Autoregressive Causal Relation: Digital Filtering Approach to Causality Measures in Frequency Domain

Tomáš Bořil and Pavel Sovka

Department of Circuit Theory in the Faculty of Electrical Engineering, Czech Technical University in Prague, Czech Republic. e-mail: borilt@gmail.com and sovka@fel.cvut.cz, phone: (+420) 224 352 288, fax: (+420) 233 339 805, mail address: Tomas Boril, CVUT FEL K13131, Technicka 2, Prague 6, 166 27, Czech Republic.

Abstract

A novel measure of the Autoregressive Causal Relation based on a multivariate autoregressive model is proposed. It reveals the strength of the connections among a simultaneous time series and also the direction of the information flow. It is defined in the frequency domain, similar to the formerly published methods such as: Directed Transfer Function, Direct Directed Transfer Function, Partial Directed Coherence, and Generalized Partial Directed Coherence. Compared to the Granger causality concept, frequency decomposition extends the possibility to reveal the frequency rhythms participating on the information flow in causal relations.

The Autoregressive Causal Relation decomposes diagonal elements of a spectral matrix and enables a user to distinguish between direct and indirect causal relations. The main advantage lies in its definition using power spectral densities, thus allowing for a clear interpretation of strength of causal relation in meaningful physical terms.

The causal measures can be used in neuroscience applications like the analysis of underlying structures of brain connectivity in neural multichannel time series during different tasks measured via electroencephalography or functional magnetic resonance imaging, or other areas using the multivariate autoregressive models for causality modeling like econometrics or atmospheric physics but this paper is focused on theoretical aspects and model data examples in order to illustrate a behavior of methods in known situations.

Keywords: Autoregressive processes, Frequency domain analysis, Brain modeling, Electroencephalography.

1. Introduction

Causal relations analysis is an important task for revealing connections among simultaneous time series as it identifies not only the strength of the relations but also the direction of the information flow.

This paper is focused on methods based on multivariate autoregressive (MVAR) models [1]. Once suitable model parameters are fitted to the data, the methods try to estimate the causal connections from these model parameters. There are many other alternative approaches to causal relations analysis, such as the concept of graphical models enabling one to distinguish between direct and indirect connections in the time series [2, 3], the concept of Transfer entropy [4], Dynamic Causal Modeling [5, 6, 7] featuring biophysical modeling using differential equations and Bayesian statistical methods for parameter estimation, Structural Equation Modeling [8], Phase Dynamics [9, 10], Phase Slope Index [11, 12], maximum likelihood models [13], etc. However, the key feature of MVAR model techniques is that it is very general, mathematically well-founded and it covers many of the most common interdependences. The natural multichannel approach allows for an analysis of the direct and indirect causal paths, and it is able to model bidirectional causal relations.

The Granger causality technique is frequently used for causal relation analysis in MVAR models [14]. The original Granger causality concept [15] reveals whether a knowledge of past values of one variable (source variable) helps to predict the current value of another variable (target variable). This approach is based only on a bivariate pairwise test classifying the strength of the causal connection by one real non-negative number in one direction (see Fig. 1a). The MVAR model is fitted to the time series and if the variance of the residual prediction error of the target variable is lower in the case where the previous values of the source variable are included into the model, then we say the source variable causes the target variable. This concept was primarily used in econometrics for identifying underlying structures of a time series such as livestock prices [16].

If the data contains more time series, two problems may appear. The first problem, a *sequential driving*, occurs when the causal connection from the first to the second variable is completely mediated via the third variable (see Fig. 1b). In this case, a pairwise test will also detect an indirect causal relation from the first to the second because it cannot distinguish between the direct and indirect connections. The second problem, a *different delay driving*, occurs when the first variable drives both the second and the third variable, but the driving of the second variable has a smaller delay than the driving of the third one (see Fig. 1c). Then, samples of the second vari-

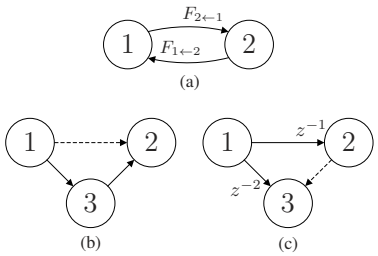

Figure 1: Typical properties of causality analysis. (a) Bidirectional connection evaluated via Granger causality with source variable 1 and target variable 2, giving causal strength $F_{2\leftarrow1}$. The analysis with source variable 2 and target variable 1 gives another result $F_{1\leftarrow2}$ as the causal relations are asymmetric, i.e., directional. (b) The sequential driving problem should be detected as direct causal relations $3 \leftarrow 1$ and $2 \leftarrow 3$ however some methods also detect the false indirect $2 \leftarrow 1$. (c) Different delay driving should be detected as direct causal relations $2 \leftarrow 1$ and $3 \leftarrow 1$ however some methods also detect the false indirect $3 \leftarrow 2$ as 2 contains some information which can help to predict future values of 3 (the z-domain delay represents the causal connection delay).

able contain information which helps predict future samples of the third variable. A pairwise method detects a causal relation from the second to the third variable and cannot distinguish that this is an indirect one. Both problems with indirect connections were solved via a multivariate extension, referred to as Conditional Granger causality [17]. The Conditional Granger causality works with MVAR model which includes all the time series in order to find the direct connections and recognize and eliminate the indirect ones.

In recent years, spectral analysis of neurophysiological data [18], such as functional magnetic resonance imaging, magnetoencephalography or electroencephalography have played an important role. To extend the basic spectral approaches, advanced techniques based on MVAR models have been used for analyzing causal connections among brain structures. The common character of these applications is to find which frequency rhythms participate in the causal connection. A number of methods for frequency decomposition have been published, with their main focus being on the analysis of the frequency rhythms participating in the information transfer. However, only a few of the methods claim to be able to distinguish between the direct and indirect connections.

Several selected methods enabling one to distinguish between the direct and indirect connections will be described. The Directed Transfer Function (DTF) [19, 20] is able to decompose the causal relations in the frequency domain but does not distinguish the direct connections from the indirect ones. The Partial Directed Coherence (PDC) [21, 22] solves this problem by only recognizing the direct connections, but uses a normalization that causes the inability to compare the strength of the coupling among variables. A lower value of the PDC may instead correspond to a stronger relation [23]. The Direct Directed Transfer Function (dDTF) [24, 25, 22, 26, 27] claims to improve the DTF in order to distinguish the indirect connections. A straightforward frequency transform of the Granger causality (originally by Geweke [28]) leads to problems when utilizing three variables, such as the occasional occurence of negative values which have no meaning in terms of causality. [29] applied a partition matrix technique to overcome the problem. However, this method is limited to three variables only. The Generalized Partial Directed Coherence (GPDC) [23, 30, 31] modifies the PDC using additional normalization to make the values more comparable. Another method, Renormalized PDC [32] provides similar outcomes as GPDC.

1.1. Problems with interpretation; new approach

Although all these methods claim to distinguish between direct and indirect causal connections in the frequency domain, the definitions and, consequently, interpretation of the results vary. [33] states: *"Unfavorable feature of PDC is its weak dependence on frequency (practically 'flat' spectrum), which does not permit to distinguish well the role of different rhythms."* and *"DTF detects not only direct but also indirect flows. This feature may be important when estimating transmissions from implanted or subdural electrodes. However in these cases Direct Directed Transfer Function which combines DTF with partial coherence, may be used."* The dDTF is described by [24]: *"The maxima of dDTF may better (in comparison to DTF) reflect a frequency of rhythm being a carrier for information flow."* The GPDC is explained by [30]: *"The square modulus of GPDC value from jth time series by ith series can be understood intuitively as the proportion of the power spectra of the jth time series, which is sent to the ith series considering the effects of the other series."* In Section 3, it is shown that this statement is not accurate and the newly proposed Autoregressive Causal Relation (ACR) that uses the power spectral density (PSD) is the same as the square modulus of GPDC only in the case of two variables with one unidirectional causal relation. [34] specifies: *"GPDC determines the correct interaction structure in terms of direct causal effects, but its absolute values lack of straightforward interpretability."* [35] mentions a possibility of misleading results of PDC, GPDC, and dDTF methods. [36] summarizes: *"Comparing the DTF, PDC and their derivatives, which of these measures is the most advantageous and accurate is still an open question."*

The aim of the ACR method proposed in this paper is to try to bring an unambiguous definition of causal relation in the frequency domain based on multivariate autoregressive models, with a clear PSD interpretation defined as a ratio of the PSD of the incoming causal component to the total PSD of the target variable.

The ACR method agrees with the concept of the suggested criteria in [37] and brings an idea how to calculate each causal component from coefficients of a general MVAR model containing multiple causal relations including the bidirectional ones.

2. Definitions

2.1. Multivariate autoregressive model

Let us consider an MVAR process consisting of k real number stationary variables with zero means (e.g., corresponding to particular simultaneous channels)

$$\mathbf{X}(t) = (X_1(t), X_2(t), \ldots, X_k(t))^T \tag{1}$$

where t stands for the sample index in a discrete-time domain. Considering an unpredictable component of the model defined as a vector of mutually uncorrelated white Gaussian noises

$$\mathbf{E}(t) = (\varepsilon_1(t), \varepsilon_2(t), \ldots, \varepsilon_k(t))^T \tag{2}$$

with zero means and variances $\sigma_1^2, \sigma_2^2, \ldots, \sigma_k^2$ respectively, the generating equations of the MVAR process of the order of p are

$$\begin{aligned} X_i(t) = {} & \varepsilon_i(t) + \\ & + a_{i1}(1) X_1(t-1) + \cdots + a_{ik}(1) X_k(t-1) + \\ & + \cdots + \\ & + a_{i1}(p) X_1(t-p) + \cdots + a_{ik}(p) X_k(t-p) \end{aligned} \tag{3}$$

for each $i = 1, \ldots, k$ where a are model coefficients (real numbers). In this way, the actual sample of each variable is described as a summation of Gaussian white noise and a linear combination of p previous values of every variable. Considering coefficient matrices

$$\mathbf{A}(n) = \begin{bmatrix} -a_{11}(n) & \cdots & -a_{1k}(n) \\ \vdots & \ddots & \vdots \\ -a_{k1}(n) & \cdots & -a_{kk}(n) \end{bmatrix} \tag{4}$$

for $n = 0, \ldots, p$, where $\mathbf{A}(0) = \mathbf{I}$ is the identity matrix, the MVAR model can be written as

$$\sum_{n=0}^{p} \mathbf{A}(n) \mathbf{X}(t-n) = \mathbf{E}(t). \tag{5}$$

The conversion of (5) into the frequency domain requires the discrete-time Fourier transform (DTFT). Matrix $\mathbf{A}(n)$ can be transformed as

$$\mathbf{A}(f) = \begin{bmatrix} A_{11}(f) & \cdots & A_{1k}(f) \\ \vdots & \ddots & \vdots \\ A_{k1}(f) & \cdots & A_{kk}(f) \end{bmatrix} \tag{6}$$

where

$$A_{ij}(f) = \delta_{ij} - \sum_{n=1}^{p} a_{ij}(n) e^{-\iota 2\pi f n} \tag{7}$$

for row index $i = 1, \ldots, k$ and column index $j = 1, \ldots, k$, where ι is the imaginary unit and $\delta_{ij} = 1$ for $i = j$, $\delta_{ij} = 0$ otherwise. Since MVAR signals are stationary random processes, i.e., finite power signals, their discrete-time Fourier transform (DTFT) does not exist in the usual sense and a spectral measure needs to be introduced. One can use the Wiener-Khinchin theorem [38] stating the PSD of an ergodic wide-sense stationary random process is the Fourier transform of the corresponding autocorrelation function.

Although many authors perform the DTFT or z-transform of (5) with stationary MVAR variables and noises leading to (12) or corresponding z-transform expression, e.g., [19, 24, 17, 29, 39], the procedure should be clarified.

Let us assume a final-length segment $X_{wi}(t)$ of one realization (sample function) of stationary process $X_i(t)$ containing N nonzero samples. The DTFT of the segment is $X_i(f)$ where $f \in [-0.5, 0.5]$ is the normalized frequency. Then, a periodogram (an estimate of the PSD of the signal) can be introduced [40]

$$\hat{S}_{ii,N}(f) = \frac{1}{N} |X_i(f)|^2 = \frac{1}{N} X_i(f) X_i^*(f) \tag{8}$$

where $*$ is transposition and complex conjugation (Hermitian transpose). The PSD of $X_i(t)$ using the Wiener-Khinchin theorem can be expressed as [41]

$$S_{ii}(f) = \lim_{N \to \infty} E\left[\hat{S}_{ii,N}(f)\right] \tag{9}$$

where E denotes an expectation value operator.

In the following equations, speaking about the Fourier transform or z-transform of the stationary processes, the final-length segments of the signals are assumed and the final expression of PSD is derived from the periodogram via (9). This approach has several advantages. The use of Fourier and z-transform instead of convolutions or correlations leads to more straightforward expressions. It allows easy interpretation by means of digital filtering. It is also compatible with the notation used by the authors of the methods which we are comparing our technique with.

Under the above stated conditions, using a DTFT vector of segmented realizations of variables (1)

$$\mathbf{X}(f) = (X_1(f), X_2(f), \ldots, X_k(f))^T \tag{10}$$

and a DTFT vector of corresponding unpredictable (mutually uncorrelated) components (2)

$$\mathbf{E}(f) = (\varepsilon_1(f), \varepsilon_2(f), \ldots, \varepsilon_k(f))^T, \tag{11}$$

the DTFT of the MVAR model can be finally expressed as

$$\mathbf{A}(f) \mathbf{X}(f) = \mathbf{E}(f). \tag{12}$$

This equation can be rewritten as [24]

$$\mathbf{X}(f) = \mathbf{A}^{-1}(f) \mathbf{E}(f) = \mathbf{H}(f) \mathbf{E}(f) \tag{13}$$

where

$$\mathbf{H}(f) = \begin{bmatrix} H_{11}(f) & \cdots & H_{1k}(f) \\ \vdots & \ddots & \vdots \\ H_{k1}(f) & \cdots & H_{kk}(f) \end{bmatrix} = \mathbf{A}(f)^{-1} \tag{14}$$

is the transfer function of the MVAR system. A PSD representation, commonly referred to as the spectral matrix, is

$$\begin{aligned} \mathbf{S}(f) = {} & \lim_{N \to \infty} E\left[\frac{1}{N} \mathbf{X}(f) \mathbf{X}^*(f)\right] = \\ = {} & \lim_{N \to \infty} E\left[\frac{1}{N} \mathbf{H}(f) \mathbf{E}(f) \mathbf{E}^*(f) \mathbf{H}^*(f)\right] = \\ = {} & \mathbf{H}(f) \mathbf{V} \mathbf{H}^*(f) \end{aligned} \tag{15}$$

where $\mathbf{V}$ is a diagonal matrix with variances $\sigma_1^2, \sigma_2^2, \ldots, \sigma_k^2$ and zero covariances between residual noises.

2.2. Instantaneous causality

The original pairwise Granger causality concept [15] defined instantaneous causality as a special form of interaction between two time series without a time delay. This permits an unknown common driving source to be omitted from the model. But the external driving source not included in the model causes a nonzero covariance between residual noises which cannot be explained by the MVAR model.

All the further discussed methods of multiple variable time series analysis are focused only on interactions involving a time delay. As mentioned in [21], the MVAR identification problem is limited to the exclusive consideration of the past values in the prediction. In the following text, stationary simultaneous time series with zero means and linear multivariate autoregressive character are assumed. All the time series that can have a causal impact on the others are included in the model, leading to zero covariances among the residual noises in the MVAR model ($\mathbf{V}$ is diagonal).

2.3. MVAR model estimation

If the time series $\mathbf{X}(t)$ satisfies stationarity, zero means and MVAR conditions then covariances between the elements of $\mathbf{E}(t)$ are zero. Coefficients $\mathbf{A}(n)$ and variances of $\mathbf{E}(t)$ can be obtained by an MVAR identification procedure. For example, by solving multichannel Yule-Walker equations [42] obtained as a result of the minimization of the prediction error variance, or by using more robust procedures such as the Vieira-Morf algorithm[1] [43, 44] based on the idea of maximum entropy [45, 46]. A detailed comparison of the MVAR estimators was performed in [47] and the Vieira-Morf method[2] was recommended as the most robust.

In order to fit the MVAR model to real data, the order p of the model has to be chosen. The Akaike Information Criterion (AIC) or the Bayesian Information Criterion (BIC) described, e.g., in [48, 44, 17] are often used for the model order estimation. However, if the MVAR conditions are not exactly met, the MVAR model is not able to describe residual noises and these criteria may fail.

2.4. Direct Directed Transfer Function (dDTF)

Partial coherence can be defined as [24]

$$\chi_{ij}^2(f) = \frac{M_{ij}^2(f)}{M_{jj}(f)\,M_{ii}(f)} \tag{16}$$

where $M_{ij}(f)$ is a minor of the spectral matrix $\mathbf{S}(f)$ obtained by removing its ith row and jth column. To analyze a causal flow from source variable $X_j(t)$ to target variable $X_i(t)$, utilizing the full frequency DTF (ffDTF) [24]

$$\eta_{ij}^2(f) = \frac{\left|H_{ij}(f)\right|^2}{\sum\limits_f \sum\limits_{m=1}^{k} \left|H_{im}(f)\right|^2} \tag{17}$$

the Direct Directed Transfer Function (dDTF)[3] [24] is

$$\delta_{ij}(f) = \chi_{ij}(f)\,\eta_{ij}(f). \tag{18}$$

The dDTF is a combination of the partial coherence and the transfer function in the direction that is supposed to differentiate the direct causal connections from the indirect ones.

2.5. Generalized Partial Directed Coherence (GPDC)

To analyze a causal flow from source variable $X_j(t)$ to target variable $X_i(t)$, the GPDC[4] is defined by [23] as

$$\mathrm{GPDC}_{ij}(f) = \frac{\frac{1}{\sigma_i}\left|A_{ij}(f)\right|}{\sqrt{\sum\limits_{m=1}^{k} \frac{1}{\sigma_m^2}\left|A_{mj}(f)\right|^2}}. \tag{19}$$

It utilizes coefficients of an MVAR model transformed into the frequency domain and focuses only on the direct connection between the two series in the given direction. The square modulus of GPDC is

$$\mathrm{GPDC}_{ij}^2(f) = \frac{\frac{1}{\sigma_i^2}\left|A_{ij}(f)\right|^2}{\sum\limits_{m=1}^{k} \frac{1}{\sigma_m^2}\left|A_{mj}(f)\right|^2}. \tag{20}$$

It can be seen that the denominator normalizes the equation so the summation of all $\mathrm{GPDC}_{ij}^2(f)$ for the given source variable $X_j(t)$ is equal to one.

2.6. Autoregressive causal relation (ACR)

2.6.1. Filtering approach to causal relations

In order to explain model (5) using the digital filtering approach, the MVAR model should be transformed into the z-domain. Let us consider a matrix

$$\mathbf{A}(z) = \begin{bmatrix} A_{11}(z) & \cdots & A_{1k}(z) \\ \vdots & \ddots & \vdots \\ A_{k1}(z) & \cdots & A_{kk}(z) \end{bmatrix} \tag{21}$$

where

$$A_{ij}(z) = \delta_{ij} - \sum_{n=1}^{p} a_{ij}(n) z^{-n} \tag{22}$$

for row index $i = 1, \ldots, k$ and column index $j = 1, \ldots, k$, where $\delta_{ij} = 1$ for $i = j$, $\delta_{ij} = 0$ otherwise. Let us introduce a z-transform of a vector of segmented realizations of variables (1)

$$\mathbf{X}(z) = (X_1(z), X_2(z), \ldots, X_k(z))^T \tag{23}$$

and a z-transform of a vector of corresponding unpredictable components (2)

$$\mathbf{E}(z) = (\varepsilon_1(z), \varepsilon_2(z), \ldots, \varepsilon_k(z))^T. \tag{24}$$

[1] Implemented, e.g., in a BioSig toolbox for Octave and Matlab available at http://biosig.sourceforge.net/

[2] Although it was incorrectly denominated as Nutall-Strand. The comment in the BioSig toolbox by the same author corrects this information.

[3] The dDTF is available in the BioSig toolbox.

[4] The GPDC is also available in the BioSig toolbox.

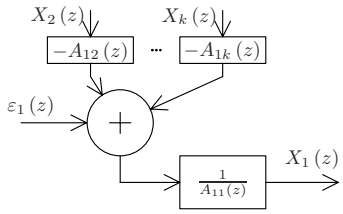

Figure 2: MVAR variable $X_1(z)$ filtering diagram.

Then the z-transform of the MVAR model (5) has the form of

$$\mathbf{A}(z)\,\mathbf{X}(z) = \mathbf{E}(z)\,. \tag{25}$$

Using this notation, let us consider an example for target variable $X_1(t)$. Applying the z-transform on (3) with $i = 1$, one can get

$$X_1(z) = \frac{1}{A_{11}(z)}\left(\varepsilon_1(z) - \sum_{j=2}^{k} A_{1j}(z)\,X_j(z)\right) \tag{26}$$

where $1/A_{11}(z)$ is an all-pole infinite impulse response (IIR) filter and $-A_{1j}(z)$ are finite impulse response (FIR) filters. The corresponding diagram is depicted in Fig. 2. The advantage of this approach is that it does not detect false indirect connections.

In general, speaking about the target variable $X_i(z)$, source variables $X_j(z)$ where $j \neq i$ are filtered with FIR filters $-A_{ij}(z)$ where

$$A_{ij}(z) = -a_{ij}(1)\,z^{-1} - \cdots - a_{ij}(p)\,z^{-p} \tag{27}$$

and all these source components including white noise $\varepsilon_i(t)$ are additionally filtered with an IIR filter

$$\frac{1}{A_{ii}(z)} = \frac{1}{1 - a_{ii}(1)\,z^{-1} - \cdots - a_{ii}(p)\,z^{-p}}. \tag{28}$$

The z-transform of the target variable is then

$$X_i(z) = \frac{1}{A_{ii}(z)}\left(\varepsilon_i(z) - \sum_{\substack{j=1\\ j\neq i}}^{k} A_{ij}(z)\,X_j(z)\right). \tag{29}$$

2.6.2. Power spectral density representation

In order to get the PSD of variable $X_i(t)$, (29) must be evaluated on the unit circle

$$X_i(f) = \frac{1}{A_{ii}(f)}\left(\varepsilon_i(f) - \sum_{\substack{j=1\\ j\neq i}}^{k} A_{ij}(f)\,X_j(f)\right). \tag{30}$$

Because of its random character, one has to derive the PSD representation of the target variable to analyze all source components. Using (8) and (9), the PSD of $X_i(t)$ can be written as[5]

$$S_{ii}(f) = \lim_{N\to\infty} E\left[\frac{1}{N} X_i(f)\,X_i^*(f)\right]. \tag{31}$$

[5]By reason of multiplication of two summations, let us alter the index j in (30) to m and n respectively.

$$\begin{aligned} S_{ii}(f) = \frac{1}{A_{ii}(f)\,A_{ii}^*(f)} \lim_{N\to\infty} E\Bigg[\frac{1}{N}\Bigg(&\varepsilon_i(f)\,\varepsilon_i^*(f) - \\ &- \varepsilon_i(f)\left(\sum_{\substack{n=1\\ n\neq i}}^{k} A_{in}^*(f)\,X_n^*(f)\right) - \\ &- \left(\sum_{\substack{m=1\\ m\neq i}}^{k} A_{im}(f)\,X_m(f)\right)\varepsilon_i^*(f) + \\ &+ \left(\sum_{\substack{m=1\\ m\neq i}}^{k} A_{im}(f)\,X_m(f)\right)\left(\sum_{\substack{n=1\\ n\neq i}}^{k} A_{in}^*(f)\,X_n^*(f)\right)\Bigg)\Bigg]. \end{aligned} \tag{32}$$

It holds

$$\begin{aligned} \lim_{N\to\infty} E\left[\tfrac{1}{N}\varepsilon_i(f)\,\varepsilon_i^*(f)\right] &= \sigma_i^2, \\ \lim_{N\to\infty} E\left[\tfrac{1}{N} X_m(f)\,X_n^*(f)\right] &= S_{mn}(f)\,. \end{aligned} \tag{33}$$

Variable $\varepsilon_i(f)$ is given by (12)

$$\varepsilon_i(f) = \sum_{m=1}^{k} A_{im}(f)\,X_m(f) \tag{34}$$

and similarly

$$\varepsilon_i^*(f) = \sum_{n=1}^{k} A_{in}^*(f)\,X_n^*(f). \tag{35}$$

With substitution

$$T_{imn}(f) = A_{im}(f)\,A_{in}^*(f)\,S_{mn}(f)\,, \tag{36}$$

one gets

$$\begin{aligned} S_{ii}(f) = \frac{1}{|A_{ii}(f)|^2}\Bigg(&\sigma_i^2 - \sum_{m=1}^{k}\sum_{\substack{n=1\\ n\neq i}}^{k} T_{imn}(f) - \\ &- \sum_{\substack{m=1\\ m\neq i}}^{k}\sum_{n=1}^{k} T_{imn}(f) + \sum_{\substack{m=1\\ m\neq i}}^{k}\sum_{\substack{n=1\\ n\neq i}}^{k} T_{imn}(f)\Bigg). \end{aligned} \tag{37}$$

From the subtraction of the last two terms, only the part of $n = i$ remains

$$\begin{aligned} S_{ii}(f) = \frac{1}{|A_{ii}(f)|^2}\Bigg(&\sigma_i^2 - \sum_{m=1}^{k}\sum_{\substack{n=1\\ n\neq i}}^{k} T_{imn}(f) - \\ &- \sum_{\substack{m=1\\ m\neq i}}^{k} A_{im}(f)\,A_{ii}^*(f)\,S_{mi}(f)\Bigg). \end{aligned} \tag{38}$$

To simplify this expression, let us define a substitution

$$U_{im}(f) = A_{im}(f) A_{ii}^{*}(f) S_{mi}(f). \tag{39}$$

Then

$$S_{ii}(f) = \\ = \frac{1}{|A_{ii}(f)|^2} \left(\sigma_i^2 - \sum_{m=1}^{k} \sum_{\substack{n=1 \\ n \neq i}}^{k} T_{imn}(f) - \sum_{\substack{m=1 \\ m \neq i}}^{k} U_{im}(f) \right). \tag{40}$$

To separate $S_{ii}(f)$ to individual causal components, one should divide the first summation into parts as follows

$$\begin{aligned} &- \sum_{m=1}^{k} \sum_{\substack{n=1 \\ n \neq i}}^{k} T_{imn}(f) = \\ &= - \sum_{\substack{m=1 \\ m \neq i}}^{k} \sum_{\substack{n=1 \\ n \neq i}}^{k} T_{imn}(f) - \sum_{\substack{n=1 \\ n \neq i}}^{k} A_{ii}(f) A_{in}^{*}(f) S_{in}(f) = \\ &= - \sum_{\substack{m=1 \\ m \neq i}}^{k} \sum_{\substack{n=1 \\ n \neq i}}^{k} T_{imn}(f) - \sum_{\substack{m=1 \\ m \neq i}}^{k} A_{ii}(f) A_{im}^{*}(f) S_{im}(f) = \\ &= - \sum_{\substack{m=1 \\ m \neq i}}^{k} \sum_{\substack{n=1 \\ n \neq i}}^{k} T_{imn}(f) - \sum_{\substack{m=1 \\ m \neq i}}^{k} U_{im}^{*}(f) = \\ &= - \sum_{\substack{m=1 \\ m \neq i}}^{k} \sum_{\substack{n=1 \\ n \neq i \\ n \neq m}}^{k} T_{imn}(f) - \sum_{\substack{m=1 \\ m \neq i}}^{k} A_{im}(f) A_{im}^{*}(f) S_{mm}(f) - \\ &\quad - \sum_{\substack{m=1 \\ m \neq i}}^{k} U_{im}^{*}(f) \end{aligned} \tag{41}$$

where $S_{mi}(f) = S_{im}^{*}(f)$ is used.

The PSD is then

$$S_{ii}(f) = \frac{1}{|A_{ii}(f)|^2} \left(\sigma_i^2 - \sum_{\substack{m=1 \\ m \neq i}}^{k} \sum_{\substack{n=1 \\ n \neq i \\ n \neq m}}^{k} T_{imn}(f) - \right. \\ \left. - \sum_{\substack{m=1 \\ m \neq i}}^{k} |A_{im}(f)|^2 S_{mm}(f) - \sum_{\substack{m=1 \\ m \neq i}}^{k} U_{im}^{*}(f) - \sum_{\substack{m=1 \\ m \neq i}}^{k} U_{im}(f) \right). \tag{42}$$

Let us use the replacements

$$- \sum_{\substack{m=1 \\ m \neq i}}^{k} \sum_{\substack{n=1 \\ n \neq i \\ n \neq m}}^{k} T_{imn}(f) = - \sum_{\substack{m=1 \\ m \neq i}}^{k-1} \sum_{\substack{n=m+1 \\ n \neq i}}^{k} 2\mathrm{Re}\left(T_{imn}(f)\right), \tag{43}$$

and

$$- \sum_{\substack{m=1 \\ m \neq i}}^{k} U_{im}(f) - \sum_{\substack{m=1 \\ m \neq i}}^{k} U_{im}^{*}(f) = - \sum_{\substack{m=1 \\ m \neq i}}^{k} 2\mathrm{Re}\left(U_{im}(f)\right). \tag{44}$$

Let us define the causal components

$$F_{i \leftarrow i}(f) = \frac{\sigma_i^2}{|A_{ii}(f)|^2}, \tag{45}$$

$$F_{i \leftarrow m,n}(f) = \frac{-2\mathrm{Re}\left(T_{imn}(f)\right)}{|A_{ii}(f)|^2}, \tag{46}$$

and

$$F_{i \leftarrow m}(f) = \frac{-|A_{im}(f)|^2 S_{mm}(f) - 2\mathrm{Re}\left(U_{im}(f)\right)}{|A_{ii}(f)|^2}. \tag{47}$$

The final form of the PSD decomposition of $S_{ii}(f)$ to individual causal components can be written in the form of

$$S_{ii}(f) = \\ = F_{i \leftarrow i}(f) + \sum_{\substack{m=1 \\ m \neq i}}^{k-1} \sum_{\substack{n=m+1 \\ n \neq i}}^{k} F_{i \leftarrow m,n}(f) + \sum_{\substack{m=1 \\ m \neq i}}^{k} F_{i \leftarrow m}(f). \tag{48}$$

The key idea is to split the PSD $S_{ii}(f)$ of the target variable $X_i(t)$ into the part $F_{i \leftarrow i}(f)$ caused by the filtration of the noise $\varepsilon_i(t)$ through the IIR filter ${A_{ii}}^{-1}(z)$, and elements caused by other variables $F_{i \leftarrow m}(f)$ and $F_{i \leftarrow m,n}(f)$. The $F_{i \leftarrow m}(f)$ part corresponds to a direct causal relation from another variable $X_m(t)$. The $F_{i \leftarrow m,n}(f)$ part is present only when two variables $X_m(t)$ and $X_n(t)$ with nonzero cross power spectral density (CPSD) $S_{mn}(f)$ both have direct causal connections to the target variable $X_i(t)$. $F_{i \leftarrow m}(f)$ and $F_{i \leftarrow m,n}(f)$ can be positive or negative so it can model both the increase and decrease of the PSD caused by the causal relations.

2.6.3. New measures

Considering (48) and according to theoretical criteria suggested in [37], we propose new measures for causal relations analysis in the frequency domain, based on the decomposition of the target variable into separate causal components.

The *absolute autoregressive causal relation* (absolute ACR) corresponds directly to the portion of the PSD of the target variable $X_i(t)$ not caused by other variables or the portion of the PSD caused by individual source variables $X_m(t)$

$$ACR_{i \leftarrow mABS}(f) = \begin{cases} F_{i \leftarrow i}(f), & \text{for } m = i \\ F_{i \leftarrow m}(f), & \text{for } m \neq i \end{cases} \tag{49}$$

and the portion of the PSD caused by couples of source variables $X_m(t)$ and $X_n(t)$, present only in the case that both variables have direct causal relation with the $X_i(t)$ (nonzero $A_{im}(f)$ and $A_{in}(f)$) and the couple of source variables has a nonzero CPSD $S_{mn}(f)$

$$ACR_{i \leftarrow m,nABS}(f) = F_{i \leftarrow m,n}(f), \quad m \neq i, n \neq i, n > m. \tag{50}$$

The key property of this approach is, as it can be seen from (48), for every target variable $X_i(t)$, the summation of its absolute

components is equal to the PSD of the variable

$$\sum_{m=1}^{k} ACR_{i\leftarrow mABS}(f) + \\ + \sum_{\substack{m=1\\ m\neq i}}^{k-1} \sum_{\substack{n=m+1\\ n\neq i}}^{k} ACR_{i\leftarrow m,nABS}(f) = S_{ii}(f),\ \forall i = 1\ldots k. \tag{51}$$

This feature, along with the ability of modeling both the increase and decrease of the PSD caused by direct causal relations, are the main advantages of the proposed measure as compared to state-of-the-art methods.

The *relative autoregressive causal relation* normalizes the absolute ACR by the PSD of the target[6] $S_{ii}(f)$ and hence, gives a quantity of the fraction of the contribution to the target variable

$$ACR_{i\leftarrow mREL}(f) = \begin{cases} \frac{F_{i\leftarrow i}(f)}{S_{ii}(f)}, & \text{for } m = i \\ \frac{F_{i\leftarrow m}(f)}{S_{ii}(f)}, & \text{for } m \neq i \end{cases} \tag{52}$$

and

$$ACR_{i\leftarrow m,nREL}(f) = \frac{F_{i\leftarrow m,n}(f)}{S_{ii}(f)},\quad m \neq i, n \neq i, n > m. \tag{53}$$

For every target variable $X_i(t)$, the summation of its relative components is equal to 1

$$\sum_{m=1}^{k} ACR_{i\leftarrow mREL}(f) + \\ + \sum_{\substack{m=1\\ m\neq i}}^{k-1} \sum_{\substack{n=m+1\\ n\neq i}}^{k} ACR_{i\leftarrow m,nREL}(f) = 1, \qquad \forall i = 1\ldots k. \tag{54}$$

This allows one to compare the strength of the individual causal components in a percentage.

2.7. Statistical evaluation

Although evaluations in this paper are primarily conducted on model data (see Sec. 4), a procedure of real data analysis using a statistical evaluation of causal measures is also outlined. Since the causal measures have a highly nonlinear relation to the time series data from which they are derived and distributions of their estimators are not well established, the use of a surrogate data method using an unwindowed Fourier transform introduced in [49] is recommended [24, 50]. The surrogate data is constructed by randomizing the phase in the Fourier spectra of the raw data, leading to data with a very similar amplitude spectra but destroyed causal relations. The causal measures are then calculated from the surrogate data. By repeating this process a number of times, an empirical estimate of a probability density function (histogram) is constructed, which corresponds to the null hypothesis that there is no causal connection.

[6] A normalization by the PSD of the source variable does not make sense in terms of dividing the total power to its parts because the sources in the MVAR model definition are copied to targets, not divided.

3. Analysis of the two-variable case

As mentioned above, the square modulus of GPDC is often interpreted as the proportion of the power spectra of the target variable caused by the causal relation [30]. To illustrate the fact that this statement may not be accurate, a comparison with the relative ACR on a two-variable case is conducted in the following.

Using (20), a general equation for causal relation in the direction $2 \leftarrow 1$ can be obtained

$$\mathrm{GPDC}^2_{21}(f) = \frac{\frac{1}{\sigma_2^2}|A_{21}(f)|^2}{\frac{1}{\sigma_1^2}|A_{11}(f)|^2 + \frac{1}{\sigma_2^2}|A_{21}(f)|^2} = \\ = \frac{\sigma_1^2|A_{21}(f)|^2}{\sigma_2^2|A_{11}(f)|^2 + \sigma_1^2|A_{21}(f)|^2}. \tag{55}$$

Using (52) (the component (53) is not present for $k = 2$) without the replacement of (44) in (47), one gets

$$ACR_{2\leftarrow 1REL}(f) = \frac{-|A_{21}(f)|^2 S_{11}(f) - U_{21}(f) - U_{21}^*(f)}{|A_{22}(f)|^2 S_{22}(f)} \tag{56}$$

where

$$U_{21}(f) = A_{21}(f) A_{22}^*(f) S_{12}(f), \tag{57}$$

this together with (15) results in

$$\begin{aligned} &ACR_{2\leftarrow 1REL}(f) = \\ &= \left[|A_{22}(f)|^2 \left(\sigma_2^2|A_{11}(f)|^2 + \sigma_1^2|A_{21}(f)|^2\right)\right]^{-1} \\ &\left[\sigma_1^2|A_{21}(f)|^2|A_{22}(f)|^2 - \sigma_2^2|A_{12}(f)|^2|A_{21}(f)|^2 + \right. \\ &+ \sigma_2^2 A_{11}(f) A_{22}(f) A_{12}^*(f) A_{21}^*(f) + \\ &\left. + \sigma_2^2 A_{11}^*(f) A_{22}^*(f) A_{12}(f) A_{21}(f)\right]. \end{aligned} \tag{58}$$

It is evident that if the model does not contain a causal relation $1 \leftarrow 2$, i.e., $A_{12}(f) = 0$, then $\mathrm{GPDC}^2_{21}(f) = ACR_{2\leftarrow 1REL}(f)$. If the model contains the causal relation $1 \leftarrow 2$, the GPDC produces a result different from the ACR. In contrast to the ACR, GPDC does not respect the effect of the feedback relation $1 \leftarrow 2$.

4. Model data illustration

In order to compare the results of the dDTF, GPDC and ACR, an advanced MVAR model used in [21, 17] was chosen (see Fig. 3)

$$\begin{aligned} X_1(t) &= \varepsilon_1(t) + 0.95\sqrt{2}X_1(t-1) - 0.9025X_1(t-2), \\ X_2(t) &= \varepsilon_2(t) + 0.5X_1(t-2), \\ X_3(t) &= \varepsilon_3(t) - 0.4X_1(t-3), \\ X_4(t) &= \varepsilon_4(t) - 0.5X_1(t-2) + 0.25\sqrt{2}X_4(t-1) + \\ &\quad + 0.25\sqrt{2}X_5(t-1), \\ X_5(t) &= \varepsilon_5(t) - 0.25\sqrt{2}X_4(t-1) + 0.25\sqrt{2}X_5(t-1) \end{aligned} \tag{59}$$

where t stands for the index of a discrete-time instance, ε_1, ε_2, ε_3, ε_4, and ε_5 are Gaussian white noises with zero means and variances 0.6, 0.5, 0.3, 0.3, and 0.6 respectively.

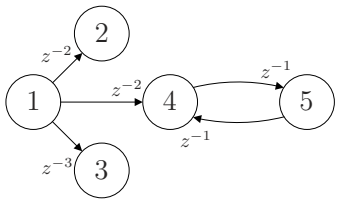

Figure 3: A diagram of the model data connection. The arrows denote causal relations, the z-domain delay represents the connection delay (the minimal value of the time shift in the equation). In addition to the bidirectional connection $4 \leftrightarrow 5$, the model also contains indirect causal connections. The sequential driving $5 \leftarrow 4 \leftarrow 1$ can cause false detection of the indirect relation $5 \leftarrow 1$. And the differently delayed drivings $2 \leftarrow 1$ and $3 \leftarrow 1$ can cause false detection of the indirect $3 \leftarrow 2$ because samples of 2 contain information, which helps with the prediction of 3, although, in fact, it is the information transferred from 1.

The absolute ACR (see Fig. 4 and 6a) shows the PSD of the contribution of the causal relation to the target variable. The main diagonal values correspond to the PSD of the variable which is not caused by other variables. The summation of all contributions to the target variable is equal to the total PSD of the variable (51), a diagonal value of the spectral matrix (15).

The advantage of the ACR is that it correctly detects only direct relations and excludes false indirect ones (e.g., $ACR_{5\leftarrow 1ABS}(f)$ or $ACR_{3\leftarrow 2ABS}(f)$ in Fig. 4 are zero) and, therefore, uncovers the correct connection diagram of causal relations among the time series.

$ACR_{1\leftarrow 1ABS}(f)$ is equal to $S_{11}(f)$ because there are no incoming causal connections to the first variable. Equation $X_1(t)$ involves a second order IIR all-pole filter with a peak representing gain of 14.51 at the normalized frequency of 0.125. As this filter processes the white noise with the variance of 0.6, the peak value of the PSD is equal to the gain squared multiplied by the variance, which is 126.2. This is the only second order IIR filter in the model and this peak propagates throughout the model according to causal connections.

Equations for $X_2(t)$ and $X_3(t)$ do not have the all-pole IIR filter, therefore $ACR_{2\leftarrow 2ABS}(f)$ and $ACR_{3\leftarrow 3ABS}(f)$ on the main diagonal (corresponding to the PSD of the variable which is not caused by other variables) are flat and are equal directly to the power spectral densities of the white noises with variances 0.5 and 0.3, respectively. The $X_1(t)$ propagates to $X_2(t)$ and $X_3(t)$ through simple FIR filters with a delay and one coefficient forming a uniform absolute value of the gain across all frequencies, therefore $ACR_{2\leftarrow 1ABS}(f)$ and $ACR_{3\leftarrow 1ABS}(f)$ are equal to $S_{11}(f)$ multiplied by the FIR gain squared, resulting in the peak values 31.56 and 20.2, respectively.

Equations for $X_4(t)$ and $X_5(t)$ have the same first order IIR all-pole filter forming a low-pass filter with a peak value of 1.547 at the zero frequency, hence $ACR_{4\leftarrow 4ABS}(f)$ and $ACR_{5\leftarrow 5ABS}(f)$ are results from filtering the white noises with this filter. The peak values at the zero frequency are equal to filter gain squared multiplied by variances of the noises, resulting in values of 0.7179 and 1.436, respectively. The situation with $X_4(t)$ and $X_5(t)$ is more complicated because of the bidirectional causal relation forming a feedback. This feedback diminishes the total PSD of the target variable at some frequencies, represented by negative values in the absolute ACR. However, the main positive peak at the normalized frequency of 0.125 is still dominant. As $X_4(t)$ has two causal inputs ($X_1(t)$ and $X_5(t)$), and because of the nonzero cross spectral density $S_{15}(f)$ caused by the sequential causal relation $5 \leftarrow 4 \leftarrow 1$, the coupled sources causal relation $ACR_{4\leftarrow 1,5ABS}(f)$ is also present but this fact does not affect the causal connections diagram at all.

The relative ACR (see Fig. 5 and 6b) is the absolute ACR normalized by the PSD of the target variable $S_{ii}(f)$, thus it depicts the relative amount of the contribution of the causal relation and the summation of all contributions is equal to 1. The main diagonal values correspond to the relative part of the power spectral density of the variable which is not caused by other variables.

The suggested use is to evaluate the absolute ACR in order to get an overview of the power of the contributions at each frequency, and then compare the values with the total PSD of the target variable via the relative ACR. The use of the relative ACR alone cannot be generally recommended because such results can return high values at frequencies where signals have very low power because of the ratio of the two values close to zero.

The dDTF (see Fig. 7) often returns shapes similar to the absolute relations (see Fig. 4), however, this is not always the case (e.g., relations $5 \leftarrow 4$ and $4 \leftarrow 5$). The value of the relation strength is difficult to compare and interpret. The dDTF falsely detects the indirect causal relation $5 \leftarrow 1$ with the same order of magnitude as the correct direct relation $5 \leftarrow 4$.

The GPDC (see Fig. 8) normalizes the values of the relation strength in the range from 0 to 1. The squared modulus of the GPDC (see Fig. 9) partially resembles the relative ACR (see Fig. 5) but appears more flattened. It reaches high values even at frequencies where signals have very low power. On the other hand, the maxima are lowered in comparison with the relative ACR. These facts complicate the interpretation of the GPDC results. In comparison to the dDTF, the advantage of the GPDC (as well as of the ACR) is that it detects only the direct causal connections, excluding the false indirect ones. This is caused by the numerator of the GPDC definition (19) which uses only the MVAR coefficients of the direct connection.

If we take the square modulus of the GPDC (see Fig. 9), the denominator of (20) is equal to the summation of all the numerators for the given j. Then, the summation of all the components for the given source is equal to 1 and the main diagonal values are the residue remaining after all the relations from the source are subsracted. This may lead to a false impression that the square modulus of the GPDC represents the absolute ACR normalized by the summation of all the absolute ACRs coming out of the common source. However, it is obvious that the numerator of (20) is not equal to the absolute ACR.

5. Conclusion

The suggested ACR measure analyzes the causal relations in the frequency domain using an MVAR model. ACR decomposes the diagonal elements of a spectral matrix into separate

causal components that show directions of influence among multivariate time series of an autoregressive character. ACR measures causality in meaningful physical terms as power contributions and can separate direct effects from indirect ones. This easily interpretable definition allows one to evaluate the performance of state-of-the-art methods. The concurrent use of both the absolute and relative ACR is always recommended for obtaining a complete representation of the absolute values of the causal relations and also the proportion of their impact on the target. The ACR is focused on the real impact of each connection, i.e., which part of the power is transferred via the connection. As the time series have random character, the power spectral density approach must be applied. This is in contrast with PDC and GPDC which only analyze the MVAR coefficients of each connection transferred into the frequency domain.

As a result of the PSD utilization, the ACR contains extra coupled causal relations from two sources together. However, these components are present only in the case when both source variables have a nonzero CPSD and both have direct causal influence to the target variable, so the presence of such components does not affect the causal relations connection diagram.

Unlike the original Granger causality concept that evaluates the strength of each causal relation with a single non-negative scalar value, the values of ACR can be positive or negative in order to model both an increase or a decrease of the PSD caused by a causal relation in the analyzed direction. Therefore, the ACR allows one not only to compare the strength of the relations, but also to analyze the effect of the relation. The GPDC and dDTF decompose the strength of causal relations into the frequency domain, however the values lack straightforward interpretability. The dDTF may not be able to recognize direct relations from indirect ones in some cases as shown in this paper. The advantage of the ACR is that the summation of the components creates the total PSD of the target variable. This allows one to compare and clearly interpret the strength of the causal components.

While this paper suggests a new methodology of interpretation of the MVAR model in terms of causal relations in the frequency domain, future research should be focused on the behavior of the signals not strictly satisfying the MVAR conditions. An elaborate study of numerical stability in such cases should also be examined.

Acknowledgment

The research described in this paper was supported in part by the Czech Grant Agency under grant No. GD102/08/H008 “Analysis and modeling of biomedical and speech signals”, by the Czech Technical University in Prague under grant No. SGS10/176/OHK3/2T/13 “Brain activity mapping and analysis”, and by the Czech Technical University in Prague under the research program No. MSM6840770012 “Transdisciplinary Research in Biomedical Engineering”.

References

[1] S. S. Pappas, A. K. Leros, S. K. Katsikas, Joint order and parameter estimation of multivariate autoregressive models using multi-model partitioning theory, Digital Signal Processing 16 (6) (2006) 782 – 795.

[2] J. Runge, J. Heitzig, V. Petoukhov, J. Kurths, Escaping the curse of dimensionality in estimating multivariate transfer entropy, Phys. Rev. Lett. 108 (2012) 258701.

[3] J. Runge, J. Heitzig, N. Marwan, J. Kurths, Quantifying causal coupling strength: A lag-specific measure for multivariate time series related to transfer entropy, Phys. Rev. E 86 (2012) 061121.

[4] T. Schreiber, Measuring information transfer, Physical Review Letters 85 (2) (2000) 461–464.

[5] K. Friston, L. Harrison, W. Penny, Dynamic Causal Modelling, NeuroImage 19 (4) (2003) 1273–1302.

[6] S. Kiebel, M. Garrido, R. Moran, C. Chen, K. Friston, Dynamic causal modeling for EEG and MEG, Human Brain Mapping 30 (6) (2009) 1866–1876.

[7] J. Daunizeau, O. David, K. Stephan, Dynamic causal modelling: A critical review of the biophysical and statistical foundations, NeuroImage 58 (2011) 312–322.

[8] R. B. Kline, Principles and Practice of Structural Equation Modeling, 2nd Edition, The Guilford Press, New York; London, 2005.

[9] M. G. Rosenblum, L. Cimponeriu, A. Bezerianos, A. Patzak, R. Mrowka, Identification of coupling direction: Application to cardiorespiratory interaction, Phys. Rev. E 65 (2002) 041909.

[10] D. A. Smirnov, B. P. Bezruchko, Estimation of interaction strength and direction from short and noisy time series, Phys. Rev. E 68 (2003) 046209.

[11] G. Nolte, A. Ziehe, V. V. Nikulin, A. Schlögl, N. Krämer, T. Brismar, K.-R. Müller, Robustly estimating the flow direction of information in complex physical systems, Physical Review Letters 100 (23) (2008) 234101–1–4.

[12] K. Sekihara, J. Owen, H. Attias, D. Wipf, S. S. Nagarajan, Estimating directions of information flow between cortical activities using phase-slope index, in: S. Supek, A. Sušac (Eds.), 17th International Conference on Biomagnetism Advances in Biomagnetism – Biomag2010, Vol. 28 of IFMBE Proceedings, Springer Berlin Heidelberg, 2010, pp. 199–202.

[13] M. Okatan, M. A. Wilson, E. N. Brown, Analyzing functional connectivity using a network likelihood model of ensemble neural spiking activity., Neural computation 17 (9) (2005) 1927–1961.

[14] M. Kor A new method to estimate short-run and long-run interaction mechanisms in interictal state, Digital Signal Processing 20 (2) (2010) 347 – 358.

[15] C. W. J. Granger, Investigating causal relations by econometric models and cross-spectral methods, Econometrica 37 (3) (1969) 424–38.

[16] R. F. Ziemer, G. S. Collins, Granger causality and U. S. Crop and Livestock prices, Southern Journal of Agricultural Economics (1984) 115–120.

[17] M. Ding, Y. Chen, S. L. Bressler, Granger causality: Basic theory and application to neuroscience, in: Handbook of Time Series Analysis, Wiley-VCH Verlag GmbH & Co. KGaA, 2006, pp. 437–460.

[18] S. Murali, V. V. Kulish, Modeling of evoked potentials of electroencephalograms: An overview, Digital Signal Processing 17 (3) (2007) 665 – 674.

[19] M. Kamiński, K. J. Blinowska, A new method of the description of the information flow in the brain structures, Biological Cybernetics 65 (3) (1991) 203–210.

[20] M. Eichler, On the evaluation of information flow in multivariate systems by the directed transfer function, Biol Cybern 94 (6) (2006) 469–482.

[21] L. A. Baccalá, K. Sameshima, Partial directed coherence: a new concept in neural structure determination, Biological Cybernetics 84 (2001) 463–474.

[22] A. Cadotte, T. Mareci, T. DeMarse, M. Parekh, R. Rajagovindan, W. Ditto, S. Talathi, D.-U. Hwang, P. Carney, Temporal lobe epilepsy: Anatomical and effective connectivity, Neural Systems and Rehabilitation Engineering, IEEE Transactions on 17 (3) (2009) 214–223.

[23] L. A. Baccalá, K. Sameshima, D. Y. Takahashi, Generalized partial directed coherence, Proc. of the 15th International Conference on Digital Signal Processing (2007) 163–166.

[24] A. Korzeniewska, M. Manczak, M. Kaminski, K. J. Blinowska, S. Kasicki, Determination of information flow direction among brain structures

by a modified directed transfer function (dDTF) method, Journal of Neuroscience Methods 125 (1-2) (2003) 195–207.
[25] H. Liang, M. Ding, R. R. Nakamura, S. L. Bressler, Causal influences in primate cerebral cortex during visual pattern discrimination, Neuroreport 11 (2000) 2875–2880.
[26] L. Astolfi, F. De Vico Fallani, F. Cincotti, D. Mattia, M. G. Marciani, S. Salinari, J. Sweeney, G. A. Miller, B. He, F. Babiloni, Estimation of effective and functional cortical connectivity from neuroelectric and hemodynamic recordings, IEEE Neural Systems and Rehabilitation Engineering 17 (3) (2009) 224–233.
[27] H. Benz, H. Zhang, A. Bezerianos, S. Acharya, N. Crone, X. Zheng, N. Thakor, Connectivity analysis as a novel approach to motor decoding for prosthesis control, Neural Systems and Rehabilitation Engineering, IEEE Transactions on 20 (2) (2012) 143 –152.
[28] J. Geweke, Measures of conditional linear dependence and feedback between time series, Journal of the American Statistical Association 79 (388) (1984) 907–915.
[29] Y. Chen, S. L. Bressler, M. Ding, Frequency decomposition of conditional Granger causality and application to multivariate neural field potential data, Journal of Neuroscience Methods 150 (2) (2006) 228–237.
[30] J. R. Sato, D. Y. Takahashi, S. M. Arcuri, K. Sameshima, P. A. Morettin, L. A. Baccalá, Frequency domain connectivity identification: An application of partial directed coherence in fMRI, Human Brain Mapping 30 (2009) 452–461.
[31] G. M. Hoerzer, S. Liebe, A. Schloegl, N. K. Logothetis, G. Rainer, Directed coupling in local field potentials of macaque v4 during visual short-term memory revealed by multivariate autoregressive models, Frontiers in Computational Neuroscience 4 (14).
[32] B. Schelter, J. Timmer, M. Eichler, Assessing the strength of directed influences among neural signals using renormalized partial directed coherence, Journal of Neuroscience Methods 179 (1) (2009) 121–130.
[33] K. J. Blinowska, Methods for localization of time-frequency specific activity and estimation of information transfer in brain, International Journal of Bioelectromagnetism 10 (1) (2008) 2–16.
[34] L. Faes, G. Nollo, Multivariate frequency domain analysis of causal interactions in physiological time series, in: A. N. Laskovski (Ed.), Biomedical Engineering Trends in Electronics, Communications and Software, InTech, 2011, pp. 403–428.
[35] S. Hu, G. Dai, G. A. Worrell, Q. Dai, H. Liang, Causality analysis of neural connectivity: Critical examination of existing methods and advances of new methods., IEEE Transactions on Neural Networks 22 (6) (2011) 829–844.
[36] M.-H. Wu, R. E. Frye, G. Zouridakis, A comparison of multivariate causality based measures of effective connectivity, Comput. Biol. Med. 41 (2011) 1132–1141.
[37] T. Bořil, P. Sovka, System interpretation of causality measures in frequency domain used in EEG analysis, 19th European Signal Processing Conference (EUSIPCO 2011) (2011) 1539–1543.
[38] N. Wiener, Generalized Harmonic Analysis, Massachusetts Institute of Technology. Dept. of Mathematics. Contribution. Ser. II, Massachusetts Institute of Technology, 1930.
[39] S. Hu, H. Liang, Causality analysis of neural connectivity: New tool and limitations of spectral granger causality, Neurocomput. 76 (1) (2012) 44–47.
[40] S. M. Kay, Modern spectral estimation: theory and application, 1st Edition, Prentice Hall, Englewood Cliffs, NJ, 1988.
[41] M. H. Hayes, Statistical Digital Signal Processing and Modeling, John Wiley & Sons, New York, 1996.
[42] R. A. Wiggins, E. A. Robinson, Recursive Solution to the Multichannel Filtering Problem, Journal of Geophysical Research 70 (1965) 1885–1891.
[43] M. Morf, A. Vieira, D. T. L. Lee, T. Kailath, Recursive Multichannel Maximum Entropy Spectral Estimation, IEEE Transactions on Geoscience Electronics 16 (1978) 85–94.
[44] S. L. Marple, Digital Spectral Analysis with Applications, Englewood Cliffs: Prentice Hall, 1987.
[45] J. P. Burg, Maximum entropy spectral analysis, in: Proceedings of the 37th Meeting Society of Exploration Geophysicists, Oklahoma City, OK, 1967.
[46] J. P. Burg, Maximum entropy spectral analysis, Ph.D. thesis, Stanford University (1975).

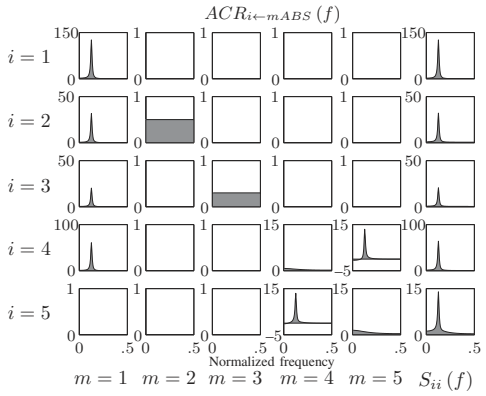

Figure 4: Absolute ACR; causal relation from source $X_m\,(t)$ to target $X_i\,(t)$. Labels of normalized frequency axis are common for the column, labels of $ACR_{i\leftarrow mABS}\,(f)$ axis are individual at each relation because of the different dynamics. The summation of all components creating the target variable including the relation from the couple of sources in Fig. 6a is equal to the PSD of the variable $S_{ii}\,(f)$ depicted in the 6th column.

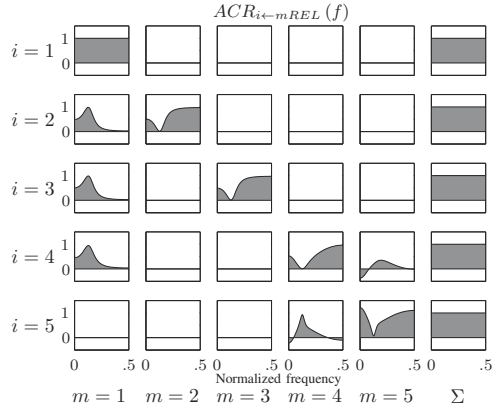

Figure 5: Relative ACR; causal relation from $X_m\,(t)$ to $X_i\,(t)$. Labels of normalized frequency axis are common for the column, labels of $ACR_{i\leftarrow mREL}\,(f)$ axis are common for the row. The summation of all components of the target variable including the relation from the couple of sources in Fig. 6b is always one (see the 6th column).

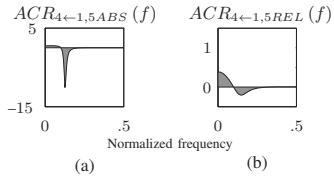

Figure 6: ACR – causal relation from a couple of sources $X_1\,(t)$ and $X_5\,(t)$ to target $X_4\,(t)$; (a) absolute ACR and (b) relative ACR. This is the only nonzero coupled causal relation because of the two direct causal connections to the same target from the sources with nonzero CPSD.

[47] A. Schlögl, A comparison of multivariate autoregressive estimators, Signal Processing 86 (9) (2006) 2426–2429.
[48] H. Akaike, A new look at the statistical model identification, Automatic Control, IEEE Transactions on 19 (6) (1974) 716–723.
[49] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, J. Doyne Farmer, Testing for nonlinearity in time series: the method of surrogate data, Physica D Nonlinear Phenomena 58 (1992) 77–94.
[50] M. Paluš, M. Vejmelka, Directionality of coupling from bivariate time series: How to avoid false causalities and missed connections, Phys. Rev. E 75 (2007) 056211.

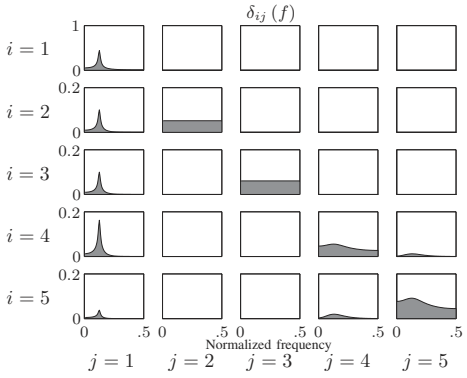

Figure 7: dDTF; causal relation from $X_j(t)$ to $X_i(t)$. Labels of normalized frequency axis are common for the column, labels of $\delta_{ij}(f)$ axis are common for the row.

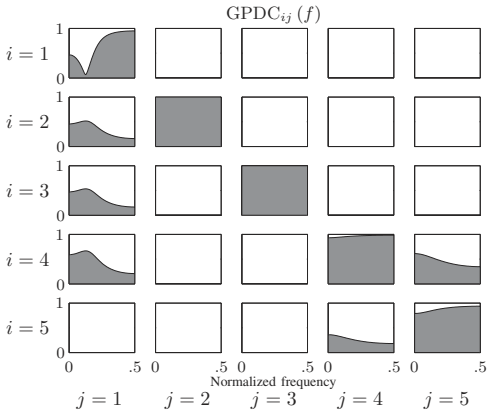

Figure 8: GPDC; causal relation from $X_j(t)$ to $X_i(t)$. Labels of normalized frequency axis are common for the column, labels of $\mathrm{GPDC}_{ij}(f)$ axis are common for the row. If we take the square modulus of GPDC, the main diagonal values are the residue remaining when all the relations from the source are subtracted (see Fig. 9).

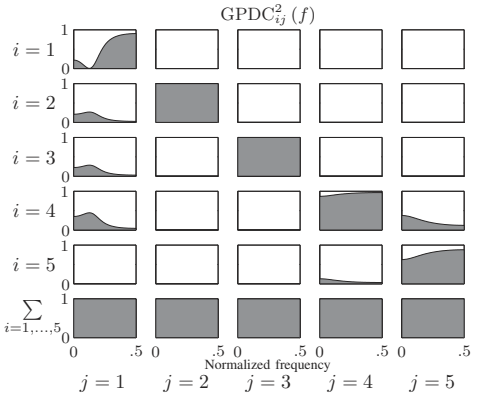

Figure 9: $\mathrm{GPDC}^2_{ij}(f)$; causal relation from $X_j(t)$ to $X_i(t)$. Labels of normalized frequency axis are common for the column, labels of $\mathrm{GPDC}^2_{ij}(f)$ axis are common for the row. The 6th row is the summation of all values in the column, always equal to one.