PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A

rsta.royalsocietypublishing.org

Research

Cite this article: Baccalá LA, de Brito CSN, Takahashi DY, Sameshima K. 2013 Unified asymptotic theory for all partial directed coherence forms. Phil Trans R Soc A 371: 20120158.
http://dx.doi.org/10.1098/rsta.2012.0158

One contribution of 13 to a Theme Issue 'Assessing causality in brain dynamics and cardiovascular control'.

Subject Areas:
biomedical engineering, statistics, applied mathematics

Keywords:
partial directed coherence, Granger causality inference, asymptotic theory, connectivity detection

Author for correspondence:
L. A. Baccalá
e-mail: baccala@lcs.poli.usp.br

† Present address: Psychology Department, Neuroscience Institute, Princeton University, Princeton, NJ 08544, USA.

Electronic supplementary material is available at http://dx.doi.org/10.1098/rsta.2012.0158 or via http://rsta.royalsocietypublishing.org.

Royal Society Publishing
Informing the science of the future

Unified asymptotic theory for all partial directed coherence forms

L. A. Baccalá[1], C. S. N. de Brito[2], D. Y. Takahashi[2,3,4,†] and K. Sameshima[4]

[1]Telecommunications and Control Department, Escola Politécnica,
[2]Bioinformatics Graduate Program,
[3]Institute of Mathematics and Statistics, and
[4]Department of Radiology and Oncology, Faculdade de Medicina, University of São Paulo, São Paulo, Brazil

This paper presents a unified mathematical derivation of the asymptotic behaviour of the three main forms of partial directed coherence (PDC). Numerical examples are used to contrast PDC, *g*PDC (generalized PDC) and *i*PDC (information PDC) as to meaning and applicability and, more importantly, to show their essential statistical equivalence insofar as connectivity inference is concerned.

1. Introduction

Partial directed coherence (PDC), a multivariate time-series technique obtained from the factorization of partial coherence [1], has become popular for evaluating the connectivity between neural structures. In the frequency domain, it closely reflects the idea of Granger causality [2], i.e. a time series $x(k)$ is Granger-caused by $y(k)$ only if knowledge of $y(k)$'s past proves helpful in predicting $x(k)$, and it is, as such, an important measure given that many neuroscience research scenarios, such as sleep staging [3], have long been linked to typical neuroelectric oscillatory behaviour.

PDC has been finding increasing applications [4–6], most of which have been carried out by comparing sample connectivities between groups classified as presenting some known disorder against normal controls. Only recently have objective trial-by-trial criteria appeared that allow the evaluation of PDC's estimator asymptotic properties. This is the case with Takahashi *et al.* [7], who confirmed a previously available connectivity hypothesis test [8] by the addition of hitherto unavailable PDC confidence intervals.

Table 1. Defining variables according to PDC type in (2.3).

variable	PDC	*g*PDC	*i*PDC
s	1	$\frac{1}{\sigma_{ii}^{1/2}}$	$\frac{1}{\sigma_{ii}^{1/2}}$
$\mathbf{S}$	$\mathcal{I}_K$	$(\mathcal{I}_K \odot \boldsymbol{\Sigma}_\mathbf{w})^{-1}$	$\boldsymbol{\Sigma}_\mathbf{w}^{-1}$

This scenario has recently been compounded by the introduction of other more general expressions for PDC, which can be more suitable under certain special circumstances. Generalized PDC (*g*PDC) [9], for example, is better suited to cases of large prediction errors or power spectral disparity between the multivariate channels. Its asymptotics have been presented without proof in reference [10].

Another, perhaps more relevant, form of PDC, termed information PDC (*i*PDC) has been introduced in references [11,12] in connection with a precise interpretation of PDC in terms of the mutual information between adequately partialized processes, thereby establishing it as a measure of direct connectivity strength.

This paper extends reference [7] to explicitly obtain the asymptotic behaviour of all PDC forms in a unified way. It starts by briefly reviewing PDC's various formulations (§2) and a statement of the asymptotic results (§3), which are numerically illustrated in §4. A final brief discussion appears in §5, leaving mathematical details to the appendix.

2. Background

Defining PDC and its variants calls for an adequately fitted multivariate autoregressive time-series (i.e. vector time-series) model of simultaneously acquired time series $\mathbf{x}(n)$ made up by $x_k(n)$, $k = 1, \ldots, K$:

$$\mathbf{x}(n) = \sum_{l=1}^{p} \mathbf{A}(l)\mathbf{x}(n-l) + \mathbf{w}(n), \tag{2.1}$$

where $\mathbf{w}(n)$ is a zero-mean white innovations process with covariance matrix $\boldsymbol{\Sigma}_\mathbf{w} = [\sigma_{ij}]$ and model order p. The $a_{ij}(l)$ coefficients from each $\mathbf{A}(l)$ matrix describe the lagged effect of the jth onto the ith series, wherefrom one can also define a frequency domain representation of (2.1) via the $\bar{\mathbf{A}}(\lambda)$ matrix whose entries are given by

$$\bar{A}_{ij}(\lambda) = \begin{cases} 1 - \sum_{l=1}^{p} a_{ij}(l)\,\mathrm{e}^{-\mathrm{j}2\pi\lambda l}, & \text{if } i = j, \\ -\sum_{l=1}^{p} a_{ij}(l)\,\mathrm{e}^{-\mathrm{j}2\pi\lambda l}, & \text{otherwise}, \end{cases} \tag{2.2}$$

where $\mathrm{j} = \sqrt{-1}$.

Denote the columns of $\bar{\mathbf{A}}(\lambda)$ by $\bar{\mathbf{a}}_j$. Then, the general expression summarizing all the forms of PDC from j to i is

$$\pi_{ij}(\lambda) = \frac{1}{s} \frac{\bar{A}_{ij}(\lambda)}{\sqrt{\bar{\mathbf{a}}_j^{\mathrm{H}}(\lambda)\mathbf{S}\bar{\mathbf{a}}_j(\lambda)}}, \tag{2.3}$$

where H denotes Hermitian transposition. Each PDC version uses its own s and $\mathbf{S}$ in (2.3) as summed up in table 1.

(a) Problem formulation

As in reference [7] for the original PDC, one must reparametrize (2.3) before applying the delta method [13], which consists of an appropriate Taylor expansion of the distribution of (2.3) with respect to $a_{ij}(r)$ and $\boldsymbol{\Sigma}_{\mathbf{w}}$, which are estimated when fitting (2.1) and whose dispersion depends on n_s, the number of available observations.

To do so, use the vec matrix-column-stacking operator on the model coefficients

$$\boldsymbol{\alpha} = \mathrm{vec}[\mathbf{A}(1)\ \mathbf{A}(2)\ \cdots\ \mathbf{A}(p)], \tag{2.4}$$

and on the residual noise covariance matrix $\boldsymbol{\sigma} = \mathrm{vec}\,\boldsymbol{\Sigma}_{\mathbf{w}}$ to produce the parameter vector $\boldsymbol{\theta}^{\mathrm{T}} = [\boldsymbol{\alpha}^{\mathrm{T}}\ \boldsymbol{\sigma}^{\mathrm{T}}]^{\mathrm{T}}$.

The delta method [13] rests on the continuous mapping nature of the parameters onto the statistics of interest (PDC) by using the asymptotic distribution parameter behaviour in terms of the number of data points n_s and the Taylor expansion of the mapping.

Lemma 2.1. *If vector time series* $\mathbf{x}(n)$ *is stationary zero-mean Gaussian, then the full model parameter vector* $\boldsymbol{\theta}$ *obeys* [14]

$$\sqrt{n_s}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \to \mathcal{N}(0, \boldsymbol{\Omega}_{\boldsymbol{\theta}}), \tag{2.5}$$

where

$$\boldsymbol{\Omega}_{\boldsymbol{\theta}} = \begin{bmatrix} \boldsymbol{\Omega}_{\boldsymbol{\alpha}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Omega}_{\boldsymbol{\sigma}} \end{bmatrix}, \tag{2.6}$$

with $\boldsymbol{\Omega}_{\boldsymbol{\alpha}} = \boldsymbol{\Gamma}_{\mathbf{x}}^{-1} \otimes \boldsymbol{\Sigma}_w$, *where* $\boldsymbol{\Gamma}_{\mathbf{x}} = E[\bar{\mathbf{x}}(n)\,\bar{\mathbf{x}}^{\mathrm{T}}(n)]$ *for*

$$\bar{\mathbf{x}}(n) = [x_1(n)\ \cdots\ x_K(n)\ \cdots\ x_1(n-p+1)\ \cdots\ x_K(n-p+1)]^{\mathrm{T}} \tag{2.7}$$

and

$$\boldsymbol{\Omega}_{\boldsymbol{\sigma}} = 2\mathbf{D}_{\mathbf{K}}\mathbf{D}_{\mathbf{K}}^{+}(\boldsymbol{\Sigma}_{\mathbf{w}} \otimes \boldsymbol{\Sigma}_{\mathbf{w}})\mathbf{D}_K^{+\mathrm{T}}\mathbf{D}_K^{\mathrm{T}}, \tag{2.8}$$

where $\mathbf{D}_K^{+}$ *is the Moore–Penrose pseudo-inverse of the standard duplication matrix.*

To simplify notation, except where essential for comprehension, i, j and λ are omitted as they are fixed with respect to (2.1) estimation.

Lemma 2.1 holds even if $\mathbf{x}(n)$ Gaussianity is relaxed [14]. In fact, finite auto- and cross-fourth-order moments of $\mathbf{w}_i(n)$ suffice with the proviso that model (2.1) adequately describes the problem.

3. Result overview

Under lemma 2.1, confidence intervals and null-hypothesis thresholds for PDC (2.3) can be computed by splitting its parameter dependence on the parameter vector $\boldsymbol{\theta}$ by considering its decomposition into numerator and denominator:

$$|\pi_{ij}(\lambda)|^2 = \pi(\boldsymbol{\theta}) = \frac{\pi_{\mathrm{n}}(\boldsymbol{\theta})}{\pi_{\mathrm{d}}(\boldsymbol{\theta})}. \tag{3.1}$$

The following results hold, thanks to the quadratic form dependence of $\pi_{\mathrm{n}}(\boldsymbol{\theta})$ and $\pi_{\mathrm{d}}(\boldsymbol{\theta})$ on $\boldsymbol{\theta}$:

1. *Confidence intervals*. For large n_s, (3.1) is asymptotically normal, i.e.

$$\sqrt{n_s}(|\hat{\pi}_{ij}(\lambda)|^2 - |\pi_{ij}(\lambda)|^2) \to \mathcal{N}(0, \gamma^2(\lambda)), \tag{3.2}$$

   where $\gamma^2(\lambda)$ is a frequency-dependent variance whose expression requires the notation introduced in the appendix.
2. *Null hypothesis threshold*. Under the null hypothesis,

$$H_0 : |\pi_{ij}(\lambda)|^2 = 0, \tag{3.3}$$

   $\gamma^2(\lambda)$ vanishes identically, and (3.2) no longer holds, requiring the next Taylor expansion term [15] in the asymptotic expression of (3.1), which has an $O(n_s^{-1})$ dependence. The

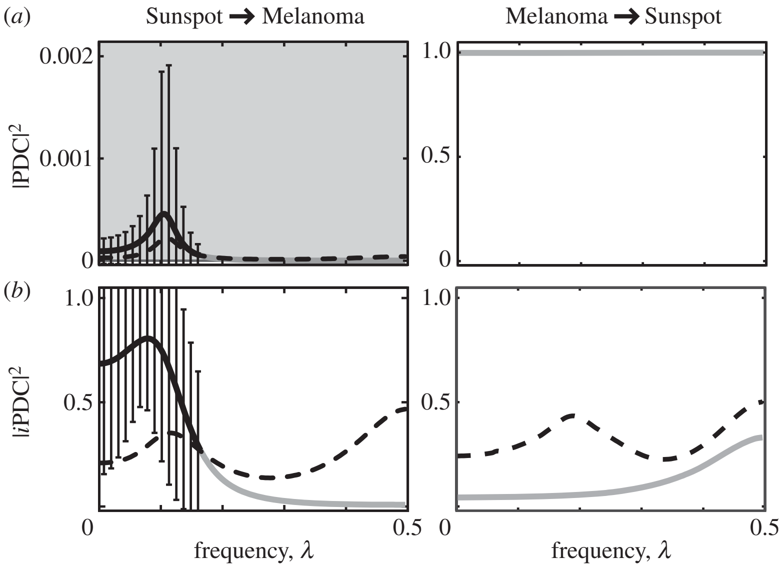

Figure 1. (*a*) Original $|\mathrm{PDC}|^2$ compared with (*b*) $|\iota\mathrm{PDC}|^2$ for example 4.1 correctly detect connection direction. Black solid lines portray significant estimates ($\alpha = 1\%$), and grey lines non-significant estimates. Original PDC Melanoma → Sunspot is below threshold even though it is close to 1. Threshold values appear as black dashed lines. Error bars Sunspot → Melanoma $|\mathrm{PDC}|^2$ depict estimated 99% confidence level interval. Note the different squared PDC magnitudes in each panel. The grey background calls attention to squared PDC magnitudes below 0.01. Conventions also apply to other figures.

resulting distribution is that of a linear combination of at most two properly weighted χ_1^2 with frequency-dependent weights:

$$n_s \bar{\mathbf{a}}_j^{\mathrm{H}}(\lambda)\mathbf{S}\bar{\mathbf{a}}_j(\lambda)(|\hat{\pi}_{ij}(\lambda)|^2 - |\pi_{ij}(\lambda)|^2) \xrightarrow{\mathrm{d}} \sum_{k=1}^{q} l_k(\lambda)\chi_1^2, \tag{3.4}$$

where $l_k(\lambda)$ (details left to the appendix) depend only on PDC's numerator $\pi_{\mathrm{n}}(\boldsymbol{\theta}(\lambda))$. PDC denominator dependence is implicit on the left-hand side of (3.4).

The χ^2 dependence comes from the quadratic behaviour of the $\boldsymbol{\theta}$ parameters, which are themselves asymptotically normal according to lemma 2.1.

4. Numerical illustrations

(a) Example 4.1: Sunspot × melanoma incidence data

Computing $|\mathrm{PDC}|^2$ for the relationship between sunspot activity and Connecticut melanoma incidence [16] over only 37 years leads to the upper panels of figure 1, whose striking feature is that the Sunspot → Melanoma PDC relationship counterintuitively turns out to be small as opposed to the Melanoma → Sunspot one. Despite its small size, closer examination reveals that Sunspot → Melanoma $|\mathrm{PDC}|^2$ is above threshold around the characteristic sunspot periodicity and that Melanoma → Sunspot $|\mathrm{PDC}|^2$, though numerically high, is below the test threshold and is, hence, not significant.

Apparent interaction size discrepancies such as this were the primary reason for introducing *g*PDC [9], and later *ι*PDC [11,12], shown in figure 1*b*, where the relative estimated sizes are more in accord with intuition and retain the correct directed inferential significance relationship. The *g*PDC results are similar to those for *i*PDC because $\boldsymbol{\Sigma}_{\mathbf{w}}$ is diagonal, and may be appreciated in reference [10].

The difference between PDC and *g*PDC/*ι*PDC stems from the different units/magnitudes of the observed quantities (sunspot number versus melanoma cases per 100 000 inhabitants). The scale invariance of *g*PDC/*ι*PDC solves this and dispenses with the common often inexplicit practice of data standardization, i.e. division of each original time series by its standard deviation (see also Baccalá *et al.* [9]).

Despite possible misleading impressions caused by the computed magnitude of the original PDC, it is important to keep in mind that connectivity inference using it is, nonetheless, correct under the appropriate statistical threshold.

Now consider the model

$$\left.\begin{aligned} x_1(n) &= 0.95\sqrt{2}x_1(n-1) - 0.9025x_1(n-2) + w_1(n),\\ x_2(n) &= 0.5x_1(n-2) + w_2(n),\\ x_3(n) &= -0.4x_1(n-3) + w_3(n),\\ x_4(n) &= -0.5x_1(n-2) + 0.25\sqrt{2}x_4(n-1) + 0.25\sqrt{2}x_5(n-1) + w_4(n)\\ \text{and}\qquad x_5(n) &= -0.25\sqrt{2}x_4(n-1) + 0.25\sqrt{2}x_5(n-1) + w_5(n) \end{aligned}\right\} \tag{4.1}$$

from reference [1] for the next two examples. In all cases, $n_s = 2000$ observations were used after a burn-in period of 1000 to ensure stationarity. The model order $p = 3$ was obtained via Akaike's information criterion.

(b) Example 4.2: Dependent inputs

Use $w_i(n) = e_i(n) + a_i e_6(n)$ in (4.1) with mutually independent white unit-variance zero-mean Gaussian time series $e_i(n)$ and realize that $e_6(n)$ simultaneously affects all series to induce correlations representing *instantaneous Granger causality* [14], something that is suited to *ι*PDC use. Different a_i produce final $x_i(n)$ that differ not only in their instantaneous coupling but also in their variances. $|\iota\text{PDC}|^2$ results for a given trial (figure 2) display correctly detected connections.

Figure 3 confirms the asymptotic statistical behaviour and is consistent with allied simulations in reference [7], for $|\text{PDC}|^2$, and in reference [10], for $|g\text{PDC}|^2$, using the same model.

(c) Example 4.3: PDC performance under coloured inputs

Now use $w_i(n) = e_i(n) + a_i e_6(n) + b_i e_7(n-1) + c_i e_7(n-2)$ in (4.1) ($e_7(n)$ is also zero-mean white unit-variance, and independent of the other $e_i(n)$). This allows investigation of the example in reference [17] by taking a_i as uniform independent random variables in the $[0, 1]$ interval and $b_i = 2$ and $c_i = 5$, leading to system inputs that are no longer white.

Contrary to PDC failure claims [17], figures 4–6 show that all PDC forms successfully infer connectivity, which is further confirmed by the time domain Wald-type Granger causality [14] *p*-value tests listed in figure 4.

The first striking difference between the various $|\text{PDC}|^2$ forms is their magnitude. The respective confidence intervals scale according to PDC magnitude.

$|\iota\text{PDC}|^2$, despite its small size, that decreases as instantaneous input power is increased, leads to correct connectivity detection. Note how the *p*-values in figure 4 decrease as $|\iota\text{PDC}|^2$ approaches the thresholds.

5. Discussion

By establishing a common PDC asymptotics framework, this paper epitomizes the effort begun in reference [7] for the original PDC and leads to rigorous *g*PDC derivation details for illustrative simulations first shown in reference [10]. The present results also apply to the *ι*PDC case [11,12].

Outside this framework is Schelter *et al.*'s PDC-related work [18], the statistical behaviour of which lies outside of the present scope.

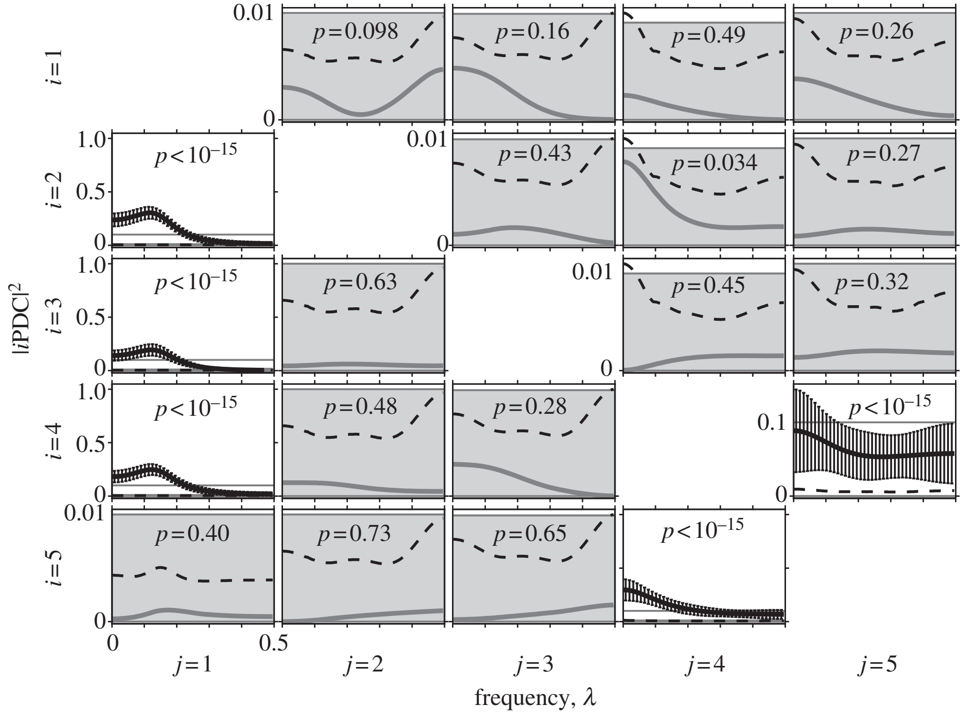

Figure 2. Illustration of $|\iota\mathrm{PDC}|^2$ results for $a_1 = 0.59$, $a_2 = 0.52$, $a_3 = 0.72$, $a_4 = 0.98$ and $a_5 = 0.66$ for example 4.2 leading to correctly detected connections from j (columns) to i (rows). Time domain Granger causality Wald tests p-values from reference [14] are also shown.

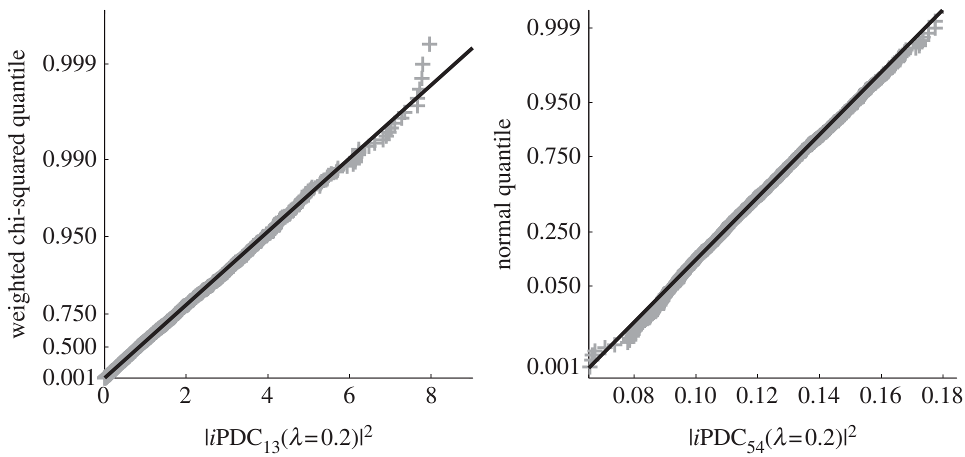

Figure 3. $|\iota\mathrm{PDC}|^2$ at $\lambda = 0.2$, example 4.2, quantile Monte Carlo performance results (grey points, 2000 realizations using a_i from figure 2) for (*a*) the non-existing connection $x_3(n) \longrightarrow x_1(n)$ and (*b*) the normally distributed existing link $x_4(n) \longrightarrow x_5(n)$. In panel (*a*) deviations from the solid theoretical line are consistent with $\alpha = 0.01$.

The derivation's unified nature implies that the PDC forms addressed herein have similar statistical behaviour and that, despite PDC magnitude issues as for the original PDC, they all correctly detect directional connections provided proper null hypothesis thresholds are used in each PDC case, and, most importantly, as long as the fitted models are good, i.e. the observed time-series durations are compatible with model orders, and the fitted residuals are white. Naturally,

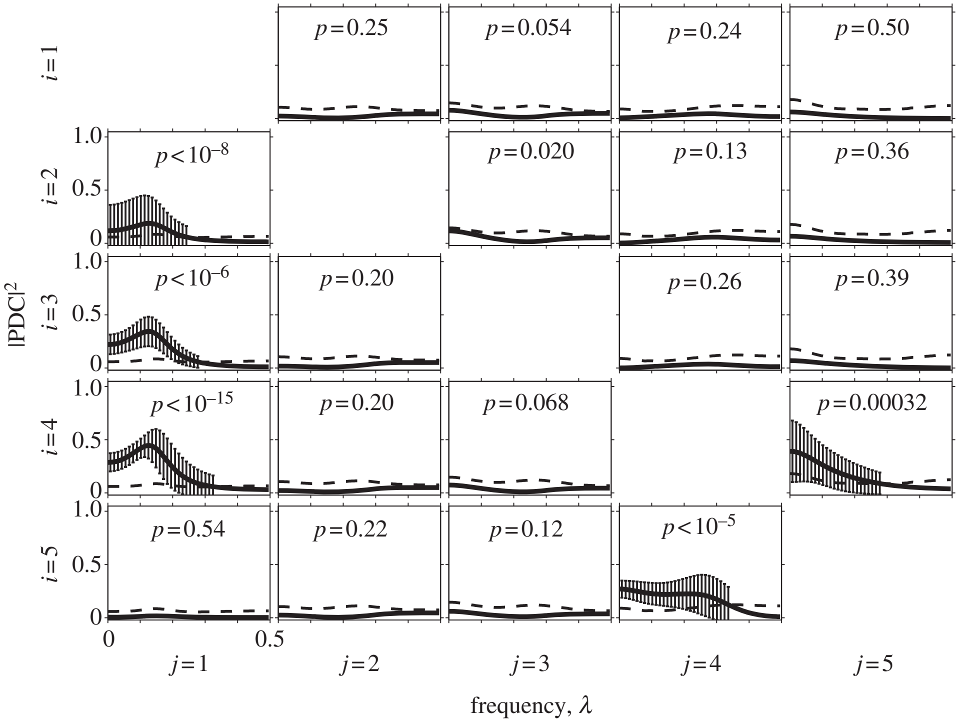

Figure 4. Original $|\mathrm{PDC}|^2$ results for example 4.3 portraying successful connectivity inference are further confirmed by p-values for the time domain Granger causality Wald tests from reference [14]. $|\mathrm{PDC}|^2$ 99% confidence intervals for the significant connections are also shown.

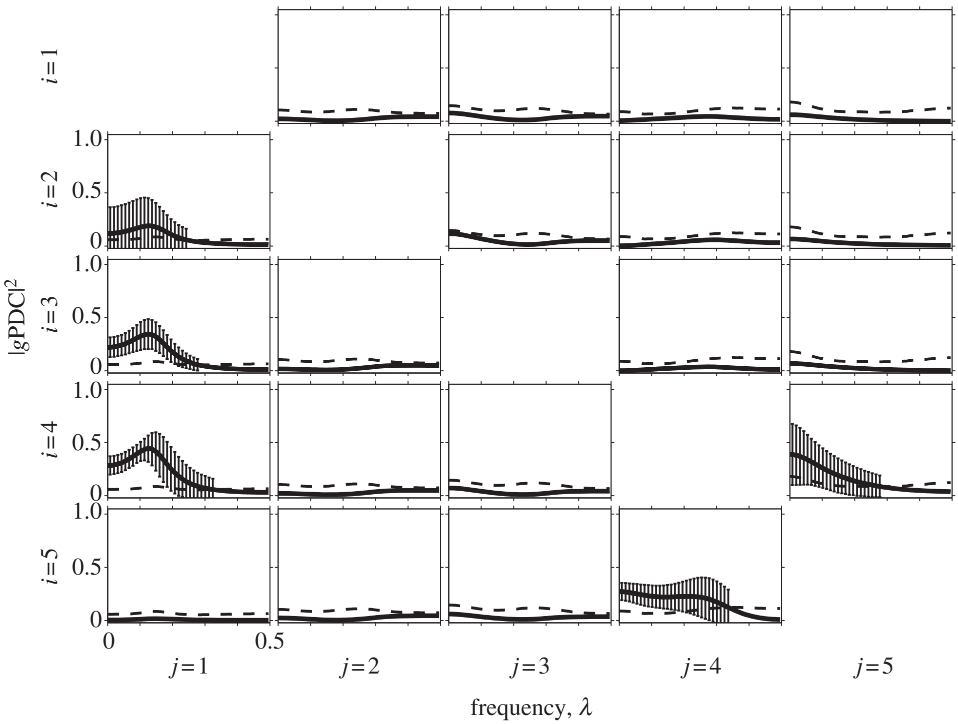

Figure 5. $|g\mathrm{PDC}|^2$ results for example 4.3 portraying successful connectivity inference. Note a confidence interval pattern similar to that of the squared PDC results (figure 4).

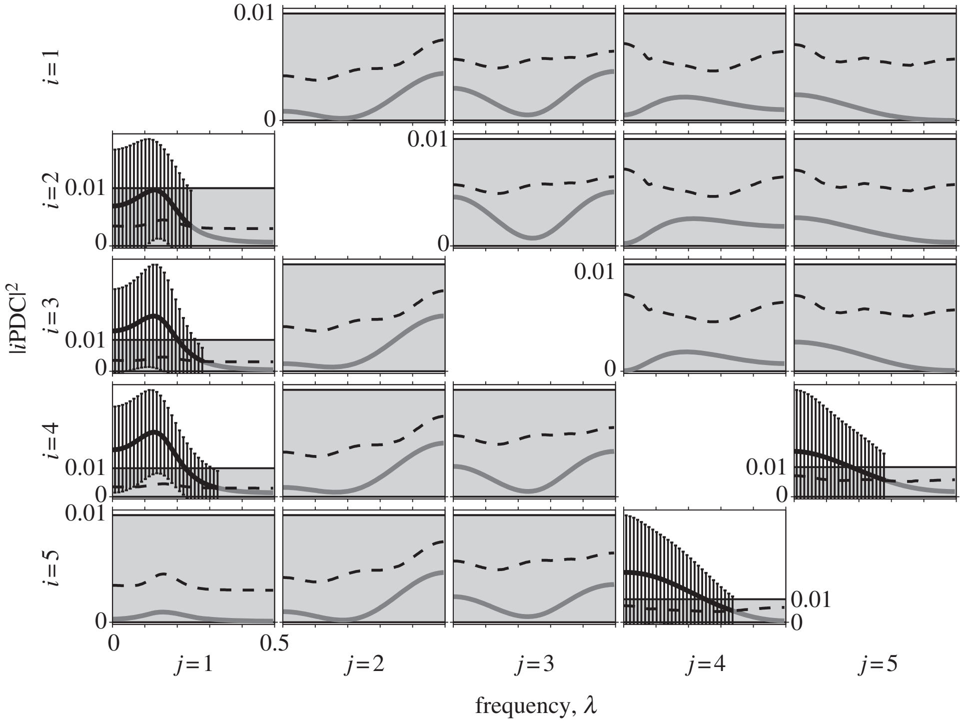

Figure 6. $|\iota\text{PDC}|^2$ results for example 4.3 portraying successful connectivity inference and confidence intervals for significant $|\iota\text{PDC}|^2$ estimates. Grey backgrounds call attention to scales below 0.01.

because of their scale invariance, *g*PDC and *i*PDC should be preferred over PDC, specially now that they, too, have objective connectivity detection criteria.

The examples in this paper were chosen to illustrate some specific points. The sunspot/melanoma example had only very few observations (merely $n_s = 37$ data points), and results were consistent even though sunspot behaviour hardly conforms to signal Gaussianity as rigorously required or to the desirable n_s large limit. This is in line with the more general validity of lemma 2.1 in a more general setting.

Example 4.2 was the most canonical one in that all theoretical premises were imposed by construction, thus permitting full appreciation of distribution quantiles. ιPDC was stressed for its objective ability to address interaction size effects in the presence of instantaneous causality.

Example 4.3 involved imposing coloured inputs whose presence can introduce biases into the estimated models. Yet, even in this case, using model parameters borrowed from other published results led to correct connectivity inference.

Because interval/threshold computations are intricate, their derivations were left to the appendix even though they constitute an important part of the present paper. Further illustrations of the examples are provided as electronic supplementary material that furthermore contains reference to the routines from the open-source AsympPDC package (available from http://www.lcs.poli.usp.br/~baccala/philosophical) used in producing the examples.

Acknowledgements. Support from FAPESP (grant no. 2005/56464-9; CInAPCe Programme), CAPES Bioinformatics Graduate Programme (D.Y.T. and C.S.N.B.), FAPESP (grant no. 2008/08171-0; D.Y.T), CNPq (grant nos 306964/2006-6 and 304404/2009-8; L.A.B.) are gratefully acknowledged.

Appendix A

In the following, function arguments are shown explicitly only when essential for comprehension and must otherwise be deduced from context.

Stating PDC in a general form requires preliminary definitions.

To include λ dependence, and avoid using complex numbers for PDC definition, consider the change of variables from $\boldsymbol{\alpha}$ to $\mathbf{a}$,

$$\mathbf{a}(\lambda) = \begin{bmatrix} \mathrm{vec}(\mathrm{Re}(\bar{\mathbf{A}}(\lambda))) \\ \mathrm{vec}(\mathrm{Im}(\bar{\mathbf{A}}(\lambda))) \end{bmatrix} = \begin{bmatrix} \mathrm{vec}(\mathcal{I}_{pK^2}) \\ \mathbf{0} \end{bmatrix} - \mathcal{C}(\lambda)\boldsymbol{\alpha}, \tag{A 1}$$

where $\mathcal{I}_N$ is an $N \times N$ identity matrix and

$$\mathcal{C}(\lambda) = \begin{bmatrix} \mathbf{C}(\lambda) \\ -\mathbf{S}(\lambda) \end{bmatrix},$$

whose blocks are $K^2 \times pK^2$ dimensional of the form

$$\mathbf{C}(\lambda) = [\mathbf{C}_1(\lambda) \cdots \mathbf{C}_p(\lambda)], \quad \mathbf{S}(\lambda) = [\mathbf{S}_1(\lambda) \cdots \mathbf{S}_p(\lambda)],$$

for

$$\mathbf{C}_r(\lambda) = \mathrm{diag}([\cos(2\pi r\lambda) \cdots \cos(2\pi r\lambda)]), \quad \mathbf{S}_r(\lambda) = \mathrm{diag}([\sin(2\pi r\lambda) \cdots \sin(2\pi r\lambda)]).$$

This reparametrizes the problem using $\bar{\boldsymbol{\theta}}(\lambda)^{\mathrm{T}} = [\mathbf{a}^{\mathrm{T}}(\lambda)\, \boldsymbol{\sigma}^{\mathrm{T}}]^{\mathrm{T}}$, i.e.

$$\pi(\bar{\boldsymbol{\theta}}) = \frac{\pi_{\mathrm{n}}(\bar{\boldsymbol{\theta}})}{\pi_{\mathrm{d}}(\bar{\boldsymbol{\theta}})} \tag{A 2}$$

with implicit i, j and λ dependence.

Lemma A.1 (α-asymptotics). *The asymptotic properties of the* $\bar{\boldsymbol{\theta}}$ *are given by*

$$\sqrt{n_s}(\hat{\bar{\boldsymbol{\theta}}} - \bar{\boldsymbol{\theta}}) \to \mathcal{N}(0, \boldsymbol{\Omega}_{\bar{\theta}}), \tag{A 3}$$

where

$$\boldsymbol{\Omega}_{\bar{\theta}} = \begin{bmatrix} \boldsymbol{\Omega}_{\mathbf{a}} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Omega}_{\sigma} \end{bmatrix} \tag{A 4}$$

with

$$\boldsymbol{\Omega}_{\mathbf{a}}(\lambda) = \mathcal{C}(\lambda)\boldsymbol{\Omega}\mathcal{C}^{\mathrm{T}}(\lambda) \tag{A 5}$$

for

$$\boldsymbol{\Omega} = \begin{bmatrix} \boldsymbol{\Omega}_{\alpha} & \boldsymbol{\Omega}_{\alpha} \\ \boldsymbol{\Omega}_{\alpha} & \boldsymbol{\Omega}_{\alpha} \end{bmatrix}.$$

The proof is straightforward by direct computation using (A 1).

To usefully write (2.3) as the (A 2) ratio, consider the $K^2 \times K^2$ matrices:

- $\mathbf{I}_{ij}$, whose entries are zero except for indices of the form $(l, m) = ((j-1)K + i, (j-1)K + i)$, which equal 1; and
- $\mathbf{I}_j$, which is non-zero only for entries whose indices are of the form (l, m) with $(j-1)K + 1 \le l = m \le jK$.

These matrices can be used to construct

$$\mathbf{I}^c_{ij} = \begin{bmatrix} \mathbf{I}_{ij} & 0 \\ 0 & \mathbf{I}_{ij} \end{bmatrix}, \quad \mathbf{I}^c_j = \begin{bmatrix} \mathbf{I}_j & 0 \\ 0 & \mathbf{I}_j \end{bmatrix},$$

which are parameter choice operators for the j to i pair of interest.

Lemma A.2. *The absolute value of squared* (2.3) *can be written as a ratio of quadratic forms*:

$$|\pi_{ij}(\lambda)|^2 = \frac{\mathbf{a}(\lambda)^{\mathrm{T}}\mathbf{I}^c_{ij}\boldsymbol{\mathcal{S}}_{\mathrm{n}}(\boldsymbol{\sigma})\mathbf{I}^c_{ij}\mathbf{a}(\lambda)}{\mathbf{a}(\lambda)^{\mathrm{T}}\mathbf{I}^c_{j}\boldsymbol{\mathcal{S}}_{\mathrm{d}}(\boldsymbol{\sigma})\mathbf{I}^c_{j}\mathbf{a}(\lambda)}, \tag{A 6}$$

where $\boldsymbol{\mathcal{S}}_{\mathrm{n}}(\boldsymbol{\sigma})$ *and* $\boldsymbol{\mathcal{S}}_{\mathrm{d}}(\boldsymbol{\sigma})$ *do not depend on* λ *and take on different values according to the desired PDC form being listed in table 2.*

Table 2. Defining variables according to PDC type in (A 6).

variable	PDC	gPDC	iPDC
$\mathcal{S}_\mathrm{n}$	$\mathcal{I}_{2K} \otimes \mathcal{I}_K$	$\mathcal{I}_{2K} \otimes (\mathcal{I}_K \odot \boldsymbol{\Sigma}_\mathbf{w})^{-1}$	$\mathcal{I}_{2K} \otimes (\mathcal{I}_K \odot \boldsymbol{\Sigma}_\mathbf{w})^{-1}$
$\mathcal{S}_\mathrm{d}$	$\mathcal{I}_{2K} \otimes \mathcal{I}_K$	$\mathcal{I}_{2K} \otimes (\mathcal{I}_K \odot \boldsymbol{\Sigma}_\mathbf{w})^{-1}$	$\mathcal{I}_{2K} \otimes \boldsymbol{\Sigma}_\mathbf{w}^{-1}$

The proof follows by direct substitution.

Thus,

$$\pi_\mathrm{n}(\boldsymbol{\theta}) = \mathbf{a}(\lambda)^\mathrm{T}\mathbf{I}_{ij}^c\boldsymbol{\mathcal{S}}_\mathrm{n}(\boldsymbol{\sigma})\mathbf{I}_{ij}^c\mathbf{a}(\lambda), \quad \pi_\mathrm{d}(\boldsymbol{\theta}) = \mathbf{a}(\lambda)^\mathrm{T}\mathbf{I}_{j}^c\boldsymbol{\mathcal{S}}_\mathrm{d}(\boldsymbol{\sigma})\mathbf{I}_{j}^c\mathbf{a}(\lambda) \tag{A 7}$$

have naturally split factors in their dependence on $\mathbf{a}$ and $\boldsymbol{\sigma}$ (table 1).

Theorem A.3 (delta method). *Let the distribution of $\mathbf{v}_n = (v_1, \ldots, v_k)^\mathrm{T}$ estimated from n observations converge in distribution as*

$$\sqrt{n}(\mathbf{v}_n - \boldsymbol{\mu}) \xrightarrow{\mathrm{d}} \mathcal{N}(0, \boldsymbol{\Sigma}_\mathbf{v}). \tag{A 8}$$

Let $g(\mathbf{v})$ be a real-valued function with continuous partials of order $m > 1$ in the neighbourhood of $\mathbf{v} = \boldsymbol{\mu}$, with all the partials of order j with $1 \le j \le m-1$ vanishing at $\mathbf{v} = \boldsymbol{\mu}$ and non-vanishing mth-order partials at $\mathbf{v} = \boldsymbol{\mu}$. Then

$$(\sqrt{n})^m(\hat{g}(\mathbf{v}_n) - g(\boldsymbol{\mu})) \xrightarrow{\mathrm{d}} \frac{1}{m!}\sum_{i_1=1}^{k}\cdots\sum_{i_m=1}^{k}\frac{\partial^m g}{\partial x_{i_1}\cdots\partial x_{i_m}}\bigg|_{x=\mu_\mathbf{v}}\prod_{j=1}^{m} Z_{i_j}, \tag{A 9}$$

with

$$\mathbf{Z} = (Z_1, \ldots, Z_k)^\mathrm{T} \sim \mathcal{N}(0, \boldsymbol{\Sigma}_\mathbf{v}). \tag{A 10}$$

From this one can readily deduce the following for large n and non-null first derivatives in (A 9).

Corollary A.4. *For a real differentiable function $g(\mathbf{v})$ asymptotically distributed as in* (A 8)*, then*

$$\sqrt{n}(\hat{g}(\mathbf{v}_n) - g(\boldsymbol{\mu})) \xrightarrow{\mathrm{d}} \mathcal{N}(0, \mathbf{g}^\mathrm{T}\boldsymbol{\Sigma}_\mathbf{v}\,\mathbf{g}) \tag{A 11}$$

is the first delta method approximation, where $\mathbf{g} = \boldsymbol{\nabla}_\mathbf{v} g$ is the gradient of $g(\mathbf{v})$ computed at $\boldsymbol{\mu}$ [15].

A.1. Confidence interval theorem

Proposition A.5. *Omitting λ and $\boldsymbol{\sigma}$ for the sake of simplicity, one has*

$$\sqrt{n_s}(|\hat{\pi}_{ij}|^2 - |\pi_{ij}|^2) \to \mathcal{N}(0, \gamma^2), \tag{A 12}$$

where n_s is the number of observations and

$$\gamma^2 = \mathbf{g_a}\boldsymbol{\Omega}_\mathbf{a}\mathbf{g}_\mathbf{a}^\mathrm{T} + \mathbf{g}_\sigma\boldsymbol{\Omega}_\sigma\mathbf{g}_\sigma^\mathrm{T}, \tag{A 13}$$

for

$$\mathbf{g_a} = 2\frac{\mathbf{a}^\mathrm{T}\mathbf{I}_{ij}^c\boldsymbol{\mathcal{S}}_\mathrm{n}\mathbf{I}_{ij}^c}{\mathbf{a}^\mathrm{T}\mathbf{I}_j^c\boldsymbol{\mathcal{S}}_\mathrm{d}\mathbf{I}_j^c\mathbf{a}} - 2\frac{\mathbf{a}^\mathrm{T}\mathbf{I}_{ij}^c\boldsymbol{\mathcal{S}}_\mathrm{n}\mathbf{I}_{ij}^c\mathbf{a}}{(\mathbf{a}^\mathrm{T}\mathbf{I}_j^c\boldsymbol{\mathcal{S}}_\mathrm{d}\mathbf{I}_j^c\mathbf{a})^2}\mathbf{a}^\mathrm{T}\mathbf{I}_j^c\boldsymbol{\mathcal{S}}_\mathrm{d}\mathbf{I}_j^c \tag{A 14}$$

and

$$\mathbf{g}_\sigma = \frac{1}{\mathbf{a}^\mathrm{T}\mathbf{I}_j^c\mathcal{S}_\mathrm{d}\mathbf{I}_j^c\mathbf{a}}[(\mathbf{I}_{ij}^c\mathbf{a})^\mathrm{T} \otimes (\mathbf{a}^\mathrm{T}\mathbf{I}_{ij}^c)]\boldsymbol{\Theta}_K\boldsymbol{\xi}_\mathrm{n} - \frac{\mathbf{a}^\mathrm{T}\mathbf{I}_{ij}^c\mathcal{S}_\mathrm{n}\mathbf{I}_{ij}^c\mathbf{a}}{(\mathbf{a}^\mathrm{T}\mathbf{I}_j^c\mathcal{S}_\mathrm{d}\mathbf{I}_j^c\mathbf{a})^2}[(\mathbf{I}_j^c\mathbf{a})^\mathrm{T} \otimes (\mathbf{a}^\mathrm{T}\mathbf{I}_j^c)]\boldsymbol{\Theta}_K\boldsymbol{\xi}_\mathrm{d}, \tag{A 15}$$

where the values of $\boldsymbol{\xi}_\mathrm{n}$ and $\boldsymbol{\xi}_\mathrm{d}$ are listed in table 3 and where $\boldsymbol{\Theta}_K = (\mathbf{T}_{2K,K} \otimes \mathbf{I}_{2K^2})(\mathbf{I}_K \otimes \mathrm{vec}(\mathbf{I}_{2K} \otimes \mathbf{I}_K))$ and $\mathbf{T}_{L,M}$ stands for the commutation matrix [14].

When $\boldsymbol{\Sigma}_\mathbf{w}$ is known *a priori* or does not need to be estimated, $\mathbf{g}_\sigma\boldsymbol{\Omega}_\sigma\mathbf{g}_\sigma^\mathrm{T}$ in (A 15) is zero.

Table 3. Defining variables according to PDC type in (A 15).

variable	PDC	*g*PDC	*i*PDC
ξ_n	0	$-\mathrm{diag}(\mathrm{vec}(\mathcal{S}_\mathrm{n}^{-2}))$	$-\mathrm{diag}(\mathrm{vec}(\mathcal{S}_\mathrm{n}^{-2}))$
ξ_d	0	$-\mathrm{diag}(\mathrm{vec}(\mathcal{S}_\mathrm{n}^{-2}))$	$-\boldsymbol{\Sigma}_\mathbf{w}^{-\mathrm{T}} \otimes \boldsymbol{\Sigma}_\mathbf{w}^{-1}$

Proof. The proof follows from theorem A.4 under lemma A.1, by computing the gradient of $\pi(\bar{\boldsymbol{\theta}})$. This is simplified as $\mathbf{a}$ and $\boldsymbol{\sigma}$ are asymptotically independent (hence $\boldsymbol{\Omega}_{\bar{\boldsymbol{\theta}}}$ is block diagonal), allowing their separate consideration.

The transpose of the required gradients can be written as

$$\frac{\partial \pi}{\partial \psi^\mathrm{T}} = \frac{1}{\pi_\mathrm{d}^2}\left[\pi_\mathrm{d}\frac{\partial \pi_\mathrm{n}}{\partial \psi^\mathrm{T}} - \pi_\mathrm{n}\frac{\partial \pi_\mathrm{d}}{\partial \psi^\mathrm{T}}\right], \tag{A 16}$$

where ψ represents either $\mathbf{a}$ or $\boldsymbol{\sigma}$ and respectively leads to $\mathbf{g}_\mathbf{a}$ and $\mathbf{g}_{\boldsymbol{\sigma}}$.

These gradients operate on the quadratic forms whose differentiation with respect to $\mathbf{a}$ yields (see reference [19], p. 175 and equation (2))

$$\frac{\partial(\mathbf{a}^\mathrm{T}\mathbf{I}^c\mathcal{S}\mathbf{I}^c\mathbf{a})}{\partial \mathbf{a}^\mathrm{T}} = 2\mathbf{a}^\mathrm{T}\mathbf{I}^c\mathcal{S}\mathbf{I}^c \tag{A 17}$$

because all appropriately valued $\mathbf{I}^c\mathcal{S}\mathbf{I}^c$ matrices are symmetric. Inserting the terms in (A 16) leads to (A 14).

The nested dependence of $\pi(\bar{\boldsymbol{\theta}})$ on $\boldsymbol{\sigma}$ calls for the use of the chain rule (see reference [19], p. 183 and equation (3)):

$$\frac{\partial(\mathbf{a}^\mathrm{T}\mathbf{I}^c\mathcal{S}\mathbf{I}^c\mathbf{a})}{\partial \boldsymbol{\sigma}^\mathrm{T}} = \frac{\partial(\mathbf{a}^\mathrm{T}\mathbf{I}^c\mathcal{S}\mathbf{I}^c\mathbf{a})}{\partial \mathcal{S}}\frac{\partial \mathcal{S}}{\partial \boldsymbol{\sigma}^\mathrm{T}}, \tag{A 18}$$

where

$$\frac{\partial(\mathbf{a}^\mathrm{T}\mathbf{I}^c\mathcal{S}\mathbf{I}^c\mathbf{a})}{\partial \mathcal{S}} = (\mathbf{I}_{ij}^c\mathbf{a})^\mathrm{T} \otimes (\mathbf{a}^\mathrm{T}\mathbf{I}_{ij}^c). \tag{A 19}$$

In its general form, $\mathcal{S} = \mathcal{I}_{2K} \otimes \bar{\mathcal{S}}$. Thus, by the chain rule (see reference [19], p. 184 and equation (11)(c)).

$$\frac{\partial \mathcal{S}}{\partial \boldsymbol{\sigma}^\mathrm{T}} = \frac{\partial \mathcal{S}}{\partial \bar{\mathcal{S}}}\frac{\partial \bar{\mathcal{S}}}{\partial \boldsymbol{\sigma}^\mathrm{T}}, \tag{A 20}$$

where

$$\frac{\partial \mathcal{S}}{\partial \bar{\mathcal{S}}} = \boldsymbol{\Theta}_K = (\mathbf{T}_{2K,K} \otimes \mathbf{I}_{2K^2})(\mathbf{I}_K \otimes \mathrm{vec}(\mathbf{I}_{2K} \otimes \mathbf{I}_K)). \tag{A 21}$$

All that is left to compute are the derivatives of $\bar{\mathcal{S}}$ in each case (table 1). For for each possible $\bar{\mathcal{S}}$ value, one has

$$\frac{\partial \mathcal{I}_K}{\partial \boldsymbol{\sigma}^\mathrm{T}} = 0, \tag{A 22}$$

which follows from lack of $\boldsymbol{\sigma}$ dependence.

When $\bar{\mathcal{S}} = \boldsymbol{\Sigma}_\mathbf{w}^{-1}$ (see reference [20], p. 183 and equation (19)) then

$$\frac{\partial \boldsymbol{\Sigma}_\mathbf{w}^{-1}}{\partial \boldsymbol{\sigma}^\mathrm{T}} = -\boldsymbol{\Sigma}_\mathbf{w}^{-\mathrm{T}} \otimes \boldsymbol{\Sigma}_\mathbf{w}^{-1}, \tag{A 23}$$

and finally let $\tilde{\mathbf{S}} = \mathcal{I}_K \odot \boldsymbol{\Sigma}_\mathbf{w}$ so that $\bar{\mathcal{S}} = \tilde{\mathbf{S}}^{-1}$ implies

$$\frac{\partial(\tilde{\mathbf{S}}^{-1})}{\partial \boldsymbol{\sigma}^\mathrm{T}} = \frac{\partial(\tilde{\mathbf{S}}^{-1})}{\partial \tilde{\mathbf{S}}}\frac{\partial \tilde{\mathbf{S}}}{\partial \boldsymbol{\Sigma}_\mathbf{w}}\frac{\partial \boldsymbol{\Sigma}_\mathbf{w}}{\partial \boldsymbol{\sigma}}, \tag{A 24}$$

which requires the chain rule twice where (see reference [19], p. 185 and equation (16))

$$\frac{\partial(\tilde{\mathbf{S}}^{-1})}{\partial\tilde{\mathbf{S}}} = -\tilde{\mathbf{S}}^{-\mathrm{T}} \otimes \tilde{\mathbf{S}}^{-1} = -(\mathcal{I}_K \odot \boldsymbol{\Sigma}_{\mathbf{w}})^{-\mathrm{T}} \otimes (\mathcal{I}_K \odot \boldsymbol{\Sigma}_{\mathbf{w}})^{-1} \tag{A 25}$$

and

$$\frac{\partial\tilde{\mathbf{S}}}{\partial\boldsymbol{\Sigma}_{\mathbf{w}}} = \mathrm{diag}(\mathrm{vec}(\mathcal{I}_K), \tag{A 26}$$

and trivially

$$\frac{\partial\boldsymbol{\Sigma}_{\mathbf{w}}}{\partial\boldsymbol{\sigma}} = \mathcal{I}_{K^2}, \tag{A 27}$$

which together in (A 24) reduce to

$$\frac{\partial\bar{\boldsymbol{\mathcal{S}}}}{\partial\boldsymbol{\sigma}} = -\mathrm{diag}(\mathrm{vec}((\mathcal{I}_K \odot \boldsymbol{\Sigma}_{\mathbf{w}})^{-2})) = -\mathrm{diag}(\mathrm{vec}(\boldsymbol{\mathcal{S}}_{\mathrm{n}}^{-2})) \tag{A 28}$$

after some additional algebra. The latter results are summarized as $\boldsymbol{\xi}_{\mathrm{n}}$ and $\boldsymbol{\xi}_{\mathrm{d}}$ quantities in table 3 and comprise (A 15).

Use of Slutsky's lemma concludes the proof by allowing the use of estimated quantities. ■

A.2. Null hypothesis test

Under the null hypothesis

$$H_0 : |\pi_{ij}|^2 = 0 \iff \mathbf{I}_{ij}^c\mathbf{a} = \mathbf{0}, \tag{A 29}$$

both (A 14) and (A 15) equal zero, and (3.2) no longer applies, so that the next Taylor term becomes necessary [15] weighted by one-half of PDC's Hessian at the point of interest with an $O(n_s^{-1})$ dependence. Via a device similar to that used in reference [7], one can show the following.

Proposition A.6. *Under* (A 29)

$$n_s(\mathbf{a}\mathbf{I}_j^c\boldsymbol{\mathcal{S}}_{\mathrm{d}}\mathbf{I}_j^c\mathbf{a})(|\hat{\pi}_{ij}|^2 - |\pi_{ij}|^2) \xrightarrow{\mathrm{d}} \sum_{k=1}^{q} l_k\chi_1^2, \tag{A 30}$$

(see (3.2)) *where* l_k *are the eigenvalues of* $\mathcal{D} = \mathbf{L}^{\mathrm{T}}\mathbf{I}_{ij}^c\boldsymbol{\mathcal{S}}_{\mathrm{n}}\mathbf{I}_{ij}^c\mathbf{L}$, *where* $\mathbf{L}$ *is the Cholesky factor of* $\bar{\boldsymbol{\Omega}}_{\mathbf{a}}$ *and* $q = \mathrm{rank}(\mathcal{D}) \leq 2$ *and reduces to* 1 *whenever* $\lambda \in \{0, \pm 0.5\}$ *or* $p = 1$ *for all* $\lambda \in [-0.5, 0.5]$.

Equation (A 30) is a linear combination of χ_1^2 variables whose weights depend on estimated parameter and covariance values appearing in $\bar{\boldsymbol{\Omega}}_{\mathbf{a}}$, which depend on λ (A 5).

Proof. Using theorem A.3 with $m = 2$ as both (A 14) and (A 15) vanish requires (A 14) and (A 15) derivation with respect to $\boldsymbol{\sigma}$ a second time. This does not alter the $\mathbf{I}_{ij}^c\mathbf{a}$ dependence and so also produces null results. The same holds when deriving (A 15) with respect to $\mathbf{a}$ because it is quadratic in $\mathbf{I}_{ij}^c\mathbf{a}$. The only non-zero surviving term comes from deriving (A 14) with respect to $\mathbf{a}$ and reduces to

$$\frac{2}{\mathbf{a}^{\mathrm{T}}\mathbf{I}_j^c\mathcal{S}_{\mathrm{d}}\mathbf{I}_j^c\mathbf{a}}\mathbf{I}_{ij}^c\mathcal{S}_{\mathrm{n}}\mathbf{I}_{ij}^c. \tag{A 31}$$

Therefore, the Hessian in (A 9) only has an upper non-zero block corresponding to (A 31) implying that only the distribution of $\mathbf{a}$ needs to be considered in writing

$$n_s(\hat{\mathbf{a}}^{\mathrm{T}}\mathbf{I}_j^c\mathcal{S}_{\mathrm{d}}\mathbf{I}_j^c\hat{\mathbf{a}})(|\hat{\pi}_{ij}|^2 - |\pi_{ij}|^2) \xrightarrow{\mathrm{d}} \mathbf{x}^{\mathrm{T}}\mathbf{I}_{ij}^c\mathcal{S}_{\mathrm{n}}\mathbf{I}_{ij}^c\mathbf{x}, \tag{A 32}$$

under theorem A.3 for $\mathbf{x} \xrightarrow{\mathrm{d}} \mathcal{N}(0, \boldsymbol{\Omega}_{\mathbf{a}})$. Slutsky's lemma concludes the first part of the proof by allowing the use of estimated quantities.

Diagonalizing $\mathbf{x}^{\mathrm{T}}\mathbf{I}_{ij}^c\mathcal{S}_{\mathrm{n}}\mathbf{I}_{ij}^c\mathbf{x}$ is done via the Cholesky factor $\mathbf{L}_{\mathbf{a}}$ from $\boldsymbol{\Omega}_{\mathbf{a}} = \mathbf{L}_{\mathbf{a}}\mathbf{L}_{\mathbf{a}}^{\mathrm{T}}$. Making $\mathbf{x} = \mathbf{L}\mathbf{y}$, where $\mathbf{L} = \mathcal{S}_{\mathrm{n}}^{1/2}\mathbf{I}_{ij}^c\mathbf{L}_{\mathbf{a}}$, yields $\mathbf{x}^{\mathrm{T}}\mathbf{I}_{ij}^c\mathcal{S}_{\mathrm{n}}\mathbf{I}_{ij}^c\mathbf{x} = \mathbf{y}^{\mathrm{T}}\mathbf{L}_{\mathbf{a}}^{\mathrm{T}}\mathbf{I}_{ij}^c\mathcal{S}_{\mathrm{n}}\mathbf{I}_{ij}^c\mathbf{L}_{\mathbf{a}}\mathbf{y} = \mathbf{y}^{\mathrm{T}}\mathbf{D}\mathbf{y}$, which means the vector elements

$\mathbf{y} = (\mathbf{L}^{\mathrm{T}}\mathbf{L})^{-1}\mathbf{L}^{\mathrm{T}}\mathbf{x}$ become mutually independent zero-mean and unit-variance so that diagonalizing $\mathbf{D} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^{\mathrm{T}}$ with $\mathbf{U}\mathbf{U}^{\mathrm{T}} = \mathbf{I}_{q\times q}$ produces

$$\mathbf{y}^{\mathrm{T}}\mathbf{D}\mathbf{y} = \sum_{k=1}^{q} l_k \mathbf{y}^{\mathrm{T}}\mathbf{u}_k\mathbf{u}_k^{\mathrm{T}}\mathbf{y} = \sum_{k=1}^{q} l_k \zeta_k^2, \tag{A 33}$$

where $\mathbf{u}_k$ is the kth column of $\mathbf{U}$. It is easy to show that the variables $\zeta_k = \mathbf{u}_k^{\mathrm{T}}\mathbf{y}$ are mutually independent, unit-variance zero-mean normal, implying ζ_k^2 are χ_1^2 random variables.

As $\operatorname{rank}(\mathbf{X}) = \operatorname{rank}(\mathbf{X}^{\mathrm{T}})$ and $\operatorname{rank}(\mathbf{XY}) \le \min(\operatorname{rank}(\mathbf{X}), \operatorname{rank}(\mathbf{Y}))$, after recalling explicit λ dependence, this implies that

$$\begin{aligned}\operatorname{rank}(\mathbf{D}) &= \operatorname{rank}(\mathbf{L}_{\mathbf{a}}^{\mathrm{T}}\mathbf{I}_{ij}^{c}\mathcal{S}_{\mathrm{n}}\mathbf{I}_{ij}^{c}\mathbf{L}_{\mathbf{a}}) = \operatorname{rank}(\mathbf{L}_{\mathbf{a}}\mathbf{L}_{\mathbf{a}}^{\mathrm{T}}\mathbf{I}_{ij}^{c}\mathcal{S}_{\mathrm{n}}\mathbf{I}_{ij}^{c}) \\ &= \operatorname{rank}(\boldsymbol{\Omega}(\lambda)_{\mathbf{a}}\mathbf{I}_{ij}^{c}\mathcal{S}_{\mathrm{n}}\mathbf{I}_{ij}^{c}) = \operatorname{rank}(\mathcal{C}(\lambda)\boldsymbol{\Omega}_{\boldsymbol{\alpha}}\mathcal{C}(\lambda)^{\mathrm{T}}\mathbf{I}_{ij}^{c}\mathcal{S}_{\mathrm{n}}\mathbf{I}_{ij}^{c})\end{aligned} \tag{A 34}$$

is upper bounded by $\operatorname{rank}(\mathbf{I}_{ij}^{c}) = 2$ and is given by $\operatorname{rank}(\mathcal{C}(\lambda)) = 1$ when $\lambda \in \{0, \pm 0.5\}$ or by $\operatorname{rank}(\mathcal{C}(\lambda)^{\mathrm{T}}\mathbf{I}_{ij}^{c}) = 1$ when $p = 1$ for all λ, thereby concluding the proof. ■

Details on obtaining the threshold values from the linear combination followed those adopted in Takahashi *et al*. [7] and references therein.

References

1. Baccalá LA, Sameshima K. 2001 Partial directed coherence: a new concept in neural structure determination. *Biol. Cybern.* **84**, 463–474. (doi:10.1007/PL00007990)
2. Granger CWJ. 1969 Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* **37**, 424–438. (doi:10.2307/1912791)
3. Steriade M, McCormick DA, Sejnowski TJ. 1993 Thalamocortical oscillations in the sleeping and aroused brain. *Science* **262**, 679–685. (doi:10.1126/science.8235588)
4. Supp GG, Schlögl A, Trujillo-Barreto N, Müller MM, Gruber T. 2007 Directed cortical information flow during human object recognition: analyzing induced EEG gamma-band responses in brain's source space. *PLoS ONE* **2**, e684. (doi:10.1371/journal.pone.0000684)
5. Cadotte AJ, Mareci TH, DeMarse TB, Parekh MB, Rajagovindan R, Ditto WL, Talathi SS, Hwang D-U, Carney PR. 2009 Temporal lobe epilepsy: anatomical and effective connectivity. *IEEE Trans. Neural Syst. Rehabil. Eng.* **17**, 214–223. (doi:10.1109/TNSRE.2008.2006220)
6. Sato JR, Takahashi DY, Arcuri SM, Sameshima K, Morettin PA, Baccalá LA. 2009 Frequency domain connectivity identification: an application of partial directed coherence in fMRI. *Hum. Brain Mapp.* **30**, 452–461. (doi:10.1002/hbm.20513)
7. Takahashi DY, Baccalá LA, Sameshima K. 2007 Connectivity inference between neural structures via partial directed coherence. *J. Appl. Stat.* **34**, 1259–1273. (doi:10.1080/02664760701593065)
8. Schelter B, Winterhalder M, Eichler M, Peifer M, Hellwig B, Guschlbauer B, Lücking CH, Dahlhaus R, Timmer J. 2006 Testing for directed influences among neural signals using partial directed coherence. *J. Neurosci. Methods* **152**, 210–219. (doi:10.1016/j.jneumeth.2005.09.001)
9. Baccalá LA, Takahashi DY, Sameshima K. 2007 Generalized partial directed coherence. In *Proc. 15th Int. Conf. Digital Signal Processing*, pp. 163–166. (doi:10.1109/ICDSP.2007.4288544)
10. de Brito CSN, Baccalá LA, Takahashi DY, Sameshima K. 2010 Asymptotic behavior of generalized partial directed coherence. In *Proc. Eng. Med. Biol. Soc. (EMBC), 2010 Annual Int. Conf. IEEE*, pp. 1718–1721. (doi:10.1109/IEMBS.2010.5626856)
11. Takahashi DY, Baccalá LA, Sameshima K. 2010 Frequency domain connectivity: an information theoretic perspective. In *Proc. Eng. Med. Biol. Soc. (EMBC), 2010 Annual Int. Conf. IEEE*, pp. 1726–1729. (doi:10.1109/IEMBS.2010.5626700)
12. Takahashi DY, Baccalá LA, Sameshima K. 2010 Information theoretic interpretation of frequency domain connectivity measures. *Biol. Cybern.* **103**, 463–469. (doi:10.1007/s00422-010-0410-x)
13. van der Vaart AW. 1998 *Asymptotic statistics*. Cambridge, UK: Cambridge University Press.
14. Lütkepohl H. 2005 *New introduction to multiple time series analysis*. New York, NY: Springer.
15. Serfling RJ. 1980 *Approximation theorems of mathematical statistics*. New York, NY: John Wiley.

16. Andrews DF, Herzberg AM. 1985 *Data: a collection of problems from many fields for the student and research worker*. New York, NY: Springer.
17. Guo S, Wu J, Ding M, Feng J. 2008 Uncovering interactions in the frequency domain. *PLoS Comput. Biol.* **4**, e1000087. (doi:10.1371/journal.pcbi.1000087)
18. Schelter B, Timmer J, Eichler M. 2009 Assessing the strength of directed influences among neural signals using renormalized partial directed coherence. *J. Neurosci. Methods* **179**, 121–130. (doi:10.1016/j.jneumeth.2009.01.006)
19. Lütkepohl H. 1996 *Handbook of matrices*. Chichester, UK: John Wiley.
20. Magnus JR, Neudecker H. 1999 *Matrix differential calculus with applications in statistics and econometrics*, 2nd edn. New York, NY: John Wiley.