A nonlinear causality measure in the frequency domain: Nonlinear partial directed coherence with applications to EEG

He, F., Billings, S. A., Wei, H. L. & Sarrigiannis, P. G.

DOI 10.1016/j.jneumeth.2014.01.013
ISSN 0165-0270
ESSN 1872-678X

Publisher: Elsevier

A Nonlinear Causality Measure in the Frequency Domain: Nonlinear Partial Directed Coherence with Applications to EEG

Fei He[a], Stephen A. Billings[a], Hua-Liang Wei[a], and Ptolemaios G. Sarrigiannis[b]

[a]Department of Automatic Control and Systems Engineering, The University of Sheffield, Sheffield, S1 3JD, U.K.

[b]Department of Clinical Neurophysiology, Sheffield Teaching Hospitals NHS Foundation Trust, Royal Hallamshire Hospital, UK.

A Nonlinear Causality Measure in the Frequency Domain: Nonlinear Partial Directed Coherence with Applications to EEG

Abstract

Background: Frequency domain Granger causality measures have been proposed and widely applied in analyzing rhythmic neurophysiological and biomedical signals. Almost all these measures are based on linear time domain regression models, and therefore can only detect linear causal effects in the frequency domain.
New Method: A frequency domain causality measure, the partial directed coherence, is explicitly linked with the frequency response function concept of linear system. By modelling the nonlinear relationships between time series using nonlinear models and employing corresponding frequency-domain analysis techniques (i.e. generalized frequency response functions), a new nonlinear partial directed coherence method is derived.
Results: The advantages of the new method are illustrated via a numerical example of a nonlinear physical system and an application to electroencephalogram signals from a patient with childhood absence epilepsy.
Comparison with Existing Method(s): The new method detects both linear and nonlinear casual effects between bivariate signals in the frequency domain, while the existing measures can only detect linear effects.
Conclusions: The proposed new method has important advantage over the classical linear measures, because detecting nonlinear dependencies has become more and more important in characterizing functional couplings in neuronal and biological systems.

Highlights:

- A new nonlinear partial directed coherence measure is derived.
- Classical partial directed coherence is linked with the concepts in the linear systems theory.
- Detection of linear and nonlinear causalities in the epilepsy EEG data.

Keywords: Granger causality, coherence, nonlinear systems, spectral analysis, epilepsy

1. Introduction

The causal relationships and the direction of information flow within multivariate time series has been well studied under the Granger causality (Granger, 1969) framework and has important applications in neuroscience (Brovelli et al., 2004; Ding et al., 2006; Hesse et al., 2003; Michalareas et al., 2013), computational biology (Feng et al., 2009; Zou and Feng, 2009) and many others. Granger causality was originally developed in the context of linear autoregressive with exogenous input (ARX) models, and has been widely studied in the time domain. More recently, since the frequency decompositions are of particular interest in analyzing rhythmic neurophysiological and biomedical signals, such as electroencephalogram (EEG), magnetoencephalogram (MEG) recordings (Brovelli et al., 2004; Michalareas et al., 2013), and circadian rhythms in gene networks (Feng et al., 2009), several spectral measures related to Granger causality have been developed (for a review see (Chicharro, 2011)), including partial directed coherence (PDC) (Baccala and Sameshima, 2001), the directed transfer function (DTF) (Kaminski and Blinowska, 1991), and Geweke's spectral Granger causality (Geweke, 1982). Both PDC and spectral Granger causality are the frequency domain counterparts of the time domain Granger causality, and an important advantage of PDC is that it does not involve any matrix inversion and therefore it is computationally more efficient and more robust. All these spectral measures are based on time domain linear regression models; however, the interactions between real signals and processes are often nonlinear. Hence, the classical linear model based causality measures cannot truly reveal the nonlinear interactions within signals, especially in the frequency domain where linear methods would fail to detect nonlinear effects such as harmonics, inter-modulations and energy transfer (Billings and Tsang, 1989; Lang and Billings, 2005). Recently, the Granger causality definition has been extended to nonlinear cases in the time domain, based on nonparametric methods (Diks and Panchenko, 2006; Hiemstra and Jones, 1994) and parametric methods using for example radial basis functions (Ancona et al., 2004), kernel methods (Marinazzo et al., 2008), local linear systems (Chen et al., 2004), and NARX models (Li et al., 2012; Zhao et al., 2013). However, it has not yet been investigated in the frequency domain, and this is the topic to be discussed in the present study.

In this paper, the linear Granger causality is first reviewed in the time and frequency domain. The focus is on the formulation of frequency domain measures especially the PDC and its link with the frequency response functions (FRFs) of linear systems. Establishing such a link is important and can generalize the PDC to the nonlinear case. As an extension of the linear regression models used in classical Granger causality analysis, the nonlinear relationship between bivariate time series can be described by using NARX models. The identified model can then be mapped to the frequency domain and a "nonlinear FRF" can be formulated. Based on the frequency domain analysis of nonlinear systems, a nonlinear partial direct coherence (NPDC) method is developed. A numerical example of a nonlinear oscillator and an application to EEG signals are investigated and used to demonstrate that the new NPDC can reveal both linear and nonlinear causal influences in signals.

2. Linear Granger Causality and Partial Directed Coherence

The time domain linear regression model of bivariate time series and classical Granger causality is first reviewed. The frequency domain causality measures and the partial directed coherence are then introduced and the link with the frequency response function of linear systems is presented.

2.1 Linear Granger Causality of Bivariate Time Series

Consider two signals (or stochastic processes) X and Y with discrete time observations $x(t)$ and $y(t)$, t = 1, 2,…, N. Jointly, the interactions of these two signals can be described by bivariate ARX models

$$x(t)=\sum_{k=1}^{q}a_{11,k}x(t-k)+\sum_{k=1}^{p}a_{12,k}y(t-k)+e_x(t) \tag{1}$$

$$y(t)=\sum_{k=1}^{q}a_{21,k}x(t-k)+\sum_{k=1}^{p}a_{22,k}y(t-k)+e_y(t) \tag{2}$$

where p and q are the model order of y and x regressors; $e_x(t)$ and $e_y(t)$ are the model prediction errors, and each is assumed serially uncorrelated over time. If the prediction error for a signal depends only on the past of its own signal, this corresponds to a univariate autoregressive (AR) model with purely terms in $x(t)$ or $y(t)$ and a special case of (1) or (2). A linear causal influence from X to Y defined by Granger can be expressed as the log ratio of the prediction error variances of the corresponding restricted (AR) and unrestricted (ARX) models

$$F_{X\to Y} = \ln\frac{\mathrm{var}(y\,|\,y^{-})}{\mathrm{var}(y\,|\,y^{-},x^{-})} = \ln\frac{\Sigma_{22}^{y^{-}}}{\Sigma_{22}} \tag{3}$$

where x^{-} and y^{-} denotes contributions from lagged input and output terms respectively; $\Sigma_{22}^{y^{-}}$ denotes the variance of e_y when there are only regression terms of Y in the second equation in (1), which relates to a univariate AR model of $y(t)$. If the signal X has casual influence on the signal Y in the Granger sense, the variance of the prediction error $\mathrm{var}(y\,|\,y^{-},x^{-})$ must be smaller than $\mathrm{var}(y\,|\,y^{-})$. The linear Granger causality $F_{X\to Y}$ is then a positive value. Similarly, the causal influence from Y to X can be defined as $F_{Y\to X} = \ln\Sigma_{11}^{x^{-}}/\Sigma_{11}$, with $\Sigma_{11}^{x^{-}}$ representing the variance of a AR model of $x(t)$ (i.e., $a_{12,k}=0$).

2.2 Causality in the Frequency Domain: Partial Directed Coherence

The linear ARX models (1) and (2) can be rewritten in matrix form and mapped to the frequency domain by Fourier transformation

$$\begin{pmatrix} A_{11}(f) & A_{12}(f) \\ A_{21}(f) & A_{22}(f) \end{pmatrix}\begin{pmatrix} X(f) \\ Y(f) \end{pmatrix} = \begin{pmatrix} E_x(f) \\ E_y(f) \end{pmatrix} \tag{4}$$

where the components of the coefficient matrix $\mathbf{A}(f)$ are $A_{lm}(f) = \delta_{lm} - \sum_{k=1}^{p(or\ q)} a_{lm,k} e^{-j2\pi k f/f_s}$ where f_s is the sampling frequency; δ_{lm} is the Kronecker delta function (i.e., $\delta_{lm}=1$ if $l=m$ and $\delta_{lm}=0$ if $l\neq m$). Equation (4) can be recast into the transfer function format

$$\begin{pmatrix} X(f) \\ Y(f) \end{pmatrix} = \begin{pmatrix} G_{xx}(f) & G_{xy}(f) \\ G_{yx}(f) & G_{yy}(f) \end{pmatrix}\begin{pmatrix} E_x(f) \\ E_y(f) \end{pmatrix} \tag{5}$$

where $\mathbf{G}(f)=\mathbf{A}^{-1}(f)$; $G_{xx}(f)=A_{22}(f)/\det(\mathbf{A})$, $G_{yy}(f)=A_{11}(f)/\det(\mathbf{A})$, $G_{xy}(f)=-A_{12}(f)/\det(\mathbf{A})$, and $G_{yx}(f)=-A_{21}(f)/\det(\mathbf{A})$.

Different frequency domain Granger causality measures, such as PDC, DTF and spectral Granger causality, can then be expressed as a function of the elements of either the coefficient matrix $\mathbf{A}(f)$ or the transfer function matrix **G**(*f*) (Baccala and Sameshima, 2001; Chicharro, 2011). For example, the "partial directed coherence" (PDC) from *X* to *Y* is defined as

$$PDC_{X\to Y}(f)=\frac{A_{21}(f)}{\sqrt{|A_{11}(f)|^2+|A_{21}(f)|^2}} \tag{6}$$

The $PDC_{X\to Y}$ describes the relative coupling strength of the interaction from a signal source (i.e. *X*) to some signal (i.e. *Y*), as compared (or normalized) to all the connections of the source to other signals. For a bivariate system, the directed interaction is described by A_{21}, and it is normalized by all the *X* related terms in the ARX models (4) (i.e. A_{21} and A_{21}). The magnitude of PDC then lies between zero and one. A very similar causality measure, the DTF, is defined from the elements of the transfer function matrix

$$DTF_{X\to Y}(f)=\frac{G_{yx}(f)}{\sqrt{|G_{yx}(f)|^2+|G_{yy}(f)|^2}} \tag{7}$$

The equivalence between PDC and DTF for bivariate analysis is provided in (Baccala and Sameshima, 2001). In this paper, the PDC is mainly investigated because it does not involve matrix inversion and is computationally more efficient and robust.

2.3 Links with the Frequency Response Functions of Linear Systems

By dividing both sides of (4) with the corresponding diagonal elements in the coefficient matrix **A**, it gives

$$\begin{pmatrix} 1 & \dfrac{A_{12}(f)}{A_{11}(f)} \\ \dfrac{A_{21}(f)}{A_{22}(f)} & 1 \end{pmatrix}\begin{pmatrix} X(f) \\ Y(f) \end{pmatrix}=\begin{pmatrix} \dfrac{E_x(f)}{A_{11}(f)} \\ \dfrac{E_y(f)}{A_{22}(f)} \end{pmatrix} \tag{8}$$

The off-diagonal elements in the transformed coefficient matrix are actually the negative frequency response functions (FRFs) of the linear (ARX) systems, if one process is treated as the input while the other is treated as the output. For example,

$$\frac{A_{21}(f)}{A_{22}(f)}=-\frac{\sum_{i=1}^{q}a_{21,i}e^{-j2\pi if/f_s}}{1-\sum_{i=1}^{p}a_{22,i}e^{-j2\pi if/f_s}}=-H_{x\to y}(f) \tag{9}$$

the FRF, $H_{x \to y}(f)$, describes the input-output relationship (i.e., with input *X* and output *Y*) of the (noise-free) "system" in the frequency domain. It is also known as the "transfer function" in the linear system theory (different from the transfer function **G** in (5)). The right-hand-side elements in (8) correspond to the spectra of the AR process of *X* and *Y*. For example,

$$\frac{E_y(f)}{A_{22}(f)} = \frac{E_y(f)}{1-\sum_{i=1}^{p} a_{22,i} e^{-j2\pi if/f_s}} = Y_{y|y^-}(f) \tag{10}$$

the spectrum of an AR process of *Y* can be regarded as a product of an error-driven FRF (i.e. $1/A_{22}(f)$) and the error spectrum $E_y(f)$. Practically $E_y(f)$ can be computed from the Fourier transform of the error sequence e_y, and for a zero mean noise or error sequence its power equals to its variance. As a result, (8) can be re-written as

$$\begin{pmatrix} 1 & -H_{y \to x}(f) \\ -H_{x \to y}(f) & 1 \end{pmatrix} \begin{pmatrix} X(f) \\ Y(f) \end{pmatrix} = \begin{pmatrix} X_{x|x^-}(f) \\ Y_{y|y^-}(f) \end{pmatrix} \tag{11}$$

This indicates that the overall spectrum of a signal in the bivariate processes often includes two parts

$$\begin{aligned} X(f) &= X_{x|x^-}(f) + H_{y \to x}(f) \cdot Y(f) = X_{x|x^-}(f) + X_{x|y^-}(f) \\ Y(f) &= Y_{y|y^-}(f) + H_{x \to y}(f) \cdot X(f) = Y_{y|y^-}(f) + Y_{y|x^-}(f) \end{aligned} \tag{12}$$

Take the expression of $Y(f)$ for example, one is an "intrinsic" part $Y_{y|y^-}(f)$ that only relates to the power from its own process, the other is a "causal" part $Y_{y|x^-}(f)$ that is a product of the FRF $H_{x \to y}(f)$ and the input spectrum $X(f)$. It is important to note that linear output terms y^- have a dual-effect with respect to $Y(f)$. That is an intrinsic effect driven only by the model prediction error e_y hence strictly $Y_{y|y^-}(f) = Y_{y|y^-_{(e)}}(f)$; and a causal effect $Y_{y|y^-_{(x)}}(f)$ that is actually integrated into the FRF $H_{x \to y}(f)$ of the ARX model, because $H_{x \to y}(f)$ also depends on parameters of the linear output terms $a_{22,i}$.

The PDC defined in (6) can then be re-expressed as a function of the FRFs of the corresponding linear ARX and AR models

$$PDC_{X\to Y}(f)=\frac{-H_{x\to y}(f)}{\sqrt{\left|A_{11}(f)/A_{22}(f)\right|^2+\left|H_{x\to y}(f)\right|^2}} \tag{13}$$

Such a linear PDC expression will be generalized to the nonlinear case, by replacing $H_{x\to y}(f)$ with a nonlinear version FRF that contains both linear and nonlinear causal effects, and generalizing the FRFs of AR models (i.e. $1/A_{22}(f)$ and $1/A_{11}(f)$) to nonlinear autoregressive (NAR) cases. These can be achieved by employing nonlinear time domain models and corresponding frequency domain analysis methods.

3. Frequency Domain Formulation of Nonlinear Causality

To develop a nonlinear Granger causality measure in the frequency domain, the nonlinear relationships in bivariate time series need to be first identified and modeled in the time domain by using NARX models as introduced in Section 3.1. The time domain model can then be mapped to the frequency domain and a "nonlinear FRF" is introduced in Section 3.2. Finally, a new nonlinear causality measure, the nonlinear partial directed coherence, is proposed in Section 3.3.

3.1 Nonlinear Time Series Analysis using NARX models

Linear Granger causality discussed in Section 2 was formulated based on linear AR or ARX models. These linear models and linear causality measures however cannot reveal any nonlinear effects that can be important in analyzing neurophysiological and biological signals. Especially in the frequency domain linear methods can never represent harmonics, inter-modulations and complex energy transfer effects that characterize nonlinear systems and which are features of many real life signals including EEG recordings.. To develop nonlinear causality measures, nonlinear models need to be used to identify the nonlinear relationships. The well-known NARX model (Billings, 2013; Leontaritis and Billings, 1985) is a natural extension of the ARX model and can describe a wide range of nonlinear dynamic systems. As an extension to the linear ARX expression (2), the polynomial NARX model with respect to Y can be expressed as

$$y(t)=\sum_{n=1}^{M}\sum_{p=0}^{n}\sum_{k_1,k_{p+q}=1}^{K}c_{p,q}(k_1,\dots,k_{p+q})\times\prod_{i=1}^{p}y(t-k_i)\prod_{i=p+1}^{p+q}x(t-k_i)+e_y(t) \tag{14}$$

where n denotes the nth-order nonlinearity of the system with a maximum order of M; and $p+q=n$, $k_i=1,\dots,K$, $\sum_{k_1,k_{p+q}=1}^{K}\equiv\sum_{k_1=1}^{K}\cdots\sum_{k_{p+q}=1}^{K}$. The number of model terms increases as the order of input and output terms (q and p) and the corresponding maximum lags (K) increase. The NARX models can typically be identified based on the forward regression with orthogonal least squares (FROLS) method (Billings et al., 1989; Chen et al., 1989). By using the FROLS algorithm, a "best" model structure is selected and the model complexity is controlled to avoid over fitting, and the model parameters are estimated. In addition, if the system under study is linear, the FROLS method automatically discards the nonlinear terms and only estimates a linear model. In cases where the system under study is stochastic with unknown colored noise, noise models should be employed to form a NARMAX model (Billings, 2013; Billings et al., 1989).

The identified model can be statistically validated using the correlation tests (Billings and Voon, 1983). Five cross-correlation functions, i.e. $\phi_{\xi\xi}(\tau)$, $\phi_{x\xi}(\tau)$, $\phi_{\xi(\xi x)}(\tau)$, $\phi_{(x^2)'\xi}$, and $\phi_{(x^2)'\xi^2}$ (with model residual ξ and $(x^2)'$ a zero-mean process of x^2), are used in conjunction with 95% confidence intervals (approximately $\pm1.96/\sqrt{N}$) to test whether the residuals are uncorrelated with all linear and nonlinear combinations of past inputs, outputs and residuals. This validation routine is applied to examples in this paper.

3.2 Frequency Domain Mapping of NARX models

As presented in Section 2.3, the output spectrum of a linear (noise-free) system $Y_{y|x^-}(f)$ is equal to the product of the input spectrum with the system FRF, i.e., $Y_{y|x^-}(f)=H_{x\to y}(f)X(f)$. The frequency domain analysis of a nonlinear system is much more complicated, and is mainly based on the concept of generalized frequency response functions (GFRFs) that extend the linear FRF to higher orders and dimensions. Following the derivations in (Billings, 2013; Lang and Billings, 1996) the output spectrum of a nonlinear system can be formulated using the output frequency response function (OFRF)

$$Y_{y|x^-}(f)=\sum_{n=1}^{M}\left(\frac{1}{\sqrt{n}}\int_{f_1+\cdots+f_n=f}H_n(f_1,\ldots,f_n)\prod_{i=1}^{n}X(f_i)df\right) \tag{15}$$

Here, the nth-order GFRF $H_n(f_1,\ldots,f_n)$ was defined as the multiple Fourier transform of the nth-order Volterra kernel, and later extended to the NARX model cases (Billings and Tsang, 1989; Jones and Billings, 1989). For a NARX model, the corresponding nth-order GFRF can be expressed in terms of the model parameters from (14) as

$$\begin{aligned}&H_n(f_1,\ldots,f_n)=\\&\frac{H_{n[x]}(f_1,\ldots,f_n)+H_{n[y]}(f_1,\ldots,f_n)+H_{n[xy]}(f_1,\ldots,f_n)}{1-\sum_{k_1=1}^{K}c_{1,0}(k_1)e^{(-j2\pi(f_1+\ldots+f_n)k_1/f_s)}}\end{aligned} \tag{16}$$

where the contributions of the pure input, output and cross-product nonlinearities, $H_{n[x]}$, $H_{n[y]}$, and $H_{n[xy]}$, are defined in the Appendix. In order to form a unified linear frequency domain expression similar to the linear system, the OFRF (15) can be reformulated into a linear-in-the-input-spectrum form

$$Y_{y|x^-}(f)=\left(H_1(f)+\sum_{n=2}^{M}\bar{H}_n(f)\right)X(f)=H_{x\to y}(f)X(f) \tag{17}$$

where

$$\bar{H}_n(f)=\frac{1}{\sqrt{n}}\left(\int_{f_1+\cdots+f_n=f}H_n(f_1,\ldots,f_n)\prod_{i=1}^{n}X(f_i)df\right)\Big/X(f) \tag{18}$$

$\bar{H}_n(f)$ can be named the nth-order averaged GFRF (AGFRF). Now $H_{x\to y}(f)$ is an overall "nonlinear FRF" (or "nonlinear transfer function") describing both the linear and nonlinear effects from X to Y. The linear FRF (9) then becomes a special case, i.e., $H_1(f)$, and corresponds to all the linear terms in a NARX model. It is worth noting that the nonlinear FRF depends on both the system properties and the input spectrum, this is a distinct difference from the linear system (Lang and Billings, 2005).

In this section, the "causal" part $Y_{y|x^-}(f)$ and FRF $H_{x\to y}(f)$ has been generalized to include both linear and nonlinear effects. This is based on an NARX model (14) but without considering the prediction error or noise (Jones and Billings, 1989). To extend the frequency domain causality measure and PDC to the nonlinear systems, the error-driven "intrinsic" part needs to be further considered and will be discussed in Section 3.3.

3.3 Nonlinear Partial Directed Coherence

As a generalization of the spectrum decomposition in the linear bivariate case (12), the overall spectrum of *y* with respect to a nonlinear NARX model (14) can still be expressed as a sum of an "intrinsic" power and a "causal" power

$$\begin{aligned} Y(f) &= Y_{y|y^-}(f) + Y_{y|x^-}(f) \\ &= Y_{y|y^-_{l(e)},y^-_{n(e)}}(f) + H_{x\to y}(f)X(f) \end{aligned} \tag{19}$$

Here the "causal" part $Y_{y|x^-}(f)$ has been analyzed and corresponds to the OFRF (17) of a noise-free NARX model. The "intrinsic" power $Y_{y|y^-}(f)$ denotes the output spectrum of a restricted NAR model, and only relates to the univariate autoregressive terms and model prediction error e_y. Importantly, the dual effects now exist in both the linear and nonlinear output terms y_l^- and y_n^-. Explicitly, the contributions to the "intrinsic" power are denoted as $y^-_{l(e)}$ and $y^-_{n(e)}$, while the contributions to the "causal" power, i.e., $y^-_{l(x)}$ and $y^-_{n(x)}$ are implicitly included in the GFRFs (16) and the nonlinear FRF. Now the "intrinsic" power can be explicitly expressed as

$$\begin{aligned} Y_{y|y^-_{l(e)},y^-_{n(e)}}(f) &= Y_{y|y^-_{l(e)}}(f) + Y_{y|y^-_{n(e)}}(f) \\ &= H_1^e(f)E(f) + \left(\sum_{n=2}^{M} \left(\frac{1}{\sqrt{n}} \int_{f_1+\cdots+f_n=f} H^e_{n[y]}(f_1,\dots f_n) \prod_{i=1}^{n} E(f_i) df \right) \Big/ E(f) \right) E(f) \\ &= \left(H_1^e(f) + \sum_{n=2}^{N} \bar{H}_n^e(f) \right) E(f) \\ &= H_y^e(f)E(f) \end{aligned} \tag{20}$$

where $E(f)$=$E_y(f)$ denotes the spectrum of the NARX model prediction error e_y in (14). The first-order error-driven GFRF is defined as $H_1^e(f)=1/(1-\sum_{k_1=1}^{K} c_{1,0}(k_1)e^{-j2\pi k_1 f/f_s})$, and the nth-order error-driven GFRFs $H_{n[y]}^e(f_1,\ldots,f_n)$ can be iteratively computed according to (28)-(30) by replacing $H_1(f)$ with $H_1^e(f)$. The AR model spectrum expression (10) becomes a special case of (20) with $1/A_{22}(f)=H_1^e(f)$. Similarly, by treating signal Y as input, X can also be modeled by a NARX model and its spectrum $X(f)$ can be expressed in terms of $Y(f)$ and $E_x(f)$. As a result, the overall frequency domain (linear and nonlinear) relationship of the two processes can be expressed in matrix form in terms of the nonlinear FRFs of the restricted NAR and unrestricted NARX models

$$\begin{pmatrix} 1 & -H_{y\to x} \\ -H_{x\to y} & 1 \end{pmatrix}\begin{pmatrix} X(f) \\ Y(f) \end{pmatrix}=\begin{pmatrix} X_{x|\bar{x}_{l(e)},\bar{x}_{n(e)}}(f) \\ Y_{y|\bar{y}_{l(e)},\bar{y}_{n(e)}}(f) \end{pmatrix}=\begin{pmatrix} H_x^e(f)E_x(f) \\ H_y^e(f)E_y(f) \end{pmatrix} \tag{21}$$

Here, $H_{x\to y}(f)$ and $H_y^e(f)$ are defined in (17) and (20) respectively. $H_{y\to x}(f)$ and $H_x^e(f)$ can be defined in a similar way by exchanging the input and output relationship. By dividing both sides of (21) with corresponding FRFs of the NAR models, i.e., $H_x^e(f)$ and $H_y^e(f)$, gives

$$\begin{pmatrix} 1/H_x^e(f) & -H_{y\to x}/H_x^e(f) \\ -H_{x\to y}/H_y^e(f) & 1/H_y^e(f) \end{pmatrix}\begin{pmatrix} X(f) \\ Y(f) \end{pmatrix}=\begin{pmatrix} E_x(f) \\ E_y(f) \end{pmatrix} \tag{22}$$

Equation (22) is a nonlinear version of the linear bivariate expression (4), and now the coefficient matrix includes both the linear and nonlinear causal information. Based on such a linear matrix representation of the nonlinear bivariate system, the nonlinear partial directed coherence (NPDC) from X to Y can then be expressed as a direct generalization of the linear PDC,

$$NPDC_{X\to Y}(f)=\frac{-H_{x\to y}/H_y^e(f)}{\sqrt{\left|1/H_x^e(f)\right|^2+\left|H_{x\to y}/H_y^e(f)\right|^2}}=\frac{-H_{x\to y}(f)}{\sqrt{\left|H_y^e(f)/H_x^e(f)\right|^2+\left|H_{x\to y}(f)\right|^2}} \tag{23}$$

The NPDC measures both linear and nonlinear causal influences from *X* to *Y*. The linear causal effects in the nonlinear system can also be expressed as a separate measure which is a function of 1st-order nonlinear FRFs of the NARX (i.e., $H_{1,x\to y}(f)$) and NAR (i.e., $H_{1,x}^e(f)$ and $H_{1,x}^e(f)$) models

$$NPDC_{X\to Y}^L(f)=\frac{-H_{1,x\to y}(f)}{\sqrt{\left|H_{1,y}^e(f)/H_{1,x}^e(f)\right|^2+\left|H_{1,x\to y}(f)\right|^2}} \tag{24}$$

Mathematically, this is the same expression with the linear PDC (13) although it is based on the linear part of a nonlinear system. In the frequency domain the measure (24) often provides similar causality profiles with that of PDC based on the linear ARX model, because for linear systems the frequency components in a causality measure would only depend on the fundamental frequencies in the input process, which is quantified either by $H_{1,x}^e(f)$ in (24) or $A_{11}(f)$ in the PDC expression (13).

4. Simulation and Application

In this section, a numerical example of a physical nonlinear oscillator and a real neurophysiological data set are given as examples to illustrate the new frequency-domain NPDC method based on time-domain NARX modelling.

4.1 A Numerical Example: Nonlinear Oscillator

The well-known Duffing oscillator with cubic nonlinearity in the stiffness is described by

$$m\ddot{y}(t)+c\dot{y}(t)+k_1y(t)+k_3y^3(t)=x(t) \tag{25}$$

where m=1, c=0.5, k_1=1 and k_3=0.3 are the parameters of the mass, damping, linear and nonlinear stiffness respectively. The input signal $x(t)$=5sin(t) represents an external excitation force and here it is a harmonic excitation. By using an explicit Euler approximation for discretization and considering noise, gives

$$\begin{cases} y(k) = (2-0.5T_s)y(k-1)-(1-0.5T_s+T_s^2)y(k-2) \\ \qquad -k_3T_s^2y^3(k-2)+T_s^2x(k-2)+e_y(k) \\ x(k)=\sum_{m=1}^{K}a_m x(k-m)+e_x(k) \end{cases} \tag{26}$$

where $e_y(k)$ and $e_x(k)$ were additive white noise sequences with zero mean and standard deviations σ_x=σ_y=0.1. The input and output were generated with a sampling time T_s=0.02 s. The sinusoidal input signal can be modeled by a 15$^{\text{th}}$-order AR model. The NARX model structure was then identified and the parameters were estimated by using a FROLS algorithm (Chen et al., 1989; Wei et al., 2004). The model is validated using the correlation tests (Billings and Voon, 1983). Based on the identified time domain model, the first, second and third order GFRFs can be computed according to (16)

$$\begin{aligned} &H_1(f)\left(1-\sum_{k_1=1}^{K}c_{1,0}(k_1)e^{-j2\pi k_1 f/f_s}\right)=\sum_{k_1=1}^{K}c_{0,1}(k_1)e^{-j2\pi k_1 f/f_s} \\ &H_2(f_1,f_2)=0 \\ &H_3(f_1,f_2,f_3)=-k_3H_1(f_1)H_1(f_2)H_1(f_3)H_1(f_1,f_2,f_3) \end{aligned} \tag{27}$$

where $c_{1,0}(1)$, $c_{1,0}(2)$, and $c_{0,1}(2)$ were the parameters of corresponding linear input and output terms given in (26). The second-order GFRF is zero because there is no second order nonlinear term in the NARX model. Although higher order GFRFs exist in this example due to the nonlinear output term, for illustration the GFRFs only up to third order were computed here. The error-driven GFRFs of the NAR model, i.e., $H_1^e(f)$ and $H_3^e(f_1,f_2,f_3)$, were also computed. As a result, the nonlinear FRFs of the identified NARX and NAR models, i.e., $H_{x\to y}(f)$ and $H_y^e(f)$, can easily be obtained. The spectra of input and output signals using a fast Fourier transform (FFT) are given in Fig. 1. The PDC based on the linear part of the NARX model and the NPDC are computed and given in Fig. 2.

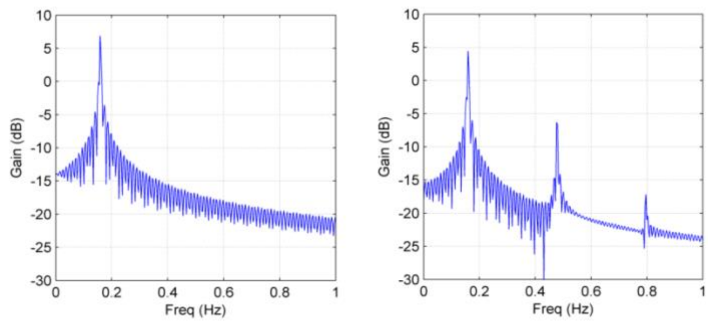

Fig. 1. The spectra of time series *X* and *Y*.

Fig. 1 shows that the input spectrum $X(f)$ only contains a single frequency component at $f_0=1/2\pi$ Hz. In the output spectrum $Y(f)$, harmonic components at $3f_0$, and $6f_0$ can be observed due to the presence of a cubic output nonlinearity in the system.

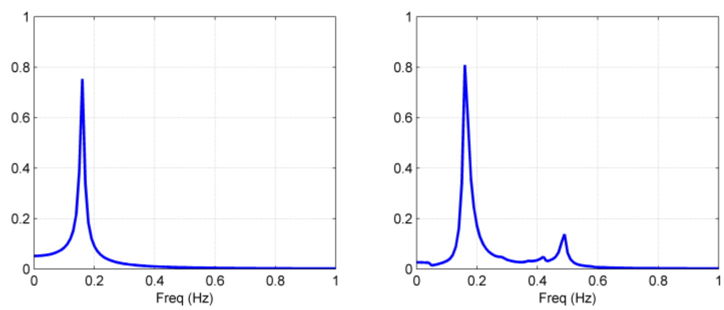

Fig. 2 The (left) PDC computed from the linear part of NARX model and (right) NPDC from *X* to *Y*

Fig. 2 indicates that the PDC based on the linear part of the NARX model identifies a strong linear causal influence from input excitation to the output of the oscillatory system at the fundamental frequency f_0. The same result can be obtained from a standard PDC based on a linear ARX model. However, these PDC measures cannot reveal the nonlinear causal effects in the frequency domain. In comparison, the proposed NPDC detects both the linear and nonlinear causal effects at fundamental and harmonic frequencies f_0 and $3f_0$. The nonlinear causal effects at other harmonic frequencies, e.g. $6f_0$, are not observed because only up to third order GFRFs were computed here. The nonlinear causalities are much lower in magnitudes compared to the linear causality at f_0 which indicates weak nonlinearity of the system and the linear effect dominates.

4.2 Application to EEG Data

The nonlinear dynamics and causalities in neurophysiological signals (e.g. EEG, MEG) have been observed and studied in the time domain (Gourevitch et al., 2006; Stam, 2005). In this section, a fragment of EEG recordings from a patient with childhood absence epilepsy is analyzed in the frequency domain. The scalp EEG signals were recorded in the department of Clinical Neurophysiology, Royal Hallamshire Hospital in Sheffield. For illustration, EEG signals from two bipolar channels, i.e. F7F3 and F8F4, located in the left and right brain were used and a recording with 20 seconds in length is shown in Fig. 3. The signals were measured at 250 Hz and normalized by subtracting their mean values.

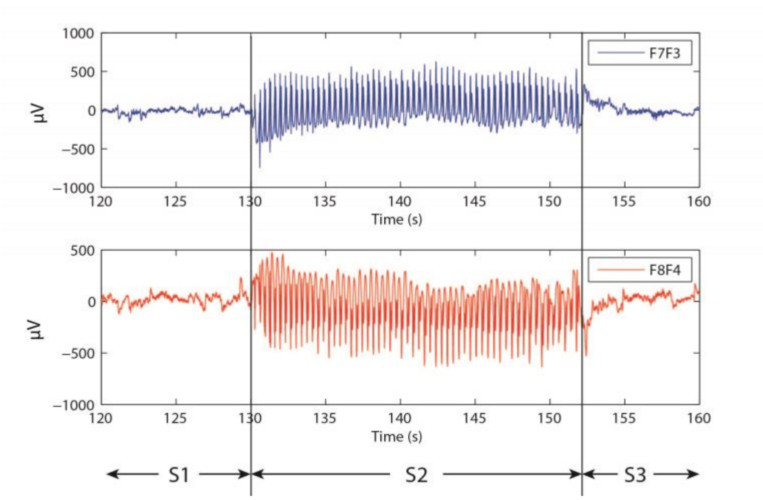

Fig. 3. The EEG recordings of two bipolar channels F7F3 and F8F4. The recordings are divided into three intervals before (S1), during (S2) and after (S3) an epileptic seizure, respectively.

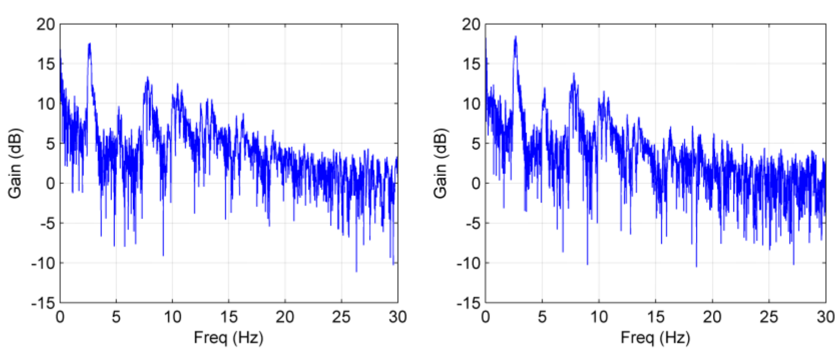

Fig. 4. Spectra of (left) input and (right) output EEG recordings.

Sustained oscillations with high amplitude are observed during the epileptic seizure (i.e. from around 130 seconds to 152 seconds); while before and after the seizure the EEG signals are with low amplitude and are relatively random. Hence the EEG recordings can naturally be divided into three intervals before, during and after the seizure (i.e. S1, S2 and S3 in Fig. 3). The spectra of the two signals over the whole sequence are provided in Fig. 4. For illustration, only the low-frequency range, i.e., 0-30 Hz, is analyzed. Significant frequency components at around 3, 5, 8, 11 and 13 Hz can be observed in both two signals. These frequency components are mainly introduced during the seizure due to sustained oscillations and the identical frequencies in both channels indicate strong linear effects because of the synchronization (Montez et al., 2006). The frequency components introduced before and after the seizure are much lower in magnitude and are actually submerged in the FFT analysis.

Now the linear and nonlinear interactions between the two EEG signals at two time intervals, i.e., during (S2) and before (S1) the seizure, are modeled individually using NARX models described in (14), where the signal from F7F3 is treated as the input X and the signal from F8F4 is the output Y. The NARX model is with second-order nonlinearity (M=2) and initially maximum lags of six (K=6) are set for both input and output terms. The identified time domain NARX models using FROLS are then mapped to the frequency domain. The linear and nonlinear causal influences from the input to the output can therefore be quantified using the linear PDC measure (24) and the overall NPDC measure (23) as shown in Fig. 5 and Fig. 6. Here, the input spectrum $X(f)$ can be approximated from the frequency mapping of a 25th-order AR model.

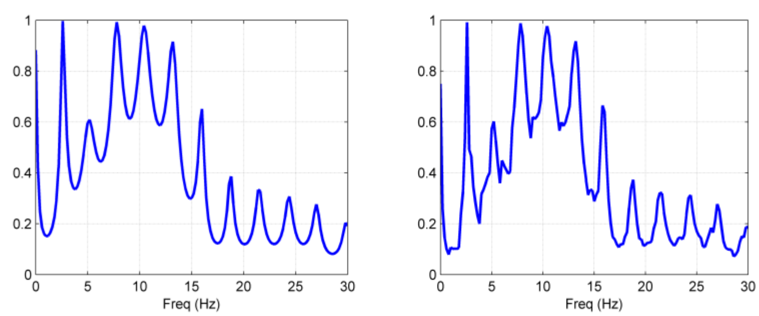

Fig. 5. The (left) linear PDC and the (right) overall NPDC from F7F3 to F8F4 with respect to the recordings during seizure (S2)

Fig. 5 shows that during the seizure the linear and the overall causality measures present very similar coherence profiles. Strong coupling strength with magnitude close to 1 are observed at around 3, 8, 11, 13 Hz, which are the frequencies identified in the spectra of the two signals. This indicates strong linear and weak nonlinear causal effects during the seizure which has been reported and understood in the time domain analysis, because of the strong synchronization during epileptic seizure (Montez et al., 2006).

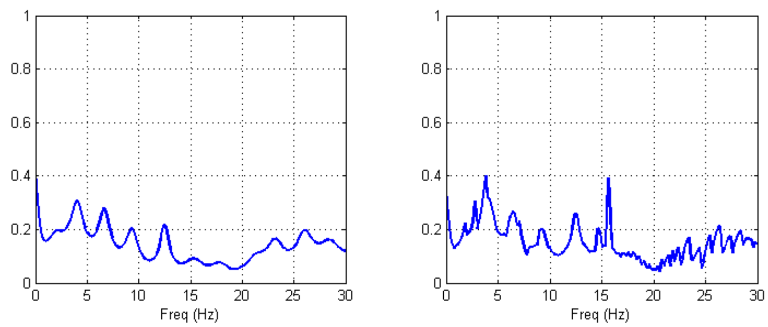

Fig. 6. The (left) linear PDC and the (right) overall NPDC from F7F3 to F8F4 with respect to the recordings before seizure (S1)

In contrast, before seizure the overall NPDC identifies extra frequency coupling components compared with the linear PDC. For example, those at around 3, 15 Hz and more components identified in high the frequency range between 20-30 Hz. This indicates that significant nonlinear causal effects exist before seizure, although the overall causality is much lower (below 0.4) in strength compared to those results during the seizure. The strong linear causality during the seizure and significant nonlinear causality in the seizure-free data confirms recent time domain causality analysis results (Zhao et al., 2013) but provides a new interpretation based on the observed frequency componenets. Similar results and conclusions can also be obtained when analyzing EEG signals from other bipolar channels or data from other patients.

5. Conclusions

In this study, the classical linear frequency-domain Granger causality measure, the PDC, is linked and interpreted by the input- and error-driven frequency response functions of linear systems. Such a new interpretation is extended to the nonlinear system cases, and a new nonlinear causality measure, the NPDC, is derived. The numerical example and an application to EEG signals with epilepsy seizure demonstrate that the proposed NPDC measure can effectively detect both linear and nonlinear causal influences and provides a more complete causality profile in the frequency domain. This is an important advantage over the classical linear measures which detect linear causalities exclusively, because detecting nonlinear dependencies has become more and more important in characterizing functional couplings in neuronal and biological systems. Due to the space limitation, only bivariate and stationary systems are investigated in this paper. Nevertheless, the methodology and the NPDC measure can be extended to multivariate and nonstationary cases, based on the frequency domain theory of multivariate nonlinear systems (Swain and Billings, 2001) and recent results on time-varying nonlinear systems (He et al., 2013).

Appendix

The contributions of the pure input, output and cross-product nonlinearities in (16) are defined as

$$
\begin{aligned}
H_{n[x]}(f_1,\dots,f_n) &= \sum_{k_1,k_n=1}^{K} c_{0,n}(k_1,\dots,k_n)e^{-j2\pi(f_1k_1+\dots+f_nk_n)/f_s} \\
H_{n[y]}(f_1,\dots,f_n) &= \sum_{p=2}^{n}\sum_{k_1,k_n=1}^{K} c_{0,n}(k_1,\dots,k_n)H_{n,p}(f_1,\dots,f_n) \\
H_{n[xy]}(f_1,\dots,f_n) &= \sum_{q=1}^{n-1}\sum_{p=1}^{n-q}\sum_{k_1,k_{p+q}=1}^{K} c_{p,q}(k_1,\dots,k_{p+q}) \\
&\quad \times H_{n-q,p}(f_1,\dots,f_{n-q})e^{-j2\pi(f_{n-q+1}k_{n-q+1}+\dots+f_{p+q}k_{p+q})/f_s}
\end{aligned}
\tag{28}
$$

The contribution of the pth-order non-linearity in $y(t)$ to the nth-order GFRF, $H_{n,p}(\cdot)$, can be recursively computed according to (Jones and Billings, 1989) as

$$
H_{n,p}(\cdot) = \sum_{i=1}^{n-p+1} H_i(f_1,\dots f_i,t)H_{n-i,p-1}(f_{i+1},\dots,f_n,t)e^{-j2\pi(f_1+\dots+f_i)k_p/f_s} \tag{29}
$$

The above recursion finishes with p=1, where the $H_{n,1}(f_1,\ldots,f_n,t)$ is defined as

$$H_{n,1}(f_1,\ldots,f_n,t) = H_n(f_1,\ldots,f_n,t)e^{-j2\pi(f_1+\ldots+f_i)k_1/f_s} \tag{30}$$

References

Ancona N, Marinazzo D, Stramaglia S. Radial basis function approach to nonlinear Granger causality of time series. Phys Rev E, 2004; 70.
Baccala LA, Sameshima K. Partial directed coherence: a new concept in neural structure determination. Biol Cybern, 2001; 84: 463-74.
Billings SA. Nonlinear system identification : NARMAX methods in the time, frequency, and spatio-temporal domains. John Wiley & Sons, 2013.
Billings SA, Chen S, Korenberg MJ. Identification of Mimo Non-Linear Systems Using a Forward-Regression Orthogonal Estimator. Int J Control, 1989; 49: 2157-89.
Billings SA, Tsang KM. Spectral-Analysis for Non-Linear Systems .2. Interpretation of Non-Linear Frequency-Response Functions. Mech Syst Signal Pr, 1989; 3: 341-59.
Billings SA, Voon WSE. Structure detection and model validity test in the identification of nonlinear systems. IEE Proceedings, Pt. D: Control Theory and Applications, 1983; 130: 193-9.
Brovelli A, Ding M, Ledberg A, Chen Y, Nakamura R, Bressler SL. Beta oscillations in a large-scale sensorimotor cortical network: directional influences revealed by Granger causality. Proceedings of the National Academy of Sciences of the United States of America, 2004; 101: 9849-54.
Chen S, Billings SA, Luo W. Orthogonal least squares methods and their application to nonlinear system identification. Int J Control, 1989; 50: 1873-96.
Chen YH, Rangarajan G, Feng JF, Ding MZ. Analyzing multiple nonlinear time series with extended Granger causality. Phys Lett A, 2004; 324: 26-35.
Chicharro D. On the spectral formulation of Granger causality. Biol Cybern, 2011; 105: 331-47.
Diks C, Panchenko V. A new statistic and practical guidelines for nonparametric Granger causality testing. J Econ Dyn Control, 2006; 30: 1647-69.
Ding M, Chen Y, Bressler S. Granger causality: basic theory and application to neuroscience. In Schelter S, Winterhalder M, Timmer J, editors. Handbook of time series analysis. Wienheim: Wiley, 2006: 438.
Feng JF, Yi DY, Krishna R, Guo SX, Buchanan-Wollaston V. Listen to Genes: Dealing with Microarray Data in the Frequency Domain. PloS one, 2009; 4.
Geweke J. Measurement of Linear-Dependence and Feedback between Multiple Time-Series. J Am Stat Assoc, 1982; 77: 304-13.
Gourevitch B, Le Bouquin-Jeannes R, Faucon G. Linear and nonlinear causality between signals: methods, examples and neurophysiological applications. Biol Cybern, 2006; 95: 349-69.
Granger CWJ. Investigating Causal Relations by Econometric Models and Cross-Spectral Methods. Econometrica, 1969; 37: 414-&.
He F, Billings SA, Wei HL, Sarrigiannis PG, Zhao Y. Spectral Analysis for Nonstationary and Nonlinear Systems: A Discrete-Time-Model-Based Approach. Biomedical Engineering, IEEE Transactions on, 2013; 60: 2233-41.
Hesse W, Moller E, Arnold M, Schack B. The use of time-variant EEG Granger causality for inspecting directed interdependencies of neural assemblies. J Neurosci Meth, 2003; 124: 27-44.
Hiemstra C, Jones JD. Testing for Linear and Nonlinear Granger Causality in the Stock Price-Volume Relation. J Financ, 1994; 49: 1639-64.
Jones JCP, Billings SA. Recursive Algorithm for Computing the Frequency-Response of a Class of Non-Linear Difference Equation Models. Int J Control, 1989; 50: 1925-40.

Kaminski MJ, Blinowska KJ. A New Method of the Description of the Information-Flow in the Brain Structures. Biol Cybern, 1991; 65: 203-10.

Lang ZQ, Billings SA. Energy transfer properties of nonlinear systems in the frequency domain. Int J Control, 2005; 78: 345-62.

Lang ZQ, Billings SA. Output frequency characteristics of nonlinear systems. Int J Control, 1996; 64: 1049-67.

Leontaritis IJ, Billings SA. Input-output parametric models for nonlinear systems, part I: deterministic nonlinear systems. Int J Control, 1985; 41: 303-28.

Li Y, Wei HL, Billings SA, Liao XF. Time-varying linear and nonlinear parametric model for Granger causality analysis (vol 85, 041906, 2012). Phys Rev E, 2012; 85.

Marinazzo D, Pellicoro M, Stramaglia S. Kernel-Granger causality and the analysis of dynamical networks. Phys Rev E, 2008; 77.

Michalareas G, Schoffelen JM, Paterson G, Gross J. Investigating causality between interacting brain areas with multivariate autoregressive models of MEG sensor data. Human brain mapping, 2013; 34: 890-913.

Montez T, Linkenkaer-Hansen K, van Dijk BW, Stam CJ. Synchronization likelihood with explicit time-frequency priors. NeuroImage, 2006; 33: 1117-25.

Stam CJ. Nonlinear dynamical analysis of EEG and MEG: Review of an emerging field. Clinical Neurophysiology, 2005; 116: 2266-301.

Swain AK, Billings SA. Generalized frequency response function matrix for MIMO non-linear systems. Int J Control, 2001; 74: 829-44.

Wei HL, Billings SA, Liu J. Term and variable selection for non-linear system identification. Int J Control, 2004; 77: 86-110.

Zhao YF, Billings SA, Wei HL, He F, Sarrigiannis PG. A new NARX-based Granger linear and nonlinear casual influence detection method with applications to EEG data. J Neurosci Meth, 2013; 212: 79-86.

Zou CL, Feng JF. Granger causality vs. dynamic Bayesian network inference: a comparative study. Bmc Bioinformatics, 2009; 10.