Preprints of the 19th World Congress
The International Federation of Automatic Control
Cape Town, South Africa. August 24-29, 2014

Frequency Domain Causality Analysis Method for Multivariate Systems in Hypothesis Testing Framework

Jing Zhang*, Dong Han*, Fan Yang*, Hao Ye*, Maoyin Chen*

**Tsinghua Laboratory for information Science and Technology,*
Department of Automation, Tsinghua University, Beijing,China
(e-mail: zhangjing06@mails.tsinghua.edu.cn;{handong2013,yangfan,haoye,mychen}@tsinghua.edu.cn)

Abstract: A variety of causality analysis methods have been proposed and used for complex large multivariate systems. In the frequency domain, partial directed coherence (PDC) is an important method. We expect that the frequency domain methods can provide a more detailed explanation of causal influence over different frequencies, but PDC provides no quantitative information to quantify the causal strength and the interpretation of causality over a specific frequency is still unexplained. Based on the statistical property of the renormalized PDC, a frequency domain causality analysis method in the hypothesis testing framework is employed in this paper to resolve these issues. In order to achieve a lower computational load, another method with a simpler definition is proposed. The interpretations of causal strength given by these two methods are shown to be consistent with that given by Granger causality, and the frequency distribution is reasonable and informative. Several simulation examples are demonstrated to illustrate the performance of these two methods.

Key words: Partial directed coherence, causality, hypothesis testing, strength, frequency domain

1. INTRODUCTION

Industrial process systems are often huge and complex, resulting in the requirement of a lot of sensors to measure and monitor the process variables. The causal relationship among time series of measurements is frequently unknown but helpful for a deeper understanding of the system structure, which is of great use for analysis and design of the process, such as fault diagnosis, modelling and alarm management (Miao et al., 2011; Yang and Xiao, 2012).

Several methods of causality identification have been proposed and widely used in many scientific and engineering areas for the recent forty years (Fan et al., 2013; Duan et al., 2013; Roebroeck et al., 2005). These methods can be classified into time domain methods, frequency domain methods, and information theory methods (Landman and Jounela, 2013).

Granger adopted and formalized Wiener's idea of causality (Wiener, 1956) in the context of linear vector auto regressive (VAR) models (Lutkepohl, 2005) and proposed a definition (later named "Granger causality") in the time domain (Granger, 1963a). Geweke (1982) then proposed a conditional Granger causality for multivariate processes to determine whether the influence between two time series is direct or indirect (mediated by a third one). In addition to these time domain methods, several frequency domain methods have also been developed, including frequency domain Granger causality (Geweke, 1984), directed transfer function (DTF) (Kaminski and Blinowska, 1991), and partial directed coherence (PDC) (Baccala and Sameshima, 2001). These frequency domain methods can further provide a measure of causality in the frequency domain through plotting the causality distribution with ω varying from 0 to π. Information theory methods are represented by transfer entropy (Schreiber, 2000).

On the other hand, frequency domain Granger causality is derived from the concept of power spectral (Geweke, 1984) and DTF is a measure of transferred energy (Kaminski and Blinowska, 1991). Unlike these methods, the interpretation of PDC is still not clear. It has been shown that PDC only aids in the structural information of the process and is able to reflect the existence of direct causality qualitatively (Gigi and Tangirala, 2010; Schelter et.al., 2009). The following PDC-related issues still need further discussions: (i) PDC provides no quantitative information to quantify the causal strength. Baccala and Sameshima (2001) stated that PDC has the ability to rank the relative interaction strength with respect to a given source signal because of normalization, but did not give a quantitative way to calculate the interaction strength or provide demonstrations. Zhang et al. (2013) discussed this issue through examples. (ii) Although PDC describes a causality distribution with ω varying from 0 to π in the frequency domain, the interpretation of causality over a specific frequency is challenging and still unexplained. It is an intuitive guess that the distribution over frequency represents how the strength of causal influence varies, yet Zhang et al. (2013) provided a counter example.

To deal with the above issues, we will propose two frequency domain causality analysis methods in the hypothesis testing framework by employing the so-called renormalized PDC (Schelter et al., 2009) and another statistical test (Schelter et al., 2005). Both methods have statistical properties similar to that of Granger causality. We show that the interpretation of causal strength based on the above two methods are consistent with that of Granger causality and transfer entropy, and the distribution over frequency is reasonable.

The rest of this paper is organized as follows. Section 2 provides a brief introduction to partial directed coherence and the renormalized PDC, and discusses two PDC-related issues through examples. In Section 3, we first interpret the Granger causality based causal strength measurement in the framework of hypothesis testing, then based on such interpretation, we give two frequency domain causality analysis methods in the same framework. Section 4 evaluates the performance of the two methods based on simulation examples, followed by concluding remarks in Section 5.

2. BRIEF INTRODUCTION TO PDC

We first introduce the concept of partial directed coherence and a few crucial but unaddressed issues related to PDC. Then a renormalized PDC proposed by Schelter et al. (2009) will be introduced.

2.1 Definition of PDC

Assume that there are n jointly stationary time series $x_1(k), x_2(k), ..., x_n(k)$.

As for partial directed coherence (PDC), a jointly stationary multivariate process can be described by an n-dimensional restraint VAR model (1) as follows, in which the model order and coefficients $\hat{a}_{ij}(r)(r=1,...,p)$ are estimated under a certain criterion, such as least squares, based on these n time series.

$$\begin{aligned} x_1(k) &= \sum_r \hat{a}_{11}(r)x_1(k-r) + \sum_r \hat{a}_{12}(r)x_2(k-r) + \cdots \\ &\quad + \sum_r \hat{a}_{1n}(r)x_n(k-r) + \hat{e}_1(k) \\ x_2(k) &= \sum_r \hat{a}_{21}(r)x_1(k-r) + \sum_r \hat{a}_{22}(r)x_2(k-r) + \cdots \\ &\quad + \sum_r \hat{a}_{2n}(r)x_n(k-r) + \hat{e}_2(k) \\ &\cdots\cdots \\ x_n(k) &= \sum_r \hat{a}_{n1}(r)x_1(k-r) + \sum_r \hat{a}_{n2}(r)x_2(k-r) + \cdots \\ &\quad + \sum_r \hat{a}_{nn}(r)x_n(k-r) + \hat{e}_n(k) \end{aligned} \tag{1}$$

Apply Z transform to (1), and let $z^{-1} = e^{-\mathrm{j}\omega}$, then the frequency response of process (1) can be written as

$$\hat{\mathbf{A}}(\omega)\mathbf{X}(\omega) = \mathbf{E}(\omega) \tag{2}$$

where

$$\hat{A}_{ij}(\omega) = -\sum_r \hat{a}_{ij}(r)e^{-\mathrm{j}\omega r}, \hat{A}_{ii}(\omega) = 1 - \sum_r \hat{a}_{ii}(r)e^{-\mathrm{j}\omega r}$$

$$\mathbf{X}(\omega) = [x_1(\omega)\ \ x_2(\omega)\ \ ...\ \ x_n(\omega)]^{\mathrm{T}}$$

$$\mathbf{E}(\omega) = [\hat{e}_1(\omega)\ \ \hat{e}_2(\omega)\ \ ...\ \ \hat{e}_n(\omega)]^{\mathrm{T}}$$

The estimated PDC $|\hat{\pi}_{ij}(\omega)|$ is defined to reflect the causality from x_j (source node) to x_i (sink node) as (Baccala and Sameshima, 2001)

$$|\hat{\pi}_{ij}(\omega)| \triangleq \frac{|\hat{A}_{ij}(\omega)|}{\sqrt{\sum_{i=1}^{n} |\hat{A}_{ij}(\omega)|^2}} \tag{3}$$

in which the following normalization property holds.

$$0 \le |\hat{\pi}_{ij}(\omega)|^2 \le 1 \text{ and } \sum_{i=1}^{n} |\hat{\pi}_{ij}(\omega)|^2 = 1 \tag{4}$$

and $|\hat{\pi}_{ij}(\omega)|$ is not zero when x_i is influenced by x_j directly.

2.2 Discussion on PDC-related issues

We now discuss the above issues (i) and (ii) in details through simulation examples.

(i) PDC only reflects the existence of direct causality qualitatively but cannot measure the strength.

Baccala and Sameshima (2001) stated that PDC has the ability to rank the relative interaction strength with respect to a given signal source because of the normalization in (3) without any formula to calculate the interaction strength or any demonstrations.

We use the following simulation example given by Zhang et al. (2013) to show the aforementioned limitations of PDC, in which the Granger causality is used as a reference.

Remark 1:
The reason why Granger causality is used as the reference is that the interpretation of causal strength based on Granger causality is explicit (Barrett and Seth, 2009). The definition of Granger causality is the formalization of Wiener's idea, which means that $F_{x_j \to x_i}$ reflects the improvement of the precision of prediction (Wiener, 1956; Barrett and Seth, 2009). Furthermore, Barrett and Seth (2009) noted that for Gaussian variables, Granger causality and transfer entropy are entirely equivalent, that is

$$F_{y \to x} = 2\Upsilon_{y \to x} \tag{5}$$

where $F_{y \to x}$ and $\Upsilon_{y \to x}$ denote the Granger causality and transfer entropy from y to x , respectively. Therefore, Granger causality in fact reflects an information-theoretic measure of causality (Barrett and Seth, 2009) and the causal strength from x_j to x_i based on Granger causality measures how much x_j contributes to improving the precision of prediction of x_i .

Example 1:
Consider a second-order process as follows, where $e_i(k)(i=1,...,5)$ are Gaussian noises with the covariance matrix set to be identity, and 5000 data points have been simulated.

$$\begin{cases} x_1(k) = 0.6x_1(k-1) + 0.4x_3(k-1) \\ \qquad +0.2x_3(k-2) + e_1(k) \\ x_2(k) = 0.6x_2(k-1) + 0.1x_2(k-2) + 0.3x_5(k-1) \\ \qquad -0.4x_5(k-2) + e_2(k) \\ x_3(k) = 0.3x_1(k-1) + 0.45x_3(k-1) - 0.24x_3(k-2) \\ \qquad +0.3x_5(k-1) + 0.2x_5(k-2) + e_3(k) \\ x_4(k) = 0.2x_1(k-1) + 0.2x_1(k-2) + 0.3x_2(k-1) \\ \qquad +0.4x_4(k-1) + 0.1x_5(k-1) + e_4(k) \\ x_5(k) = 0.4x_2(k-1) - 0.5x_5(k-1) \\ \qquad +0.15x_5(k-2) + e_5(k) \end{cases} \tag{6}$$

Since no formula was given by Baccala and Sameshima (2001) to calculate the strength based on PDC $|\hat{\pi}_{ij}(\omega)|$, we use $S_{\text{PDC}}(x_j \to x_i) = \int_0^\pi |\pi_{ij}(\omega)| \, \mathrm{d}\omega$ and $S_{\text{PDC}}(x_j \to x_i) = \int_0^\pi |\pi_{ij}(\omega)|^2 \, \mathrm{d}\omega$ to measure the causal strength from x_j to x_i, which are common indices to measure power in many scientific areas, such as signal processing.

We consider the causal strengths from x_1 to both x_3 and x_4, i.e. $S_{\text{PDC}}(x_1 \to x_3)$ and $S_{\text{PDC}}(x_1 \to x_4)$, which are given in Table 1.

It can be seen that $S_{\text{PDC}}(x_1 \to x_3)$ is slightly less than $S_{\text{PDC}}(x_1 \to x_4)$, which shows that the interaction strength from x_1 to x_4 is stronger than that from x_1 to x_3 according to the viewpoint of Baccala and Sameshima (2001), that is, PDC has the ability to rank the relative interaction strength with respect to a given signal source.

Table 1. Causal strengths $S_{\text{PDC}}(x_1 \to x_3)$ and $S_{\text{PDC}}(x_1 \to x_4)$

Definition	S_{13}	S_{14}
$S_{\text{PDC}}(x_j \to x_i) = \int_0^\pi \lvert\pi_{ij}(\omega)\rvert \, \mathrm{d}\omega$	0.8537	0.8598
$S_{\text{PDC}}(x_j \to x_i) = \int_0^\pi \lvert\pi_{ij}(\omega)\rvert^2 \, \mathrm{d}\omega$	0.2699	0.3490

For comparison, the Granger causalities $F_{x_1 \to x_3}$ and $F_{x_1 \to x_4}$ are given in Table 2 as a reference.

Table 2. Results of $F_{x_1 \to x_3}$ and $F_{x_1 \to x_4}$

$F_{x_1 \to x_3}$	$F_{x_1 \to x_4}$	Threshold
0.2099	0.1088	0.0027

It can be seen that $F_{x_1 \to x_4} < F_{x_1 \to x_3}$, which shows that the causal strength from x_1 to x_3 is stronger than that from x_1 to x_4. Therefore the conclusion about the interaction strength given by PDC contradicts that given by Granger causality. Since Granger causality has the ability to reflect the strength between different pairs of time series according to Remark 1, we have to conclude that PDC only reflects the existence of direct causality qualitatively but cannot measure the strength.

(ii)The interpretation of the distribution of PDC in the frequency domain is still unexplained.

The expression of PDC (3) is directly related to the frequency ω, which means that, compared to the time domain methods, PDC ought to further provide a detailed measure of causality with respect to different frequencies. It is an intuitive guess that the distribution over frequency represents how the strength of causal influence varies. The following example discussed by Zhang et al. (2013), however, provides a counter example.

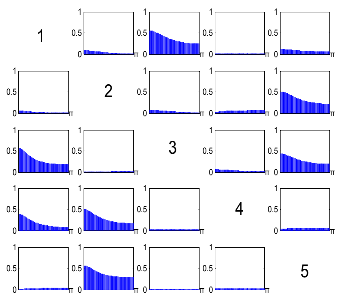
Example 2:

Consider two first-order VAR processes as follows:

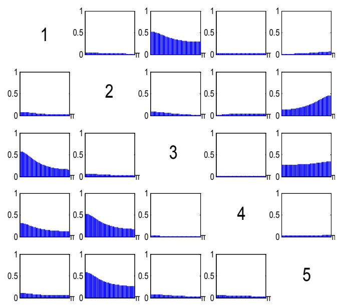
$$\begin{cases} x_1(k) = 0.6x_1(k-1) + 0.4x_3(k-1) + e_1(k) \\ x_2(k) = 0.6x_2(k-1) + 0.3x_5(k-1) + e_2(k) \\ x_3(k) = 0.3x_1(k-1) + 0.45x_3(k-1) + 0.3x_5(k-1) + e_3(k) \\ x_4(k) = 0.2x_1(k-1) + 0.3x_2(k-1) + 0.4x_4(k-1) + e_4(k) \\ x_5(k) = 0.4x_2(k-1) + 0.5x_5(k-1) + e_5(k) \end{cases} \tag{7}$$

$$\begin{cases} x_1(k) = 0.6x_1(k-1) + 0.4x_3(k-1) + e_1(k) \\ x_2(k) = 0.6x_2(k-1) + 0.3x_5(k-1) + e_2(k) \\ x_3(k) = 0.3x_1(k-1) + 0.45x_3(k-1) + 0.3x_5(k-1) + e_3(k) \\ x_4(k) = 0.2x_1(k-1) + 0.3x_2(k-1) + 0.4x_4(k-1) + e_4(k) \\ x_5(k) = 0.4x_2(k-1) - 0.5x_5(k-1) + e_5(k) \end{cases} \tag{8}$$

The corresponding PDC calculated based on simulation data are given in Fig. 1, which are sorted as matrix plots, where $|\hat{\pi}_{ij}(\omega)|$ is displayed in the i th row and the j th column.



Frequency ω (0 to π)
(a) PDC of process (7)



Frequency ω (0 to π)
(b) PDC of process (8)

Fig. 1. Plots to show the disadvantages of PDC

There exist some interesting differences between the results of the two processes. Equations (7) and (8), where the two models are almost the same except the sign of the coefficient of $x_5(k-1)$ in the last equations (the equations about the evolution of $x_5(k)$). It is obvious that whether the coefficient of $x_5(k-1)$ is positive or negative does not affect the qualitative causal relationship, because it only affects the dynamics of x_5 itself. However, the PDC distributions in Fig. 1 appear quite different for $|\hat{\pi}_{i5}(i \neq 5)|$: the value of $|\hat{\pi}_{i5}(i \neq 5)|$ monotonically decreases in Fig. 1(a), while in Fig. 1(b) $|\hat{\pi}_{i5}(i \neq 5)|$ increases as ω grows. It has been demonstrated that the above difference is caused by normalization in (3) by Zhang et al. (2013).

This example shows that the distribution of PDC cannot reveal the strength of causal influence with ω varying from 0 to π and its implication in the frequency domain seems unexplained.

2.3 Brief Introduction to Renormalized PDC

Baccala and Sameshima (2001) asserted that PDC reflects the relative rather than the absolute strength of influence because of the normalization, making the causality given by PDC be influenced by the number of other signals that are influenced by the same source signals, that is, the causality $x_j \to x_i$ may change if more (or less) signals are influenced by x_j . To tackle this disadvantage, a renormalized PDC (RPDC) was proposed by Schelter et al. (2009), which avoids normalization in its definition.

Since we will utilize a statistical property of RPDC in this paper, we next give a brief introduction to RPDC.

Notation 1:

Let $*_{i,j}$ denote the element in the i th row and the j th column of matrix $*$.

Let $\mathbf{\Sigma}$ and $\mathbf{R}$ denote the covariance matrices of the noise $\mathbf{e}(k)=[e_1(k),...,e_n(k)]^T$ and process $\mathbf{x}(k)=[x_1(k),...,x_n(k)]^T$ respectively. Define $\mathbf{H} \triangleq \mathbf{R}^{-1}$ and

$$\hat{\mathbf{X}}_{ij}(\omega) \triangleq \begin{pmatrix} \mathrm{Re}(\hat{A}_{ij}(\omega)) \\ \mathrm{Im}(\hat{A}_{ij}(\omega)) \end{pmatrix} \tag{9}$$

$$\mathbf{V}_{ij}(\omega) \triangleq \sum_{t,l=1}^{p} H_{jj}(t,l)\Sigma_{ii} \begin{pmatrix} \cos(t\omega)\cos(l\omega) & \cos(t\omega)\sin(l\omega) \\ \sin(t\omega)\cos(l\omega) & \sin(t\omega)\sin(l\omega) \end{pmatrix} \tag{10}$$

The index RPDC is defined as (Schelter et al. 2009)

$$\hat{\lambda}_{ij}(\omega) \triangleq N\hat{\mathbf{X}}_{ij}(\omega)'\hat{\mathbf{V}}_{ij}(\omega)^{-1}\hat{\mathbf{X}}_{ij}(\omega) \tag{11}$$

where N is the number of data points and $\hat{\mathbf{V}}_{ij}(\omega)$, in which $\mathbf{\Sigma}$ and $\mathbf{R}$ need to be estimated based on $\mathbf{e}(k)$ and $\mathbf{x}(k)$, is an estimate of $\mathbf{V}_{ij}(\omega)$.

Schelter et al. (2009) provided the following important proposition.

Proposition 1(Schelter et al., 2009)*:*

Under the null hypothesis of $|A_{ij}(\omega)|^2=0$, for $p \geq 2$ and $\omega \neq 0 \bmod \pi$, the RPDC $\hat{\lambda}_{ij}(\omega)$ follows an approximate χ^2 distribution with two degrees of freedom as N tends to infinity. When $p=1$ or $\omega = 0 \bmod \pi$, the RPDC $\hat{\lambda}_{ij}(\omega)$ with $\hat{\mathbf{V}}_{ij}(\omega)^{-1}$ being the generalized inverse of $\hat{\mathbf{V}}_{ij}(\omega)$ is approximately a χ^2 distribution with one degree of freedom as N tends to infinity.

3. TWO FREQUENCY DOMAIN CAUSALITY ANALYSIS METHODS IN HYPOTHESIS TESTING FRAMEWORK

As mentioned in Section 2.2, two PDC-related issues still remain unaddressed. However, Granger causality as a time domain method can shed some light for us to deal with these issues. We will firstly point out that Granger causality can be explained as a hypothesis test problem in essence, and the value of $F_{x_j \to x_i}$ actually corresponds to an index in the hypothesis test problem. Motivated by this, two frequency domain causality analysis methods will be proposed with the assistance of RPDC (Schelter et al., 2009) and another statistical test (Schelter et al., 2005), respectively.

3.1 Interpretation of Granger Causality Based Causal Strength Measurement in the Hypothesis Testing Framework

It has been stated that under the null hypothesis of $F_{x_j \to x_i}=0$, N times the Granger causality $NF_{x_j \to x_i}$ is a χ^2 distribution with p degree of freedom as N tends to infinity (Whittle, 1953; Granger, 1963b), where N is the number of data points and p denotes the model order.

Therefore the procedure of using Granger causality to detect the existence of causality can be interpreted in the following framework of hypothesis test. First set a significance level α_0 :

$$\alpha_0 = \mathrm{Prob}\left\{NF_{x_j \to x_i} > \chi^2_{p,\alpha_0}\right\} \tag{12}$$

which in fact determines the false alarm rate under the null hypothesis of $F_{x_j \to x_i}=0$. If $\hat{F}_{x_j \to x_i} > \chi^2_{p,\alpha_0}/N$, the null hypothesis should be rejected (Li, 2006).

When Granger causality is used to measure the strength between each pair of time series, it has been concluded that a larger $\hat{F}_{x_j \to x_i}$ implies a stronger causal strength from x_j to x_i , as explained in Remark 1. In the following, we will interpret its meaning in the framework of hypothesis test, which forms a basis for the discussion in Section 3.2.

Each specific estimated $\hat{F}^0_{x_j \to x_i}$ corresponds to a probability which represents the probability that $F_{x_j \to x_i}$ takes value in a small neighborhood of $\hat{F}^0_{x_j \to x_i}$ under the null hypothesis of $F_{x_j \to x_i}=0$. Since the probability density decreases along with

$F_{x_j \to x_i}$, the larger $\hat{F}^0_{x_j \to x_i}$ is, the smaller probability that it happens under the null hypothesis is. Thus using Granger causality to measure the causal strength from x_j to x_i has its physical meaning in the sense of hypothesis testing.

3.2 Proposed Method 1: Frequency Domain Causality Analysis Based on RPDC

According to the discussion in Remark 1 and Section 3.1, Granger causality $F_{x_j \to x_i}$, which follows a χ^2 distribution, is a proper tool to measure the causality in the time domain, and the way of applying Granger causality in causality analysis can be interpreted as a typical hypothesis test.

Considering that RPDC defined in the frequency domain is also approximately a χ^2 distribution according to Proposition 1, its statistical property offers us a good way to deal with PDC-related issues (i) and (ii).

In this section, similar to Granger causality based causality analysis, a frequency domain causality analysis method will be presented. We will apply RPDC as a tool to perform causality analysis in the frequency domain, including measuring the causal strength at different frequencies and measuring that between different pairs of time series, to solve the aforementioned issues of PDC.

Remark 2:
It is worth noting that RPDC defined in the frequency domain was proposed by Schelter et al. (2009) to tackle another disadvantage of PDC rather than issues (i) and (ii), that is, the causality given by PDC will be influenced by the number of other signals that are influenced by the same source signals.

Similar to the idea in Section 3.1, the causal strength from x_j to x_i at frequency ω (with ω varying from 0 to π) is defined as

$$S_{\text{method1}}(x_j \to x_i \mid \omega) \triangleq \hat{\lambda}_{ij}(\omega) = N\hat{\mathbf{X}}_{ij}(\omega)'\hat{\mathbf{V}}_{ij}(\omega)^{-1}\hat{\mathbf{X}}_{ij}(\omega) \quad (13)$$

If $S_{\text{method1}}(x_j \to x_i \mid \omega)$ exceeds the threshold, which is set to be χ^2_{2,α_0} for model order $p \geq 2$ or χ^2_{1,α_0} for $p = 1$ (α_0 is a given significance level that can be commonly chosen as a small value, such as 0.05 (Schelter et al., 2009)), it is concluded that there exists evident causal influence from x_j to x_i at frequency ω.

Definition (13) measures the existence and strength of causality between a given pair of time series at a certain frequency. Similar to the analysis of Granger causality in Section 3.1, a larger estimated $S^0_{\text{method1}}(x_j \to x_i \mid \omega)$ corresponds to a smaller probability that $\hat{\lambda}_{ij}(\omega)$ takes value in a small neighborhood of $\hat{\lambda}^0_{ij}(\omega)$ under the null hypothesis in its hypothesis test problem, which means that if $S_{\text{method1}}(x_j \to x_i \mid \omega_1) > S_{\text{method1}}(x_j \to x_i \mid \omega_2) > \text{threshold}$, the causal strength from x_j to x_i at frequency ω_1 is stronger than that at frequency ω_2. Thus in this sense, the distribution of $S_{\text{method1}}(x_j \to x_i \mid \omega)$ is able to describe the frequency distribution of causal strength.

In addition, the causal strength from x_j to x_i based on the causality analysis method is defined as

$$S_{\text{method1}}(x_j \to x_i) \triangleq \frac{1}{2\pi}\int_0^{\pi} \hat{\lambda}_{ij}(\omega)\,\mathrm{d}\,\omega \quad (14)$$

If $S_{\text{method1}}(x_j \to x_i)$ exceeds the threshold, which is set to be $\frac{1}{2\pi}\int_0^{\pi}\chi^2_{2,\alpha_0}\,\mathrm{d}\,\omega$ for $p \geq 2$ or $\frac{1}{2\pi}\int_0^{\pi}\chi^2_{1,\alpha_0}\,\mathrm{d}\,\omega$ for $p = 1$, we can conclude that there exists evident causal influence from x_j to x_i.

Definition (14), which takes the whole frequency domain into consideration, measures the existence and strength of the causality between each pair of time series. Although $S_{\text{method1}}(x_j \to x_i)$ does not follow a χ^2 distribution, it is still intuitive that if $S_{\text{method1}}(x_j \to x_i) > S_{\text{method1}}(x_b \to x_a)$, the causal strength from x_j to x_i is stronger than that from x_b to x_a.

In Section 5, simulation examples will be given to show the effectiveness and advantages of this method.

3.2 Proposed Method 2: Frequency Domain Causality Analysis Based on $N\,|\,\hat{A}_{ij}(\omega)\,|^2 / \hat{C}_{ij}(\omega)$

The definitions in (13) and (14) are somewhat complex and lead to a relative high computational load. In this section, we will replace RPDC with a simpler statistic test given by Schelter et al. (2005), which also follows a χ^2 distribution, to achieve a lower computational load.

We first introduce this statistical test.

Proposition 2 (Schelter et al., 2005)*:*
Under the null hypothesis of $|A_{ij}(\omega)|^2 = 0$, the distribution of $N\,|\,\hat{A}_{ij}(\omega)\,|^2 / \hat{C}_{ij}(\omega)$ is approximately a χ^2 distribution with one degree of freedom. Here, $\hat{C}_{ij}(\omega)$ is the estimate of $C_{ij}(\omega)$ defined as

$$C_{ij}(\omega) = \Sigma_{ii}\Big[\sum_{t,l=1}^{p} H_{jj}(t,l)(\cos(t\omega)\cos(l\omega) + \sin(t\omega)\sin(l\omega))\Big]$$

where Σ_{ii} and $H_{jj}(t,l)$ have the same meaning as in (10).

Proposition 2 was used by Schelter et al. (2009) to propose a significance level for testing nonzero PDC at a certain frequency in order to solve the over fitting problems in model estimation.

In this section, by replacing the RPDC statistic test in (13) and (14) with $N|\hat{A}_{ij}(\omega)|^2/\hat{C}_{ij}(\omega)$, we give the following causality analysis method which can lead to a lower computational load.

The causal strength from x_j to x_i at frequency ω (with ω varying from 0 to π) is defined as

$$S_{\text{method2}}(x_j \to x_i \mid \omega) \triangleq \hat{\gamma}_{ij}(\omega) = N|\hat{A}_{ij}(\omega)|^2/\hat{C}_{ij}(\omega) \qquad (15)$$

If $S_{\text{method2}}(x_j \to x_i \mid \omega)$ exceeds the threshold, which is set to be χ^2_{1,α_0} (α_0 is a given significance level that can be commonly chosen as a small value, such as 0.05), it is concluded that there exists causal influence from x_j to x_i at frequency ω.

In addition, the causal strength from x_j to x_i based on this method is defined as

$$S_{\text{method2}}(x_j \to x_i) \triangleq \frac{1}{2\pi}\int_0^{\pi} \hat{\gamma}_{ij}(\omega)\,\mathrm{d}\omega \qquad (16)$$

If $S_{\text{method2}}(x_j \to x_i)$ exceeds the threshold, which is set to be $\frac{1}{2\pi}\int_0^{\pi} \chi^2_{1,\alpha_0}\,\mathrm{d}\omega$, we conclude that there exists causal influence from x_j to x_i.

The physical meanings of this method are similar to those given in Section 3.2, thus being omitted here.

Since the definitions are simper with scalar operation, the method has a lower computational load.

Remark 3:

Although $\hat{C}_{ij}(\omega)$ in (15) and (16), in which $\mathbf{\Sigma}$ and $\mathbf{R}$ need to be estimated based on $\mathbf{e}(k)$ and $\mathbf{x}(k)$, is an estimate of true $C_{ij}(\omega)$, there exists the same approximation when estimating $V_{ij}(\omega)$ in method 1 given in Section 3. It will be shown that the performance of this method is as good as that of method 1.

4. SIMULATION STUDY

In this section, based on several simulation examples, the two frequency domain causality analysis methods given in Section 3 are shown to have the ability to measure the causality in the frequency domain.

4.1 Example 1

Consider the second-order process (6) given in section 2.2. According to (6), the real causal relationship among the five variables is shown in Fig. 2.

The causal strengths between each pair of time series based on Granger causality and the two methods, i.e. $F_{x_j \to x_i}$, $S_{\text{method1}}(x_j \to x_i)$, and $S_{\text{method2}}(x_j \to x_i)$, are calculated and summarized in Table 3(a), 3(b), and 3(c), respectively. The strength from x_j to x_i which exceeds the threshold is given in the i th row and the j th column (the blank implies that the calculated strength does not exceed the threshold, leading to the conclusion of no causality).

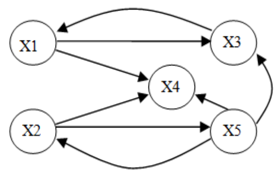


Fig. 2 Causal relationship among the five variables

Table 3 Causal strengths in process (6)

(a) Causal strengths from x_j to x_i based on $F_{x_j \to x_i}$

		0.2002		
				0.3932
0.2099				0.3056
0.1088	0.1175			0.0998
	0.1951			

(b) Causal strengths from x_j to x_i based on $S_{\text{method1}}(x_j \to x_i)$

		24.5012		
				39.2280
27.7010				32.0942
16.2684	19.6765			15.7450
	22.9304			

(b) Causal strengths from x_j to x_i based on $S_{\text{method2}}(x_j \to x_i)$

		33.0849		
				43.8977
38.8698				39.5011
22.1023	26.9955			19.9758
	32.8947			

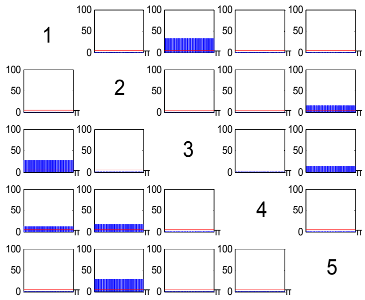
It can be seen that the two methods proposed in this paper can detect the existence of causality in the process correctly. Unlike PDC shown in Table 1, the calculated strengths based on them are consistent with that given by Granger causality, which means that a larger $F_{x_j \to x_i}$ corresponds to a larger $S_{\text{method1}}(x_j \to x_i)$ and a larger $S_{\text{method2}}(x_j \to x_i)$.

Hence both of the two methods presented in this paper can deal with issue (i) mentioned in Section 2.2, because they are able to detect the existence of direct causality among a set of time series correctly, and can rank the causal strengths.

4.2 Example 2

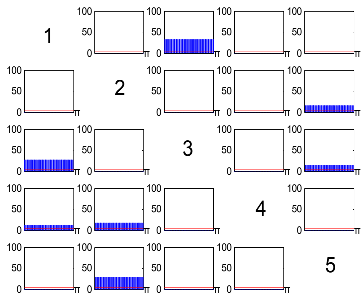
Consider the two first-order VAR processes (7) and (8) given in Section 2.2.

The corresponding $S_{\text{method1}}(x_j \to x_i \mid \omega)$ are given in Figs. 3(a) and 3(b) and $S_{\text{method2}}(x_j \to x_i \mid \omega)$ are given in Figs. 4(a) and 4(b), respectively. Unlike the distributions of PDC given in Fig. 1, the distributions of $S_{\text{method1}}(x_j \to x_i \mid \omega)$ or $S_{\text{method2}}(x_j \to x_i \mid \omega)$ of these two processes are roughly similar as expected. Especially, in Fig. 1, the PDC distributions appear quite different for $|\hat{\pi}_{i5}(i \neq 5)|$, that is, $|\hat{\pi}_{i5}(i \neq 5)|$ monotonically decreases in Fig. 1(a), while in Fig. 1(b) $|\hat{\pi}_{i5}(i \neq 5)|$ increases as ω grows. In Fig. 3 and Fig. 4, however, the distributions of $S_{\text{method1}}(x_5 \to x_i \mid \omega)$ and $S_{\text{method2}}(x_5 \to x_i \mid \omega)$ of the two processes are similar.



Frequency ω (0 to π)

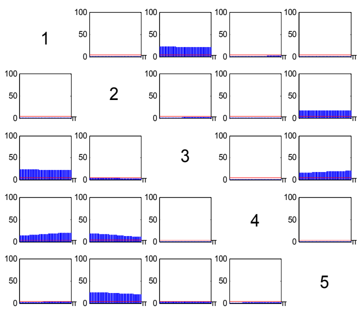
(a) $S_{\text{method1}}(x_j \to x_i \mid \omega)$ of process (7)



Frequency ω (0 to π)

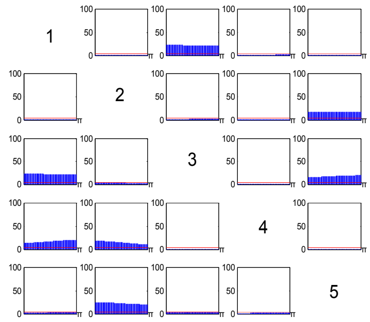
(b) $S_{\text{method1}}(x_j \to x_i \mid \omega)$ of process (8)

Fig. 3 Distribution of $S_{\text{method1}}(x_j \to x_i \mid \omega)$ over frequency



Frequency ω (0 to π)

(a) $S_{\text{method2}}(x_j \to x_i \mid \omega)$ of process (7)



Frequency ω (0 to π)

(b) $S_{\text{method1}}(x_j \to x_i \mid \omega)$ of process (8)

Fig. 4 Distribution of $S_{\text{method2}}(x_j \to x_i \mid \omega)$ over frequency

The example shows that the interpretation of distribution of $S_{\text{method1}}(x_j \to x_i \mid \omega)$ or $S_{\text{method2}}(x_j \to x_i \mid \omega)$ over frequency is more reasonable than that of PDC.

4.3 Example 3

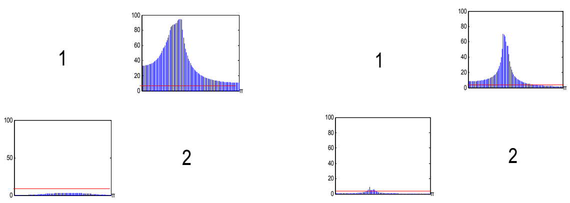
We further discuss the interpretation of the distributions of $S_{\text{method1}}(x_j \to x_i \mid \omega)$ or $S_{\text{method2}}(x_j \to x_i \mid \omega)$ in the frequency domain.

Consider a second-order VAR process as follows:

$$\begin{cases} x_1(k) = 0.3x_1(k-1) - 0.15x_1(k-2) \\ \qquad\quad + 0.4x_2(k-1) + 0.2x_2(k-2) + e_1(k) \\ x_2(k) = 0.2x_2(k-1) + 0.2x_2(k-2) + e_2(k) \end{cases} \tag{17}$$

where $e_i(k)(i=1,2)$ are Gaussian noises with the covariance matrix of the noises set to be identity, and 5000 data points have been simulated.

Fig.5 gives the corresponding calculated $S_{\text{method1}}(x_j \to x_i \mid \omega)$ and $S_{\text{method2}}(x_j \to x_i \mid \omega)$, of which the values both appear to be maximal around 0.4π.



Frequency ω (0 to π) Frequency ω (0 to π)

(a) (b)

Fig. 5. Results of the two methods

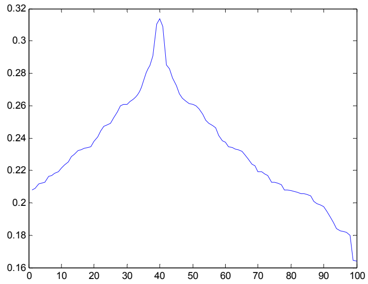
In order to discuss the interpretation of these distributions the frequency domain, the following second-order process (18) is used

$$\begin{cases} x_1(k) = 0.3x_1(k-1) - 0.15x_1(k-2) + 0.4x_2(k-1) \\ \qquad\quad + 0.2x_2(k-2) + e_1(k) \\ x_2(k) = 0.2x_2(k-1) + 0.2x_2(k-2) + e_2(k) + \sin(k\omega) \end{cases} \tag{18}$$

where $e_i(k)(i=1,2)$ are Gaussian noises and the covariance matrix of the noises is set to be identity, too.

We choose ω from 0 to π with interval 0.01π and calculate Granger causality $F_{x_2 \to x_1}$ at each frequency. The averaged $F_{x_2 \to x_1}$ of 50 random simulation experiments at each frequency is given in Fig. 6.

The distribution of Granger causality $F_{x_2 \to x_1}$ in the frequency domain is consistent with the above two and the maximum of $F_{x_2 \to x_1}$ is also at frequency around 0.4π.



Frequency $\omega / 0.01\pi$ (0 to π)

Fig. 6 Averaged $F_{x_2 \to x_1}$ versus frequency

The above examples 2 and 3 show that the distributions of the two methods given in this paper are reasonable and can reveal the strength, which reflects how much x_j contributes to improving the precision of prediction of x_i, at different frequencies.

5. CONCLUDING REMARKS

In this paper, we have proposed two frequency domain causality analysis methods in the hypothesis testing framework to solve the two existing issues in PDC. The main contribution of the two methods includes: (i) An index to quantify the causal strength and the threshold are defined; thus the methods can not only detect the causality among the time series, but also have the ability to measure the causal strength. The interpretation of causal strength is shown to be consistent with that of Granger causality and transfer entropy. (ii) The distribution over frequency is meaningful and can reveal how the causal strength varies.

ACKNOWLEDGEMENTS

The work was supported by National Natural Science Foundation of China 61290324 and 61203119.

REFFRENCES

Baccala, L.A., Sameshima, K. (2001). Partial directed coherence: a new concept in neural structure determination. *Biol. Cybern*, 84, 463-474.

Barrett, A.B., Seth, A.K. (2009). Granger causality and transfer entropy are equivalent for Gaussian variables. *Physical Review Letters*, 103, 238701.

Duan, P., Yang, F., Shah, S.L., Chen, T. (2013). Direct causality detection via the transfer entropy approach. *Control Systems Technology*, PP(99), 1.

Fan, N., Yang, F., Ye, H. (2013). Application of PDC method in causality analysis of chemical process variables. *Journal of Tsinghua University*, 53(2), 210-214.

Granger, C.J.W. (1963a). Investigating causal relations by econometric models and cross-spectral methods. *Journal of the Econometric Society*, 37, 424-438.

Granger, C.J.W. (1963b). Economic processes involving feedback. *Information and Control*, 6, 28-48.

Geweke, J.F. (1982). Measures of linear dependence and feedback between multiple time series. *Journal of the American Statistical Association*, 77 (378), 304-313

Geweke, J.F. (1984). Measures of conditional linear dependence and feedback between time series. *Journal of the American Statistical Association*, 79, 907-915.

Gigi, S., Tangirala, A.K. (2010). Quantitative analysis of directional strengths in jointly stationary linear multivariate processes. *Biol.Cybern*, 103, 119-133.

Kaminski, M.J., Blinowska, K.J. (1991). A new method of the description of the information flow in the brain structures. *Biol .Cybern*, 65, 203-210.

Lutkepohl, H. (2005). New introduction to multiple time series analysis. *Springer*, New York.

Li, J. (2006). Testing Granger causality in the presence of threshold effects. *International Journal of Forecasting*, 22(4), 771-780.

Landman, R., Jounela, S.L. (2013). Data-Driven causal analysis and its application on a large-scale board machine. *18th Nordic Process Control Workshop*, Oulu, Finland.

Miao, X.Y., Wu, X., Li, R., Chen, K.W., Yao, L. (2011). Altered connectivity pattern of hubs in default-mode network with Alzheimer's disease: an Granger causality modeling approach. *PLoS One*, 6(10), e25546.

Roebroeck, A., Formisano, E., Goebel, R. (2005). Mapping directed influence over the brain using Granger causality and fMRI. *Neuroimage*, 25(1), 230-242.

Schreiber, T. (2000). Measuring information transfer. *Physical Review Letters*, 82(2), 461-464.

Schelter, B., Witerhalder, M., Michael, E., Martin, P., Bernhard, H., Brigitte, G. et al. (2005). Testing for directed influences among neural signals using partial directed coherence. *Journal of Neuroscience Methods*, 152, 210-129.

Schelter, B., Timmer, J., Eichler, M. (2009). Assessing the strength of directed influences among neural signals using renormalized partial directed coherence. *Journal of Neuroscience Methods*, 179, 121-130.

Whittle, P. (1953). The analysis of multiple stationary time series. *J. Royal Statist Soc. B.*, 15, 125-139.

Wiener, N. (1956). The theory of prediction. in Beckenbach EF (ed) *Modern Mathematics for Engineers*, chapter 8, McGraw-Hill, New York.

Yang, F., Xiao, D. (2012). Progress in root cause and fault propagation analysis of large-scale industrial processes. *J. Control Sci. Eng*, 478373.

Zhang, J., Yang, F., Ye, H. (2013). Quantitative analysis of partial directed coherence in jointly stationary multivariate processes. *The 23th Chinese Process Control Conference,* Guiyang, China.