Biol Cybern (2006) 94: 469–482
DOI 10.1007/s00422-006-0062-z

ORIGINAL PAPER

Biological Cybernetics

Michael Eichler

On the evaluation of information flow in multivariate systems by the directed transfer function

Received: 18 October 2005 / Accepted 14 February 2006 / Published online: 17 March 2006

Abstract The directed transfer function (DTF) has been proposed as a measure of information flow between the components of multivariate time series. In this paper, we discuss the interpretation of the DTF and compare it with other measures for directed relationships. In particular, we show that the DTF does not indicate multivariate or bivariate Granger causality, but that it is closely related to the concept of impulse response function and can be viewed as a spectral measure for the total causal influence from one component to another. Furthermore, we investigate the statistical properties of the DTF and establish a simple significance level for testing for the null hypothesis of no information flow.

Keywords Directed transfer function · Granger causality · Impulse response function · Transfer function · Multivariate time series · Significance test

1 Introduction

The identification of information flow and causal influences in complex multi-variable systems is an important problem in neuroscience as well as in many other scientific areas. For instance, signals reflecting neural activity such as electroencephalographic (EEG) or local field potential (LFP) recordings have been used to learn patterns of interactions between brain areas that are activated during certain tasks and thus to improve our understanding of neural processing of information (e.g., Schack et al. 1999; Liang et al. 2000).

One commonly used approach for inferring causal relationships from such temporally structured data is based on vector autoregressive (VAR) models and the concept of Granger causality. This concept of causality, introduced by Granger (1969), is based on the common sense perception that causes always precede their effects in time: if one time series causes another series, knowledge of the former series should help to predict future values of the latter series. Although Granger (1969, 1980) always stressed the need to include all relevant information in an analysis to avoid so-called spurious causalities, much of the literature on Granger causality has been concerned with the analysis of relationships between only two time series. Consequently, relationships among multiple time series are still quite frequently investigated using bivariate Granger causality, that is, analysing pairs of time series separately (e.g., Goebel et al. 2003; Hesse et al. 2003; Brovelli et al. 2004).

In order to determine the directional influences between the components in a multivariate system by a full multivariate frequency-domain based method, Kamiński and Blinowska (1991) introduced the directed transfer function (DTF). The usefulness of the DTF method has been demonstrated in many articles (e.g., Kamiński et al. 2001; Veeramani et al. 2003; Blinowska et al. 2004; Kamiński 2005), and it has been applied, for example, to localize epileptic loci (Franaszczuk and Bergey 1998), to determine LFP propagation between brain structures of animals in different behavioural states (Korzeniewska et al. 1997), to investigate EEG activity propagation in different sleep stages (Kamiński et al. 1997), and to study epileptogenesis (Medvedev and Willoughby 1999). However, the DTF is defined in terms of the spectral transfer function, which makes its interpretation in terms of causal influences between the components difficult. In particular, its relation to the concept of Granger causality remains unclear although there have been attempts to establish such a relationship. Kamiński et al. (2001) claimed that the DTF is equivalent to the concept of bivariate Granger causality. On the other hand, Kuś et al. (2004), Blinowska et al. (2004) and Kamiński (2005) show that DTF differs from bivariate measures of Granger causality; alternatively, they suggest that DTF interprets Granger causality in a multivariate sense.

The aim of the present paper is to enhance the interpretability of the DTF both theoretically and practically. Firstly, we discuss the interpretation of the DTF as a measure of information flow between processes and, in particular, clarify how it relates to the concepts of bivariate and multivariate Granger causality. Secondly, we investigate the statistical properties of

M. Eichler
Institut für Angewandte Mathematik, Universität Heidelberg,
Im Neuenheimer Feld 294, 69120 Heidelberg, Germany
E-mail: eichler@statlab.uni-heidelberg.de

the DTF. So far, there exists only a simulation-based approach for the assessment of the statistical significance of the DTF, which makes use of the method of surrogate data (Kamiński et al. 2001). We derive the asymptotic distribution of the DTF under the null hypothesis of no information flow and propose a pointwise significance level that can be easily implemented.

2 Information flow in multivariate systems

2.1 Moving average representation and impulse response function

Let $X(t) = \big(X_1(t), \ldots, X_d(t)\big)'$ denote a multivariate time series from d data channels. For the theoretical discussion in this section, we assume that $X = \{X(t)\}$ is a weakly stationary and purely non-deterministic multivariate time series with mean zero. Then X has an infinite moving average representation

$$X(t) = \sum_{u=0}^{\infty} \mathbf{b}(u)\, e(t-u), \tag{1}$$

where $\mathbf{b}(u)$ is a square-summable sequence of $d \times d$ matrices, $\mathbf{b}(0) = \mathbf{I}$ is the identity matrix and $e = \{e(t)\}$ is a white noise process with mean zero and non-singular covariance matrix Σ (e.g., Brockwell and Davis 1991). The sequence $\mathbf{b}(u)$ is called the *impulse response function* of the linear system given by Eq. 1 and describes how the output series X of the system is related to the input series e. More precisely, the coefficient $\mathbf{b}_{ij}(u)$ measures the response of the variable X_i at time t to a random shock of unit size at variable X_j at time $t-u$.

2.2 Vector autoregressions and Granger causality

In applications in neuroscience, time series are more commonly evaluated by use of VAR models, which represent a time series X at time t in terms of its previous values $X(t-u)$, $u > 0$ and a random component $e(t)$. More precisely, let X be a multivariate time series with moving average representation (Eq. 1). Then the spectral density matrix $\mathbf{f}(\lambda)$ of X exists for almost all frequencies $\lambda \in [-\pi, \pi]$ and is given by

$$\mathbf{f}(\lambda) = (2\pi)^{-1}\, \mathbf{B}(\lambda)\, \Sigma\, \mathbf{B}(\lambda)^*, \tag{2}$$

where

$$\mathbf{B}(\lambda) = \sum_{u=0}^{\infty} \mathbf{b}(u)\, \mathrm{e}^{-\mathrm{i}\lambda u} \tag{3}$$

is the Fourier transform of the impulse response function $\mathbf{b}(u)$ and $\mathbf{A}^*$ denotes the conjugate transpose of the matrix $\mathbf{A}$. In the sequel, we additionally assume that the spectral density matrix $\mathbf{f}(\lambda)$ satisfies the boundedness condition

$$c\,\mathbf{I} \le \mathbf{f}(\lambda) \le c'\,\mathbf{I} \quad \text{for all } \lambda \in [-\pi, \pi] \tag{4}$$

and some constants $c' > c > 0$. Here, $\mathbf{A} \le \mathbf{B}$ for matrices $\mathbf{A}$ and $\mathbf{B}$ indicates that $\mathbf{B} - \mathbf{A}$ is non-negative definite. Then the time series X is said to be invertible and has a VAR representation

$$X(t) = \sum_{u=1}^{\infty} \mathbf{a}(u)\, X(t-u) + e(t), \tag{5}$$

where $\mathbf{a}(u)$ is again a square-summable sequence of $d \times d$ matrices and $e = \{e(t)\}$ is the white noise process in Eq. 1.

The autoregressive representation of X is closely related to the concept of (linear) *Granger causality*, which is a fundamental tool to describe the causal relationship between time series. According to the original definition of Granger (1969), one time series X causes another series Y, if the one-step ahead prediction of Y based on the past of Y and that of any relevant auxiliary variables Z can be improved (in the mean square sense) by adding the past of X to the set of predictor variables. In the context of multivariate time series X with autoregressive representation (Eq. 5), this leads to the following equivalent definition: one component X_i linearly *Granger-causes* another component X_j if the coefficients $\mathbf{a}_{ji}(u)$ do not vanish uniformly for all lags u (e.g., Sims 1980; Hsiao 1982; Toda and Philipps 1993; Hayo 1999; Dufour and Renault 1998). Thus, linear Granger causality describes the direct linear effect of one component X_i on another component X_j. In this paper, we do not consider any forms of non-linear Granger causality and, for simplicity, will use the term Granger causality in the restricted meaning of linear Granger causality.

One drawback of the concept of Granger causality as a measure for causal relationships among multiple time series is the fact that it depends on the series X that is available for the analysis. It is well known that omission of important relevant variables can lead to so-called spurious causalities and, thus, to a wrong identification of the underlying causal structure. Despite this fact, much of the literature on Granger causality has been concerned with the analysis of relationships between only two time series (or two vector time series) and, as a consequence, relationships among multiple time series are still quite frequently investigated using bivariate Granger causality; examples involving EEG signals or time-resolved functional magnetic resonance imaging (fMRI) recordings can be found in Kamiński et al. (2001), Goebel et al. (2003) and Hesse et al. (2003).

For the purposes of this paper, we will therefore also consider the concept of *bivariate Granger causality*. To this end, let X be a d-variate weakly stationary time series satisfying Eq. 5 and let X_i and X_j be two components of X. Then the bivariate subprocess (X_i, X_j) is again a weakly stationary time series and has an autoregressive representation

$$\begin{aligned} X_i(t) &= \textstyle\sum_{u=1}^{\infty} \tilde{\mathbf{a}}_{ii}(u)\, X_i(t-u) \\ &\quad + \textstyle\sum_{u=1}^{\infty} \tilde{\mathbf{a}}_{ij}(u)\, X_j(t-u) + \tilde{e}_i(t), \\ X_j(t) &= \textstyle\sum_{u=1}^{\infty} \tilde{\mathbf{a}}_{ji}(u)\, X_i(t-u) \\ &\quad + \textstyle\sum_{u=1}^{\infty} \tilde{\mathbf{a}}_{jj}(u)\, X_j(t-u) + \tilde{e}_j(t), \end{aligned} \tag{6}$$

where $\tilde{e}(t) = (\tilde{e}_i(t), \tilde{e}_j(t))$ is a white noise process with some non-singular covariance matrix $\tilde{\Sigma}$. Based on this

representation, we say that the series X_i *bivariately Granger-causes* the series X_j if the coefficients $\tilde{\mathbf{a}}_{ji}(u)$ are not zero uniformly for all lags u.

To illustrate the difference between the multivariate and bivariate concepts of Granger causality, we consider the following trivariate system. Let

$$\begin{aligned} X_1(t) &= \alpha\, X_2(t-1) + e_1(t), \\ X_2(t) &= \beta\, X_3(t-1) + e_2(t), \\ X_3(t) &= e_3(t), \end{aligned} \tag{7}$$

where $e_v(t)$, $v = 1, 2, 3$, are independent and identically normally distributed with mean zero and variance σ^2. Then X_3 Granger-causes X_2, which in turn Granger-causes X_1, whereas X_3 does not Granger-cause X_1 (all with respect to the full trivariate series $X_{\{1,2,3\}}$). On the other hand, in a bivariate autoregressive representation of X_1 and X_3, we have

$$\begin{aligned} X_1(t) &= \alpha\,\beta\, X_3(t-2) + \tilde{e}_1(t), \\ X_3(t) &= \tilde{e}_3(t) \end{aligned}$$

with $\tilde{e}_1(t) = e_1(t) + \alpha\, e_2(t-1)$ and $\tilde{e}_3(t) = e_3(t)$. From this representation, we find that X_3 bivariately Granger-causes X_1.

In general, the relationship between the two notions of multivariate and bivariate Granger causality is more complicated than in this example and, in most cases, an analytic derivation of the bivariate representation would be very difficult to obtain. However, Eichler (2005) has presented graphical conditions for relating the two concepts to each other based on a graphical approach for describing Granger-causal relationships in multiple time series (e.g., Eichler 2002).

In practice, VAR models of finite order p – that is, with $\mathbf{a}(u) = 0$ for all lags $u > p$ – are used to approximate the time series of interest. Such models are particularly suitable for describing systems with some kind of oscillating behaviour. For EEG signals, the validity of the VAR modelling approach has been demonstrated, for example, by Franaszczuk et al. (1985), Pijn et al. (1991, 1997), Stam et al. (1999), Blinowska and Malinowski (1991) and Achermann et al. (1994); applications to time-resolved fMRI recordings are shown, for example, in Goebel et al. (2003), Harrison et al. (2003) and Valdés-Sosa (2004). For the general theory on VAR time series models, we refer to Lütkepohl (1993) and Reinsel (2003).

2.3 Directed transfer function

Many biomedical time series such as EEG signals are characterized in terms of their frequency properties. It is therefore important to examine the relationships among multiple time series also in the frequency domain. The frequency-domain analysis of weakly stationary time series X is based on the spectral representation of X, which is given by

$$X(t) = \int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}\lambda t}\, \mathrm{d}Z_X(\lambda) = \int_{-\pi}^{\pi} \mathbf{B}(\lambda)\, \mathrm{e}^{\mathrm{i}\lambda t}\, \mathrm{d}Z_e(\lambda), \tag{8}$$

where $Z_X(\lambda)$ and $Z_e(\lambda)$ are two random processes on $[-\pi, \pi]$ with mean zero and orthogonal increments (e.g., Brockwell and Davis 1991). In this representation, the complex-valued random increments $\mathrm{d}Z_X(\lambda)$ and $\mathrm{d}Z_e(\lambda)$ indicate the frequency components of the time series X and the white noise process e, respectively, at frequency λ. From Eq. 8, we find that these are related by

$$\mathrm{d}Z_X(\lambda) = \mathbf{B}(\lambda)\, \mathrm{d}Z_e(\lambda). \tag{9}$$

Thus, the complex-valued function $\mathbf{B}(\lambda)$ describes how the frequency components of the input process – here the white noise process e – are transformed by the linear filter in Eq. 1 to the frequency components of the output process X. In particular, the entry $\mathbf{B}_{ij}(\lambda)$ measures the response of variable X_i to sinusoidal random shocks of frequency λ at variable X_j. The function $\mathbf{B}(\lambda)$ is therefore called the *transfer function* (e.g., Brockwell and Davis 1991) or the *frequency response function* (e.g., Chatfield 2003) of X.

From Eq. 9, it is clear that the transfer function $\mathbf{B}(u)$ can be used as a measure for directional information flow in multivariate time series. Kamiński and Blinowska (1991) proposed the DTF, which is a normalized version of the transfer function. It is given by

$$\gamma_{ij}^2(\lambda) = \frac{|\mathbf{B}_{ij}(\lambda)|^2}{\sum_k |\mathbf{B}_{ik}(\lambda)|^2} \tag{10}$$

and describes the ratio of the influence of component X_j on component X_i to all the influences on component X_i. Due to the normalization, the DTF takes values in [0, 1]. For the comparison of the information flow for different target processes or between different experiments, Kamiński et al. (2001) and Kamiński (2005) suggested also a non-normalized version of the DTF, which is given by

$$\theta_{ij}^2(\lambda) = |\mathbf{B}_{ij}(\lambda)|^2 \tag{11}$$

and basically measures the amplitude of the complex-valued transfer function $\mathbf{B}_{ij}(\lambda)$.

We note that the transfer function $\mathbf{B}(\lambda)$ and hence the DTF $\theta_{ij}^2(\lambda)$ can be computed also from the coefficients in the autoregressive representation (Eq. 5). Let

$$\mathbf{A}(\lambda) = \sum_{u=1}^{\infty} \mathbf{a}(u)\, \mathrm{e}^{-\mathrm{i}\lambda u}$$

be the Fourier transform of the autoregressive coefficients $\mathbf{a}(u)$ and define $\bar{\mathbf{A}}(\lambda) = \mathbf{I} - \mathbf{A}(\lambda)$. Then $\mathbf{B}(\lambda)$ and $\mathbf{A}(\lambda)$ are related by

$$\mathbf{B}(\lambda) = \bar{\mathbf{A}}(\lambda)^{-1} = \big(\mathbf{I} - \mathbf{A}(\lambda)\big)^{-1}.$$

Furthermore, expanding the inverse $\bar{\mathbf{A}}(\lambda)^{-1} = \big(\mathbf{I} - \mathbf{A}(\lambda)\big)^{-1}$ as a geometric series, we obtain

$$\begin{aligned} \mathbf{B}_{ij}(\lambda) = \mathbf{A}_{ij}(\lambda) &+ \sum_{k=1}^{d} \mathbf{A}_{ik}(\lambda)\, \mathbf{A}_{kj}(\lambda) \\ &+ \sum_{k_1,k_2=1}^{d} \mathbf{A}_{ik_1}(\lambda)\, \mathbf{A}_{k_1k_2}(\lambda)\, \mathbf{A}_{k_2j}(\lambda) + \ldots . \end{aligned} \tag{12}$$

This shows that the transfer function $\mathbf{B}_{ij}(\lambda)$ accumulates the information flow from direct pathways – measured by $\mathbf{A}_{ij}(\lambda)$ – as well as from indirect pathways via components $X_{k_1}, \ldots, X_{k_r}$.

Finally, we note that information flow among multiple time series can also be described directly by the entries in the matrix $\bar{\mathbf{A}}(\lambda)$. This leads to the *partial directed coherence* (PDC) introduced by Sameshima and Baccalá (1999) and Baccalá and Sameshima (2001). Using a different normalization from that of Kamiński and Blinowska (1991), they defined the PDC as

$$\pi_{ij}(\lambda) = \frac{|\bar{\mathbf{A}}_{ij}(\lambda)|}{\sqrt{\sum_k |\bar{\mathbf{A}}_{kj}(\lambda)|^2}}. \tag{13}$$

Alternatively, one might also consider $|\bar{\mathbf{A}}_{ij}(\lambda)|$ as a non-normalized version of the PDC.

3 Relationship between DTF and Granger causality

From the definition of the DTF, it is clear that the DTF $\gamma_{ij}^2(\lambda)$ is directly related to the transfer function $\mathbf{B}(\lambda)$ and, thus, to the impulse response function $\mathbf{b}(u)$ and the moving average representation of X. Similarly, it follows from Eq. 13 that the PDC $\pi_{ij}(\lambda)$ vanishes for all frequencies λ if and only if $\mathbf{a}_{ij}(u) = 0$ for all lags u. Consequently, the PDC can be viewed as a frequency-domain measure for multivariate Granger causality (see Sameshima and Baccalá 1999; Baccalá and Sameshima 2001).

In contrast, the relationship between the DTF and the notion of Granger causality remains unclear. For a bivariate time series $X = (X_1, X_2)$, it can be shown that

$$\theta_{21}^2(\lambda) = |\mathbf{B}_{21}(\lambda)|^2 = \frac{|\mathbf{A}_{21}(\lambda)|^2}{|\det \mathbf{A}(\lambda)|^2}.$$

Therefore, the DTF $\gamma_{21}^2(\lambda)$ vanishes if and only if X_1 does not Granger-cause X_2. However, for the general case of multivariate time series, there exists no similar relationship between the off-diagonal entries in $\mathbf{B}(\lambda)$ and in $\mathbf{A}(\lambda)$. Furthermore, the literature about the DTF provides conflicting views about this topic. For instance, Kamiński et al. (2001) claimed that the DTF can be interpreted in terms of bivariate Granger causality. On the other hand, Blinowska et al. (2004) and Kuś et al. (2004) argue against the use of bivariate measures of Granger causality and in favour of the DTF as a multivariate measure of Granger causality, and Kamiński (2005) suggest that the DTF can interpret Granger causality in a multivariate sense.

Part of this discussion about the role of DTF as a measure of Granger causality seems to be due to the fact that the concept of Granger causality is closely related to the time domain and the notion of predictability, and that it is not immediately clear how to describe this concept in the frequency domain. However, it should be noted that Granger causality in principle is a binary relation: one process Granger-causes another process, multivariately or bivariately, or it does not. Consequently, it is natural to require that any frequency-domain quantities that measure the Granger-causal effect of a process X_i on another process X_j vanish for all frequencies if and only if X_i does not Granger-cause X_j. Under this premise, we discuss the relationship between the DTF and the two concepts of multivariate and bivariate Granger causality. For illustration, we consider the following trivariate system. Suppose that $X = (X_1, X_2, X_3)$ is a weakly stationary time series with autoregressive representation

$$\begin{aligned} X_1(t) &= \alpha X_2(t-1) + \beta X_3(t-2) + e_1(t), \\ X_2(t) &= \gamma X_3(t-1) + e_2(t), \\ X_3(t) &= e_3(t), \end{aligned} \tag{14}$$

where $\{e_1(t)\}, \{e_2(t)\}$ and $\{e_3(t)\}$ are uncorrelated white noise processes with variances $\mathrm{var}\big(e_i(t)\big) = \sigma_i^2$, $i = 1, 2, 3$. The dependence structure of this process is also depicted by the path diagram in Fig. 1; it indicates that X_2 Granger-causes X_1 while X_3 Granger-causes both X_2 and X_1.

For the derivation of the DTF of X, we note that

$$\bar{\mathbf{A}}(\lambda) = \begin{pmatrix} 1 & -\alpha\, e^{\mathrm{i}\lambda} & -\beta\, e^{\mathrm{i}2\lambda} \\ 0 & 1 & -\gamma\, e^{\mathrm{i}\lambda} \\ 0 & 0 & 1 \end{pmatrix}.$$

Then inversion of $\bar{\mathbf{A}}(\lambda)$ yields

$$\mathbf{B}(\lambda) = \begin{pmatrix} 1 & \alpha\, e^{\mathrm{i}\lambda} & (\beta + \alpha\gamma)\, e^{\mathrm{i}2\lambda} \\ 0 & 1 & \gamma\, e^{\mathrm{i}\lambda} \\ 0 & 0 & 1 \end{pmatrix}. \tag{15}$$

Comparing Eqs. 14 and 15, we find that the DTF from X_2 to X_1 is non-zero if and only if X_2 Granger-causes X_1; the same holds for the DTF from X_3 to X_2. In contrast, the non-normalized DTF from X_3 to X_1 is given by

$$\theta_{13}^2(\lambda) = (\beta + \alpha\,\gamma)^2.$$

We consider the following two situations.

1. Suppose that $\beta = 0$. This implies that X_3 does not Granger-cause X_1, but the DTF still shows a non-zero effect of X_3 on X_1 if $\alpha\gamma \neq 0$. This is due to the fact that the DTF – unlike measures of multivariate Granger causality, which describe only the direct effect of X_3 on X_1 – also takes into account the indirect effect mediated by X_2 as indicated by Eq. 12.
2. Suppose that $\beta = -\alpha\,\gamma \neq 0$, that is, the direct and indirect effect of X_3 on X_1 cancel out. In this case, the total effect of X_3 on X_1 is zero and the DTF vanishes for all frequencies. On the other hand, X_3 has a non-zero direct effect on X_1 and, thus, X_3 Granger-causes X_1. The situation is depicted in Fig. 2a.

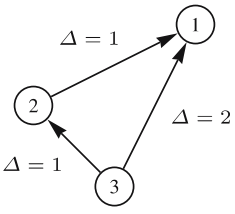

Fig. 1 Path diagram associated with the autoregressive representation of the process $X(t)$ in Eq. 14

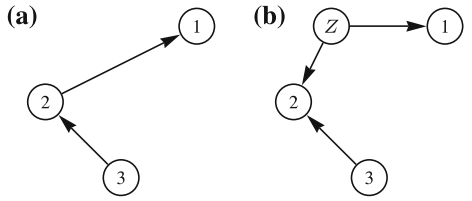

Fig. 2 Relation between DTF and multivariate Granger causality. **a** path diagram of three-dimensional process with $\theta^2_{13}(\lambda) \not\equiv 0$ and X_3 does not Granger-cause X_1. **b** path diagram of three-dimensional process with one latent variable Z for which $\theta^2_{13}(\lambda) \equiv 0$ and X_3 Granger-causes X_1 with respect to (X_1, X_2, X_3)

It could be argued that the second situation is negligible in practice since the condition $\beta = -\alpha\,\gamma$ describes only a null set in the parameter space of the VAR model (Eq. 14). However, such constraints arise naturally if latent variables are present – a problem that cannot be ruled out particularly for applications in neuroscience. As an example, we consider the trivariate process $X = (X_1, X_2, X_3)$ given by

$$\begin{aligned} X_1(t) &= \alpha_1 Z(t-2) + e_1(t), \\ X_2(t) &= \alpha_2 Z(t-1) + \gamma X_3(t-1) + e_2(t), \\ X_3(t) &= e_3(t), \end{aligned} \tag{16}$$

where $\{Z(t)\}$, $\{e_1(t)\}$, $\{e_2(t)\}$ and $\{e_3(t)\}$ are uncorrelated white noise processes with variance one. The structure is shown in Fig. 2b. In this system, the process Z represents an unobserved variable with a common effect on variables X_1 and X_2. Simple calculations show that the process X has an autoregressive representation of the form (Eq. 14) with coefficients $\alpha = \alpha_1\alpha_2/(1+\alpha_2^2)$ and $\beta = -\alpha\gamma$, that is, the observed process X fulfills the above constraint for all parameter values of the underlying process (X, Z).

In this example, the DTF indicates correctly that there is no information flow from X_3 to X_1, whereas measures of Granger causality wrongly detect a direct effect of X_3 on X_1. Situations of this kind are known as type I spurious causality (Hsiao 1982). Since variable Z is not available for predicting $X_1(t)$, the value of $X_2(t-1)$ serves as a proxy for $Z(t-2)$, while X_3 is used as a purifying variable to eliminate the noise in $X_2(t-1)$. The example suggests that the DTF—measuring only the total information flow from one variable to another – is not affected by such type I spurious causalities. This is, however, not true in general as we will see later.

For a discussion of the relationship between DTF and bivariate Granger causality, we consider again the trivariate process specified by Eq. 14. Simple evaluations show that the bivariate subprocess (X_1, X_2) has autoregressive representation

$$\begin{aligned} X_1(t) &= \left(\alpha + \frac{\beta\gamma\sigma_3^2}{\sigma_2^2 + \gamma^2\sigma_3^2}\right) X_2(t-1) + \tilde{e}_1(t), \\ X_2(t) &= \tilde{e}_2(t), \end{aligned} \tag{17}$$

where

$$\tilde{e}_1(t) = e_1(t) + \frac{\beta\sigma_2^2}{\sigma_2^2 + \gamma^2\sigma_3^2}\, e_3(t-2) - \frac{\beta\gamma\sigma_3^2}{\sigma_2^2 + \gamma^2\sigma_3^2}\, e_2(t-1)$$

and $\tilde{e}_2(t) = e_2(t) + \gamma\, e_3(t-1)$. Furthermore, calculations show that $\mathrm{cov}\left(\tilde{e}_1(t), \tilde{e}_2(s)\right) = 0$ for all t, s and hence that $\tilde{e}_1(t)$ and $\tilde{e}_2(t)$ are two uncorrelated white noise processes with mean zero and variances $\tilde{\sigma}_1^2 = \sigma_1^2 + \beta^2(1+\gamma^2)\sigma_2^2\sigma_3^2 / (\sigma_2^2 + \gamma^2\sigma_3^2)$ and $\tilde{\sigma}_2^2 = \sigma_2^2 + \gamma^2\sigma_3^2$, respectively. Similarly, we find that the autoregressive representation of the subprocess (X_1, X_3) is given by

$$\begin{aligned} X_1(t) &= \left(\alpha + \beta\gamma\right) X_3(t-2) + \tilde{e}_1(t), \\ X_3(t) &= \tilde{e}_3(t), \end{aligned} \tag{18}$$

where $\tilde{e}_1(t) = e_1(t) + \alpha e_2(t)$ and $\tilde{e}_3(t) = e_3(t)$. Obviously, $\tilde{e}_1(t)$ and $\tilde{e}_3(t)\big)$ are two uncorrelated white noise processes with mean zero and variances $\sigma_1^2 + \alpha^2\sigma_2^2$ and σ_3^2, respectively.

Comparing Eqs. 18 and 15, we find that X_2 bivariately Granger-causes X_1 if and only if the DTF from X_2 to X_1 is identical to zero. Like the DTF, the concept of bivariate Granger causality accumulates direct and indirect information flow from one variable to another. However, it does not eliminate effects due to confounding as shown by the bivariate representation in Eq. 17. We consider the following two situations.

1. Suppose that $\alpha = 0$. Then Eq. 17 implies that X_2 bivariately Granger-causes X_1 although X_2 has neither a direct nor an indirect effect on X_1 (see Fig. 3). The DTF from X_2 to X_1 indicates correctly that there is no information flow from X_2 to X_1.
2. Suppose that $\alpha = -\beta\gamma\sigma_3^2/(\sigma_2^2 + \gamma^2\sigma_3^2)$. In this case, the direct effect from X_2 on X_1 and the effect due to confounding by X_3 cancel out and X_2 does not bivariately Granger-cause X_1. On the other hand, both the DTF and the measures of multivariate Granger causality eliminate the common effect by X_3 and identify the causal link from X_2 to X_1.

Again, such parameter constraints may arise naturally in situations where the system is affected by latent variables. As an example, we consider the following system. Suppose that $X = (X_1, X_2, X_3)$ is a weakly stationary time series given by

$$\begin{aligned} X_1(t) &= \beta_1\, Z_1(t-2) + e_1(t), \\ X_2(t) &= \gamma_1\, Z_2(t-2) + e_2(t), \\ X_3(t) &= \beta_2\, Z_1(t-1) + \gamma_2\, Z_2(t-1) + e_3(t), \end{aligned} \tag{19}$$

where $\{Z_1(t)\}$, $\{Z_2(t)\}$, $\{e_1(t)\}$, $\{e_2(t)\}$ and $\{e_3(t)\}$ are uncorrelated white noise processes with variance one. Here, Z_1

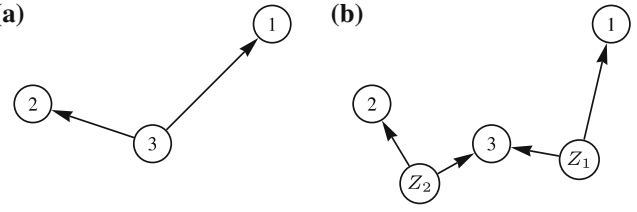

Fig. 3 Relation between DTF and bivariate Granger causality. **a** path diagram of three-dimensional process with $\theta^2_{12}(\lambda) \equiv 0$ and X_2 bivariately Granger-causes X_1. **b** path diagram of three-dimensional process with two latent variables Z_1 and Z_2 for which $\theta^2_{12}(\lambda) \not\equiv 0$ and X_2 does not bivariately Granger-cause X_1

and Z_2 are two unobserved latent variables. The structure of the system is depicted in Fig. 3b.

Straightforward, but lengthy algebraic manipulations show that $X = (X_1, X_2, X_3)$ has a trivariate autoregressive representation as in Eq.14 with coefficients

$$\alpha = -\gamma_1 \gamma_2 \beta_1 \beta_2/\tau,$$
$$\beta = \beta_1 \beta_2 (1 + \gamma_1^2)/\tau,$$
$$\gamma = \gamma_1 \gamma_2/\sigma_3^2,$$

where $\tau = 1 + \beta_2^2 + \gamma_1^2 + \gamma_2^2 + \beta_2^2 \gamma_1^2$, and variances $\sigma_3^2 = 1 + \beta_2^2 + \gamma_2^2$ and $\sigma_2^2 = 1 + \gamma_1^2 (1 + \beta_2^2)/\sigma_3^2$. Further calculations show that the system satisfies the constraint

$$\alpha = -\beta\gamma\sigma_3^2/(\sigma_2^2 + \gamma^2\sigma_3^2)$$

for all values of the parameters β_1, β_2, γ_1 and γ_2.

In this example, conditioning on X_3 in a multivariate analysis of the process X induces a type I spurious causality from X_2 to X_1 regardless of whether DTF or multivariate Granger causality is used for the analysis. The difference between this and the previous example of type I spurious causality is that according to the autoregressive representation of X the third variable X_3 seems to act as a confounder and not as a mediating variable for X_2 and X_1. Both DTF and multivariate Granger causality correct for any confounding by other variables included in the analysis and, consequently, fail to detect that there is no information flow from X_2 to X_3. In contrast, a bivariate analysis of X_1 and X_2 shows that X_2 does not bivariately Granger-cause X_1 and thus provides a correct description of the connectivity among the variables.

Summarizing the results of this section, we note that the DTF and the concepts of bivariate and multivariate Granger causality are different descriptions of the relationships among multiple time series that concentrate on different aspects of the relationships. In particular, the DTF measures the total effect of one series on another, whereas multivariate Granger causality is concerned only with the direct effect that one series has on another series. In contrast, bivariate Granger causality cannot distinguish between causal effects – direct or indirect – and confounding due to other variables and thus seems less suitable for the description of information flow among multiple time series.

Some of the theoretical examples in this section involved latent variables. In practice, the possible existence of latent variables, which despite the best efforts of experimentalists to include all important variables cannot be ruled out completely, poses a serious problem for the identification of the connectivity and the information flow in multivariate systems. Obviously, if we rely on a single measure of information flow, we have no basis to decide whether the obtained results are correct or not. However, the examples have shown that the combined use of the three discussed measures of information flow may provide additional information that indeed allows identifying spurious causalities induced by latent variables; for example, we have seen that bivariate Granger causality can be used to detect spurious causality of type I. For such a combined analysis, it becomes important to understand the properties of the different measures and the conditions under which they succeed or fail. We note that Eichler (2005) more generally suggested to consider also Granger causality with respect to arbitrary subprocesses in order to identify the causal structure of systems with latent variables.

4 Statistical properties of the DTF

In practice, the identification of directed relationships is based on estimates of the DTF. These can be simply obtained by substituting estimates for the autoregressive coefficients into the definition of the DTF. However, the statistical properties of these estimators have not yet been investigated although there exists a simulation-based approach for the statistical assessment of significance (Kamiński et al., 2001). In the following, we establish a simple pointwise significance level for testing whether for given indices i and j and frequency λ the DTF $\gamma_{ij}(\lambda)$ differs significantly from zero.

For the derivation, we assume that the spectral density matrix $\mathbf{f}(\lambda)$ of X satisfies the boundedness condition (4). Under this assumption, the function $\hat{\mathbf{B}}(\lambda) = (\mathbf{I} - \hat{\mathbf{A}}(\lambda))^{-1}$ has a Taylor expansion in a neighbourhood of $\mathbf{A}(\lambda)$ given by

$$\hat{\mathbf{B}}(\lambda) = \mathbf{B}(\lambda) - \mathbf{B}(\lambda)\big(\hat{\mathbf{A}}(\lambda) - \mathbf{A}(\lambda)\big)\mathbf{B}(\lambda) + \mathbf{r}(\lambda), \tag{20}$$

where the remainder term $\mathbf{r}(\lambda)$ tends to zero in probability with a rate faster than $T^{-1/2}$ and thus is negligible compared with the first two terms. The estimate $\hat{\mathbf{A}}(\lambda)$ depends on the parameter estimates $\hat{\mathbf{a}}_{kl}(u)$, which are, for large T, approximately normally distributed with mean $\mathbf{a}_{kl}(u)$ and covariances satisfying

$$\lim_{T\to\infty} T \operatorname{cov}\big(\hat{\mathbf{a}}_{kl}(u), \hat{\mathbf{a}}_{mn}(v)\big) = \mathbf{H}_{ln}(u, v)\, \Sigma_{km},$$

where $\mathbf{H}_{ln}(u, v)$ are entries of the inverse $\mathbf{H} = \mathbf{R}^{-1}$ of the covariance matrix $\mathbf{R}$ of the process (Lütkepohl, 1993). The covariance matrix $\mathbf{R}$ is composed of $d \times d$ submatrices

$$\mathbf{R}(u, v) = \begin{pmatrix} \mathbf{R}_{11}(u, v) & \cdots & \mathbf{R}_{1d}(u, v) \\ \vdots & \ddots & \vdots \\ \mathbf{R}_{d1}(u, v) & \cdots & \mathbf{R}_{dd}(u, v) \end{pmatrix}$$

with entries $\mathbf{R}_{ln}(u, v) = \operatorname{cov}\big(X_l(t - u), X_n(t - v)\big)$ for $l, n = 1, \ldots, d$ and $u, v = 1, \ldots, p$; the matrix $\mathbf{H}$ is partitioned similarly. Based on the asymptotic distribution of the autoregressive estimates $\hat{\mathbf{a}}_{ij}(u)$, we show in the appendix that for large sample sizes T under the null hypothesis of $|\mathbf{B}_{ij}(\lambda)|^2 = 0$ the distribution of

$$\frac{T}{C_\lambda} \big|\hat{\mathbf{B}}_{ij}(\lambda)\big|^2 \tag{21}$$

with

$$C_\lambda = \sum_{u,v=1}^{p} \Big[\big(\mathbf{B}(\lambda)'\mathbf{H}(u, v)\overline{\mathbf{B}(\lambda)}\big)_{jj}\big(\mathbf{B}(\lambda)\Sigma\mathbf{B}(\lambda)^*\big)_{ii} \times \big(\cos(u\lambda)\cos(v\lambda) + \sin(u\lambda)\sin(v\lambda)\big)\Big] \tag{22}$$

can be approximated by that of a weighted average of two independent χ^2 distributed random variables with one degree of freedom. Thus the critical value of the distribution is bounded by the critical value of a χ^2 distribution with one degree of freedom. This leads to a significance level for the non-normalized DTF $\theta_{ij}^2(\lambda)$.

Next, we consider the complex-valued function

$$\hat{\gamma}_{ij}(\lambda) = \frac{\hat{\mathbf{B}}_{ij}(\lambda)}{\sqrt{\sum_k \left|\hat{\mathbf{B}}_{ik}(\lambda)\right|^2}},$$

from which we obtain the estimate $\hat{\gamma}_{ij}^2(\lambda)$ for the normalized by taking the squared absolute value. This function is nonlinear in the parameter estimates $\hat{\mathbf{a}}_{kl}(u)$. Taylor expansion of $\hat{\gamma}_{ij}(\lambda)$ about $\mathbf{a}_{kl}(u)$ yields

$$\hat{\gamma}_{ij}(\lambda) = \frac{\hat{\mathbf{B}}_{ij}(\lambda)}{\sqrt{\sum_k \left|\mathbf{B}_{ik}\right|^2}} + \mathbf{B}_{ij}(\lambda) r_1(\lambda) + r_2(\lambda) \tag{23}$$

with $|r_1(\lambda)| \leq C\|\hat{\mathbf{a}} - \mathbf{a}\|$ and $|r_2(\lambda)| \leq C\|\hat{\mathbf{a}} - \mathbf{a}\|^2$. Here, $\mathbf{a} = \big(\mathbf{a}(1), \ldots, \mathbf{a}(p)\big)$ and $\hat{\mathbf{a}}$ denotes the corresponding estimate. Under the hypothesis of $|\mathbf{B}_{ij}(\lambda)|^2 = 0$, the second term $\mathbf{B}_{ij}(\lambda) R_1$ vanishes and we have

$$\hat{\gamma}_{ij}^2(\lambda) = \frac{|\hat{\mathbf{B}}_{ij}(\lambda)|^2}{\sum_k \left|\mathbf{B}_{ik}(\lambda)\right|^2} + r(\lambda),$$

where the remainder $r(\lambda)$ is of order $\|\hat{\mathbf{a}} - \mathbf{a}\|^3$ and is thus negligible compared with the first term. Therefore the α-significance level for the normalized DTF $\gamma_{ij}^2(\lambda)$ can be approximated by

$$\frac{\hat{C}_\lambda \, \chi^2_{1,1-\alpha}}{T \sum_k |\hat{\mathbf{B}}_{ik}(\lambda)|^2}, \tag{24}$$

where $\chi^2_{1,1-\alpha}$ is the $1-\alpha$ quantile of the χ^2 distribution with one degree of freedom and $\hat{C}_\lambda$ is an estimate of the constant in Eq. 22 obtained by substituting estimates $(\hat{\mathbf{R}}^{-1})(u, v)$, $\hat{\Sigma}$ and $\hat{\mathbf{B}}(\lambda)$ for $\mathbf{H}(u, v)$, Σ and $\mathbf{B}(\lambda)$, respectively.

The significance level in Eq. 24 for the normalized DTF $\gamma_{ij}^2(\lambda)$ depends on the frequency λ through the constant C_λ and the normalizing factor $\sum_k |\mathbf{B}_{ik}(\lambda)|^2$. The latter compensates for the effects of normalization, that is, significance depends not on the relative but on the absolute strength—as measured by the non-normalized DTF—of the information flow at a particular frequency. The effects of the dependence on C_λ are far more difficult to predict. Unlike in the case of the PDC (Schelter et al. 2005), the factor C_λ depends not only on the variances $\mathbf{R}$ and Σ, but also on the transfer function $\mathbf{B}(\lambda)$ and, thus, on the DTF itself. We will examine this dependence briefly in the next section by simulation studies. We note that the asymptotic distribution of the non-normalized DTF under the null hypothesis also depends on the frequency, which underlines the importance of non-constant significance levels that adapt to the local variation in distribution of the estimates.

One drawback of the proposed significance level is that it is only a pointwise level, which on average is exceeded for a small number of frequencies even under the null hypothesis of no information flow. However, the estimates $\hat{\gamma}_{ij}^2(\lambda)$ (and $\hat{\theta}_{ij}^2(\lambda)$) of the DTF at different frequencies λ are highly correlated since the DTF is a smooth function of the autoregressive coefficients $\mathbf{a}_{ij}(1), \ldots, \mathbf{a}_{ij}(p)$, which are themselves correlated. Therefore, it seems hardly possible to derive a uniform non-constant significance level for the DTF. We note that for similar reasons it if rather common in other applications to use pointwise levels.

5 Applications

The usefulness of the DTF as a measure of information flow in multivariate systems has been demonstrated in many articles (e.g., Kamiński and Blinowska 1991; Kamiński et al. 2001; Blinowska et al. 2004). In this section, we will, therefore, examine only the effectiveness of the pointwise significance level that has been derived in the previous section.

5.1 Simulated examples

In the first simulation study, we consider a simple system of three processes given by the VAR process

$$X(t) = \mathbf{a}(1)\, X(t-1) + \mathbf{a}(2)\, X(t-2) + e(t)$$

with coefficient matrices

$$\mathbf{a}(1) = \begin{pmatrix} \frac{4}{15} & \frac{2}{5} & 0 \\ \frac{2}{5} & \frac{5}{28} & 0 \\ 0 & \frac{2}{5} & \frac{5}{28} \end{pmatrix} \quad \text{and} \quad \mathbf{a}(2) = \begin{pmatrix} -\frac{1}{4} & -\frac{1}{5} & 0 \\ \frac{1}{5} & -\frac{1}{9} & 0 \\ 0 & \frac{1}{5} & -\frac{1}{9} \end{pmatrix} \tag{25}$$

and covariance matrix $\Sigma = \mathbf{I}$. In this system, the first two components exhibit a reciprocal interaction while the third component is influenced unidirectionally by the second process. The coupling scheme for this system is depicted in Fig. 4.

To study the properties of the significance test, we generated samples of length $T = 1{,}000$ and fitted an autoregressive model of order $p = 4$ to the simulated data. Figure 5 shows estimates of the spectral densities and the non-normalized DTF for one of these samples. As expected, the DTF indicates substantial information flow for the three direct links $3 \to 2$, $2 \to 1$ and $1 \to 2$ in the system. Additionally, it shows a smaller information flow from X_3 to X_1, which corresponds to the indirect link $3 \to 2 \to 1$. On the other hand, the DTFs $\theta_{21}^2(\lambda)$ and $\theta_{31}^2(\lambda)$ stay below the pointwise significance level and thus indicate that there is no

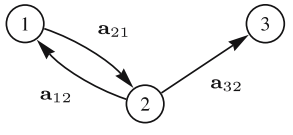

Fig. 4 Path diagram and coefficients for the three-dimensional autoregressive process in the first simulation example

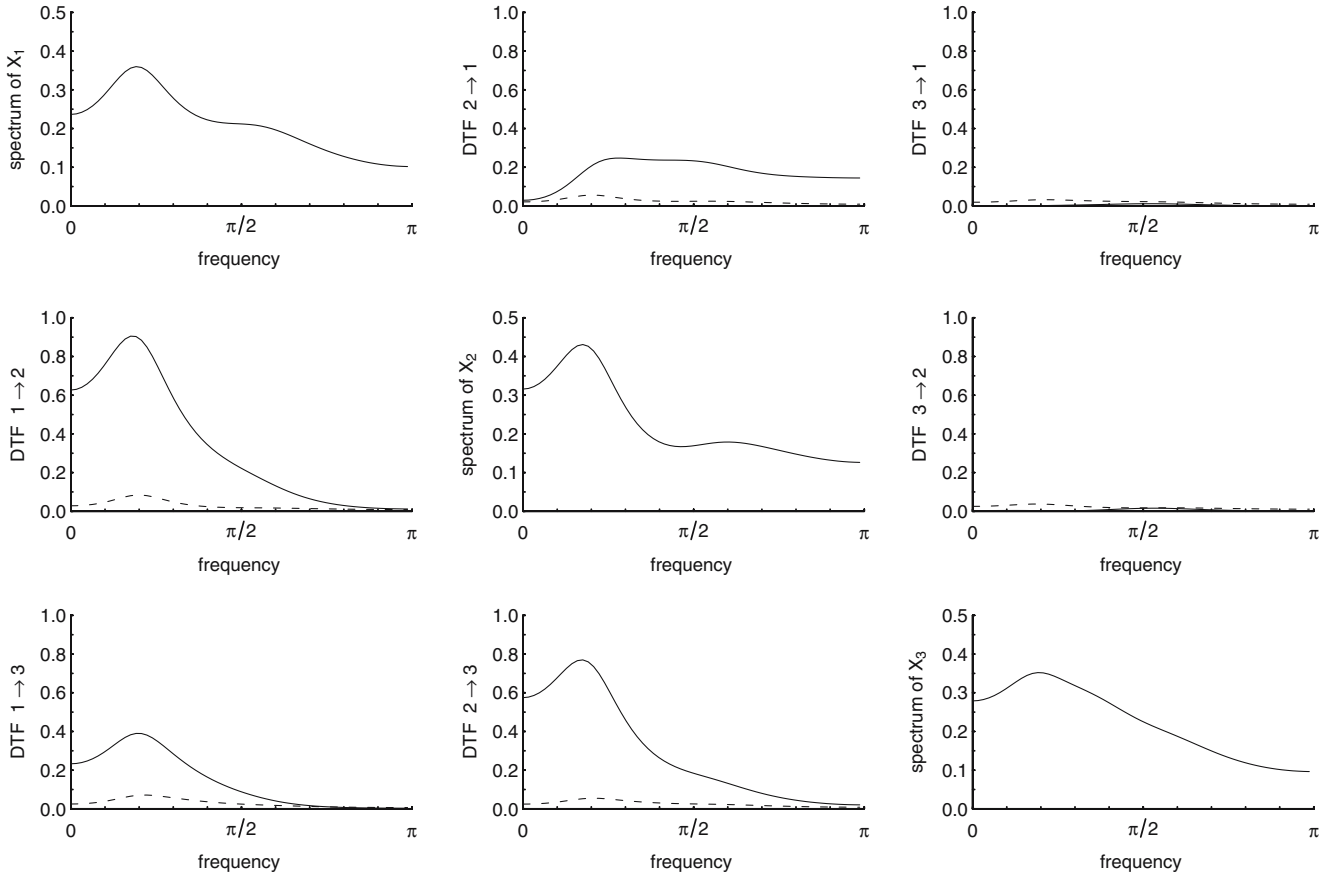

Fig. 5 Results for a three-dimensional linear system with the coupling scheme given by Fig. 4: estimated spectral densities (*on diagonal*) and non-normalized DTF (*off-diagonals*). The *dashed lines* signify pointwise 95% test bounds for the hypothesis that the DTF is zero

significant information flow from X_1 to X_2 and from X_1 to X_3, respectively. We note that the analysis of the normalized DTF $\gamma_{ij}^2(\lambda)$ leads to very similar results (not shown).

Figure 6 shows the pointwise rejection rates out of 10,000 samples. The first two plots give the rejection rates for the DTF from X_3 to X_2 and the DTF from X_3 to X_1, which are both uniformly zero. In each case, the null hypothesis of no information flow is rejected in 2–5% of all samples, that is, the actual size of the proposed pointwise significance test turns out to be smaller than the nominal size of 5%. The reason for this conservative behaviour of the test is that we used the 95 % quantile of the χ^2 distribution with one degree of freedom as an upper bound in our construction of the significance level in Eq. 24. The remaining four plots demonstrate the ability of the significance test to detect the alternative, that is, the presence of information flow. Only for frequencies for which the value of the DTF is close to zero, the rejection rate falls below 90%. We note that there is no clear relationship between the value of the DTF and the rejection rate, which is again due to fact that the significance level is not constant and, in particular, depends on the transfer function $\mathbf{B}(\lambda)$.

To illustrate the dependence of the significance level on the transfer function $\mathbf{B}(\lambda)$, we performed a second simulation study based on the bivariate autoregressive process given by

$$\begin{aligned} X_1(t) &= \frac{7}{8}X_1(t-1) - \frac{4}{5}X_1(t-2) \\ &\quad + \frac{4}{5}X_1(t-3) - \frac{1}{2}X_1(t-4) + e_1(t), \\ X_2(t) &= \frac{1}{20}X_1(t-1) + \frac{1}{20}X_1(t-2) + e_2(t), \end{aligned} \tag{26}$$

with $\Sigma = \mathbf{I}$. The estimates of the power spectra and the non-normalized as well as the normalized DTF obtained from simulated data ($T = 1{,}000$) for this process are presented in Fig. 7. First, we note that the spectral density of the first component (Fig. 7a) exhibits a strong peak at frequency $\lambda = 0.4$ and a smaller peak at $\lambda = 1.8$. Next, we find also a strong peak at frequency $\lambda = 0.4$ in the plots of the non-normalized DTFs $\theta_{21}^2(\lambda)$ and $\theta_{12}^2(\lambda)$ (Fig. 7b). The peaks in both plots are of about equal height, but while the DTF from X_1 to X_2 clearly exceeds the significance level for frequencies less than $\lambda = \pi/2$, the DTF from X_2 to X_1 is enveloped by the pointwise significance level and, thus, indicates correctly that there is no significant information flow between the two processes in that direction. We note that the curve of the significance level for $\theta_{21}^2(\lambda)$ resembles the power spectrum of the first component, which shows that it depends strongly on $\hat{\mathbf{B}}_{11}(\lambda)$.

Finally, comparing the non-normalized and the normalized DTF (Fig. 7c), we find that after normalization the peak

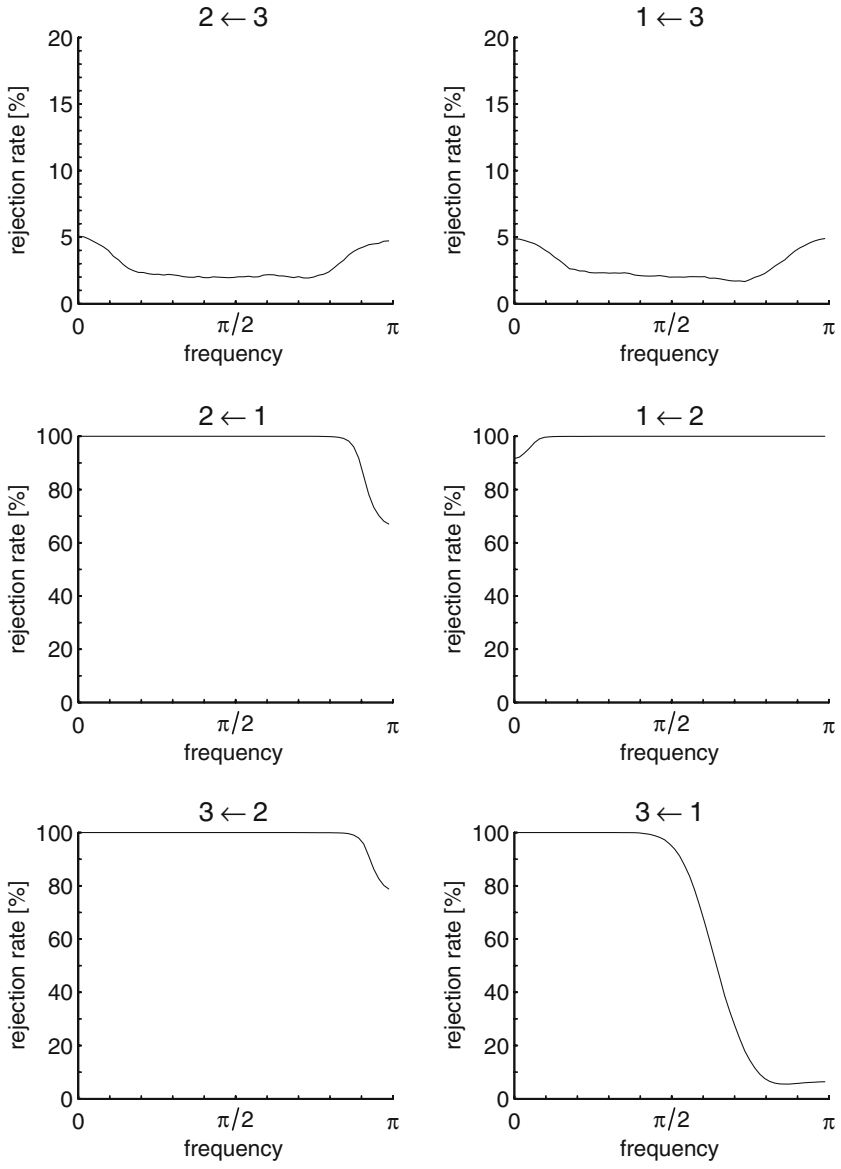

Fig. 6 Efficiency of the DTF-based pointwise significance test: rejection rates out of 10,000 replications for the linear system with coupling scheme given by Fig. 4

in the DTF from X_2 to X_1 has vanished and the corresponding significance level is now almost constant. In contrast, the normalized DTF from X_1 to X_2 does not seem much affected by the normalization, and the significance level still varies over frequency.

Kamiński et al. (2001) and Kamiński (2005) have advocated the use of the non-normalized DTF in cases where the information flow for different target processes or between different experiments are compared. The last example has shown that the variation of the non-normalized DTF can vary greatly over frequency in particular if the signals have strong oscillating components. It, thus, demonstrates the importance of a non-constant significance level that can adapt to such changes in the variation of the estimates over frequency.

5.2 Application to neuronal spike trains

In this example, we analyse neuronal spike train data recorded from the lumbar spinal dorsal horn of a pentobarbital-anaesthetized rat during noxious stimulation. The firing times of ten neurons were recorded simultaneously by a single electrode with an observation time of 100 s. The data have been described in detail in Sandkühler and Eblen-Zajjur (1994); the connectivity among the recorded neurons has been analysed previously by partial correlation analysis (Eichler et al. 2003) and partial directed correlations (Dahlhaus and Eichler 2002).

For the present analysis, the spike trains were converted to binary time series of length $T = 19{,}999$ and a VAR

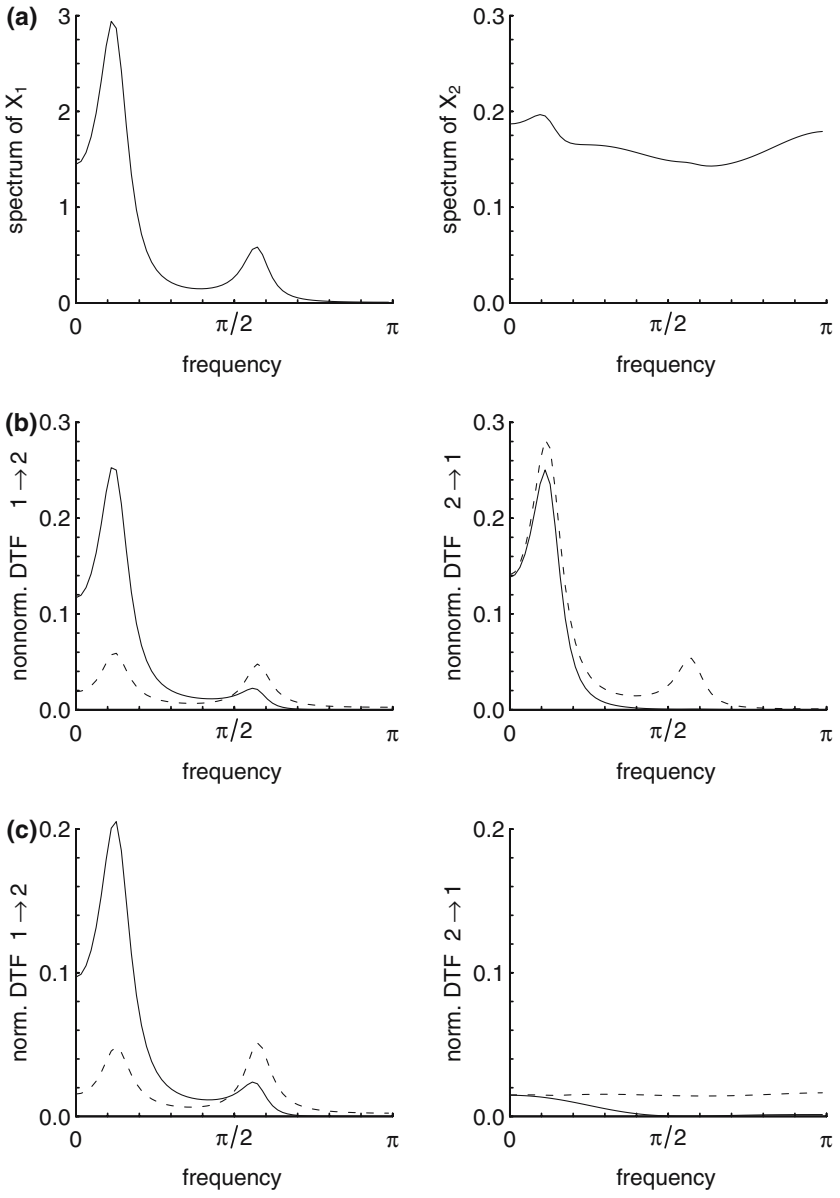

Fig. 7 Results for the bivariate autoregressive process in Eq. 26: estimates of **a** spectral densities, **b** non-normalized DTF and **c** normalized DTF. The *dashed lines* signify pointwise 95% test bounds for the hypothesis that the DTF is zero

model of order $p = 100$ was fitted to the resulting series. Figure 8 displays on the diagonal the estimated spectra for five neuronal spike trains. The strong peaks in the spectra for neurons 1 and 2 indicate that these neurons show rhythmic discharges at 5 Hz; similarly, neuron 5 fires rhythmically at 7.5 Hz.

In order to identify the information flow between these five neurons, we have estimated the non-normalized DTF (Fig. 8, off-diagonals). Due to the strong peaks in the spectra of the spike trains, the significance level for the DTF exhibits also strong peaks at the major frequencies of the involved spike trains. As a result, some of the clearly visible peaks in the DTF, most notably in that from neuron 2 to neuron 1 or that from neuron 4 to neuron 5, are marked as non-significant. Summarizing we find information flow from neuron i to neuron j whenever $i < j$ except for neurons 4 and 5, which is in correspondence with the results of Eichler et al. (2003) and Dahlhaus and Eichler (2002).

For comparison, we have also computed the normalized DTF, which is depicted in Fig. 9. Here, the pointwise significance level for the DTF exhibits in general less peaks than in the case of the non-normalized DTF, but the DTF from neuron 3 to neuron 2 or those from neuron 5 to neurons 4 and 1 still show small peaks that are enveloped by the significance level and, thus, judged non-significant. We conclude that consideration of the normalized version of the DTF does not protect against the occurrence of non-significant peaks in the DTF.

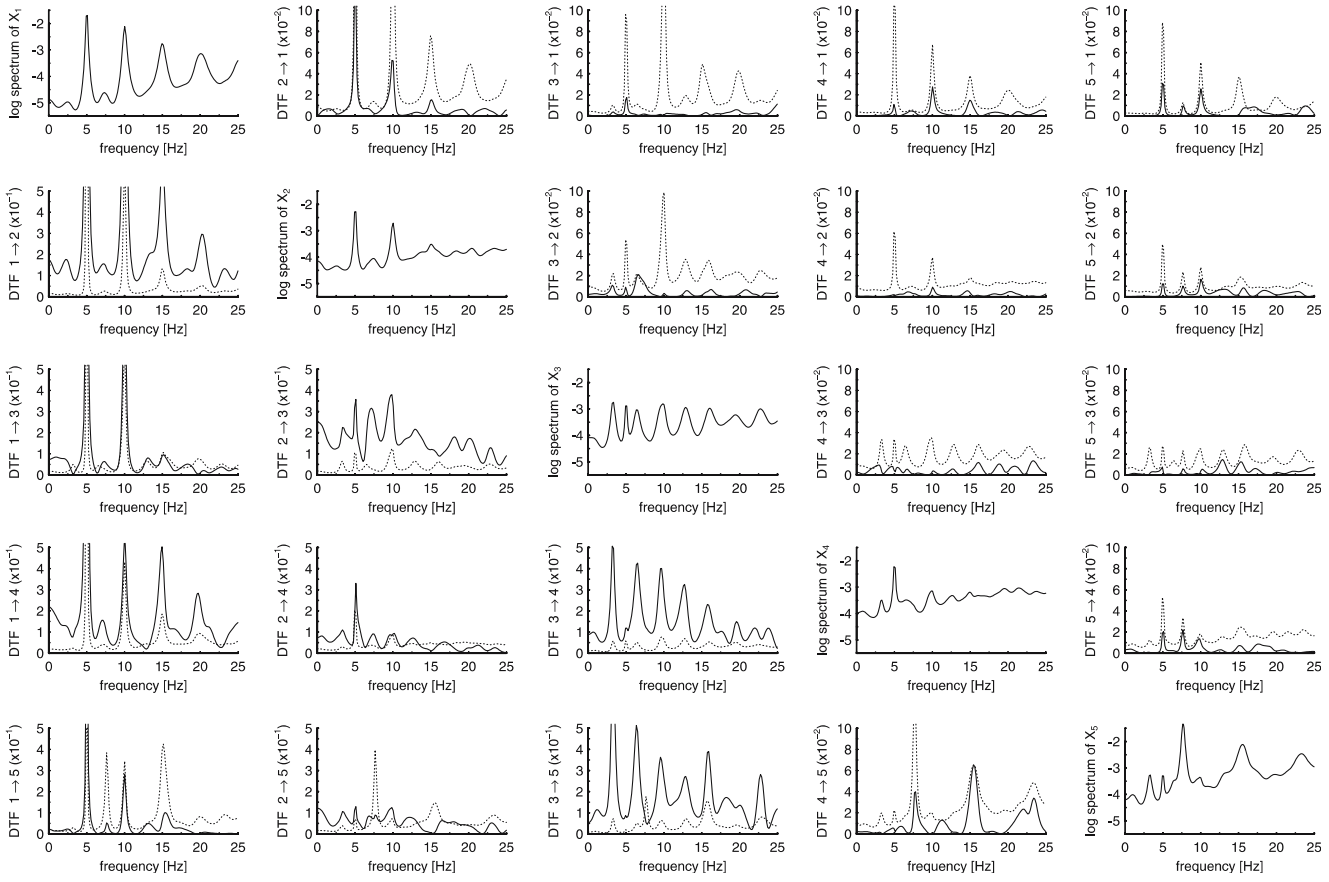

Fig. 8 Results for neuronal spike train data: estimates of log-spectral densities (*on diagonal*) and non-normalized DTF (*off-diagonals*). The *dotted lines* signify pointwise 95% test bounds for the hypothesis that the DTF is zero

6 Conclusion

In this paper, we have discussed the theoretical interpretation of the DTF as a measure of information flow in multivariate systems. We have shown that the DTF measures the response of one target component to sinusoidal inputs in another component and, thus, is closely related to the impulse response function in the time domain. Furthermore, we have shown by a number of examples that the DTF does not provide necessary or sufficient conditions for either bivariate or multivariate Granger causality. This means that the DTF cannot be used as a measure for either concept of Granger causality. The DTF and bivariate as well as multivariate Granger causality focus on different aspects of the connectivity structure and, therefore, should be seen as complementary tools for the description of the relationships among multiple time series. The DTF method seems to be most suitable as a spectral measure of the total causal influence that one component exerts over another component.

While the usefulness of the DTF as a measure of information flow in multivariate systems has been demonstrated in many studies, the statistical properties have not yet been discussed thoroughly. In this paper, we have derived the asymptotic distribution of the DTF under the null hypothesis of no information flow and have proposed a pointwise significance level that forms an upper bound for the true unknown critical value of the asymptotic distribution.

The significance level is a pointwise level that adapts to local variations in the distribution of the DTF. Simulations have shown that the variation especially of the non-normalized DTF may increase greatly at frequencies where peaks are present in the corresponding power spectra. In such cases, the non-constant significance level adjusts appropriately and, thus, prevents false detection of information flow due to spurious peaks in the DTF. Furthermore, the proposed significance level is easy to implement and fast to compute. Thus, it enables a widespread use in the statistical evaluation of estimates of the DTF in applications in neuroscience and other fields.

7 Appendix

In this appendix, we establish the asymptotic distribution of the estimator for the non-normalized DTF $\theta_{ij}^2(\lambda) = |\mathbf{B}_{ij}(\lambda)|^2$. For the technical derivation, we need some notation from matrix theory: vec($\mathbf{A}$) stands for the vector resulting from stacking the columns of the matrix $\mathbf{A}$ on top of each other

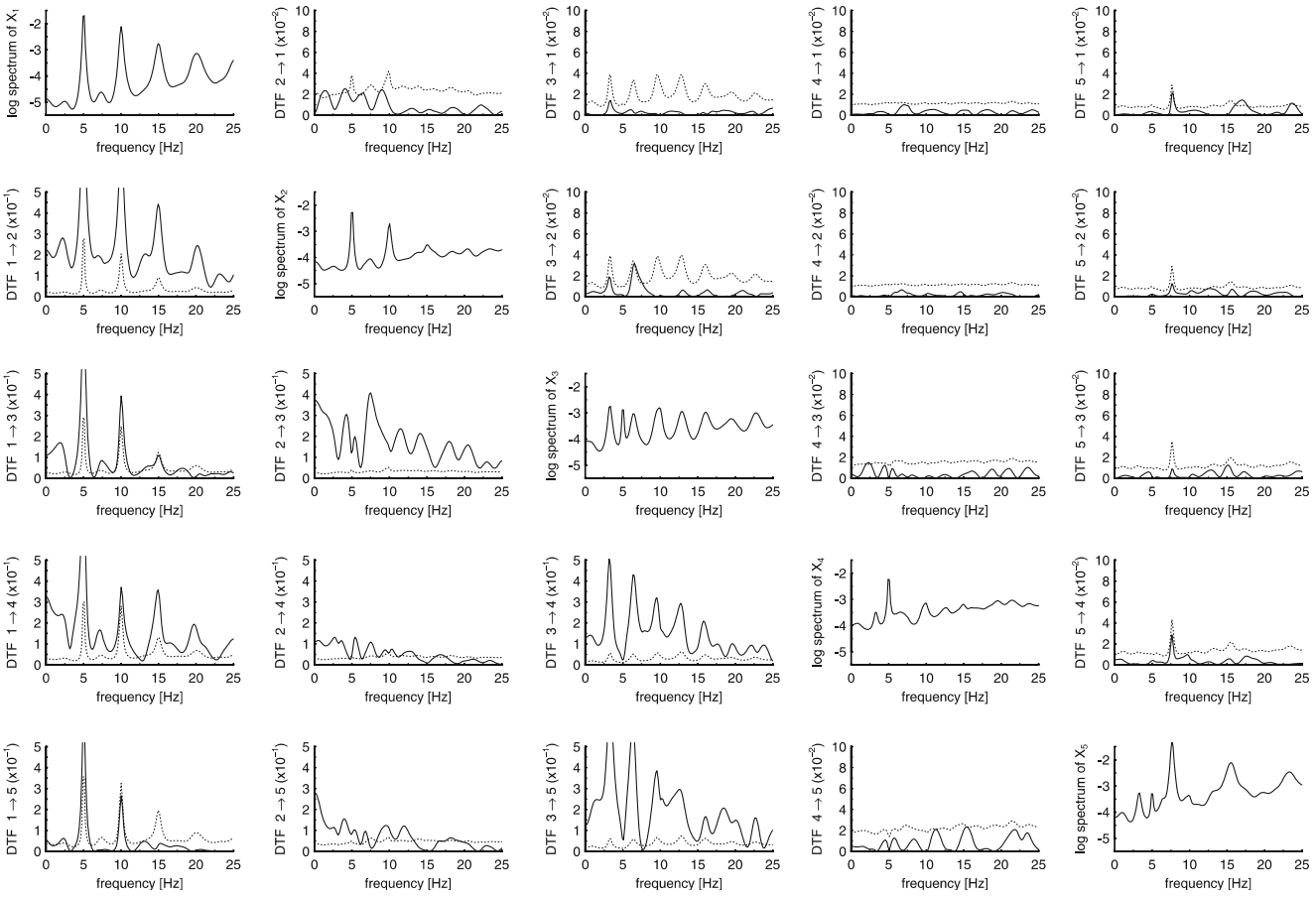

Fig. 9 Results for neuronal spike train data: estimates of log-spectral densities (*on diagonal*) and normalized DTF (*off-diagonals*). The *dotted lines* signify pointwise 95% test bounds for the hypothesis that the DTF is zero

and $\mathbf{A} \otimes \mathbf{B}$ denotes the Kronecker product of two matrices $\mathbf{A}$ and $\mathbf{B}$. For details we refer to Harville (1997).

Let $\hat{\mathbf{a}}(u), u = 1, \ldots, p$ be some commonly used estimator for the autoregressive coefficients such as the least squares or Yule-Walker estimator. Then it is well known (e.g., Lütkepohl 1993) that, for $u = 1, \ldots, p$, the vectors $\operatorname{vec}\big(\hat{\mathbf{a}}(u)\big)$ are jointly asymptotically normally distributed with means $\operatorname{vec}\big(\mathbf{a}(u)\big)$ and covariances satisfying

$$\lim_{T\to\infty} T \operatorname{cov}\big[\operatorname{vec}\big(\hat{\mathbf{a}}(u)\big), \operatorname{vec}\big(\hat{\mathbf{a}}(v)\big)\big] = \mathbf{H}(u, v) \otimes \Sigma.$$

The estimation of the DTF is based on the Fourier transform $\hat{\mathbf{A}}(\lambda) = \sum_{u=1}^{p} \hat{\mathbf{a}}(u)\, \mathrm{e}^{-\mathrm{i}\lambda u}$. Since it is complex valued, we consider its real and imaginary part separately. Then the above limiting distribution for the parameter estimates implies that

$$X(\lambda) = \sqrt{T} \begin{pmatrix} \operatorname{vec}\big(\operatorname{Re}[\hat{\mathbf{A}}(\lambda) - \mathbf{A}(\lambda)]\big) \\ \operatorname{vec}\big(\operatorname{Im}[\hat{\mathbf{A}}(\lambda) - \mathbf{A}(\lambda)]\big) \end{pmatrix}$$

is asymptotically normally distributed with mean zero and covariance matrix

$$\tilde{\mathbf{V}}(\lambda) = \sum_{u,v=1}^{p} \left[\begin{pmatrix} \cos(u\lambda)\cos(v\lambda) & \cos(u\lambda)\sin(v\lambda) \\ \sin(u\lambda)\cos(v\lambda) & \sin(u\lambda)\sin(v\lambda) \end{pmatrix} \otimes \mathbf{H}(u, v) \otimes \Sigma \right].$$

We note that for $p \geq 2$ and $\lambda \neq 0 \mod \pi$ the matrix $\tilde{\mathbf{V}}(\lambda)$ has full rank and hence is positive definite.

Next, we consider the entry $\hat{\mathbf{B}}_{ij}(\lambda)$ of the inverse matrix $\hat{\mathbf{B}}(\lambda) = \big(\mathbf{I} - \hat{\mathbf{A}}(\lambda)\big)^{-1}$. From the Taylor expansion in (20), we obtain that under the null hypothesis of $\mathbf{B}_{ij}(\lambda) = 0$ the entry $\hat{\mathbf{B}}_{ij}(\lambda)$ can be approximated by

$$\hat{\mathbf{B}}_{ij}(\lambda) = \beta(\lambda)' \operatorname{vec}\big(\hat{\mathbf{A}}(\lambda) - \mathbf{A}(\lambda)\big) + r(\lambda),$$

where $\beta(\lambda)$ is the vector with elements $\mathbf{B}_{ik}(\lambda)\mathbf{B}_{lj}(\lambda)$ for $k, l = 1, \ldots, d$ and $r(\lambda)$ is a remainder term that converges to zero in probability with a rate faster than $T^{-1/2}$ and hence is negligible compared with the main term. Treating real and imaginary parts separately, we find that

$$Y(\lambda) = \begin{pmatrix} \sqrt{T} \operatorname{Re} \hat{\mathbf{B}}_{ij}(\lambda) \\ \sqrt{T} \operatorname{Im} \hat{\mathbf{B}}_{ij}(\lambda) \end{pmatrix} \tag{27}$$

has the same limiting distribution as $\Phi(\lambda)\, X(\lambda)$, where

$$\Phi(\lambda) = \begin{pmatrix} \operatorname{Re} \beta(\lambda)' & -\operatorname{Im} \beta(\lambda)' \\ \operatorname{Im} \beta(\lambda)' & \operatorname{Re} \beta(\lambda)' \end{pmatrix},$$

and thus is asymptotically normally distributed with mean zero and covariance matrix

$$\mathbf{V}(\lambda) = \Phi(\lambda)\, \tilde{\mathbf{V}}(\lambda)\, \Phi(\lambda)'$$

with $\tilde{\mathbf{V}}(\lambda)$ defined as above. For $p \geq 2$ and $\lambda \neq 0 \mod \pi$ the matrix $\mathbf{V}(\lambda)$ is positive definite and thus can be factorized as $\mathbf{V}(\lambda) = \mathbf{Q}(\lambda)\mathbf{D}(\lambda)\mathbf{Q}(\lambda)'$, where $\mathbf{Q}(\lambda)$ is some orthogonal matrix and $\mathbf{D}(\lambda)$ is the diagonal matrix of the eigenvalues of $\mathbf{V}(\lambda)$. It follows that

$$T\,|\hat{\mathbf{B}}_{ij}(\lambda)|^2 = Y(\lambda)'Y(\lambda) \approx Z'\,\mathbf{D}(\lambda)\,Z = \mathbf{D}_{11}(\lambda)\,Z_1^2 + \mathbf{D}_{22}(\lambda)\,Z_2^2$$

for some standard normally distributed $Z = (Z_1, Z_2)'$. Simple calculations show that the eigenvalues of $\mathbf{V}(\lambda)$ satisfy

$$\frac{\mathbf{D}_{ii}(\lambda)}{C_\lambda} = \frac{1}{2} \pm \sqrt{\frac{1}{4} - \frac{\det \mathbf{V}(\lambda)}{C_\lambda^2}} > 0,$$

where $C_\lambda = \mathbf{V}_{11}(\lambda) + \mathbf{V}_{22}(\lambda)$ is the constant in Eq. 22. It follows that for large sample sizes $T\,|\hat{\mathbf{B}}_{ij}(\lambda)|^2/C_\lambda$ has approximately the same distribution as the weighted average of two independent χ^2 distributed random variables with one degree of freedom.

Now suppose that $\lambda = 0 \mod \pi$. Then $\mathbf{B}(\lambda)$ and $\mathbf{A}(\lambda)$ both are real-valued. Similar arguments as above show that, again under the null hypothesis of $\mathbf{B}_{ij}(\lambda) = 0$, $\sqrt{T}\,\hat{\mathbf{B}}_{ij}(\lambda)$ is asymptotically normally distributed with mean zero and variance

$$\beta(\lambda)'\,\tilde{\mathbf{V}}(\lambda)\,\beta(\lambda) = C_\lambda,$$

where

$$\tilde{\mathbf{V}}(\lambda) = \sum_{u,v=1}^{p} \mathbf{H}(u, v) \otimes \Sigma$$

is the asymptotic variance of $\mathrm{vec}\left(\hat{\mathbf{A}}(\lambda)\right)$. Consequently, the ratio $T\,|\hat{\mathbf{B}}_{ij}(\lambda)|^2/C_\lambda$ is asymptotically χ^2-distributed with one degree of freedom.

Finally, we consider the case $p = 1$, that is, $\hat{\mathbf{A}}(\lambda) = \hat{\mathbf{a}}(1)\,e^{-\mathrm{i}\lambda}$. It follows that $Y(\lambda)$ in (27) is of the form

$$Y(\lambda) = \sqrt{T}\,\tilde{\Phi}(\lambda)\ \mathrm{vec}\left(\hat{\mathbf{a}}(1) - \mathbf{a}(1)\right),$$

where

$$\tilde{\Phi}(\lambda) = \begin{pmatrix} \mathrm{Re}\,\beta(\lambda)'\cos(\lambda) - \mathrm{Im}\,\beta(\lambda)'\sin(\lambda) \\ \mathrm{Im}\,\beta(\lambda)'\cos(\lambda) + \mathrm{Re}\,\beta(\lambda)'\sin(\lambda) \end{pmatrix},$$

and hence has asymptotic variance

$$\mathbf{V}(\lambda) = \tilde{\Phi}(\lambda)\left[\mathbf{H}(1, 1) \otimes \Sigma\right]\tilde{\Phi}(\lambda)'.$$

We note that $\tilde{\Phi}(\lambda) = 0$ if and only if $\beta(\lambda)'\,e^{\mathrm{i}\lambda} = 0$, which contradicts the invertibility of the matrix $\mathbf{B}(\lambda) \otimes \mathbf{B}(\lambda)'$, of which $\beta(\lambda)$ is a column vector. Thus $\tilde{\Phi}(\lambda)$ has at least rank 1. Since $\mathbf{H}(1, 1) \otimes \Sigma$ is positive definite, it follows that the 2×2 matrix $\mathbf{V}(\lambda)$ is non-negative definite with at least rank 1. Using again the factorization $\mathbf{V}(\lambda) = \mathbf{Q}(\lambda)\mathbf{D}(\lambda)\mathbf{Q}(\lambda)'$, we find that $T\,|\hat{\mathbf{B}}_{ij}(\lambda)|^2/C_\lambda$ asymptotically has the same distribution as the weighted average of two independent χ^2 distributed random variables with one degree of freedom. In particular, the asymptotic distribution becomes a χ^2-distribution with one degree of freedom if $\mathbf{V}(\lambda)$ has rank 1 and, hence, one eigenvalue of $\mathbf{V}(\lambda)$ is zero.

References

Achermann P, Hartmann R, Gunzinger A, Guggenbühl W, Borbély AA (1994) All night sleep and artificial stochastic control have similar correlation dimension. Electroencephalogr Clin Neurophysiol 90:384–387

Baccalá LA, Sameshima K (2001) Partial directed coherence: a new concept in neural structure determination. Biol Cybern 84:463–474

Blinowska KJ, Kuś R, Kamiński M (2004) Granger causality and information flow in multivariate processes. Phys Rev E 70:050902(R)

Blinowska KJ, Malinowski M (1991) Non-linear and linear forecasting of the EEG time series. Biol Cybern 66:159–165

Brockwell PJ, Davis RA (1991) Time series: theory and methods, 2nd edn. Springer Berlin Heidelberg, New York

Brovelli A, Ding M, A L, Chen Y, Nakamura R, Bressler SL (2004) Beta oscillations in a large-scale sensorimotor cortical network: directional influences revealed by granger causality. Proc Nat Acad Sci USA 101:9849–9854

Chatfield C (2003) The analysis of time series: an introduction, 6th edn. Chapman & Hall/CRC, Boca Raton

Dahlhaus R, Eichler M (2002) Causality and graphical models for multivariate time series and point processes. In: Hutten H, Kroesl P (eds) IFMBE Proceedings EMBEC 2002, vol 3(2), pp 1430–1431

Dufour JM, Renault E (1998) Short run and long run causality in time series: theory. Econometrica 66:1099–1125

Eichler M (2002) Granger-causality and path diagrams for multivariate time series. Journal of Econometrics (to appear)

Eichler M (2005) A graphical approach for evaluating effective connectivity in neural systems. Philos Trans R Soc B 360:953–967

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

Franaszczuk PJ, Bergey GK (1998) Application of the directed transfer function method to mesial and lateral onset temporal lobe seizures. Brain Topogr 11:13–21

Franaszczuk PJ, Blinowska KJ, Kowalczyk M (1985) The application of parametric multichannel spectral estimates in the study of electrical brain activity. Biol Cybern 51:239–247

Goebel R, Roebroeck A, Kim DS, Formisano E (2003) Investigating directed cortical interactions in time-resolved fMRI data using vector autoregressive modeling and Granger causality mapping. Magn Reson Imaging 21:1251–1261

Granger CWJ (1969) Investigating causal relations by econometric models and cross-spectral methods. Econometrica 37:424–438

Granger CWJ (1980) Testing for causality, a personal viewpoint. J Econ Dynam Control 2:329–352

Harrison L, Penny WD, Friston KJ (2003) Multivariate autoregressive modeling of fMRI time series. Neuroimage 4:1477–1491

Harville DA (1997) Matrix algebra from a statistician's perspective. Springer Berlin Heidelberg, New York

Hayo B (1999) Money-output Granger causality revisited: an empirical analysis of EU countries. Appl Econ 31:1489–1501

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

Hsiao C (1982) Autoregressive modeling and causal ordering of econometric variables. J Econ Dynam Control 4:243–259

Kamiński M (2005) Determination of transmission patterns in multichannel data. Philos Trans R Soc B 360:947–952

Kamiński M, Blinowska KJ, Szelenberger W (1997) Topographic analysis of coherence and propagation of EEG activity during sleep and wakefulness. Electroencephalogr Clin Neurophysiol 102: 216–277

Kamiński M, Ding M, Truccolo WA, Bressler SL (2001) Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance. Biol Cybern 85:145–157

Kamiński MJ, Blinowska KJ (1991) A new method of the description of the information flow in the brain structures. Biol Cybern 65:203–210

Korzeniewska A, Kasicki S, Kamiński M, Blinowska KJ (1997) Information flow between hippocampus and related structures during various types of rat's behaviour. J Neurosci Methods 73:49–60

Kuś R, Kamiński M, Blinowska KJ (2004) Determination of EEG activity propagation: pair-wise versus multichannel estimate. IEEE Trans Biomed Eng 51:1501–1510

Liang H, Ding M, Nakamura R, Bressler SL (2000) Causal influences in primate cerebral cortex during visual pattern discrimination. NeuroReport 11:2875–2880

Lütkepohl H (1993) Introduction to multiple time series analysis. Springer Berlin Heidelberg, New York

Medvedev A, Willoughby JO (1999) Autoregressive modeling of the EEG in systemic kainic acid-induced epileptogenesis. Int J Neurosci 97:149–167

Pijn JPM, Van Neerven DN, Noest A, Lopes de Silva FH (1991) Chaos or noise in EEG signals: dependence on state and brain site. Electroencephalogr Clin Neurophysiol 79:371–381

Pijn JPM, Velis DN, van der Heyden MJ, DeGoede J, van Velen CWM, Lopes de Silva FH (1997) Nonlinear dynamics of epileptic seizure on basis of intracranial EEG recordings. Brain Topogr 9:249–270

Reinsel GC (2003) Elements of multivariate time series analysis, 2nd edn. Springer Berlin Heidelberg, New York

Sameshima K, Baccalá LA (1999) Using partial directed coherence to describe neuronal ensemble interactions. J Neurosci Methods 94:93–103

Sandkühler J, Eblen-Zajjur AA (1994) Identification and characterization of rhythmic nociceptive and non-reciceptive spinal dorsal horn neurons in the rat. Neuroscience 61:991–1006

Schack B, Rappelsberger P, Weiss S, Möller E (1999) Adaptive phase estimation and its application in EEG analysis of word processing. J Neurosci Methods 93:49–59

Schelter B, Winterhalder M, Eichler M, Peifer M, Hellwig B, Guschlbauer B, Lücking CH, Dahlhaus R, Timmer J (2005) Testing for directed influences in neuroscience using partial directed coherence. J Neurosci Methods (in press)

Sims CA (1980) Macroeconomics and reality. Econometrica 48:1–4

Stam C, Pijn JPM, Suffczynski P, Lopes da Silva FH (1999) Dynamics of the human alpha rhythm: evidence for non-linearity. Clin Neurophysiol 110:1801–1813

Toda HY, Philipps PCB (1993) Vector autoregressions and causality. Econometrica 61:1367–1393

Valdés-Sosa PA (2004) Spatio-temporal autoregressive models defined over brain manifolds. Neuroinformatics 2:239–250

Veeramani B, Narayanan K, Prasad A, Spanias A, Iasemidis LD (2003) On the use of directed transfer function for nonlinear systems. In: Hamza MH (ed) Simulation and modelling. IASTED/ACTA Press, Calgary, pp 270–274