GENERALIZED PARTIAL DIRECTED COHERENCE

Luiz A. Baccalá[*†]

Escola Politécnica,
Universidade de São Paulo,
Av. Prof. Luciano Gualberto, Tr.3 ,#158,
São Paulo, SP, Brazil

K. Sameshima, D. Y. Takahashi

Faculdade de Medicina,
Universidade de São Paulo,
Av. Dr. Arnaldo 455,
São Paulo, SP, Brazil

ABSTRACT

This paper re-examines the definition of **partial directed coherence** (PDC) which was recently introduced as a linear frequency-domain quantifier of the multivariate relationship between simultaneously observed time series for application in functional connectivity inference in neuroscience. The present reappraisal aims at improving PDC's performance under scenarios that involve severely unbalanced predictive modelling errors (innovations noise). The present modification turns out to be more robust in estimating imprecisions associated with finite time series samples.

Index Terms— Neural connectivity, Granger Causality, partial directed coherence, hypothesis testing, variance stabilization

1. INTRODUCTION

In [1], we introduced partial directed coherence (PDC) as a frequency domain description of the directed linear relationship between pairs of time series $x_i(n)$ and $x_j(n)$ when observed in conjunction with a set of other time series. The main motivation for PDC's introduction was its potential in disclosing crucial aspects of functional connectivity in neuroscience due to the central role played by neural rhythms (α, β, γ, etc) that are of paramount physiologic relevance [2]. Functional connectivity issues are important in neuroscience because they offer the possibilty of describing the dynamics of the interrelations between neural structures directly from the measurement of neuroelectric data [3, 4].

Conceptually, PDC is a generalization to the case of multiple time-series of Saito and Harashima's 'directed coherence' (DC) [5] which was also introduced in the context of analysing neural data and which was aimed at pinpointing the direction of information flow between isolated *pairs* of time series in a frequency domain representation of the notion of Granger Causality [6] (whereby a time-series $x_2(n)$ is said to *Granger cause* $x_1(n)$ if consideration of $x_2(n)$'s past implies significant improvement in the mean-squared prediction error $x_1(n)$) to the scenario when more than just pairs of time series are simultaneously analyzed. PDC is not the only generalization of DC to multiple time series [7, 8] (see [9] for a comparison). However, because PDC is based on the notion of partial coherence [10] its chief property is to provide a description of the mutual interaction between pairs of time series after deducting the effect of other simultaneously observed time series. As such, PDC's interest is heavily associated with highlighting the direct interaction between time series pairs that cannot be attributed to the evolution of other simultaneously observed series.

If one assumes that a set of simultaneously observed time series

$$\mathbf{x}(n) = [x_1(n)\ \dots\ x_N(n)\ ]^T$$

is adequately represented by a Multivariate Autoregressive Model of order p:

$$\mathbf{x}(n) = \sum_{k=1}^{p} \mathbf{A}_k \mathbf{x}(n-k) + \mathbf{w}(n) \tag{1}$$

where $\mathbf{A}_k$ comprise the coefficients $a_{ij}(k)$ that relate the ijseries at lag k (describing the interactions between time series pairs over time) and where

$$\mathbf{w}(n) = [w_1(n)\ \dots\ w_N(n)]^T$$

is the vector of model innovations (zero mean and with covariance matrix $\mathbf{\Sigma}_w$) leading to PDC expressed as [1]:

$$\pi_{ij}(f) = \frac{\bar{A}_{ij}(f)}{\sqrt{\sum\limits_{k=1}^{N} \bar{A}_{kj}(f)\bar{A}^*_{kj}(f)}} \tag{2}$$

where f is the normalized frequency in the interval $[-.5, .5]$ where

$$\bar{A}_{ij}(f) = \delta_{ij} - \sum_{k=1}^{p} a_{ij}(k) e^{-2j\pi f k}, \tag{3}$$

[*]Grants FAPESP 02/06925-1 and CNPq 305576/2003-8.
[†]Emails to baccala@lcs.poli.usp.br

1-4244-0882-2/07/$20.00 ©2007 IEEE

for $\delta_{ij} = 1$ whenever $i = j$ and $\delta_{ij} = 0$ otherwise and $j = \sqrt{-1}$.

The central problem in connectivity analysis is to examine the hypothesis

$$H_0 : \pi_{ij}(f) = 0, \tag{4}$$

which, if rejected, implies the existence of a direct connection leaving $x_j(n)$ to $x_i(n)$ that cannot be explained by other simultaneously observed time series. This happens because $a_{ij}(k) \neq 0$ must also hold at least for some $1 \leq k \leq p$. In fact, likelihood ratio tests over $a_{ij}(k)$ form the core of time domain approaches to connectivity appraisal under the notion of Granger Causality [8, 11].

For the remainder of this discussion, it may be useful to keep in mind that Granger Causality is not reciprocal since $a_{ij}(k) = 0$ does not imply $a_{ji}(k) = 0$, and that it is this lack of reciprocity that allows detection of possibly uni/bidirectional interactions and allows one to speak of directional connectivity.

In proposing PDC, we intended to provide a frequency domain description that decoupled interactions rooted in the sole effect of the past of the interacting time series, upon which Granger Causality [11] rests, from interactions of instantaneous nature (also known as Granger instantaneous causality [11] which are imposed by the off-diagonal terms of $\mathbf{\Sigma}_w$. This lead to definition (2) forsaking $\mathbf{\Sigma}_w$. This was actually in a simplification of a prior development which considered functions of $\mathbf{\Sigma}_w$'s terms as possible weighing functions for $\bar{A}_{ij}(f)$ in alternatives to (2) (described as PDC factor in [1]). At the time PDC was introduced, [1], the use of weighing terms in (2) lacked what now is the clear motivation considered herein.

This paper proceeds by examining a simple case where accuracy issues make the use of weighing functions essential and lead to a natural generalized definition for PDC that shares the important property of scale invariance that also characterizes some other common measures of mutual interaction like ordinary coherence and partial coherence.

2. MOTIVATING EXAMPLE

The case in point can be illustrated via the following simple model

$$\begin{bmatrix} x(n) \\ y(n) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{bmatrix} \begin{bmatrix} x(n-1) \\ y(n-1) \end{bmatrix} + \begin{bmatrix} w_x(n) \\ w_y(n) \end{bmatrix} \tag{5}$$

whose Granger causality structure is such that $y(k)$ causes $x(k)$ since $a_{12} \neq 0$ but not otherwise. In other words, this is so because the past of $y(n)$ influences $x(n)$ but not the other way round. Obviously, this should be true regardless of the "units" of either series, or of whether they are subject to static gains, i. e., if $y(n)$ is amplified by α, one would have

$$\alpha y(n) = a_{22}\alpha y(n-1) + \alpha w_y(n) \tag{6}$$

whence

$$x(n) = a_{11}x(n-1) + \frac{a_{12}}{\alpha}u(n-1) + w_x(n) \tag{7}$$

for $u(n) = \alpha y(n)$, which leaves the causality structure strictly unaltered.

Yet, as a simple calculation of PDC, as defined in (2), reveals:

$$|\pi_{xu}(f)|^2 = \frac{\left|\frac{a_{12}}{\alpha}\right|^2}{|(1 - a_{22}e^{j2\pi f})|^2 + \left|\frac{a_{12}}{\alpha}\right|^2} \to 0 \tag{8}$$

as α grows. This means that rejection of the null hypothesis of lack of Granger causality from $u(n)$ to $x(n)$, i.e. being able to reject $a_{12} = 0$, requires statistically reliable computations whose inherent variance needs to be lower than the absolute value of

$$\left|\frac{a_{12}}{\alpha}\right|^2$$

itself. This dependence on time series dynamic ranges as modified by gains obscures PDC's ability to correctly pinpoint the direction of information flow.

To remedy this, we propose the following simple modified version of PDC for this case:

$$\left|\pi^{(w)}_{xu}(f)\right|^2 = \frac{\frac{\sigma_u^2}{\sigma_x^2}\left|\frac{a_{12}}{\alpha}\right|^2}{|1 - a_{22}e^{j2\pi f}|^2 + \frac{\sigma_u^2}{\sigma_x^2}\left|\frac{a_{12}}{\alpha}\right|^2} \tag{9}$$

which is invariant to eventual gains affecting $y(n)$ as $\sigma_u = \alpha\sigma_y$ leads to

$$\left|\pi^{(w)}_{12}(f)\right|^2 = \frac{\frac{1}{\sigma_1^2}a_{12}^2}{\frac{1}{\sigma_2^2}|1 - a_{22}e^{j2\pi f}|^2 + \frac{1}{\sigma_1^2}|a_{12}|^2} \tag{10}$$

where the use of σ_i^2 refers to the variances of the innovations processes $w_i(n)$.

3. THE GENERALIZED DEFINITION

To circumvent the numerical problem associated with time series scaling, we define the new partial directed coherence estimator as:

$$\pi^{(w)}_{ij}(f) = \frac{\frac{1}{\sigma_i}\bar{A}_{ij}(f)}{\sqrt{\sum_{k=1}^{N}\frac{1}{\sigma_k^2}\bar{A}_{kj}(f)\bar{A}^*_{kj}(f)}} \tag{11}$$

Table 1. Number of realizations in which $|PDC_{12}|^2 \geq 0.1$ is a function of time series lengths (n_s) for both unaltered (first 2 columns) and pre-normalized signals (last 2 columns indicated by an overbar)

n_s	PDC_{12}	$PDC_{12}^{(w)}$	$\overline{PDC}_{12}$	$\overline{PDC}_{12}^{(w)}$
128	30	2	20	0
256	7	0	4	0
512	9	0	9	0
1024	8	0	20	0

whence it follows that

$$\left|\pi_{ij}^{(w)}(f)\right|^2 \leq 1. \tag{12}$$

and

$$\sum_{i=1}^{N} \left|\pi_{ij}^{(w)}(f)\right|^2 = 1 \tag{13}$$

Note that the new definition (11) preserves the normalizations (12,13) that also hold for the original PDC definition (2) (see [1]).

3.1. Numerical Illustration

As an example, consider the following first order model where $a_{21}(1) = 0$ implies absence of Granger causality from $x_1(n)$ to $x_2(n)$ and while all other parameters equal .5 but where innovations covariance matrix $\mathbf{\Sigma}_w$ is diagonal with distinct variance values: $\sigma_1^2 = 1$ and $\sigma_2^2 = 10$.

Use of the original definition of PDC in (2) is contrasted (Fig. 1a) to that of the present weighted PDC generalization in (11) (Fig. 1b) which is readily seen as having much smaller variability under the null hypothesis of lack of Granger causality. The results are for 100 simulated realizations comprising $n_s = 128$ time samples.

4. DISCUSSION

It has been suggested that prior normalization of $x_i(k)$ by their variances (before model estimation) could be used to circumvent the dynamic range effect that motivated the generalization suggested in (11). This is clearly not the case as observed in Table 1, where little improvement is observed if pre-normalized time series data are used in lieu of raw data. Because $PDC^{(w)}$ is invariant to scale, (see Fig. 2), prior time series data normalization results essentially superfluous.

One of the main advantages of $PDC^{(w)}$ is its hugely reduced variability vis-à-vis PDC for short time series (Table 1). This implies increased power when testing for Granger causality; this may be key for models involving

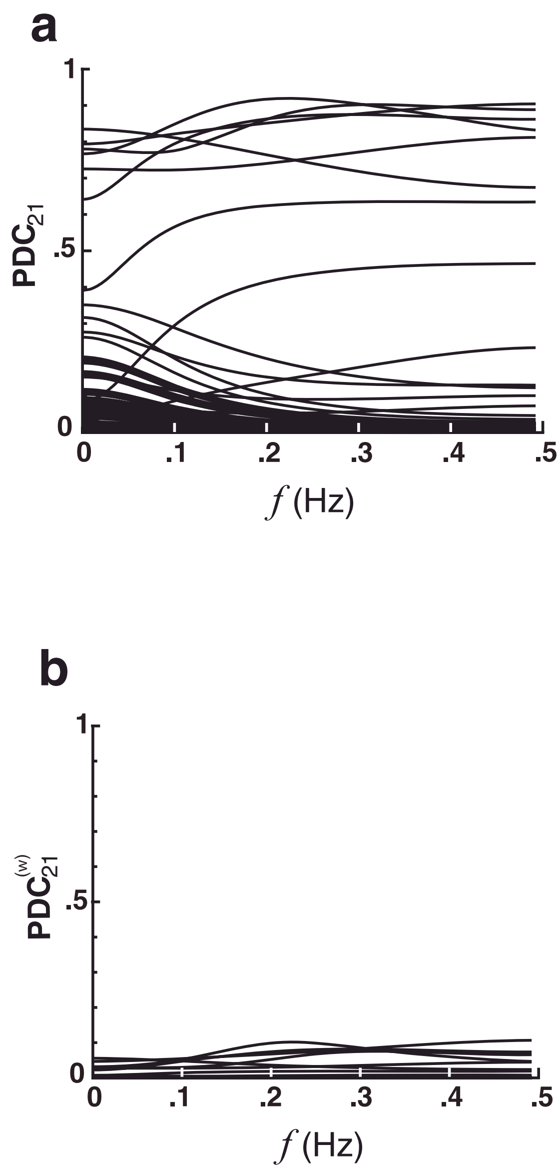

Fig. 1. (a) Illustration of the originally defined PDC as a function of frequency f for 100 time series realizations of $n_s = 128$ data points in a test of lack of Granger causality from $x_1(n)$ to $x_2(n)$ under the condition of distinct innovations noise variances described in Section 3.1 and (b) $PDC^{(w)}$ variability for the same data as in (a).

many time series which to some extent are always subject to model misspecification, at least for some of the variables involved, resulting in distinct estimated innovations noise variances.

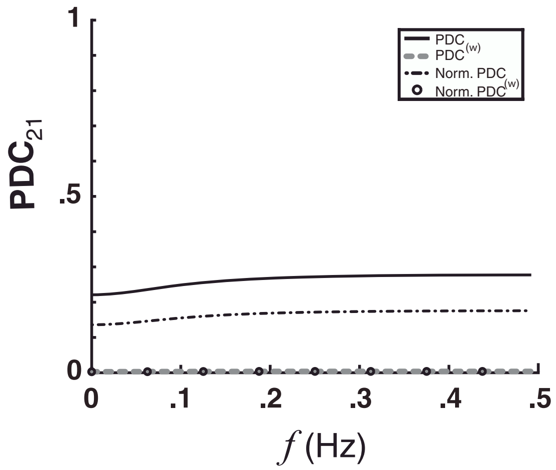

Fig. 2. Comparison between PDC and $PDC^{(w)}$ with and without time-series data pre-normalization.

In conclusion what (11) achieves is an instance of variance stabilization of the frequency domain representation of Granger causality via PDC. This fact can be shown to be of extreme importance in bootstrap-based approaches to testing for connectivity (4) (see [12]) as until very recently no knowledge of the asymptotic distribution of the original PDC (2) was available [13, 14] whose developments can be easily generalized to include (11) using the methods described in [14].

One should note that variance stabilization of the present kind is also observed when DTF (*directed transfer function* introduced by [7]) is replaced by the variance weighted version proposed in [8].

5. REFERENCES

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

[2] E. Başar, *Brain Function and Oscillations: Brain Oscillations. Principles and Approaches.*, Springer-Verlag, Berlin, 1998.

[3] E. E. Fanselow, K. Sameshima, L. A. Baccalá, and M. A. L. Nicolelis, "Thalamic bursting in rats during different awake behavioral states," *Proc. Natl. Acad. Sci. USA*, vol. 98, pp. 15330–15335, 2001.

[4] L. A. Baccalá, M. Y. Alvarenga, K. Sameshima, C. L. Jorge, and L. H. Castro, "Graph theoretical characterization and tracking of the effective neural connectivity during episodes of mesial temporal epileptic seizure.," *J. Integr. Neurosci.*, vol. 3, pp. 379–395, 2005.

[5] Y. Saito and H. Harashima, "Tracking of information within multichannel eeg record - causal analysis in eeg," in *Recent Advances in EEG and EMG Data Processing*, N. Yamaguchi and K. Fujisawa, Eds., pp. 133–146. Elsevier, Amsterdam, 1981.

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

[7] M. J. Kaminski and K. J. Blinowska, "A new method of the description of the information flow in the brain structures," *Biol. Cybern.*, vol. 65, pp. 203–210, 1991.

[8] L. A. Baccalá, K. Sameshima, G. Ballester, A. C. Valle, and C. Timo-Iaria, "Studying the interaction between brain structures via directed coherence and granger causality.," *Applied Sig. Process.*, vol. 5, pp. 40–48, 1998.

[9] L. A. Baccalá and K. Sameshima, "Overcoming the limitations of correlation analysis for many simultaneously processed neural structures," *Prog. Brain Res.*, vol. 130, pp. 33–47, 2001.

[10] S. Bendat and A. G. Piersol, *Random Data: Analysis and Measurement Procedures*, John Wiley, New York, second edition, 1986.

[11] H. Lütkepohl, *Introduction to Multiple Time Series Analysis*, Springer-Verlag, Berlin, 1993.

[12] L. A. Baccalá, D. Y. Takahashi, and K. Sameshima, "Computer intensive testing for the influence between time series," in *Handbook of Time Serie Analysis*, M. Wiltenhalter, B. Schelter, and J. Timmer, Eds. John Wiley, Berlin, 2006.

[13] B. Schelter, M. Winterhalder, M. Eichler, M. Peifer, B. Hellwig, G. Guschlbauer, C.H. B. Lucking, R. Dahlhaus, and J. Timmer, "Testing for directed influences among neural signals using partial directed coherence," *Journal of Neuroscience Methods*, vol. 152, pp. 210–219, 2006.

[14] D. Y. Takahashi, L. A. Baccalá, and K. Sameshima, "Connectivity inference via partial directed coherence: Asymptotic results.," *Journal of Applied Statistics*, p. In Press, 2006.