# PERFORMANCE STUDY OF CAUSALITY MEASURES

*T. Bořil, P. Sovka*

Department of Circuit Theory, Faculty of Electrical Engineering,

Czech Technical University in Prague

**Abstract**

**Analysis of dynamic relations in human brain has became an important task in neurophysiology in recent time. Unlike a coherence analysis used for revealing cooperating parts in the brain, causality measures make possible not only to evaluate the strength of relations but also a direction of their influence. Well-accepted Granger causality and other derived causality measures based on multivariate autoregressive models have been used to analyse directions of information flow in multichannel time series in econometrics and neurophysiology for a long time. Recently, a new method Phase Slope Index (PSI) based on completely different approach has been proposed. This paper compares conditional Granger causality with PSI, shows some important issues with PSI and proposes an idea how to interpret results of both methods used on the same data set.**

## 1 Introduction

In the late `60s of the 20th century, econometrics started examination of causal relations among time series in order to create economic models and predict a behavior of variables like agricultural prices. The basic idea of causality was formulated by Wiener [1]. If the prediction of one time series could be improved by knowledge of past values of another one, then we say the second series has a causal influence on the first one. Granger formalized it in the scope of autoregressive models later [2]. A whole new area opens for Granger causality in the form of neurophysiology at the end of 20th century. Functional magnetic resonance imaging (fMRI), electroencephalography (EEG) and magnetoencephalography (MEG) bring data which meets conditions of causality analysis very well and together with progress of computational technology it is possible to find causal relations in the large data of brain activity records, e.g. [3, 4].

Since the time the Granger causality was formulated, several other methods measuring causal relations has been proposed based on similar idea – Directed Transfer Function (DTF) [5], direct Directed Transfer Function (dDTF) [6], Partial Directed Coherence (PDC) [7] and Generalized Partial Directed Coherence (GPDC) [8]. All these methods are based on multivariate autoregressive (MVAR) models and can be understood as a frequency domain extension to the Granger causality. Recently, a method using completely different approach called Phase-Slope Index (PSI) has been proposed [9, 10]. It evaluates a slope of a phase of cross-spectra between two time series and it behaves better than Granger causality in noisy conditions in some cases when the Granger causality gives fake albeit significant results.

This paper compares conditional Granger causality (CGC) with Phase Slope Index, shows some important issues with PSI like inability to detect bidirectional connections and proposes an idea how to interpret results of both methods used on the same data set.

## 2 Conditional Granger Causality

Let's consider three stationary stochastic AR processes $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{z}$. To examine the causal influence from $\boldsymbol{y}$ to $\boldsymbol{x}$, we represent $\boldsymbol{x}$ as MVAR model of $\boldsymbol{x}$ and $\boldsymbol{z}$ [11]:

$$x(t)=\sum_{j=1}^{m}\alpha_{1j}x(t-j)+\sum_{j=1}^{m}\beta_{1j}z(t-j)+\varepsilon_{xz}(t),\quad \mathrm{var}(\boldsymbol{\varepsilon}_{xz})=\Sigma_{xz} \tag{1}$$

where prediction error $\boldsymbol{\varepsilon}_{xz}$ is white noise. Then we represent $\boldsymbol{x}$ as MVAR model of $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{z}$:

$$x(t)=\sum_{j=1}^{m}a_{1j}x(t-j)+\sum_{j=1}^{m}b_{1j}y(t-j)+\sum_{j=1}^{m}c_{1j}z(t-j)+\varepsilon_{xyz}(t),\quad \mathrm{var}(\boldsymbol{\varepsilon}_{xyz})=\Sigma_{xyz} \tag{2}$$

where prediction error $\varepsilon_{xyz}$ is white noise. Granger causality from $\boldsymbol{y}$ to $\boldsymbol{x}$ conditional on $\boldsymbol{z}$ is:

$$F_{\mathbf{y}\to\mathbf{x}|\mathbf{z}} = \ln\frac{\Sigma_{xz}}{\Sigma_{xyz}}. \tag{3}$$

When the causal influence from $\boldsymbol{y}$ to $\boldsymbol{x}$ is entirely mediated by $\boldsymbol{z}$, $b_{1j}$ is uniformly zero and $\Sigma_{XZ} = \Sigma_{XYZ}$, $F_{\mathbf{y}\to\mathbf{x}|\mathbf{z}} = 0$.

## 3 Phase Slope Index

PSI [9] is insensitive to mixtures of independent sources, gives meaningful results even if the phase spectrum is not linear, and properly weights contributions from different frequencies. PSI is based on a slope of the phase of cross-spectra between two time series. The idea behind using phase slope is that interactions require some time and if the speed at which different waves travel is similar, then the phase difference between sender and recipient increases with frequency and we expect a positive slope of the phase spectrum. This quantity is defined as

$$\tilde{\boldsymbol{\psi}}_{ij} = \Im\left(\sum_{f\in F} C_{ij}^{*}(f)\, C_{ij}(f+\delta f)\right) \tag{4}$$

where $C_{ij}(f) = S_{ij}(f)/\sqrt{S_{ii}(f)S_{jj}(f)}$ is the complex coherency, $S_{ij}(f) = \left\langle \hat{y}_i(f)\,\hat{y}_j^{*}(f)\right\rangle$ is the cross-spectral matrix between two time series for channels $i$ and $j$, $\langle\cdot\rangle$ denotes expectation value, $\delta f$ is the frequency resolution, $\Im(\cdot)$ denotes taking the imaginary part and $F$ is the set of frequencies over which the slope is summed. Finally, it is convenient to normalize $\tilde{\boldsymbol{\psi}}$ by an estimate of its standard deviation $\boldsymbol{\psi} = \tilde{\boldsymbol{\psi}}/\mathrm{std}(\tilde{\boldsymbol{\psi}})$, with $\mathrm{std}(\tilde{\boldsymbol{\psi}})$ being estimated by the jackknife method.

## 4 Experiments & Results

We have generated 6000 samples of five stable MVAR signals with causal influences $\boldsymbol{v}_1\to\boldsymbol{v}_2$, $\boldsymbol{v}_2\to\boldsymbol{v}_3$, $\boldsymbol{v}_1\to\boldsymbol{v}_4$ and bidirectional causal influence $\boldsymbol{v}_4\leftrightarrow\boldsymbol{v}_5$ (Fig. 1) in order to model sequential driving problem (indirect causal connections $\boldsymbol{v}_1\to\boldsymbol{v}_3$ and $\boldsymbol{v}_1\to\boldsymbol{v}_5$), differently delayed driving problem (indirect causality $\boldsymbol{v}_2\to\boldsymbol{v}_4$ caused by longer delay in $\boldsymbol{v}_1\to\boldsymbol{v}_4$ than in $\boldsymbol{v}_1\to\boldsymbol{v}_2$), leading to indirect sequentially driven connection $\boldsymbol{v}_2\to\boldsymbol{v}_5$) and bidirectional causality problem ($\boldsymbol{v}_4\leftrightarrow\boldsymbol{v}_5$):

$$\begin{aligned}
v_1(t) &= 0.9v_1(t-1) - 0.3v_1(t-2) + \varepsilon_1(t),\\
v_2(t) &= 0.8v_2(t-1) - 0.5v_2(t-2) + 0.16v_1(t-1) + \varepsilon_2(t),\\
v_3(t) &= -0.2v_3(t-1) - 0.4v_3(t-2) - 0.57v_2(t-1) + \varepsilon_3(t),\\
v_4(t) &= -0.3v_4(t-1) + 0.2v_4(t-2) + 0.5v_1(t-2) + 0.4v_5(t-1) + \varepsilon_4(t),\\
v_5(t) &= 0.1v_5(t-1) + 0.3v_5(t-2) + 0.7v_4(t-1) + \varepsilon_5(t)
\end{aligned} \tag{5}$$

where $t$ stands for an index in a discrete time instance, $\varepsilon_1$, $\varepsilon_2$, $\varepsilon_3$, $\varepsilon_4$ and $\varepsilon_5$ are Gaussian white noises with zero means and variances var($\varepsilon_1$) = 1, var($\varepsilon_2$) = 0.4, var($\varepsilon_3$) = 0.7, var($\varepsilon_4$) = 0.25 and var($\varepsilon_5$) = 0.8.

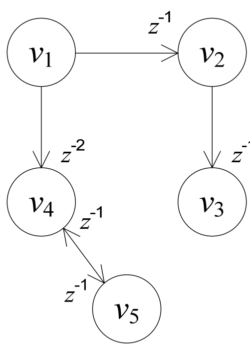



Figure 1: Scheme of MVAR signal with causal connections.

Table 1: CONDITIONAL GRANGER CAUSALITY $F_{vi \to vj}$, where $i$ corresponds to row, $j$ to column

| $F_{vi \to vj}$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| 1 | - | 0.1129 | 0.0005 | 0.7975 | 0.0015 |
| 2 | 0.0017 | - | 0.2749 | 0.0005 | 0.0010 |
| 3 | 0.0024 | 0.0024 | - | 0.0005 | 0.0010 |
| 4 | 0.0009 | 0.0007 | 0.0007 | - | 0.1793 |
| 5 | 0.0010 | 0.0008 | 0.0003 | 0.4087 | - |

Table 2: PHASE SLOPE INDEX $\psi_{vi \to vj}$, where $i$ corresponds to row, $j$ to column

| $\psi_{vi \to vj}$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| 1 | 0 | 5.6215 | 3.9274 | 30.1385 | 12.7616 |
| 2 | –5.6215 | 0 | 8.5329 | 2.4573 | 2.6105 |
| 3 | –3.9274 | –8.5329 | 0 | 0.8838 | 1.2180 |
| 4 | –30.1385 | –2.4573 | –0.8838 | 0 | 5.1342 |
| 5 | –12.7616 | –2.6105 | –1.2180 | –5.1342 | 0 |

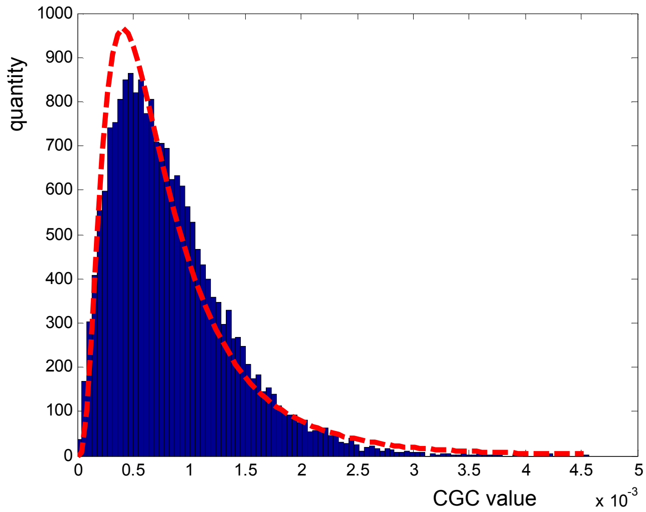
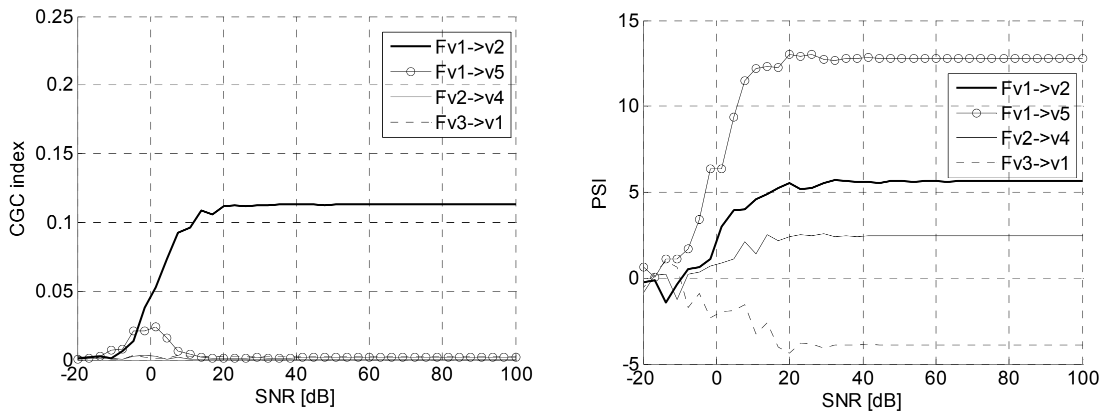



Figure 2: Histogram of CGC values for 1000 permutations of surrogate data.




(a) (b)

Figure 3: Causality analysis of data with additive Gaussian noise via (a) CGC and (b) PSI.

Conditional Granger causality between all pairs of variables was calculated for MVAR model order 5 (Table 1). Significance threshold was estimated using a random permutation test consisting of creating 1000 surrogate data sets via repeated random permutations of samples in each channel of input data in order to destroy the causal connections and calculation of CGC values of each surrogate data set leading to probability density of CGC values corresponding to null hypothesis of no causal connection (Fig. 2). This empirical distribution was fitted to log-normal distribution and corresponding CGC significance threshold for significance level $P < 0.001$ was estimated to 0.0060. Looking at Table 1, it is obvious the CGC has detected all direct causal connections and the rest of values are bellow the significance threshold value.

Phase Slope Index was calculated for each pair of channels across all frequency bins assuming epochs of 200 samples divided into 100 samples segments with 50 % overlap within each epoch (Table 2). According to [9], results with absolute value greater than 2 should be accepted as statistically significant. The results of PSI should be read differently than in the case of CGC where a positive value means connection in the direction and zero means no connection in the direction. In the case of PSI, zero means no connection in both directions, a positive value means connection in this direction and a negative value means a connection in the opposite direction. One can see the PSI recognizes all causal connections except the $\boldsymbol{v}_5 \rightarrow \boldsymbol{v}_4$ and it also cannot distinguish between direct and indirect connections.

In the next step, an additive Gaussian white noise was mixed into the data with various signal-to-noise ratios (SNR) in the range from 100 dB to –20 dB (Fig. 3). Only four connections were selected for clarity, specifically the weakest direct causality $\boldsymbol{v}_1 \rightarrow \boldsymbol{v}_2$, the strongest indirect causality $\boldsymbol{v}_1 \rightarrow \boldsymbol{v}_5$ (sequential driving problem) and $\boldsymbol{v}_2 \rightarrow \boldsymbol{v}_4$ (differently delayed driving problem) and the strongest from no causal connections $\boldsymbol{v}_3 \rightarrow \boldsymbol{v}_1$. Both methods works well up to 15 dB SNR, although PSI cannot distinguish direct and indirect connections.

## 5 Discussion

As shown in [10], PSI can behave better in some noisy conditions where Granger causality detects fake significant causalities. On the other hand, PSI cannot distinguish direct and indirect connections and cannot detect bidirectional causal connections which are very common in neurophysiological measurements.

Conditional Granger causality has a tendency to detect causal connections even in the case when they are not present. This fact has motivated us to propose an interpretation of results obtained by these two methods to combine the better properties of each. PSI can serve as an confirmation of unidirectional connections. All possible combinations of results of both methods on the same data set are shown in Table 3 where 0 stands for nonsignificant value, + or – stands for positive or negative significant value.

Table 3: COMBINATION OF CGC AND PSI ON THE SAME DATA SET AND INTERPRETATION OF RESULTS

| *CGC* 1→2 | *CGC* 2→1 | *PSI* 1→2 | *Decision* |
|---|---|---|---|
| 0 | 0 | 0 | conformity, no causal relation or nonlinear relation |
| 0 | 0 | +/– | probably indirect connection |
| + | 0 | + | conformity, unidirectional connection |
| 0 | + | – | conformity, unidirectional connection |
| + | + | 0/+/– | ? – bidirectional connection or CGC fails |
| + | 0 | 0/– | ? – CGC fails or strong indirect connection influence |
| 0 | + | 0/+ | ? – CGC fails or strong indirect connection influence |

## 6 Conclusion

We have performed experiments with causality measures in order to examine their properties on artificial data to get known their behavior in various tasks. Although the Phase Slope Index has some promising properties in noisy conditions, it fails to detect bidirectional connections which are often present among sources of brain activity. It also cannot distinguish between direct and indirect causalities. We have proposed a procedure how to interpret the results of conditional Granger causality and PSI on the same data set in order to get the best properties of each method.

## Acknowledgement

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

## References

[1] N. Wiener. *The theory of prediction*. In: Berkenbach, E. F. (Ed.), Modern mathematics for engineers. McGraw-Hill, New York, 1956.

[2] C. W. J. Granger. *Investigating causal relations by econometric models and cross-spectral methods*. Econometrica 37, 424-438, 1969.

[3] H. Liang, M. Ding, R. Nakamura, S. L. Bressler. *Causal influences in primate cerebral cortex during visual pattern discrimination*. Neuroreport, vol. 11, issue 13, 2875-2880, 2000.

[4] A. Brovelli, M. Ding, A. Ledberg, Y. Chen, R. Nakamura, S. Bressler. *Beta oscillations in a large-scale sensorimotor cortical network: Directional influences revealed by Granger causality*. Proc. Natl. Acad. Sci. USA, vol. 101, no. 26, p. 9849 – 9854, 2004.

[5] M. J. Kamiński, K. J. Blinowska. *A new method of the description of the information flow in the brain structures*. Biological cybernetics, 65(3): 203-10, 1991.

[6] A. Korzeniewska, M. Manczak, M. J. Kamiński, K. J. Blinowska, S. Kasicki. *Determination of information flow direction between brain structures by a modified Directed Transfer Function method (dDTF)*. J. Neurosci. Meth. 125, 195–207, 2003.

[7] L. A. Baccalá, K. Sameshima. *Partial directed coherence: a new concept in neural structure determination*. Biol. Cybern., vol. 84, 463-474, 2001.

[8] L. A. Baccalá, K. Sameshima, D. Y. Takahashi. *Generalized partial directed coherence*. Proc. of the 2007 15th Intl. Conf. on Digital Signal Processing, 2007.

[9] G. Nolte, A. Ziehe, V. V. Nikulin, A. Schlögl, N. Krämer, T. Brismar, K. R. Müller. *Robustly estimating the flow direction of information in complex physical systems*. Physical Review Letters 100: 234101, 2008.

[10] G. Nolte, A. Ziehe, N. Krämer, F. Popescu, K. R. Müller. *Comparison of Granger causality and phase slope index*. JMLR Workshop and Conference Proceedings 6 (NIPS 2008): 267-276, 2010.

[11] M. Ding, Y. Chen, S. L. Bressler. *Granger Causality: Basic theory and application to neuroscience*. In Winterhalder, M., Schelter, B., Timmer, J. (eds.) Handbook of Time Series Analysis, Wiley, Chichester, 2006.

---

Ing. Tomáš Bořil
e-mail: borilt@gmail.com

Prof. Ing. Pavel Sovka, CSc.
e-mail: sovka@fel.cvut.cz