21

Multivariate Frequency Domain Analysis of Causal Interactions in Physiological Time Series

Luca Faes and Giandomenico Nollo
Department of Physics and BIOtech Center
University of Trento
Italy

1. Introduction

A common way of obtaining information about a physiological system is to measure one or more signals from the system, consider their temporal evolution in the form of numerical time series, and obtain quantitative indexes through the application of time series analysis techniques. While historical approaches to time series analysis were addressed to the study of single signals, recent advances have made it possible to study collectively the behavior of several signals measured simultaneously from the considered system. In fact, multivariate (MV) time series analysis is nowadays extensively used to characterize interdependencies among multiple signals collected from dynamical physiological systems. Applications of this approach are ubiquitous, for instance, in neurophysiology and cardiovascular physiology (see, e.g., (Pereda et al., 2005) and (Porta et al., 2009) and references therein). In neurophysiology, the time series to be analyzed are obtained, for example, sampling electroencephalographic (EEG) or magnetoencephalographic (MEG) signals which measure the temporal dynamics of the electro-magnetic fields of the brain as reflected at different locations of the scalp. In cardiovascular physiology, the time series are commonly constructed measuring at each cardiac beat cardiovascular and cardiorespiratory variables such as the heart period, the systolic/diastolic arterial pressure, and the respiratory flow. It is well recognized that the application of MV analysis to these physiological time series may provide unique information about the coupling mechanisms underlying brain dynamics and cardiovascular control, and may also lead to the definition of quantitative indexes useful in medical settings to assess the degree of mechanism impairment in pathological conditions.

MV time series analysis is not only important to detect *coupling*, i.e., the presence or absence of interactions, between the considered time series, but also to identify driver-response relationships between them. This problem is a special case of the general question of assessing *causality*, or cause-effect relations, between (sub)systems, processes or phenomena. The assessment of coupling and causality in MV processes is often performed by linear time series analysis approaches, i.e. approaches in which a linear model is supposed to underlie the generation of temporal dynamics and interactions of the considered signals (Kay, 1988; Gourevitch et al., 2006). While non-linear methods are continuously under development

(Pereda et al., 2005; Faes et al., 2008), the traditional linear approach remains of great interest for the study of physiological signals, mainly because it has the important advantage to be strictly connected to the frequency-domain representation of multichannel data. Indeed, physiological signals such as the brain and cardiovascular ones are rich of oscillatory content and thus lend themselves to spectral representation. Typical examples of physiological rhythms are the EEG dynamics, typically observed within the well-bounded frequency bands from delta to gamma (Nunez, 1995), and the cardiovascular oscillations, characterized by spectral peaks within the so-called low frequency (LF, ~0.1 Hz) and high frequency (HF, synchronous with respiratory activity) bands (Akselrod et al., 1981). As a consequence, the linear frequency-domain evaluation of coupling and causality constitutes an eligible approach to characterize the interdependence among specific oscillations manifested within the same frequency band in two or more physiological signals.

While an unique and universally accepted definition of causality does not exist, in time series analysis inference about cause-effect relationships is commonly based on the notion introduced by Nobel Prize winning Clive Granger (Granger, 1969). Granger causality was mathematically formalized within a linear time-domain framework widely applied in economy and finance but rapidly spread to other fields including the analysis of physiological time series. This notion of causality is defined in terms of predictability and exploits the direction of the flow of time to achieve a causal ordering of dependent processes. The definition may be contextualized in a different way for bivariate (based on two signals only) and MV (based on more than two signals) analysis; in the MV formulation, a distinction between direct causality from one series to another and indirect causality (i.e., causality between two series mediated by other series) is achieved (Faes et al., 2010b). Moreover, while the most intuitive definition of causality accounts for lagged effects only (i.e., effects of the past of a time series on the present of another), the concept of instantaneous causality, describing influences which occur within the same lag, is crucial for the evaluation of causal relationships among processes (Lutkepohl, 1993). Finally, the different facets of the concept of causality may be related to the concept of coupling between two processes, according to which the presence or absence of an interaction is detected and measured, but the directionality of such interaction is not elicited.

The notions of causality and coupling are commonly formalized in the context of a MV autoregressive (MVAR) representation of the available time series, which allows to derive time- and frequency-domain pictures of these concepts respectively through the model coefficients and through their spectral representation. Accordingly, several frequency domain measures of causality and coupling have been introduced and applied in recent years. Coupling is traditionally investigated by means of the coherence (Coh) and the partial coherence (PCoh), classically known, e.g., from Kay (1988) or (Bendat & Piersol, 1986). Measures able to quantify causality in the frequency domain have been proposed more recently: the most used are the directed transfer function (DTF) (Kaminski & Blinowska, 1991), the directed coherence (DC) (Baccala et al., 1998), and the partial directed coherence (PDC) (Baccala & Sameshima, 2001). All these measures have been used extensively for the analysis of physiological time series, and applications showing their usefulness for the interpretation of interaction mechanisms among, e.g., EEG rhythms or cardiovascular oscillations, are plentiful in the literature (see, for instance, (Porta et al., 2002; Schlogl & Supp, 2006; Astolfi et al., 2007; Faes & Nollo, 2010a)). Despite this, several issues have to be taken into account for their correct utilization. While the relationships existing among these indices are generally understood, and most of the properties linking these measures to the different concepts of causality and

coupling are known, an organic joint description and contextualization in relation to the underlying time domain concepts is lacking. Also for this reason, the interpretation of frequency-domain coupling and causality measures is not always straightforward, and this may lead to an erroneous description of connectivity and related mechanisms. Examples of ambiguities emerged in the interpretation of these measures are the debates about the ability of PCoh to measure some forms of causality (Albo et al., 2004; Baccala & Sameshima, 2006), and about the specific kind of causality which is reflected by the DTF and DC measures (Kaminski et al., 2001; Baccala & Sameshima, 2001; Eichler, 2006). An aspect which is perhaps more problematic regards the structure of the model used to represent the data prior to computation of the frequency domain measures, which commonly accounts for lagged but not for instantaneous effects among the series. Despite this, the significance of instantaneous correlations among the series is almost never tested in practical applications, and the possible effects on coupling and causality measures of forsaking such correlations have not been investigated thoroughly. Very recent studies have suggested that neglecting instantaneous interactions in the model representation may lead to heavily modified connectivity patterns (Hyvarinen et al., 2008; Faes & Nollo, 2010b).

The mission of this chapter is to enhance the theoretical interpretability of the available frequency domain measures of coupling and causality derived from the MVAR representation of multiple time series. To this end, a common framework for the definition of Coh, PCoh, DC/DTF, and PDC is provided on the basis of the frequency domain MVAR representation, and is exploited to relate the various measures to each other as well as to the specific coupling or causality definitions which they underlie. The chapter is structured as follows: Sect. 2 presents a comprehensive definition of the various forms of causality and coupling that can be observed in MV processes; Sect. 3 particularizes these definitions for standard MVAR processes, derives the corresponding frequency domain measures of coupling and causality, and discusses their interpretation; Sect. 4 proposes an extended MVAR representation to be used in the presence of significant instantaneous correlations in the observed process, whereby novel frequency domain causality measures are defined and compared to the existing ones; Sect. 5 briefly discuss the practical application of the measures on physiological time series; and Sect. 6 concludes the chapter.

2. Causality and coupling in multivariate processes

Let us consider M stationary stochastic processes y_m, m=1,...,M. Without loss of generality we assume that the processes are real-valued, defined at discrete time ($y_m=\{y_m(n)\}$; e.g., are sampled versions of the continuous time processes $y_m(t)$, taken at the times $t_n=nT$, with T the sampling period) and have zero mean ($E[y_m(n)]=0$, where $E[\cdot]$ is the statistical expectation operator). A MV closed loop process is defined as:

$$y_m(n)=f_m(Y_m,\acute{Y}_l \,|\, l\neq m)+w_m(n)\ , \quad l,m=1,...,M, \tag{1}$$

where f_m is the function linking the set of the p past values of the m-th process, collected in $Y_m=\{y_m(n\text{-}1),...,y_m(n\text{-}p)\}$, as well as the sets of the present and the p past values of all other processes, collected in $\acute{Y}_l=\{y_l(n),Y_l\}=\{y_l(n),y_l(n\text{-}1),...,y_l(n\text{-}p)\}$, $l\neq m$, to the present value $y_m(n)$, and w_m is a white noise process describing the error in the representation. Given two processes y_i and y_j, i,j=1,...,M, different definitions of causality and coupling between the processes may be defined as discussed in the following, and summarized in Table 1.

		Strictly causal MVAR representation	Extended MVAR representation
DIRECT			
a) direct causality	$y_j \rightarrow y_i$	PDC, $\pi_{ij}(f)$	PDC, $\tilde{\pi}_{ij}(f)$
b) extended direct causality	$y_j \dot{\rightarrow} y_i$	-	ePDC, $\chi_{ij}(f)$
c) direct coupling	$y_i \leftrightarrow y_j$	PCoh, $\Pi_{ij}(f)$	
DIRECT+INDIRECT			
a) causality	$y_j \Rightarrow y_i$	DC, $\gamma_{ij}(f)$	DC, $\tilde{\gamma}_{ij}(f)$
b) extended causality	$y_j \dot{\Rightarrow} y_i$	-	eDC, $\xi_{ij}(f)$
c) coupling	$y_i \Leftrightarrow y_j$	Coh, $\Gamma_{ij}(f)$	

Table 1. Frequency domain measures of causality and coupling between two processes y_i and y_j of a multivariate closed loop process. Note that causality and direct causality measure lagged effects only, while extended causality and extended direct causality measure combined instantaneous and lagged effects.

Denoting as $Z_j=\{Y_l \mid l=1,...,M, l\neq j\}$ the set of the past values of all processes except y_j, *direct causality* from y_j to y_i, $y_j \rightarrow y_i$, exists if the prediction of $y_i(n)$ based on Z_j and Y_j is better than the prediction of $y_i(n)$ solely based on Z_j. *Causality* from y_j to y_i, $y_j \Rightarrow y_i$, exists if a cascade of direct causality relations $y_j \rightarrow y_m \cdots \rightarrow y_i$ occurs for at least one value m in the set (1,…,M); if $m=i$ or $m=j$ causality reduces to direct causality. This last case is obvious for a bivariate closed loop process (M=2), where only one definition exists and agrees with the notion of Granger causality (Granger, 1969) involving only the relations between two processes. For multivariate processes ($M\geq 3$) the definition of direct causality agrees with the notion of *prima facie* cause introduced in (Granger, 1980); the definition of causality is a generalization including also causal indirect effects between two processes, i.e., effects mediated by one or more other processes in the MV closed loop.

While the definitions provided above are based on the exclusive consideration of lagged effects from one series to another, the interactions modeled in (1) consider also the possible instantaneous effects, i.e. effects which occur within the same lag. If we consider the directed interaction from y_j to y_i, lagged causality (with lag $k\geq 1$) occurs if $y_j(n-k)$ is useful to predict $y_i(n)$, while instantaneous causality (with lag k=0) occurs if $y_j(n)$ is useful to predict $y_i(n)$. These two concepts may be combined together to provide extended causality definitions as follows. Denoting as $Z_{ij}=\{Y_i, \acute{Y}_l \mid l=1,...,M, l\neq j, l\neq i\}$ the set of the past values of y_i and the present and past values of all other processes except y_j, *extended direct causality* from y_j to y_i, $y_j \dot{\rightarrow} y_i$, exists if the prediction of $y_i(n)$ based on Z_{ij} and $\acute{Y}_j$ is better than the prediction of $y_i(n)$ solely based on Z_{ij}. *Extended causality* from y_j to y_i, $y_j \dot{\Rightarrow} y_i$, exists if a cascade of extended direct causality relations $y_j \dot{\rightarrow} y_m \cdots \dot{\rightarrow} y_i$ occurs for at least one value m in the set (1,…,M); again, if $m=i$ or $m=j$ extended causality reduces to extended direct causality.

Definitions of coupling between two processes are derived from the causality definitions as follows. *Direct coupling* between y_i and y_j, $y_i \leftrightarrow y_j$, exists if $y_i \dot{\rightarrow} y_m$ and $y_j \dot{\rightarrow} y_m$; while the most obvious case is when $m=i$ or $m=j$, two processes are considered as directly coupled also when they both directly cause a third common process ($m\neq i$, $m\neq j$). *Coupling* between y_i and

y_j, $y_i \Leftrightarrow y_j$, exists if $y_m \dot{\Rightarrow} y_i$ and $y_m \dot{\Rightarrow} y_j$; again, coupling may arise when the of the two processes causes the other ($m=i$ or $m=j$), or when both processes are caused by other common processes ($m\neq i$, $m\neq j$). Thus, the coupling definitions generalize the concept of causality accounting for both forward and backward interactions between two processes.
An illustrative example of the described causality and coupling relations is reported in Fig. 1. In the diagrams, the set of interactions is represented with a network where nodes correspond to processes and connecting arrows depict direct causality relations.

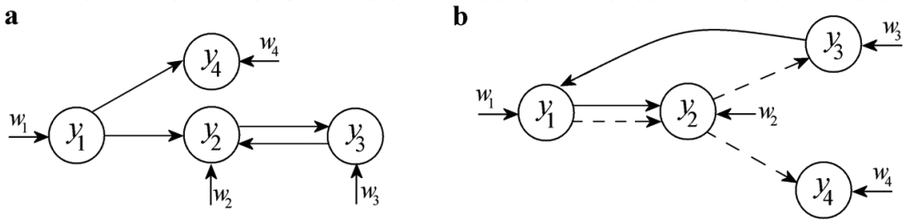


Fig. 1. Examples of networks of interacting processes exhibiting only lagged interactions (**a**) and combined instantaneous and lagged interactions (**b**). Lagged and instantaneous effects are depicted with solid and dashed arrows, respectively.

Fig. 1a shows a network of M=4 interacting processes in which only lagged effects from one process to another are present. In this situation, extended causality reduces to causality due to the absence of instantaneous effects. The direct causality relations imposed in the net are $y_1 \rightarrow y_2$, $y_2 \rightarrow y_3$, $y_3 \rightarrow y_2$, and $y_1 \rightarrow y_4$. Since direct causality is a condition sufficient for causality, we observe also $y_1 \Rightarrow y_2$, $y_2 \Rightarrow y_3$, $y_3 \Rightarrow y_2$, and $y_1 \Rightarrow y_4$; moreover, the cascade $y_1 \rightarrow y_2 \rightarrow y_3$ determines an indirect effect such that causality $y_1 \Rightarrow y_3$ exists. Direct coupling follows from direct causality, so that $y_1 \leftrightarrow y_2$, $y_2 \leftrightarrow y_3$, and $y_1 \leftrightarrow y_4$, but is also caused by the common driving exerted by y_1 and y_3 on y_2, so that $y_1 \leftrightarrow y_3$. Finally, coupling is present between each pair of processes: $y_1 \Leftrightarrow y_2$, $y_2 \Leftrightarrow y_3$, $y_1 \Leftrightarrow y_4$, and $y_1 \Leftrightarrow y_3$ result from the causality relations, while $y_2 \Leftrightarrow y_4$ and $y_3 \Leftrightarrow y_4$ result from the common driving exerted by y_1 respectively on y_2 and y_4, and on y_3 and y_4. In Fig. 1b, instantaneous effects are considered together with lagged ones. In this case, direct causality occurs only when lagged effects are present, i.e., over the directions $y_1 \rightarrow y_2$, $y_3 \rightarrow y_1$. Extended direct causality follows from lagged and/or instantaneous direct causality, so that we have $y_1 \dot{\rightarrow} y_2$, $y_2 \dot{\rightarrow} y_3$, $y_2 \dot{\rightarrow} y_4$, and $y_3 \dot{\rightarrow} y_1$. As no indirect lagged causality is present, causality follows exclusively from direct causality, i.e. $y_1 \Rightarrow y_2$, $y_3 \Rightarrow y_1$. On the contrary, extended causality is observed very often because of the existence of several cascades of instantaneous and/or lagged effects: we observe indeed $y_1 \dot{\Rightarrow} y_2$, $y_1 \dot{\Rightarrow} y_3$, $y_1 \dot{\Rightarrow} y_4$, $y_2 \dot{\Rightarrow} y_1$, $y_2 \dot{\Rightarrow} y_3$, $y_2 \dot{\Rightarrow} y_4$, $y_3 \dot{\Rightarrow} y_1$, $y_3 \dot{\Rightarrow} y_2$, $y_3 \dot{\Rightarrow} y_4$. Direct coupling follows from extended direct causality: $y_1 \leftrightarrow y_2$, $y_2 \leftrightarrow y_3$, $y_2 \leftrightarrow y_4$, $y_3 \leftrightarrow y_1$ (no common driving of two processes on a third one is observed). Finally, coupling is detected between all pairs of processes as the network is fully connected (i.e., there are no isolated groups of processes).
While causality definitions cannot be explored by means of conventional statistical operators, the concepts of coupling and direct coupling may be quantified through standard analysis of the correlation structure of the observed processes. Specifically, defining as $\mathbf{Y}(n)=[y_1(n) \cdots y_M(n)]^T$ the observed $M\times 1$ vector process, as $\mathbf{R}(k)=E[\mathbf{Y}(n)\mathbf{Y}^T(n-k)]$ its $M\times M$ correlation matrix evaluated at lag k, and as $\mathbf{P}(k)=\mathbf{R}(k)^{-1}$ the inverse correlation matrix, coupling $y_i \Leftrightarrow y_j$ and direct coupling $y_i \leftrightarrow y_j$ are quantified at the time lag k respectively by the correlation coefficient and the partial correlation coefficient (Whittaker, 1990):

$$\rho_{ij}(k)=\frac{r_{ij}(k)}{\sqrt{r_{ii}(k)r_{jj}(k)}},\ \eta_{ij}(k)=-\frac{p_{ij}(k)}{\sqrt{p_{ii}(k)p_{jj}(k)}}, \tag{2}$$

where $r_{ij}(k)$ and $p_{ij}(k)$ are the *i-j* elements of $\mathbf{R}(k)$ and $\mathbf{P}(k)$. The correlation and partial correlation coefficients are normalized measures of the linear interdependence existing between $y_i(n)$ and $y_j(n-k)$, and of the linear interdependence between $y_i(n)$ and $y_j(n-k)$ after removing the effects of all remaining processes. As such, ρ_{ij} and η_{ij} quantify the correlation and the "direct correlation" (i.e., the correlation that cannot be accounted for by the influence of any other process) between y_i and y_j. To identify the frequency-domain analogous of these two coefficients, we consider the spectral representation of the vector process $\mathbf{Y}(n)$, which is provided by the $M\times M$ spectral density matrix $\mathbf{S}(f)$, defined as the Fourier transform (FT) of the correlation matrix $\mathbf{R}(k)$. The spectral matrix contains the spectrum of $y_i(n)$, $S_{ii}(f)$, and the cross-spectrum between $y_i(n)$ and $y_j(n)$, $S_{ij}(f)$, as diagonal and off-diagonal terms, respectively ($i,j=1,\ldots,M$). In analogy with the time domain definitions, the spectral matrix and its inverse, $\mathbf{P}(f)=\mathbf{S}(f)^{-1}$, are exploited to provide frequency-domain measures of coupling and direct coupling, respectively through the *coherence* (Coh) and the *partial coherence* (PCoh) functions (Bendat & Piersol, 1986):

$$\Gamma_{ij}(f)=\frac{S_{ij}(f)}{\sqrt{S_{ii}(f)S_{jj}(f)}},\ \Pi_{ij}(f)=\frac{P_{ij}(f)}{\sqrt{P_{ii}(f)P_{jj}(f)}}, \tag{3}$$

As the functions in (3) are complex-valued, their squared modulus is commonly used to measure the strength of coupling and direct coupling in the frequency domain. Specifically, the magnitude-squared Coh $|\Gamma_{ij}(f)|^2$ measures the strength of the linear, non-directed interactions between the processes y_i and y_j as a function of frequency, being 0 in case of uncoupling and 1 in case of full coupling. The squared PCoh $|\Pi_{ij}(f)|^2$ measures the strength of the direct, non-directed interaction between y_i and y_j, i.e. the strength of the interaction remaining after subtracting the effect of the remaining processes. We stress that, due to the symmetrical nature of these measures, they cannot provide information about causality; such an information may be extracted, as explained in the following, from the coefficients of a parametric representation of the time series.

3. Causality and coupling in MVAR processes

3.1 Time domain definitions

The joint multivariate process $\mathbf{Y}(n)$ can be represented as the output of a MV linear shift-invariant filter (Kay, 1988):

$$\mathbf{Y}(n)=\sum_{k=-\infty}^{\infty}\mathbf{H}(k)\mathbf{U}(n-k), \tag{4}$$

where $\mathbf{U}(n)=[u_1(n)\cdots u_M(n)]^T$ is a vector of M zero-mean input processes and $\mathbf{H}(k)$ is the $M\times M$ filter impulse response matrix. A particular case of the general model in (4), extensively used in time series analysis, is the MV autoregressive (MVAR) model (Kay, 1988):

$$\mathbf{Y}(n)=\sum_{k=1}^{p}\mathbf{A}(k)\mathbf{Y}(n-k)+\mathbf{U}(n), \tag{5}$$

where p is the model order, defining the maximum lag used to quantify interactions. The input process $\mathbf{U}(n)$, also called innovation process, is assumed to be composed of white and uncorrelated noises; this means that the correlation matrix of $\mathbf{U}(n)$, $\mathbf{R}_{\mathbf{U}}(k)=\mathrm{E}[\mathbf{U}(n)\mathbf{U}^{\mathrm{T}}(n-k)]$, is zero for each lag $k>0$, while it is equal to the covariance matrix $\mathbf{\Sigma}=\mathrm{cov}(\mathbf{U}(n))$ for $k=0$. One major benefit of the representation in (5) is that it allows to interpret properties of the joint description of the processes $y_m(n)$ —like coupling or causality— in terms of the estimated coefficients $\mathbf{A}(k)$. In fact, the i-j element of $\mathbf{A}(k)$, $a_{ij}(k)$, quantifies the causal linear interaction effect occurring at lag k from y_j to y_i. As a consequence, the definitions of causality and coupling provided above for a general closed-loop MV process can be specified for a MVAR process in terms of the off-diagonal elements of $\mathbf{A}(k)$ as follows: $y_j \rightarrow y_i$ if $a_{ij}(k)\neq 0$ for at least one $k=1,\ldots,p$; $y_j \Rightarrow y_i$ if $a_{m_s m_{s-1}}(k_s)\neq 0$ for at least a set of $L\geq 2$ values for m_s (with $m_0=j$, $m_{L-1}=i$) and a set of lags $k_0,\ldots,k_{L-1}$ with values in $(1,\ldots,p)$; $y_i \leftrightarrow y_j$ if $a_{mi}(k_1)\neq 0$ for at least one k_1 or $a_{mj}(k_2)\neq 0$ for at least one k_2; $y_i \Leftrightarrow y_j$ if $a_{m_s m_{s-1}}(k_s)\neq 0$ for at least a set of $L\geq 2$ values for m_s (either with $m_0=m$, $m_{L-1}=i$ or with $m_0=m$, $m_{L-1}=j$) and a set of lags $k_0,\ldots,k_{L-1}$. Thus, causality and coupling relations are found when the pathway relevant to the interaction is active, i.e., is described by nonzero coefficients in $\mathbf{A}$. Note that the extended definitions of causality and direct causality cannot be tested from the coefficients of the MVAR model (5), as the model does not describe instantaneous interactions. We refer to Sect. 3.3 to see how the MVAR coefficients may be related to causality and coupling effects in an illustrative example.

3.2 Frequency domain definitions

The spectral representation of a MVAR process is derived considering the FT of the representations in (4) and (5), which yield respectively the equations $\mathbf{Y}(f)=\mathbf{H}(f)\mathbf{U}(f)$ and $\mathbf{Y}(f)=\mathbf{A}(f)\mathbf{Y}(f)+\mathbf{U}(f)$, where $\mathbf{Y}(f)$ and $\mathbf{U}(f)$ are the FTs of $\mathbf{Y}(n)$ and $\mathbf{U}(n)$, and the $M\times M$ transfer matrix and coefficient matrix are defined in the frequency domain as:

$$\mathbf{H}(f)=\sum_{k=-\infty}^{\infty}\mathbf{H}(k)e^{-j2\pi f kT},\ \mathbf{A}(f)=\sum_{k=1}^{p}\mathbf{A}(k)e^{-j2\pi f kT}. \tag{6}$$

Comparing the two spectral representations, it is easy to show that the coefficient and transfer matrices are linked by: $\mathbf{H}(f)=[\mathbf{I}-\mathbf{A}(f)]^{-1}=\bar{\mathbf{A}}(f)^{-1}$. This relation is useful to draw the connection between the cross-spectral density matrix $\mathbf{S}(f)$ and its inverse $\mathbf{P}(f)$, as well as to derive frequency domain estimates of coupling and causality in terms of the MVAR representation. Indeed, the following factorizations hold for a MVAR process (Kay, 1988):

$$\mathbf{S}(f)=\mathbf{H}(f)\mathbf{\Sigma}\mathbf{H}^{\mathrm{H}}(f),\ \mathbf{P}(f)=\bar{\mathbf{A}}^{\mathrm{H}}(f)\mathbf{\Sigma}^{-1}\bar{\mathbf{A}}(f), \tag{7}$$

where the superscript $^{\mathrm{H}}$ stands for the Hermitian transpose. The $(i$-$j)$th elements of $\mathbf{S}(f)$ and $\mathbf{P}(f)$ can be represented in the compact form:

$$S_{ij}(f)=\mathbf{h}_i(f)\mathbf{\Sigma}\mathbf{h}_j^{\mathrm{H}}(f),\ P_{ij}(f)=\bar{\mathbf{a}}_i^{\mathrm{H}}(f)\mathbf{\Sigma}^{-1}\bar{\mathbf{a}}_j(f), \tag{8}$$

where $\mathbf{h}_i(f)$ is the i-th row of the transfer matrix ($\mathbf{H}(f)=[\mathbf{h}_1(f)\cdots\mathbf{h}_M(f)]^{\mathrm{T}}$) and $\bar{\mathbf{a}}_i(f)$ is the i-th column of the coefficient matrix ($\bar{\mathbf{A}}(f)=[\bar{\mathbf{a}}_1(f)\cdots\bar{\mathbf{a}}_M(f)]$). Under the assumption that the input white noises are uncorrelated even at lag zero, their covariance $\mathrm{cov}(\mathbf{U}(n))$ reduces to the diagonal matrix $\mathbf{\Sigma}=\mathrm{diag}(\sigma^2_i)$, and its inverse to the matrix $\mathbf{\Sigma}^{-1}=\mathrm{diag}(1/\sigma^2_i)$ which is diagonal as well (σ^2_i is the variance of u_i). In this specific case, (8) factorize into:

$$S_{ij}(f)=\sum_{m=1}^{M}\sigma_m^2 H_{im}(f)H_{jm}^*(f),\ P_{ij}(f)=\sum_{m=1}^{M}\frac{1}{\sigma_m^2}\overline{\mathbf{A}}_{mi}^*(f)\overline{\mathbf{A}}_{mj}(f) \tag{9}$$

The usefulness of the factorizations in (9) is in the fact that they allow to decompose the frequency domain measures of coupling and direct coupling previously defined into terms eliciting the directional information from one process to another. Substituting (8) and (9) into (3), the Coh between y_i and y_j can be factored as:

$$\Gamma_{ij}(f)=\frac{\mathbf{h}_i(f)\boldsymbol{\Sigma}\mathbf{h}_j^{\mathrm{H}}(f)}{\sqrt{\mathbf{h}_i(f)\boldsymbol{\Sigma}\mathbf{h}_i^{\mathrm{H}}(f)}\sqrt{\mathbf{h}_j(f)\boldsymbol{\Sigma}\mathbf{h}_j^{\mathrm{H}}(f)}}=\sum_{m=1}^{M}\frac{\sigma_m H_{im}(f)}{\sqrt{S_{ii}(f)}}\frac{\sigma_m H_{jm}^*(f)}{\sqrt{S_{jj}(f)}}=\sum_{m=1}^{M}\gamma_{im}(f)\gamma_{jm}^*(f), \tag{10}$$

where the last term contains the so-called *directed coherence* (DC). Thus, the DC from y_j to y_i is defined as (Baccala et al., 1998):

$$\gamma_{ij}(f)=\frac{\sigma_j H_{ij}(f)}{\sqrt{\sum_{m=1}^{M}\sigma_m^2\left|H_{im}(f)\right|^2}}. \tag{11}$$

Note that the directed transfer function (DTF) defined in (Kaminski & Blinowska, 1991) is a particularization of the DC in which all input variances are all equal ($\sigma^2_1=\sigma^2_2=\cdots=\sigma^2_M$) so that they cancel each other in (11). The factorization in (10) justifies the term DC, as $\gamma_{ij}(f)$ can be interpreted as a measure of the influence of y_j onto y_i, as opposed to $\gamma_{ji}(f)$ which measures the interaction occurring over the opposite direction from y_i to y_j. Further interpretation of the DC in terms of coupling strength is achieved considering its normalization properties:

$$0\le\left|\gamma_{ij}(f)\right|^2\le 1\ ,\ \sum_{m=1}^{M}\left|\gamma_{im}(f)\right|^2=1. \tag{12}$$

The inequality in (12) indicates that the squared DC $|\gamma_{ij}(f)|^2$ measures a normalized coupling strength, being 0 in the absence of directed coupling from y_j to y_i at the frequency f, and 1 in the presence of full coupling. The equality indicates that $|\gamma_{ij}(f)|^2$ measures the coupling strength from y_j to y_i as the normalized proportion of $S_{ii}(f)$ which is due to y_j, i.e. is transferred from u_j via the transfer function $H_{ij}(f)$. Indeed, combining (9) and (12) it is easy to show that the spectrum of the process y_i may be decomposed as:

$$S_{ii}(f)=\sum_{m=1}^{M}S_{i|m}(f)\ ,\ S_{i|m}(f)=\left|\gamma_{im}(f)\right|^2 S_{ii}(f). \tag{13}$$

where $S_{i|m}(f)$ is the part of $S_{ii}(f)$ due to y_m; $S_{i|i}(f)$ measures the part of $S_{ii}(f)$ due to none of the other processes, which is quantified in normalized units by the squared DC $|\gamma_{ii}(f)|^2$. Note that the useful decomposition in (13) does not hold for the DTF, unless all input variances are equal to each other so that the DC reduces to the DTF. For this reason, in the following we will consider the DC only, as it provides a similar, but more general and interpretable in terms of power content, measure of frequency domain causality.

In a similar way to that followed to decompose the Coh, the PCoh defined in (3) can be factored, using (8) and (9), as:

$$\Pi_{ij}(f)=\frac{\bar{\mathbf{a}}_i^{\mathrm{H}}(f)\boldsymbol{\Sigma}^{-1}\bar{\mathbf{a}}_j(f)}{\sqrt{\bar{\mathbf{a}}_i^{\mathrm{H}}(f)\boldsymbol{\Sigma}^{-1}\bar{\mathbf{a}}_i(f)}\sqrt{\bar{\mathbf{a}}_j^{\mathrm{H}}(f)\boldsymbol{\Sigma}^{-1}\bar{\mathbf{a}}_j(f)}}=\sum_{m=1}^{M}\frac{\frac{1}{\sigma_m}\overline{A}_{mj}(f)}{\sqrt{P_{jj}(f)}}\frac{\frac{1}{\sigma_m}\overline{A}_{mi}^{*}(f)}{\sqrt{P_{ii}(f)}}=\sum_{m=1}^{M}\pi_{mj}(f)\pi_{mi}^{*}(f), \quad (14)$$

where the last term contains the *partial directed coherence* (PDC) functions. The PDC from y_j to y_i is thus defined as (Baccala et al., 2007):

$$\pi_{ij}(f)=\frac{\frac{1}{\sigma_i}\overline{A}_{ij}(f)}{\sqrt{\sum_{m=1}^{M}\frac{1}{\sigma_m^2}\left|\overline{A}_{mj}(f)\right|^2}}. \quad (15)$$

As suggested by the factorization in (14), the PDC extracts the directional information from the PCoh, and is thus a measure of the direct directed interaction occurring from y_j to y_i at the frequency f. The normalization properties for the squared modulus of the PDC are:

$$0\le\left|\pi_{ij}(f)\right|^2\le 1\ ,\ \sum_{m=1}^{M}\left|\pi_{mj}(f)\right|^2=1\ , \quad (16)$$

suggesting that $|\pi_{ij}(f)|^2$ quantifies the interaction from y_j to y_i as the normalized proportion of $P_{jj}(f)$ which is which is sent to y_i, via the coefficients $\bar{A}_{ji}(f)$. Indeed, we have that:

$$P_{jj}(f)=\sum_{m=1}^{M}P_{j\to m}(f)\ ,\ P_{j\to m}(f)=\left|\pi_{mj}(f)\right|^2P_{jj}(f), \quad (17)$$

where $P_{j\to m}(f)$ is the part of $P_{jj}(f)$ sent to y_m; in particular, $P_{j\to j}(f)$ measures the part of $P_{jj}(f)$ which is not sent to the other processes, and is expressed in normalized terms by the squared PDC $|\pi_{jj}(f)|^2$. The quantity which we denote as PDC was named "generalized PDC" in (Baccala et al., 2007), while the original version of the PDC (Baccala & Sameshima, 2001) was not including inner normalization by the input noise variances. Our definition (15) follows directly from the decomposition in (14); besides, this definition shares with the Coh, PCoh and DC functions the desirable property of scale-invariance, contrary to the original PDC that may be affected by different amplitudes for the considered signals.

Although both DC and PDC may be regarded as frequency domain descriptors of causality, there are important differences between these two estimators. First, as the DC and the PDC are factors in the decomposition of Coh and PCoh, respectively, they measure causality and direct causality in the frequency domain. In fact, the PDC $\pi_{ij}(f)$ is nonzero if and only if direct causality $y_j\to y_i$ exists, because the numerator of (15) contains, with $i\neq j$, the term $\overline{A}_{ij}(f)$, which is nonzero only when $a_{ij}(k)\neq 0$ for some k and is uniformly zero when $a_{ij}(k)=0$ for each k. As to the DC, one can show that, expanding $\mathbf{H}(f)=\bar{\mathbf{A}}(f)^{-1}$ as a geometric series, the transfer function $H_{ij}(f)$ contains a sum of terms each one related to one of the (direct or indirect) transfer paths connecting y_j to y_i (Eichler, 2006). Hence, the DC $\gamma_{ij}(f)$ is nonzero whenever any path connecting y_j to y_i is significant, i.e., when causality $y_j\Rightarrow y_i$ occurs. Another important difference between DC and PDC is in the normalization: as seen in (12) and in (16), $\gamma_{ij}(f)$ is normalized with respect to the structure that receives the signal, while $\pi_{ij}(f)$ is normalized with respect to the structure that sends the signal. Summarizing, we can

state that the DC measures causality as the amount of information flowing from y_j to y_i through all (direct and indirect) transfer pathways, relative to the total inflow entering the structure at which y_i is measured; the PDC measures direct causality as the amount of information flowing from y_j to y_i through the direct transfer pathway only, relative to the total outflow leaving the structure at which y_j is measured. We note that this dual interpretation highlights advantages and disadvantages of both measures. The DC has a meaningful physical interpretation as it measures causality as the amount of signal power transferred from one process to another, but cannot distinguish between direct and indirect causal effects measured in the frequency domain. Conversely, the PDC reflects clearly the underlying interaction structure as it provides a one-to-one representation of direct causality, but is hardly useful as a quantitative measure because its magnitude quantifies the information flow through the inverse spectral matrix elements (which do not find easy interpretation in terms of power spectral density).

3.3 Theoretical example

To discuss the properties and compare the behavior of the frequency domain measures of causality and coupling summarized in Table 1, we consider the MVAR vector process of order p=2, composed by M=4 processes, generated by the equations:

$$\begin{cases} y_1(n) = 2\rho\cos(2\pi f_1)y_1(n-1) - \rho^2 y_1(n-2) + u_1(n) \\ y_2(n) = y_1(n-1) + 0.5y_3(n-1) + u_2(n) \\ y_3(n) = 2\rho\cos(2\pi f_3)y_3(n-1) - \rho^2 y_3(n-2) + 0.5y_2(n-1) + 0.5y_2(n-2) + u_3(n) \\ y_4(n) = y_1(n-2) + u_4(n) \end{cases}, \tag{18}$$

with ρ=0.9, f_1=0.1 and f_3=0.3, where the inputs $u_i(n)$, i=1,2,3, are fully uncorrelated and with variance σ^2_i. The process (18) is one of the possible MVAR realizations of the diagram of Fig. 1a. The coupling and causality relations emerging from the diagram, discussed in Sect. 2, can be interpreted here in terms of the MVAR coefficients set in (18). In fact, the nonzero off-diagonal values of the coefficient matrix ($a_{21}(1)$=1, $a_{23}(1)$=0.5, $a_{32}(1)$=0.5, $a_{32}(2)$=0.5, $a_{41}(2)$=1) determine direct causality and causality among the processes —and consequently direct coupling and coupling— in agreement with the definitions particularized at the end of Sect. 3.1. For instance, $a_{21}(1)$ and $a_{32}(2)$=0.5 determine direct causality $y_1 \rightarrow y_2$ and $y_2 \rightarrow y_3$ as well as causality $y_1 \Rightarrow y_2$, $y_2 \Rightarrow y_3$ (direct interaction) and $y_1 \Rightarrow y_3$ (indirect interaction). The diagonal values of the coefficient matrix determine autonomous oscillations in the processes. Indeed, the values set for $a_{ii}(k)$, $a_{ii}(1)=2\rho\cos(2\pi f_i)$, $a_{ii}(2)=-\rho^2$, generate complex-conjugate poles with modulus ρ and phases $\pm 2\pi f_i$ for the process y_i (the sampling period is implicitly assumed to be T=1). In this case, narrow-band oscillations at 0.1 Hz and 0.2 Hz are set for y_1 and y_3.

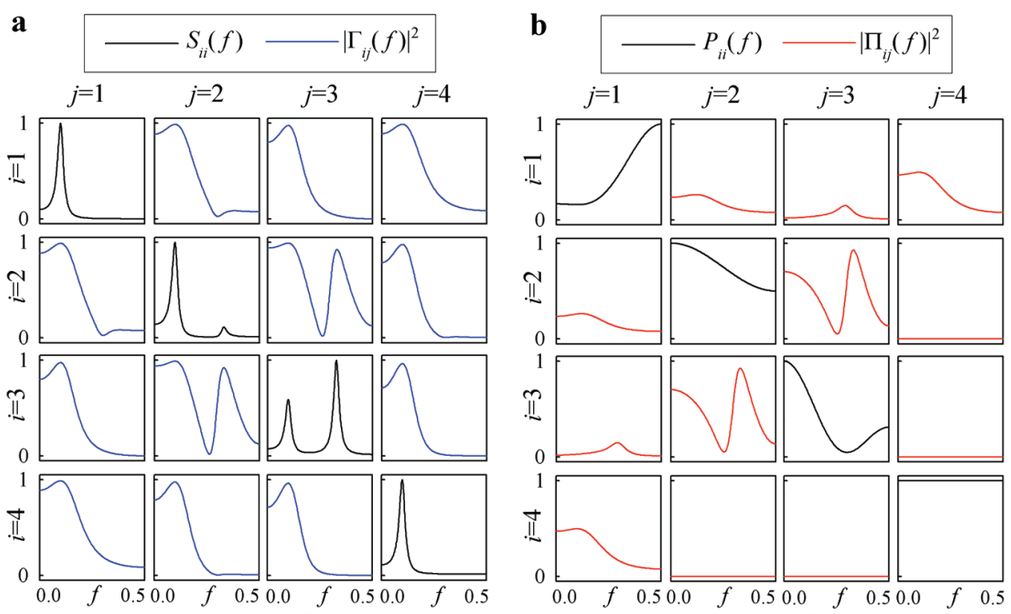
The trends of spectral and cross-spectral density functions are reported in Fig. 2.

The spectra of the four processes, reported as diagonal plots in Fig. 2a (black), exhibit clear peaks at the frequency of the two imposed oscillations: the peaks at ~0.1 Hz and ~0.3 Hz are dominant for y_1 and y_3, respectively, and appear also in the spectra of the remaining processes according to the imposed causal information transfer. The inverse spectra, computed as the diagonal elements of the inverse spectral matrix $\mathbf{P}(f)$, are also reported as diagonal plots in Fig. 2b (black). Off diagonal plots of Fig. 2a and Fig. 2b depict respectively the trends of the squared magnitudes of Coh and PCoh; note the symmetry of the two

functions ($\Gamma_{ij}(f)=\Gamma^*_{ji}(f)$, $\Pi_{ij}(f)=\Pi^*_{ji}(f)$), reflecting the fact that they measure coupling and direct coupling but cannot account for directionality of the considered interaction. As expected, Coh is nonzero for each pair of processes, thus measuring the full connectivity of the considered network. PCoh is clearly nonzero whenever a direct coupling relation exists ($y_1 \leftrightarrow y_2$, $y_2 \leftrightarrow y_3$, $y_1 \leftrightarrow y_4$, $y_1 \leftrightarrow y_3$), and is uniformly zero between y_2 and y_4 and between y_3 and y_4 where no direct coupling is present.

Figs. 3 and 4 depict the decomposition of spectra and inverse spectra, as well as the trends of DC and PDC functions resulting from these decompositions.

Fig. 3a provides a graphical representation of (13), showing how the spectrum of each process can be decomposed into power contributions related to all processes; normalizing these contributions one gets the squared modulus of the DC, as depicted in Fig. 3b. In the example, the spectrum of y_1 is decomposed in one part only, deriving from the same process. This indicates that none part of the power of y_1 is due to the other processes. The absence of external contributions is reflected by the null profiles of the DC from y_1 to y_2, to y_3 and to y_4 seen in Fig. 3b; as a result, the squared DC $|\gamma_{11}(f)|^2$ has a flat unitary profile. On the contrary, the decompositions of y_i, with i=2,3,4, results in contributions from the other processes, so that the squared DC $|\gamma_{ij}(f)|^2$ is nonzero for some $j \neq i$, and the squared DC $|\gamma_{ii}(f)|^2$ is not uniformly equal to 1 as a result of the normalization condition. In particular, we observe that the power of the peak at f_1=0.1 Hz is entirely due to y_1 for all processes, determining very high values of the squared DC in the first column of the matrix plot in Fig. 3b, i.e., $|\gamma_{i1}(f_1)|^2=1$; this behavior represents in the frequency domain the causality relations imposed from y_1 to all other processes. The remaining two causality relations, relevant to the bidirectional interaction between y_2 and y_3, concern the oscillation at f_2=0.3 Hz, which is



$S_{ii}(f)$: spectrum of the process y_i; $P_{ii}(f)$: inverse spectrum of y_i; $\Gamma_{ij}(f)$: coherence between y_j and y_i; $\Pi_{ij}(f)$: partial coherence between y_j and y_i.

Fig. 2. Spectral functions and frequency domain coupling measures for the theoretical example (18).

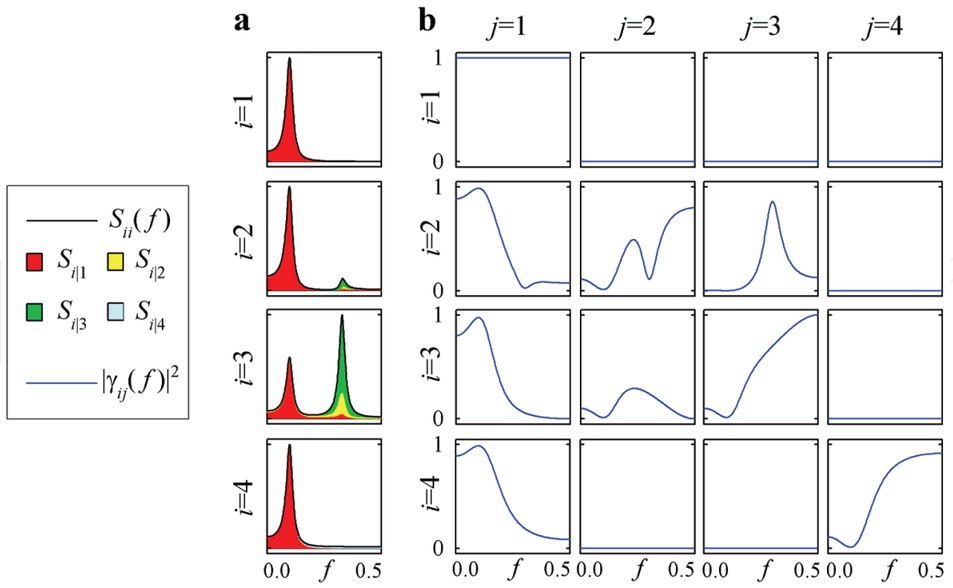


Fig. 3. Decomposition of the power spectrum of each process y_i in (18), $S_{ii}(f)$, into contributions from each process y_j ($S_{i|j}$, shaded areas in each plot) (**a**), and corresponding DC from y_j to y_i, $\gamma_{ij}(f)$ (**b**) depicted for each $i,j=1,...,M$.

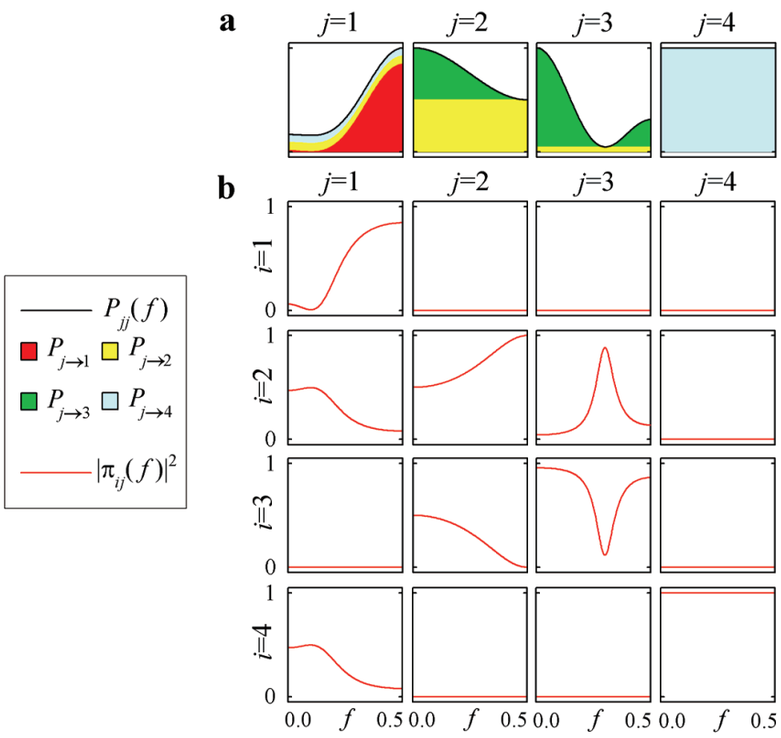


Fig. 4. Decomposition of the inverse power spectrum of each process y_j in (18), $P_{jj}(f)$, into contributions towards each process y_i ($P_{j\to i}$, shaded areas in each plot) (**a**), and corresponding PDC from y_j to y_i, $\pi_{ij}(f)$ (**b**) depicted for each $i,j=1,...,M$.

generated in y_3 and then transmitted forward and backward to y_2. This behavior is reflected by the spectral decomposition of the peak at ~0.3 Hz of the spectra of y_2 and y_3, which shows contributions from both processes, and then in the nonzero profiles of the DCs $|\gamma_{23}(f)|^2$ and $|\gamma_{32}(f)|^2$ (with peaks at ~0.3 Hz). To complete the frequency domain picture of causality, we observe that the DC is uniformly zero along all directions over which no causality is imposed in the time domain (e.g., compare Fig. 3b with Fig. 1a).

A specular interpretation to the one given above holds for the decomposition of the inverse spectra into absolute and normalized contributions sent to all processes, which are depicted for the considered example in the area plot of Fig. 4a and in the matrix PDC plot of Fig. 4b, respectively. The difference is that now contributions are measured as outflows instead as inflows, are normalized to the structure sending the signal instead to that receiving the signal, and reflect the concept of direct causality instead that of causality. With reference to the proposed example, we see that the inverse spectrum of y_1 is decomposed into contributions flowing out towards y_2 and y_4 (yellow and cyan areas underlying $P_{11}(f)$ in Fig. 4a), which are expressed in normalized units by the squared PDCs $|\pi_{21}(f)|^2$ and $|\pi_{41}(f)|^2$. While y_2 and y_3 affect each other (absolute units: $P_{2\to3}\neq0$, $P_{3\to2}\neq0$; normalized units: $|\pi_{32}|^2\neq0$ and $|\pi_{23}|^2\neq0$) without being affected by the other processes ($|\pi_{i2}|^2=0$, $|\pi_{i3}|^2=0$, i=1,4), y_4 does not send information to any process ($P_{4\to i}=0$, $|\pi_{i4}|^2=0$, i=1,2,3). As can be easily seen comparing Fig. 4 with Fig. 1a, the profiles of $P_{j\to i}$ and $|\pi_{ij}|^2$ provide a frequency domain description, respectively in absolute and normalized terms, of the imposed pattern of direct causality. We note also that all inverse spectra contain a contribution coming from the same process, which describes the part of $P_{jj}(f)$ which is not sent to any of the other processes ($P_{j\to j}$ in Fig. 4a). After normalization, this contribution is quantified by the PDC $|\pi_{jj}|^2$, as depicted by the diagonal plots of Fig. 4b.

4. Causality and coupling in the presence of instantaneous interactions

4.1 MVAR processes with instantaneous effects

The MVAR model defined in (5) is a strictly causal model, in the sense that it accounts only for lagged effects, i.e. the effects of the past of a time series on another series, while instantaneous effects (i.e., effects of $y_j(n)$ on $y_i(n)$) are not described by any model coefficient. The problem with this model representation is that any zero-lag correlation among the observed series y_i, when present, cannot be described by the model because $\mathbf{A}(k)$ is defined only for positive lags (k=0 is not considered in (5)). Neglecting instantaneous effects in the MVAR representation of multiple processes implies that any zero-lag correlation among the processes is translated into a correlation among the model inputs (Lutkepohl, 1993). As a result, the input covariance matrix $\mathbf{\Sigma}$=cov($\mathbf{U}(n)$) is not diagonal. We will see in the next subsections that, since non-diagonality of $\mathbf{\Sigma}$ contradicts the assumptions of spectral factorization, the presence of significant instantaneous interactions may be detrimental for the estimation of causality and direct causality through the DC and PDC estimators.

As an alternative to using the strictly causal model (5), the multivariate process $\mathbf{Y}(n)$ can be described including instantaneous effects into the interactions allowed by the model. This is achieved considering the extended MVAR process (Faes & Nollo, 2010b):

$$\mathbf{Y}(n)=\sum_{k=0}^{p}\mathbf{B}(k)\mathbf{Y}(n-k)+\mathbf{W}(n)\,, \tag{19}$$

where $\mathbf{W}(n)=[w_1(n),...,w_M(n)]^T$ is a vector of zero-mean uncorrelated white noise processes with diagonal covariance matrix $\mathbf{\Lambda}=\text{diag}(\lambda^2_i)$. The difference with respect to strictly causal MVAR modelling as in (5) is that now the lag variable k takes the value 0 as well, which brings instantaneous effects from $y_j(n)$ to $y_i(n)$ into the model in the form of the coefficients $b_{ij}(0)$ of the matrix $\mathbf{B}(0)$. In the extended MVAR model, absence of correlation among the noise inputs w_i, i.e. diagonality of $\mathbf{\Lambda}=\text{cov}(\mathbf{W}(n))$, is guaranteed by the presence of the instantaneous effects. Thus, the assumption that the input is a white noise vector process is always fulfilled by the extended representation.

The relation between the strictly causal representation and the extended representation can be established moving the term $\mathbf{B}(0)\mathbf{Y}(n)$ from the right to the left side of (19) and then left-multiplying both sides by by the matrix $\mathbf{L}^{-1}=[\mathbf{I}-\mathbf{B}(0)]$. The comparison with (5) yields:

$$\mathbf{A}(k)=\mathbf{L}\mathbf{B}(k)=[\mathbf{I}-\mathbf{B}(0)]^{-1}\mathbf{B}(k), \tag{20}$$

$$\mathbf{U}(n)=\mathbf{L}\mathbf{W}(n), \quad \mathbf{\Sigma}=\mathbf{L}\mathbf{\Lambda}\mathbf{L}^T \ . \tag{21}$$

These relationships indicate that the two representations coincide in the absence of instantaneous effects, and that the assumption of uncorrelated inputs is not satisfied in the presence of instantaneous effects. In fact, in the model (5) the input noise covariance $\mathbf{\Sigma}$ is not diagonal whenever $\mathbf{B}(0)\neq 0$ (and $\mathbf{L}\neq\mathbf{I}$). If the matrix $\mathbf{B}(0)$ has all zero entries we have $\mathbf{L}=\mathbf{I}$ and the model (19) reduces to (5) ($\mathbf{A}(k)=\mathbf{B}(k)$, $\mathbf{U}(n)=\mathbf{W}(n)$, $\mathbf{\Sigma}=\mathbf{\Lambda}$). By contrast, the existence of $\mathbf{B}(0)\neq 0$ makes coefficients $\mathbf{B}(k)$ differ from $\mathbf{A}(k)$ at each lag k. This property is crucial as it says that different patterns of causality may be found depending on whether instantaneous effects are included or not in the MVAR model used to represent the available data set.

Contrary to the strictly causal MVAR model which may describe only lagged interactions, the extended MVAR representation allows to detect any type of interaction defined in Sect. 2 and Table 1 from the elements $b_{ij}(k)$ of the matrix coefficients $\mathbf{B}(k)$. Specifically, direct causality $y_j \rightarrow y_i$ and extended direct causality $y_j \dot{\rightarrow} y_i$ are detected if $b_{ij}(k)\neq 0$ for at least one $k=1,\ldots,p$, and for at least one $k=0,1,\ldots,p$, respectively. Causality $y_j \Rightarrow y_i$ and extended causality $y_j \dot{\Rightarrow} y_i$ are detected if $b_{m_s m_{s-1}}(k_s)\neq 0$ for a set of lags $k_0,...,k_{L-1}$ with values in $(1,...,p)$ and with values in $(0,1,...,p)$, respectively. Direct coupling $y_i \leftrightarrow y_j$ is detected if $b_{ji}(k)\neq 0$ and/or $b_{ij}(k)\neq 0$ for at least one $k=0,1,\ldots,p$. Coupling $y_i \Leftrightarrow y_j$ is detected if $b_{m_s m_{s-1}}(k_s)\neq 0$ for a set of $L\geq 2$ values for m_s (either with $m_0=j$, $m_{L-1}=i$ or with $m_0=i$, $m_{L-1}=j$) and a set of lags $k_0,...,k_{L-1}$.

4.2 Frequency domain analysis of extended MVAR processes

The spectral representation of an extended MVAR process is obtained taking the FT of (19) to yield $\mathbf{Y}(f)=\mathbf{B}(f)\mathbf{Y}(f)+\mathbf{W}(f)$, where

$$\mathbf{B}(f)=\mathbf{B}(0)+\sum_{k=1}^{p}\mathbf{B}(k)e^{-j2\pi f kT} \tag{22}$$

is the frequency domain coefficient matrix. The representation evidencing input-output relations is $\mathbf{Y}(f)=\mathbf{G}(f)\mathbf{W}(f)$, where the transfer matrix is given by $\mathbf{G}(f)=[\mathbf{I}-\mathbf{B}(f)]^{-1}=\overline{\mathbf{B}}(f)^{-1}$. Given these representations, the spectral matrix $\mathbf{S}(f)$ and its inverse $\mathbf{P}(f)$ are expressed for the extended MVAR model as:

$$\mathbf{S}(f)=\mathbf{G}(f)\mathbf{\Lambda}\mathbf{G}^H(f), \ \mathbf{P}(f)=\overline{\mathbf{B}}^H(f)\mathbf{\Lambda}^{-1}\overline{\mathbf{B}}(f), \tag{23}$$

By means of some matrix algebra involving the spectral representations of (5) and (19), as well as (21), it is easy to show that the spectral matrix (and its inverse as well) resulting in (23) are exactly the same as those obtained in (7). This demonstrates the equivalence of the spectral representation for strictly causal MVAR processes and extended MVAR processes. Consequently, also the concepts of coupling and direct coupling are equivalent for the two process representation, since Coh and PCoh estimated as in (3) depend exclusively on the elements of $\mathbf{S}(f)$ and $\mathbf{P}(f)$. For this reason, a single estimator for Coh and PCoh is indicated in Table 1. A substantial difference between the conditions without and with instantaneous effects arises when coupling relations are decomposed to infer causality. We remark that the original formulation of DC and PDC holds fully only under the assumption of uncorrelation of the input processes, leading to diagonality of $\mathbf{\Sigma}$ and $\mathbf{\Sigma}^{-1}$. When such an assumption is not fulfilled, the spectral factorizations in (9) do not hold anymore and the DC and PDC may become unable to identify causality and direct causality in the frequency domain. On the contrary, since the extended MVAR representation leads to diagonal input covariance matrices $\mathbf{\Lambda}$ and $\mathbf{\Lambda}^{-1}$ by construction, the factorizations in (9) are valid (using $\mathbf{B}(f)$ and $\mathbf{G}(f)$ as coefficient and transfer matrices in place of $\mathbf{A}(f)$ and $\mathbf{H}(f)$) still in the presence of instantaneous interactions among the observed processes. In particular, the following factorizations hold for the Coh:

$$\Gamma_{ij}(f)=\frac{\mathbf{g}_i(f)\mathbf{\Lambda}\mathbf{g}_j^{\mathrm{H}}(f)}{\sqrt{\mathbf{g}_i(f)\mathbf{\Lambda}\mathbf{g}_i^{\mathrm{H}}(f)}\sqrt{\mathbf{g}_j(f)\mathbf{\Lambda}\mathbf{g}_j^{\mathrm{H}}(f)}}=\sum_{m=1}^{M}\frac{\lambda_m G_{im}(f)}{\sqrt{S_{ii}(f)}}\frac{\lambda_m G_{jm}^{*}(f)}{\sqrt{S_{jj}(f)}}=\sum_{m=1}^{M}\xi_{im}(f)\xi_{jm}^{*}(f) \tag{24}$$

and the PCoh:

$$\Pi_{ij}(f)=\frac{\overline{\mathbf{b}}_i^{\mathrm{H}}(f)\mathbf{\Lambda}^{-1}\overline{\mathbf{b}}_j(f)}{\sqrt{\overline{\mathbf{b}}_i^{\mathrm{H}}(f)\mathbf{\Lambda}^{-1}\overline{\mathbf{b}}_i(f)}\sqrt{\overline{\mathbf{b}}_j^{\mathrm{H}}(f)\mathbf{\Lambda}^{-1}\overline{\mathbf{b}}_j(f)}}=\sum_{m=1}^{M}\frac{\frac{1}{\lambda_m}\overline{B}_{mj}(f)}{\sqrt{P_{jj}(f)}}\frac{\frac{1}{\lambda_m}\overline{B}_{mi}^{*}(f)}{\sqrt{P_{ii}(f)}}=\sum_{m=1}^{M}\chi_{mj}(f)\chi_{mi}^{*}(f), \tag{25}$$

where $\mathbf{g}_i(f)$ is the i-th row of $\mathbf{H}(f)$ and $\overline{\mathbf{b}}_i(f)$ is the i-th column of $\overline{\mathbf{B}}(f)$. The last terms of (24) and (25) contain the so-called *extended DC* (eDC) and *extended PDC* (ePDC), which are defined, for the extended MVAR model including instantaneous effects, respectively as:

$$\xi_{ij}(f)=\frac{\lambda_j G_{ij}(f)}{\sqrt{S_{ii}(f)}}=\frac{\lambda_j G_{ij}(f)}{\sqrt{\sum_{m=1}^{M}\lambda_m^2\left|G_{im}(f)\right|^2}}, \tag{26}$$

and as (Faes & Nollo, 2010b):

$$\chi_{ij}(f)=\frac{\frac{1}{\lambda_i}\overline{B}_{ij}(f)}{\sqrt{P_{jj}(f)}}=\frac{\frac{1}{\lambda_i}\overline{B}_{ij}(f)}{\sqrt{\sum_{m=1}^{M}\frac{1}{\lambda_m^2}\left|\overline{B}_{mj}(f)\right|^2}}. \tag{27}$$

The normalization conditions in (12) and (16) keep holding for the eDC and the ePDC defined in (26) and in (27). Hence, the squared eDC and ePDC $|\xi_{ij}(f)|^2$ and $|\chi_{ij}(f)|^2$ maintain their meaning of normalized proportion of $S_{ii}(f)$ which comes from y_j, and

normalized proportion of $P_{jj}(f)$ which is sent to y_i, respectively, even in the presence of significant zero-lag interactions among the observed processes. In other words, we have that the meaningful decompositions in (13) and (17) are always valid for the eDC and the ePDC, respectively. On the contrary, these decompositions hold for the DC and the PDC only if instantaneous effects are negligible; when they are not, DC and PDC can still be computed through (11) and (15) but, since the factorizations in (9) are not valid when $\mathbf{\Sigma}$ is not diagonal, their numerator is no more a factor in the decomposition of Coh and PCoh, and their denominator is no more equal to $S_{ii}(f)$ or $P_{jj}(f)$. As we will see in a theoretical example in the next subsection, these limitations may lead to erroneous interpretations of causality and direct causality in the frequency domain.

When we use the extended measures in (26) and (27), the information which flows from y_i to y_j is both lagged (k>0) and instantaneous (k=0), because it is measured in the frequency domain by the function $\mathbf{B}(f)$ which incorporates both $\mathbf{B}(0)$ and $\mathbf{B}(k)$ with k>0. Therefore, eDC and ePDC measure in the frequency domain the concepts of extended causality and extended direct causality, respectively (see Table 1). If we want to explore lagged causality in the presence of zero-lag interactions, we have to exclude the coefficients related to the instantaneous effects from the desired spectral causality measure. Hence, we set:

$$\widetilde{\mathbf{B}}(f)=\overline{\mathbf{B}}(f)+\mathbf{B}(0)=\mathbf{I}-\sum_{k=1}^{p}\mathbf{B}(k)e^{-j2\pi fkT} \quad , \quad \widetilde{\mathbf{G}}(f)=\widetilde{\mathbf{B}}(f)^{-1} \tag{28}$$

and then we define the following DC and PDC functions:

$$\widetilde{\gamma}_{ij}(f)=\frac{\lambda_j\widetilde{G}_{ij}(f)}{\sqrt{\sum_{m=1}^{M}\lambda_m^2\left|\widetilde{G}_{im}(f)\right|^2}} \; , \; \widetilde{\pi}_{ij}(f)=\frac{\frac{1}{\lambda_i}\widetilde{B}_{ij}(f)}{\sqrt{\sum_{m=1}^{M}\frac{1}{\lambda_m^2}\left|\widetilde{B}_{mj}(f)\right|^2}} \; . \tag{29}$$

Since they are derived exclusively from time domain matrices of lagged effects, $\widetilde{\gamma}_{ij}(f)$ and $\widetilde{\pi}_{ij}(f)$ measure respectively causality and direct causality in the frequency domain (see Table 1). We stress that, even though measuring the same kind of causality, the DC and PDC given in (29) are different from the corresponding functions given in (11) and (15), because the presence of instantaneous effects leads to different estimates of the coefficient matrix, or of the transfer matrix, using strictly causal or extended MVAR models. Only in the absence of instantaneous effects ($\mathbf{B}(0)$=0) DC and PDC estimated by the two models are the same, and are also equivalent to eDC and ePDC.

4.3 Theoretical example

In this section we compare the behavior of the different measures of frequency domain causality in a MVAR processes with imposed connectivity patterns including instantaneous interactions. The process is defined according to the interaction diagram of Fig. 1b, and is generated by the equations:

$$\begin{cases} y_1(n)=2\rho_1\cos(2\pi f_1)y_1(n-1)-\rho_1^2y_1(n-2)-0.4y_3(n-1)+w_1(n) \\ y_2(n)=y_1(n)+0.2y_1(n-1)-\rho_2^2y_2(n-2)+w_2(n) \\ y_3(n)=0.8y_2(n)+w_3(n) \\ y_4(n)=0.6y_2(n)+w_4(n) \end{cases} \; , \tag{30}$$

with ρ_1=0.95, f_1=0.125, ρ_2=0.8, and where the inputs $w_i(n)$, i=1,…,4 are fully uncorrelated white noise processes with variance equal to 1 for w_1 and w_4, equal to 2 for w_2, and equal to 8 for w_3. The diagonal values of the coefficient matrix are set to generate autonomous oscillations at ~0.125 Hz and ~0.25 Hz for y_1 and y_2, respectively. The nonzero off-diagonal coefficients set the direct directed interactions, which are exclusively instantaneous from y_2 to y_3 and from y_2 to y_4, exclusively lagged from y_3 to y_1, and mixed instantaneous and lagged from y_1 to y_2. The coupling and causality relations resulting from this scheme are described in detail in Sect. 2.1, with reference to Fig. 1b.

The MVAR process (30) is suitably described by the MVAR model with instantaneous effects of Fig. 5a, in which the use of coefficients describing both instantaneous and lagged effects allows to reproduce identically both the set of interactions imposed in (30) and the connectivity pattern of Fig. 1b. On the contrary, when a strictly causal MVAR process in the form of (5) is used to describe the same network, the resulting model is that of Fig. 5b. The strictly causal structure in Fig. 5b results from the application of (20) and (21) to the extended structure, leading to different values for the coefficients. As seen in Fig. 5, the result is an overestimation of the number of active direct pathways, and a general different estimation of the causality patterns. For instance, while in the original formulation (30) and in the extended MVAR representation of Fig. 5a direct causality is present only from y_1 to y_2 and from y_3 to y_1, a much higher number of direct causality relations is erroneously represented in Fig. 5b: in some cases instantaneous effects are misinterpreted as lagged (e.g., from y_2 to y_3 and from y_2 to y_4), in some other spurious connections appear (e.g., from y_1 to y_3 and from y_1 to y_4). The misleading connectivity pattern of Fig. 5b is the result of the impossibility for the model (5) to describe instantaneous effects. In the strictly causal representation, these effects are translated into the input covariance matrix (according to (20)): indeed, not only the input variances are different, but also cross-correlations between the input processes arise; in this case, we have σ^2_{12}=1, σ^2_{13}=0.8, σ^2_{14}=0.6, σ^2_{23}=2.4, σ^2_{24}=1.8, σ^2_{34}=1.44, whereas λ^2_{ij}=0 for each $i \neq j$.

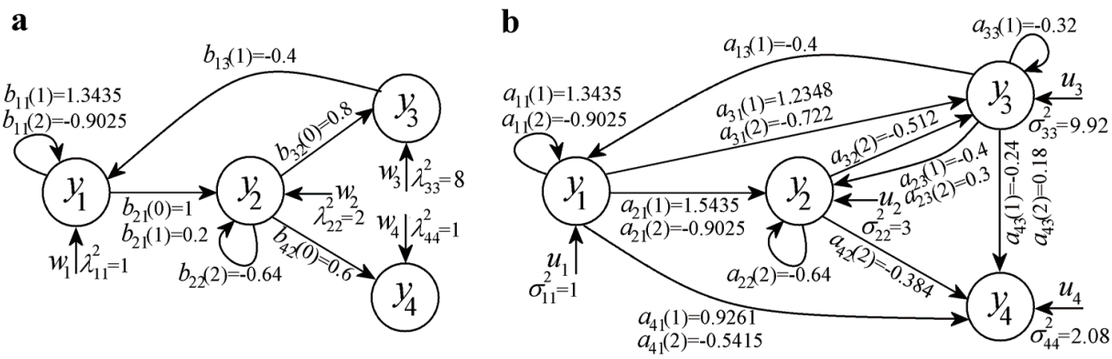
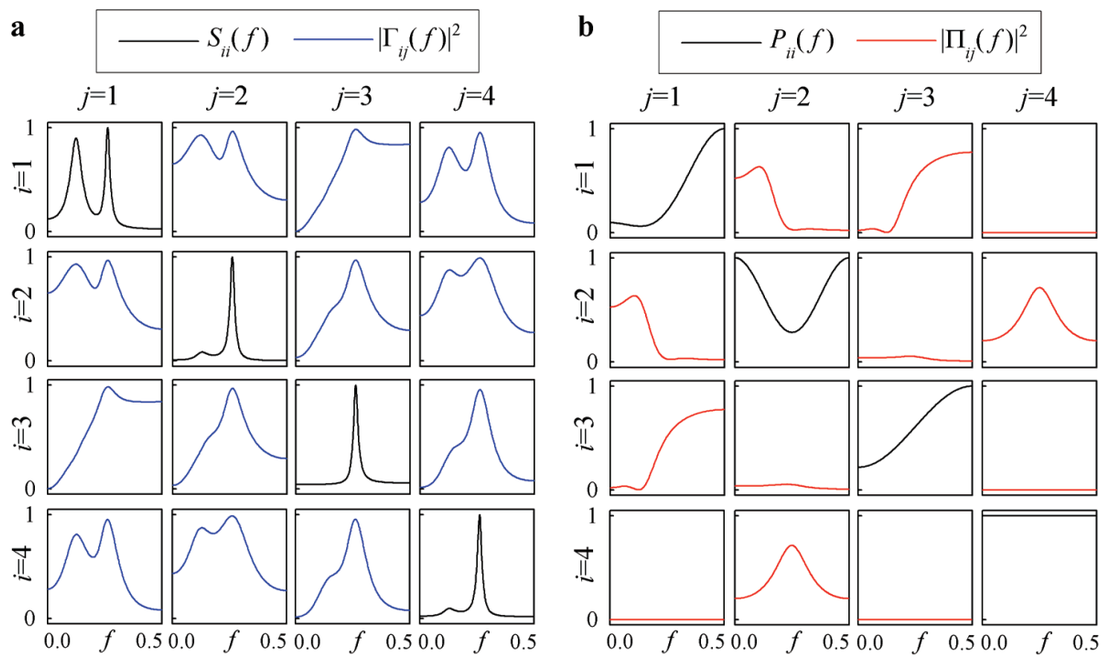


Fig. 5. Extended MVAR representation (**a**) and strictly causal MVAR representation (**b**) of the MVAR process generated by the equations in (30).

Fig. 6 reports the spectral and cross-spectral functions of the MVAR process (30). We remark that, due to the equivalence of eqs. (7) and (22), the profiles of spectra and inverse spectra, as well as of Coh and PCoh, perfectly overlap when calculated either from the strictly causal or from the extended MVAR representation. Despite this, these profiles are readily interpretable from the definitions of coupling and direct coupling only when the extended

representation is used, while they do not describe the strictly causal representation. For instance, the PCohs reported in Fig. 6b have a one-to-one correspondence with the extended MVAR diagram of Fig. 5a, as a nonzero PCoh is shown in Fig. 6b exactly when a direct coupling is described by the coefficients in Fig. 5a (i.e., between y_1 and y_2, y_2 and y_3, y_2 and y_4, and y_1 and y_3). On the contrary, the PCoh profiles do not explain the direct coupling interactions which may be inferred from the strictly causal model in Fig. 5b: e.g., the nonzero coefficients $a_{41}(1)$ and $a_{41}(2)$ suggest the existence of direct coupling $y_1 \leftrightarrow y_4$,while such a coupling is not reflected by nonzero values of the PCoh $\Pi_{14}(f)$.



$S_{ii}(f)$: spectrum of the process y_i; $P_{ii}(f)$: inverse spectrum of y_i; $\Gamma_{ij}(f)$: coherence between y_j and y_i; $\Pi_{ij}(f)$: partial coherence between y_j and y_i.

Fig. 6. Spectral functions and frequency domain coupling measures for the theoretical example (30).

The problems of using the strictly causal MVAR representation in the presence of instantaneous effects become even more severe when one aims at disentangling the coupling relations to measure causality in the frequency domain. In this case, the spectral representations closely reflect the time domain diagrams, but —quite for this reason— only the extended spectral profiles are correct while the strictly causal one may be strongly misleading. This is demonstrated in Figs. 7 and 8, depicting respectively the frequency domain evaluation of causality and direct causality for the considered theoretical example.

As shown in Fig. 7, the extended MVAR representation of the considered process yields a frequency-domain connectivity pattern which is able to describe all and only the imposed direct connections: the PDC correctly portrays (lagged) direct causality from y_1 to y_2 and from y_3 to y_1, being zero over all other directions (black dashed curves in Fig. 7a); the ePDC portrays all extended causality relations, being nonzero only from y_1 to y_2 (mixed instantaneous and lagged effect), from y_3 to y_1 (lagged effect), as well as from y_2 to y_3 and from y_2 to y_4

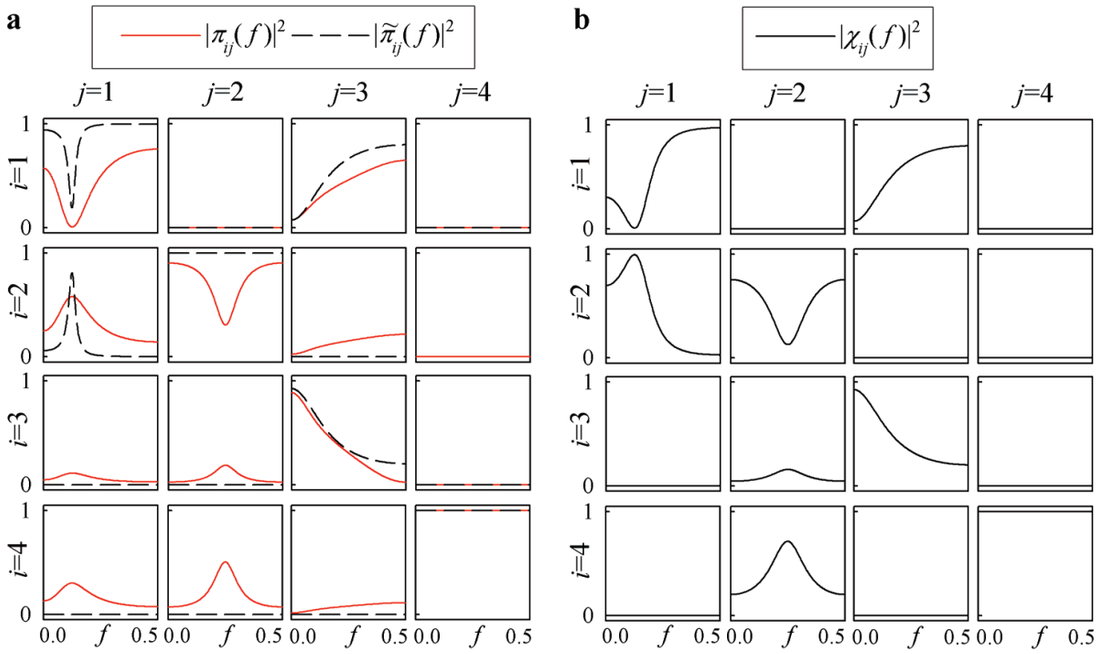


Fig. 7. Diagrams of squared PDC for the strictly causal MVAR representation ($|\pi_{ij}(f)|^2$) and the extended representation ($|\tilde{\pi}_{ij}(f)|^2$) (**a**), and of squared ePDC ($|\chi_{ij}(f)|^2$, **b**), for the theoretical example (30).

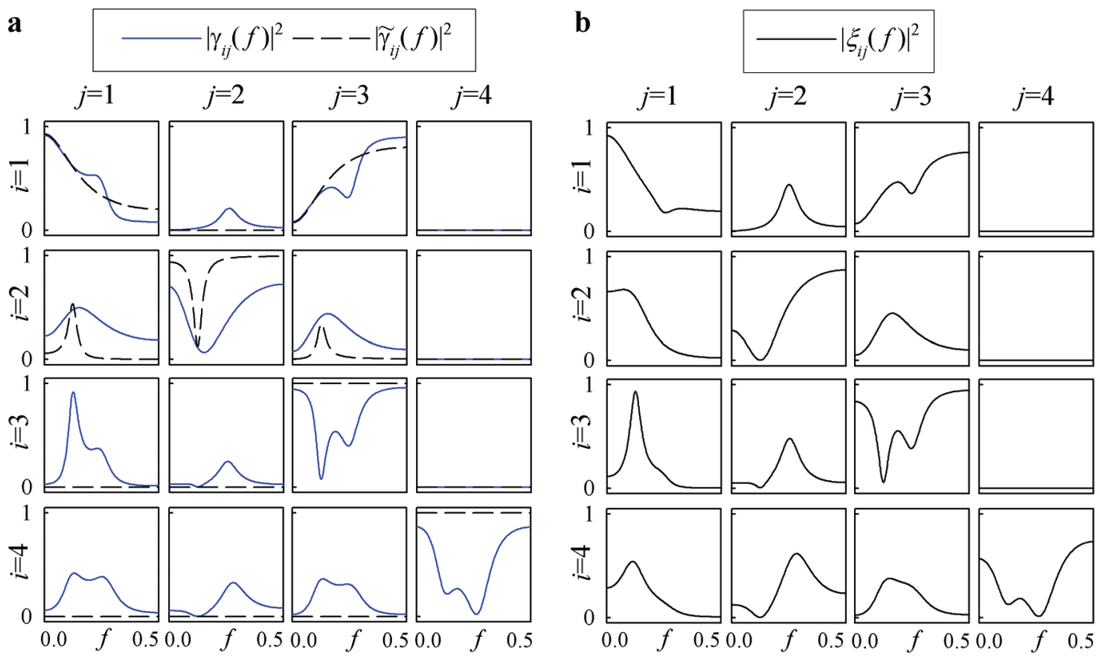


Fig. 8. Diagrams of squared DC for the strictly causal MVAR representation ($|\gamma_{ij}(f)|^2$) and the extended representation ($|\tilde{\gamma}_{ij}(f)|^2$) (**a**), and of squared eDC ($|\xi_{ij}(f)|^2$, **b**), for the theoretical example (30).

(instantaneous effect) (Fig. 7b). On the contrary, utilization of the strictly causal MVAR representation leads to erroneous interpretation of lagged direct causality. Indeed, as seen in Fig. 7a (red solid curves), the PDC interprets as lagged the instantaneous connections from y_2 to y_3 and from y_2 to y_4. Moreover, spurious causality effects are measured, as the PDC is nonzero from y_1 to y_3, from y_1 to y_4 and from y_3 to y_4, albeit no direct effects (neither lagged nor instantaneous) are imposed in (30) over these directions.

A similar situation occurs when causality and extended causality are studied in the frequency domain through DC-based functions. Fig. 8a shows that the pattern of causality relations imposed in (30) (i.e., $y_1 \Rightarrow y_2$, $y_3 \Rightarrow y_1$, and $y_3 \Rightarrow y_2$) is not reflected by the DC measured from the strictly causal model through eq. (11). The DC profile (blue solid curves) describes indeed many other causal effects besides the two correct ones; precisely, all (direct or indirect) causality relations emerging from the diagram of Fig. 5b are measured with nonzero values of the DC and thus interpreted as lagged causal effects. These effects are actually due to instantaneous interactions, and thus should not be represented by the DC as it is a measure of lagged causality only. The correct representation is given using the DC measured from the extended MVAR model through eq. (29): in this case, the only nonzero squared DCs are those measured over the three directions with imposed causality, while the squared DC is zero over all other directions (black dashed curves in Fig. 8a). The relations of causality emerging thanks to the instantaneous effects are detected by the eDC computed through (29) and plotted in Fig. 8b, which is able to measure also instantaneous effects in addition to the lagged ones. Thus, we see that a correct frequency domain representation of causality and extended causality is given by the DC and eDC functions derived from the extended MVAR representation of the considered process.

5. Practical analysis

5.1 Model identification

The practical application of the theoretical concepts described in the previous sections is based on considering the available set of time series measured from the physiological system under analysis, $\{y_m(n), m=1,\ldots,M; n=1,\ldots N\}$, as a finite-length realization of the vector stochastic process describing the evolution of the system over time. Hence, the descriptive equations of the MVAR processes (5) and (19) are seen as a model of how the observed data have been generated. To obtain the various frequency domain functions measuring causality and coupling, estimation algorithms have to be applied to the observed time series for providing estimates of the model coefficients, which are then used in the generating equations in place of the true unknown coefficient values. Obviously, the estimates will never be the exact coefficients, and consequently the frequency domain measures estimated from the real data will always be an approximation of the true functions. The goodness of the approximation depends on practical factors such as the data length, and on the type and parameters of the procedure adopted for the identification of the model coefficients. In the following, we describe some of the possible approaches to identify and validate the MVAR models in (5) and (19) from experimental data.

Identification of the strictly causal MVAR model (5) can be performed with relative ease by means of classic regression methods. The several existing MVAR estimators (see, e.g., (Kay, 1988) or (Lutkepohl, 1993) for detailed descriptions) are all based on the principle of minimizing the prediction error, i.e. the difference between actual and predicted data. A simple, consistent and asymptotically efficient estimator is the MV least squares method. It

is based first on representing (5) through the compact representation $\mathbf{Y}=\mathbf{AZ}+\mathbf{U}$, where $\mathbf{A}=[\mathbf{A}(1)\cdots\mathbf{A}(p)]$ is the $M\times pM$ matrix of the unknown coefficients, $\mathbf{Y}=[\mathbf{Y}(p+1)\cdots\mathbf{Y}(N)]$ and $\mathbf{U}=[\mathbf{U}(p+1)\cdots\mathbf{U}(N)]$ are $M\times(N-p)$ matrices, and $\mathbf{Z}=[\mathbf{Z}_1\cdots\mathbf{Z}_p]$ is a $pM\times(N-p)$ matrix having $\mathbf{Z}_i=[\mathbf{Y}(i)\cdots\mathbf{Y}(N-p+i-1)]$ as i-th row block of (i=1,…M). Given this representation, the method estimates the coefficient matrices through the well known least squares formula: $\hat{\mathbf{A}}=\mathbf{YZ}^T[\mathbf{ZZ}^T]^{-1}$, and the input process as the residual time series: $\hat{\mathbf{U}}=\hat{\mathbf{A}}\mathbf{Z}-\mathbf{Y}$. As to model order selection, one common approach is to set the order p at the value for which the Akaike figure of merit (Akaike, 1974), defined as $AIC(p)=N\log\det\boldsymbol{\Sigma}+M^2p$, reaches a minimum within a predefined range of orders (typically from 1 to 30). While the presented model identification and order selection methods have good statistical properties, more accurate approaches exist; e.g., we refer the reader to (Schlogl, 2006) for a comparison of different MVAR estimators, and to (Erla et al., 2009) for an identification approach combining MVAR coefficient estimation and order selection.

The identification of the extended MVAR model (19) is much less straightforward, because the estimation of instantaneous causality is hard to extract from the covariance information (which is, *per se*, non-directional). In principle, availing of an estimate of the instantaneous effects, described by the matrix $\mathbf{B}(0)$ in the representation (19), identification of the extended MVAR model would follow from that of the strictly causal model describing the same data. Indeed, we recall from (20) and (21) that lagged coefficients and residuals may be estimated for the extended model as $\hat{\mathbf{B}}(k)=[\mathbf{I}-\mathbf{B}(0)]\hat{\mathbf{A}}(k)$ and $\hat{\mathbf{W}}(n)=[\mathbf{I}-\mathbf{B}(0)]\hat{\mathbf{U}}(n)$. Hence, the key for extended MVAR identification is to find the matrix $\mathbf{B}(0)$ which satisfies the instantaneous model $\mathbf{U}(n)=\mathbf{LW}(n)=[\mathbf{I}-\mathbf{B}(0)]^{-1}\mathbf{W}(n)$, and then to use it together with estimates of $\mathbf{A}(k)$ and $\mathbf{U}(n)$ to estimate $\mathbf{W}(n)$ and $\mathbf{B}(k)$ for each $k\geq1$.

The basic problem with the instantaneous model is that it is strictly related to the zero-lag covariance structure of the observed data and, as such, it suffers from lack of identifiability. In other words, there may be several combinations of $\mathbf{L}$ (or, equivalently, $\mathbf{B}(0)$) and $\mathbf{W}(n)$ which result in the same $\mathbf{U}(n)$, and thus describe the observed data $\mathbf{Y}(n)$ equally well. The easiest way to solve this ambiguity is to impose *a priori* the structure of instantaneous causation, i.e. to set the direction (though not the strength) of the instantaneous transfer paths. Mathematically, this can be achieved determining the mixing matrix $\mathbf{L}$ and the diagonal input covariance of the extended model, $\boldsymbol{\Lambda}=\mathrm{cov}(\mathbf{W}(n))$, by application of the Cholesky decomposition to the estimate of input covariance of the strictly causal model, $\boldsymbol{\Sigma}=\mathrm{cov}(\mathbf{U}(n))$ (Faes & Nollo, 2010b). While this decomposition agrees with (21), the resulting $\mathbf{L}$ is is a lower triangular matrix, and $\mathbf{B}(0)$ is also lower triangular with null diagonal. To fulfill this constraint in practical applications, the observed time series have to be ordered in a way such that, for each $j<i$, instantaneous effects are allowed from $y_j(n)$ to $y_i(n)$ ($b_{ij}(0)\neq0$) but not from $y_i(n)$ to $y_j(n)$ ($b_{ji}(0)=0$). This can be done in some applications by means of physical considerations, e.g., based on the temporal order of the events of interest for each measured variable. Following this approach, frequency domain causality was assessed between heart period, arterial pressure and respiration variability in a recent study (Faes & Nollo, 2010a). Nevertheless, if similar prior knowledge is not available, as happens e.g., in the analysis of EEG data, other ways have to be followed to overcome problem of identifiability of the extended model. The most promising one seems that based on exploiting non-Gaussianity in the identification. Indeed, (Shimizu et al., 2006) demonstrated that, if the input process $\mathbf{W}(n)$ has a non-Gaussian distribution, no prior knowledge on the network structure is needed to identify the instantaneous model. The proposed algorithm

estimates $\mathbf{B}(0)$ in two steps (see (Shimizu et al., 2006) for details): first, independent component analysis (ICA) is performed on the estimated residuals $\mathbf{U}(n)$, finding a mixing matrix $\mathbf{M}$ which represents an unordered and non-normalized version of $\mathbf{L}$; second, an appropriate row-permutation followed by normalization is applied to $\mathbf{M}^{-1}$ to get an estimate of $\mathbf{L}^{-1}$, and thus of $\mathbf{B}(0)=\mathbf{I}-\mathbf{L}^{-1}$. This algorithm has been exploited to demonstrate identifiability of the extended MVAR model in (Hyvarinen et al., 2008). More recently, the full algorithm has been proposed to estimate frequency domain causality, and used to assess patterns of directional cortical connectivity from multichannel EEG (Faes et al., 2010a).

5.2 Validation

Although model identification allows to find the estimates of coefficients and input covariances which easily lead to the desired frequency-domain measures of coupling and causality, validation steps need to be performed to guarantee a correct interpretation of the obtained results. In fact, proper checking has to be performed, in both time and frequency domains, in order to confirm the suitability of the estimated model for describing the observed data and to assess the significance of the estimated coupling strengths.

Model validation refers to the use of a range of diagnostic tools which are available for checking the adequacy of the estimated structure. While validation tests are rarely used in the practical application of tools like the DC or the PDC, they constitute actually fundamental safeguards against drawing erroneous inferences consequently to model misspecification. A striking example is the undervaluation of the impact of instantaneous effects in the assessment of frequency domain causality, which we demonstrated in this chapter and might be avoided by proper checking of the significance of instantaneous correlations in experimental multichannel data. While a thorough description of the statistical tools for model validation is beyond the scope of this work, here we remark the model assumptions that must be verified prior to frequency domain analysis. The identification of strictly causal MVAR models should result in temporally uncorrelated and mutually independent residuals $\hat{\mathbf{U}}(n)$. These assumptions may be checked, e.g., using the Ljung-Box portmanteau test for whiteness and the Kendall's τ test for independence. Independence of the residuals has to be checked particularly at zero lag; when this assumption is violated, one should allow instantaneous effects to be modeled moving from the strictly causal to the extended MVAR representation. If performed without incorporating prior assumptions in the model formulation, the extended MVAR identification should result also in non-Gaussian residuals $\hat{\mathbf{W}}(n)$; non-Gaussianity may be tested, e.g., by the Jarque-Bera test for non-normal distribution. All the diagnostic tests mentioned here are excellently reviewed in (Lutkepohl, 1993).

Another issue of great practical importance is the assessment of the significance of any frequency domain index of coupling or causal coupling. Due to practical estimation problems, nonzero values are indeed likely to occur at some frequencies even in the case of absence of a true interaction between the two considered processes. This problem is commonly faced by means of statistical hypothesis testing procedures based on setting a threshold for significance at the upper limit of the confidence interval of the considered index, where confidence intervals are based on the sampling distribution of the index computed under the null hypothesis of absence of interaction. Comparing at each specific frequency the estimated index with the threshold allows rejection or acceptance of the null hypothesis, and thus detection or denial of coupling or causality, according to the predetermined level of significance. The sampling distribution in the absence of coupling or

causality may be derived theoretically or empirically: theoretical approaches are elegant and computationally more efficient, empirical ones are more general and free of approximations. For a detailed description of the two approaches we refer to (Koopmans, 1974; Schelter et al., 2006; Eichler, 2006), where the statistical properties of Coh, DC/DTF, and PDC are studied and corresponding analytical threshold for significance are derived, and to (Faes et al., 2004; Faes et al., 2010b), where empirical significance levels are estimated for the same functions on the basis of surrogate data generation procedure devised specifically for each function.

6. Conclusion

In this chapter we have discussed the theoretical interpretation of the most common frequency-domain measures of coupling (Coh, PCoh) and causality (DC, PDC) in MV time series. We have shown that: (i) each of these measures reflects a specific time-domain definition of coupling or causality (see Table 1); (ii) while Coh and PCoh are symmetric measures, they can be decomposed into factors eliciting directionality, these factors being exactly the DC and the PDC; (iii) the squared modulus of the DC and the PDC measure respectively the fraction of power of the output signal which is received from the input signal, normalized to all incoming contributions, and of inverse power of the input signal which is sent to the output signal, normalized to all outgoing contributions. The picture emerging from these results suggests that both DC and PDC should be computed for the complete inference of frequency-domain causality in MV time series: DC measures causality in meaningful physical terms as power contributions, but cannot separate direct effects from indirect ones; PDC determines the correct interaction structure in terms of direct causal effects, but its absolute values lack of straightforward interpretability.

In addition, we emphasized the necessity extending the MVAR modeling approach traditionally used to assess frequency-domain causality whenever the time resolution of the measurements is lower than the time scale of the lagged influences occurring in the observed MV process. In such a case, the interpretation of the lagged effects may change considerably if instantaneous effects are not included in the model. We showed how the traditional DC and PDC computed from the strictly causal model may lead to misleading connectivity patterns, while the correct interpretation is obtained defining the two functions from the coefficients of the extended model. Moreover, we introduced novel frequency-domain measures of causality and direct causality (eDC and ePDC) which were shown to reflect extended causality definitions combining both lagged and instantaneous causality from one signal to another.

7. References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automation and Control*, Vol. 19, pp. 716-723

Akselrod, S.; Gordon, D.; Ubel, F. A.; Shannon, D. C.; Berger, A. C. & Cohen, R. J. (1981). Power spectrum analysis of heart rate fluctuation: a quantitative probe of beat-to-beat cardiovascular control. *Science*, Vol. 213, pp. 220-222

Albo, Z.; Di Prisco, G. V.; Chen, Y.; Rangarajan, G.; Truccolo, W.; Feng, J.; Vertes, R. P. & Ding, M. (2004). Is partial coherence a viable technique for identifying generators of neural oscillations? *Biological Cybernetics*, Vol. 90, No. 5, pp. 318-326

Astolfi, L.; Cincotti, F.; Mattia, D.; Marciani, M. G.; Baccala, L. A.; De Vico, F. F.; Salinari, S.; Ursino, M.; Zavaglia, M.; Ding, L.; Edgar, J. C.; Miller, G. A.; He, B. & Babiloni, F. (2007). Comparison of different cortical connectivity estimators for high-resolution EEG recordings. *Human Brain Mapping*, Vol. 28, No. 2, pp. 143-157

Baccala, L.; Sameshima, K. & Takahashi, D. Y. (2007). Generalized partial directed coherence, *Proceedings of the 2007 15th International Conference on Digital Signal Processing*, pp. 163-166, Cardiff, UK

Baccala, L. A. & Sameshima, K. (2001). Partial directed coherence: a new concept in neural structure determination. *Biological Cybernetics*, Vol. 84, No. 6, pp. 463-474

Baccala, L. A. & Sameshima, K. (2006). Comments on 'Is partial coherence a viable technique for identifying generators of neural oscillations?': Why the term 'Gersch Causality' is inappropriate: common neural structure inference pitfalls. *Biological Cybernetics*, Vol. 95, No. 2, pp. 135-141

Baccala, L. A.; Sameshima, K.; Ballester, G.; Valle, A. C. & Timo-Iaria, C. (1998). Studying the interaction between brain structures via directed coherence and Granger causality. *Applied Signal Processing*, Vol. 5, pp. 40-48

Bendat, J. S. & Piersol, A. G. (1986). *Random data: analysis and measurement procedures*, John Wiley & Sons, New York

Eichler, M. (2006). On the evaluation of information flow in multivariate systems by the directed transfer function. *Biological Cybernetics*, Vol. 94, No. 6, pp. 469-482

Erla, S.; Faes, L. & Nollo, G. (2009). Robust estimation of partial directed coherence by the vector optimal parameter search algorithm, *Proceeding of the 4-th IEEE-EMBS Conference on Neural Engineering*, pp. 6280-6283, Antalya, Turkey

Faes, L.; Erla, S.; Tranquillini, E.; Orrico, D. & Nollo, G. (2010a). An identifiable model to assess frequency-domain Granger causality in the presence of significant instantaneous interactions, *Proceedings of the 32th Annual International Conference of the IEEE-EMBS*, pp. 1699-1702, Buenos Aires, Argentina

Faes, L. & Nollo, G. (2010a). A method to assess frequency domain causality in the presence of instantaneous effects and its application to short term cardiovascular variability. *Methods of Information in Medicine*, Vol. 49, No. 5, doi: 10.3414/ME09-02-0030

Faes, L. & Nollo, G. (2010b). Extended causal modelling to assess partial directed coherence in multiple time series with significant instantaneous interactions. *Biological Cybernetics*, in press; doi: .1007/s00422-010-0406-6.

Faes, L.; Nollo, G. & Chon, K. H. (2008). Assessment of Granger causality by nonlinear model identification: application to short-term cardiovascular variability. *Annals of Biomedical Engineering*, Vol. 36, No. 3, pp. 381-395

Faes, L.; Pinna, G. D.; Porta, A.; Maestri, R. & Nollo, G. (2004). Surrogate data analysis for assessing the significance of the coherence function. *IEEE Transactions on Biomedical Engineering*, Vol. 51, No. 7, pp. 1156-1166

Faes, L.; Porta, A. & Nollo, G. (2010b). Testing Frequency Domain Causality in Multivariate Time Series. *IEEE Transactions on Biomedical Engineering*, Vol. 57, No. 8, pp. 1897-1906

Faes, L. & Nollo, G. (2010a). Assessing frequency domain causality in cardiovascular time series with instantaneous interactions. *Methods of Information in Medicine*, Vol. 49, pp. 453-457

Faes, L. & Nollo, G. (2010b). Extended causal modelling to assess partial directed coherence in multiple time series with significant instantaneous interactions. *Biological Cybernetics*, Vol. 103, No. 5, pp. 387-400

Gourevitch, B.; Bouquin-Jeannes, R. L. & Faucon, G. (2006). Linear and nonlinear causality between signals: methods, examples and neurophysiological applications. *Biological Cybernetics*, Vol. 95, No. 4, pp. 349-369

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, Vol. 37, pp. 424-438

Granger, C. W. J. (1980). Testing for causality: a personal viewpoint. *J. Econom. Dynam. Control*, Vol. 2, pp. 329-352

Hyvarinen, A.; Shimizu, S. & Hoyer, P. O. (2008). Causal modelling combining instantaneous and lagged effects: an identifiable model based on non-Gaussianity, *Proceedings of the 25th International Conference on Machine Learning*, pp. 1-8, Helsinki

Kaminski, M.; Ding, M.; Truccolo, W. A. & Bressler, S. L. (2001). Evaluating causal relations in neural systems: granger causality, directed transfer function and statistical assessment of significance. *Biological Cybernetics*, Vol. 85, No. 2, pp. 145-157

Kaminski, M. J. & Blinowska, K. J. (1991). A new method of the description of the information flow in the brain structures. *Biological Cybernetics*, Vol. 65, No. 3, pp. 203-210

Kay, S. M. (1988). *Modern spectral estimation. Theory & application*, Prentice Hall, Englewood Cliffs, New Jersey

Koopmans, L. H. (1974). *The spectral analysis of time series*, Academic Press, New York

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

Nunez, P. L. (1995). *Neocortical dynamics and human EEG rhythms*, Oxford University Press, USA

Pereda, E.; Quiroga, R. Q. & Bhattacharya, J. (2005). Nonlinear multivariate analysis of neurophysiological signals. *Progress in Neurobiology*, Vol. 77, No. 1-2, pp. 1-37

Porta, A.; Aletti, F.; Vallais, F. & Baselli, G. (2009). Multimodal signal processing for the analysis of cardiovascular variability. *Philosophical Transactions of the Royal Society of London Series A-Mathematical Physical and Engineering Sciences*, Vol. 367, No. 1887, pp. 391-409

Porta, A.; Furlan, R.; Rimoldi, O.; Pagani, M.; Malliani, A. & van de Borne, P. (2002). Quantifying the strength of the linear causal coupling in closed loop interacting cardiovascular variability signals. *Biological Cybernetics*, Vol. 86, No. 3, pp. 241-251

Schelter, B.; Winterhalder, M.; Eichler, M.; Peifer, M.; Hellwig, B.; Guschlbauer, B.; Lucking, C. H.; Dahlhaus, R. & Timmer, J. (2006). Testing for directed influences among neural signals using partial directed coherence. *Journal of Neuroscience Methods*, Vol. 152, No. 1-2, pp. 210-219

Schlogl, A. (2006). A comparison of multivariate autoregressive estimators. *Signal Processing*, Vol. 86, pp. 2426-2429

Schlogl, A. & Supp, G. (2006). Analyzing event-related EEG data with multivariate autoregressive parameters. *Progress in Brain Research*, Vol. 159, pp. 135-147

Shimizu, S.; Hoyer, P. O.; Hyvarinen, A. & Kerminen, A. (2006). A linear non-Gaussian acyclic model for causal discovery. *Journal of Machine Learning Research*, Vol. 7, pp. 2003-2030

Whittaker, J. (1990). *Graphical models in applied multivariate statistics*, Wiley, Chichester

Biomedical Engineering, Trends in Electronics, Communications and Software

Edited by Mr Anthony Laskovski

ISBN 978-953-307-475-7
Hard cover, 736 pages
Publisher InTech
Published online 08, January, 2011
Published in print edition January, 2011

Rapid technological developments in the last century have brought the field of biomedical engineering into a totally new realm. Breakthroughs in materials science, imaging, electronics and, more recently, the information age have improved our understanding of the human body. As a result, the field of biomedical engineering is thriving, with innovations that aim to improve the quality and reduce the cost of medical care. This book is the first in a series of three that will present recent trends in biomedical engineering, with a particular focus on applications in electronics and communications. More specifically: wireless monitoring, sensors, medical imaging and the management of medical information are covered, among other subjects.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Luca Faes and Giandomenico Nollo (2011). Multivariate Frequency Domain Analysis of Causal Interactions in Physiological Time Series, Biomedical Engineering, Trends in Electronics, Communications and Software, Mr Anthony Laskovski (Ed.), ISBN: 978-953-307-475-7, InTech, Available from: http://www.intechopen.com/books/biomedical-engineering-trends-in-electronics-communications-and-software/multivariate-frequency-domain-analysis-of-causal-interactions-in-physiological-time-series

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen