. On oscillation-hysteresis in a simple triode generator. Phil Mag 1922;43:700–19.

Volkmann J, Joliot M, Mogilner A, Ioannides AA, Lado F, Fazzini E, Ribary U, Llinás R. Central motor loop oscillations in Parkinsonian resting tremor revealed by magnetoencephalography. Neurology 1996;46:1359–70.