International Journal of Bioelectromagnetism *http://www.ijbem.org*
Vol. 10, No. 1, pp. 31 - 36, 2008

Partial Directed Coherence Asymptotics for VAR Processes of Infinite Order

Daniel Y Takahashi[a,b], Luiz A Baccalá[a], Koichi Sameshima[a,b]
[a]*University of São Paulo, São Paulo, Brazil,* [b]*IINN-ELS/IEP-HSL, São Paulo, Brazil*

Correspondence: K Sameshima, Departamento de Radiologia, Faculdade de Medicina, University of São Paulo, CEP 01246-903, São Paulo, Brazil. E-mail: ksameshi@usp.br, phone +(55)11-30698139.

Abstract. Currently available asymptotic confidence interval/null hypothesis threshold results for partial directed coherence (PDC) are strictly valid for vector autoregressive (VAR) processes of finite order p. The present paper discusses the extension to more general situations when VAR(p) are used as approximations to more general VAR processes of infinite order, in which case the order of the best fitted model becomes a function of the number of available data points and the resulting confidence intervals must be rescaled appropriately.

Keywords: Partial Directed Coherence; Neural Connectivity; Granger Causality; VAR model order

1. Introduction

Introduced recently as a tool for pinpointing the direct link between neural structures [Baccalá and Sameshima, 2001], PDC is based on the factorization of partial coherence [Bendat and Piersol, 1986] by means of fitting VAR models to multivariate neural activity data. The growing literature on PDC [Sameshima and Baccalá, 1999; Fanselow et al., 2001; Baccalá et al., 2004, Supp et al. 2005; Astolfi et al., 2006; Astolfi et al., 2007; Supp et al. 2007] and other connectivity inference methods [Kaminski and Blinowska, 1991; Kaminski et al., 2001; Yamashita et al., 2005] has a visible application oriented bias. Instead, the theoretical investigation of their statistical adequacy is relatively scarce even in ideal situations amenable to analytical investigation.

Rigorous asymptotic confidence interval results for PDC have been obtained recently [Takahashi et al., 2007] by assuming that the underlying data generating processes were VAR of finite known order, a fact that is seldom valid in practice. In fact, in many cases, even simple vector moving average data generating mechanisms are physiologically plausible and are mathematically equivalent to VAR models of infinite order [Lütkepohl, 1993]. In addition, VAR processes of infinite order arise in more general situations making their consideration important from a practical standpoint.

Here we discuss how to generalize the former rigorous results to encompass the infinite order VAR process case under mild conditions that guarantees PDC convergence by fitting VAR models whose order depends on the signal duration in terms of the number of available time samples.

The present paper also states the allied results (modified from [Schelter et al. 2006; Takahashi et al., 2007] for null hypothesis testing of zero connectivity.

After notational preliminaries, the formal results are stated without proof followed by a numerical example.

2. Preliminaries

Consider that $\mathbf{x}(k)=[x_1(k),\ldots,x_N(k)]^T$, $1\le k\le n$, is a simultaneously observed jointly second-order stationary time series with VAR(p) representation, *i.e.*,

$$\mathbf{x}(k)=\sum_{r=1}^{p}\mathbf{A}_r\mathbf{x}(k-r)+\mathbf{w}(k), \qquad (1)$$

where the $a_{ij}(r)$ coefficient in $\mathbf{A}_r$ describes the linear relationship between time series $x_i(k)$ and $x_j(k)$ at the r -th past lag, and $\mathbf{w}(k)$ represent the driving innovations.

The allied partial directed coherence from $x_j(k)$ to $x_i(k)$ is given by [Sameshima and Baccalá, 1999; Baccalá and Sameshima, 2001] as

$$\pi_{ij}(\lambda) = \frac{\overline{A}_{ij}(\lambda)}{\sqrt{\sum_{n=1}^{N} \overline{A}_{nj}(\lambda)\overline{A}_{nj}^{*}(\lambda)}}, \tag{2}$$

where

$$\overline{A}_{ij}(\lambda) = \delta_{ij} - \sum_{r=1}^{p} a_{ij}(r)\exp(-2\sqrt{-1}\pi\lambda r), \tag{3}$$

with $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ otherwise.

To compactly state the asymptotic statistical behaviours for (2) when the true data generation process is a stable, potentially infinite order VAR process for the sample size n (duration), approximated by (1), it is convenient to introduce the following notation:

$$\overline{\boldsymbol{\alpha}}(\lambda) = vec(\mathbf{I}) - \sum_{r=1}^{p} vec(\mathbf{A}_r)\exp(-2\sqrt{-1}\pi\lambda r),$$

where $\mathbf{I}$ is an $N \times N$ identity matrix and vec stands for the usual matrix column stacking operator. Also let

$$\mathbf{a}(\lambda) = \begin{bmatrix} Re(\overline{\boldsymbol{\alpha}}(\lambda)) \\ Im(\overline{\boldsymbol{\alpha}}(\lambda)) \end{bmatrix}.$$

Also let

$$\mathbf{I}_{ij}^{c} = \begin{bmatrix} \mathbf{I}_{ij} & 0 \\ 0 & \mathbf{I}_{ij} \end{bmatrix}$$

where the $N^2 \times N^2$ matrix $\mathbf{I}_{ij}$ is made by zeros except for the entry $(l,m) = ((j-1)N + i, (j-1)N + i)$, which equals 1.

Likewise

$$\mathbf{I}_{j}^{c} = \begin{bmatrix} \mathbf{I}_{j} & 0 \\ 0 & \mathbf{I}_{j} \end{bmatrix}$$

contains $N^2 \times N^2$ blocks $\mathbf{I}_j$ with zeros except for $(l,m) : (j-1)N + 1 \le l = m \le jN$.

This allows rewriting $|\pi_{ij}(\lambda)|^2$ as a ratio of quadratic forms of real variables

$$|\pi_{ij}(\lambda)|^2 = \frac{\mathbf{a}^T(\lambda)\mathbf{I}_{ij}^{c}\mathbf{a}(\lambda)}{\mathbf{a}^T(\lambda)\mathbf{I}_{j}^{c}\mathbf{a}(\lambda)},$$

Consider also the $N^2 \times pN^2$ dimensional matrices

$$\mathbf{C}(\lambda) = [\mathbf{C}_1(\lambda) \ldots \mathbf{C}_p(\lambda)]$$

and

$$\mathbf{S}(\lambda) = [\mathbf{S}_1(\lambda) \ldots \mathbf{S}_p(\lambda)],$$

such that

$$\mathbf{C}_r(\lambda) = diag([cos(2\pi r\lambda) \ldots cos(2\pi r\lambda)])$$

and

$$\mathbf{S}_r(\lambda) = diag([sin(2\pi r\lambda) \ldots sin(2\pi r\lambda)]).$$

Finally

$$\mathsf{C}(\lambda)=\begin{bmatrix}\mathbf{C}(\lambda) & \mathbf{0}\\ \mathbf{0} & -\mathbf{S}(\lambda)\end{bmatrix}.$$

3. Results

Under the condition that the $\mathbf{x}(k)$ multivariate data is produced by a potentially infinite order VAR process with absolutely summable $\mathbf{A}_r$, i.e.,

$$\sum_{r=1}^{\infty}\|\mathbf{A}_r\|\leq\infty$$

and is canonically represented as a vector moving average process (VMA)

$$x(t)=\sum_{r=0}^{\infty}\mathbf{H}_r w(t-r)$$

for

$$det\left[\sum_{r=0}^{\infty}\mathbf{H}_r z^r\right]\neq 0 \text{ for } |z|\leq 1 \text{ and } \sum_{r=0}^{\infty}r^{1/2}\|\mathbf{H}_r\|\leq\infty.$$

then

Proposition 1 *If a finite order VAR(p_n) process is fitted by multivariate least squares so that its order p_n depends upon the sample size n under the conditions that*

$$p_n\to\infty,\qquad p_n^3/n\to 0,\qquad \sqrt{n}\sum_{r=p_{n+1}}^{\infty}\|\mathbf{A}_r\|\to 0\quad \text{as } n\to\infty.$$

and $|\pi_{ij}(\lambda)|^2\neq 0$ *nor* $|\pi_{ij}(\lambda)|^2\neq 1$.

Then PDC's estimator $|\hat{\pi}_{ij}(\lambda)|^2$ obtained by substituting the estimated values in (2) and (3) is consistent and asymptotically distributed as

$$\sqrt{n-p_n}\,\gamma^{-1}(\hat{\mathbf{a}}(\lambda))(|\hat{\pi}_{ij}(\lambda)|^2-|\pi_{ij}(\lambda)|^2)\xrightarrow{d}\mathsf{N}(0,1),$$

for

$$\gamma^2(\hat{\mathbf{a}}(\lambda))=\mathbf{G}(\hat{\mathbf{a}}(\lambda))^T\hat{\overline{\mathbf{\Omega}}}(\lambda)\mathbf{G}(\hat{\mathbf{a}}(\lambda))$$

and

$$\mathbf{G}(\hat{\mathbf{a}}(\lambda))=2\mathbf{I}_{ij}\hat{\mathbf{a}}(\lambda)\frac{1}{\hat{\mathbf{a}}^T(\lambda)\mathbf{I}_j\hat{\mathbf{a}}(\lambda)}-2\mathbf{I}_j\hat{\mathbf{a}}(\lambda)\frac{\hat{\mathbf{a}}^T(\lambda)\mathbf{I}_{ij}\hat{\mathbf{a}}(\lambda)}{(\hat{\mathbf{a}}^T(\lambda)\mathbf{I}_j\hat{\mathbf{a}}(\lambda))^2}.$$

In the above equations

$$\overline{\mathbf{\Omega}}(\lambda)=\mathsf{C}(\lambda)\left(\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}\otimes\mathbf{\Gamma}_{\mathbf{x}}(0)^{-1}\otimes\mathbf{\Sigma}_{\mathbf{w}}\right)\mathsf{C}^T(\lambda),$$

where $\mathbf{\Gamma}_{\mathbf{x}}(0)$ and $\mathbf{\Sigma}_{\mathbf{w}}$ stand, respectively, for the autocovariance matrix of $\mathbf{x}(k)$ and the covariance matrix of $\mathbf{w}(k)$. The hat above the variables indicates the associated least-squares estimates.

Proposition 2 *When $|\pi_{ij}(\lambda)|^2=0$ or $|\pi_{ij}(\lambda)|^2=1$, under the conditions of Proposition 1,*

$$(n-p_n)(\hat{\mathbf{a}}^T(\lambda)\mathbf{I}_j\hat{\mathbf{a}}(\lambda))(|\hat{\pi}_{ij}(\lambda)|^2-|\pi_{ij}(\lambda)|^2)\xrightarrow{d}\sum_{k=1}^{q}l_k(\lambda)\chi_1^2,$$

where $l_k(\lambda)$ are the eigenvalues of $\mathbf{D}(\lambda)=\mathbf{L}^T(\lambda)\mathbf{I}_{ij}\mathbf{L}(\lambda)$, in which the matrix $\mathbf{L}(\lambda)$ is the

Choleski factor in

$$\overline{\mathbf{\Omega}}(\lambda) = \mathbf{L}(\lambda)\mathbf{L}^T(\lambda),$$

where $q = rank(\mathbf{D}) = 2$ *, unless* $\lambda \in \{0, \pm 0.5\}$ *or* $p = 1$ *leading to* $q = 1$.

The proofs of these results employ the asymptotics in [Lütkepohl, 1993] and follow the procedures adopted in [Takahashi et al., 2007]. The validity of Proposition 2 for $|\pi_{ij}(\lambda)|^2 = 1$ is a new result.

Numerical Example

To illustrate the validity of the results, an invertible VMA process with an infinite order VAR process representation was simulated. Proposition 1 was used to compute the approximate confidence interval and Proposition 2 to calculate the test statistics under the null hypothesis of $|\pi_{ij}(\lambda)|^2 = 0$. The simulated VMA is given by:

$$\begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} w_1(k) \\ w_2(k) \end{bmatrix} + \begin{bmatrix} .7 & .9 \\ 0 & .4 \end{bmatrix} \begin{bmatrix} w_1(k-1) \\ w_2(k-1) \end{bmatrix},$$

where $w_i(k)$ are mutually uncorrelated standard Gaussian innovations and corresponds to the VAR(∞)

$$\begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} = \sum_{m=1}^{\infty} \begin{bmatrix} -(-.7)^m & 3((-.4)^m - (-.7)^m) \\ 0 & -(-.4)^m \end{bmatrix} \begin{bmatrix} x_1(k-m) \\ x_2(k-m) \end{bmatrix} + \begin{bmatrix} w_1(k) \\ w_2(k) \end{bmatrix},$$

wherefrom the only existing unidirectional connectivity is clearly from $x_2(k)$ to $x_1(k)$.

The computed approximate confidence intervals and tests for the nullity of PDC for $n = 20, 200, 2000$ observed points are shown in Figure 1. Note that in practice, the conditions on the growth of p_n are difficult to verify and criteria like AIC must be used. Thus, in the present example, the computed confidence intervals and statistical tests were calculated for the VAR(p) model whose order p was estimated by AIC and lead to reasonable results even for as few as 20 data samples.

A computational MATLAB program that calculates PDC and its statistics for a given data set can be obtained at **http://www.lcs.poli.usp.br/~baccala/pdc/**.

4. Discussion

As should possibly be expected, the present result compared with [Takahashi et al., 2007] leads to slightly slacker confidence intervals, whose slackness depends on p_n. If n is large the new limits obtained are virtually identical to the old ones in [Takahashi et al., 2007].

The assumption that time series data results from a finite order VAR process is very restrictive, if not unrealistic, thus the rigorous inclusion of infinite order VAR processes considerably enhances the theoretical scope of applicability of the available statistical asymptotic results [Schelter et al., 2006; Takahashi et al., 2007]. Furthermore, these results show the consistency of applying these considerations to finite order VAR(p) processes without requiring explicit *a priori* knowledge of its order. This is essentially equivalent to a nonparametric result in the sense that the finite parameter model, with an increasing number of data points, is used to estimate an infinite parameter model.

Also of note is that this result can justify the use of VAR models in the inference of the relationship between some deterministic nonlinear processes as in [Schelter et al., 2006] if a very large number of points is available and a VAR model of sufficiently high order is fitted. This possibility is theoretically justified because these processes can be represented as jointly nondeterministic linear processes [Priestley, 1981].

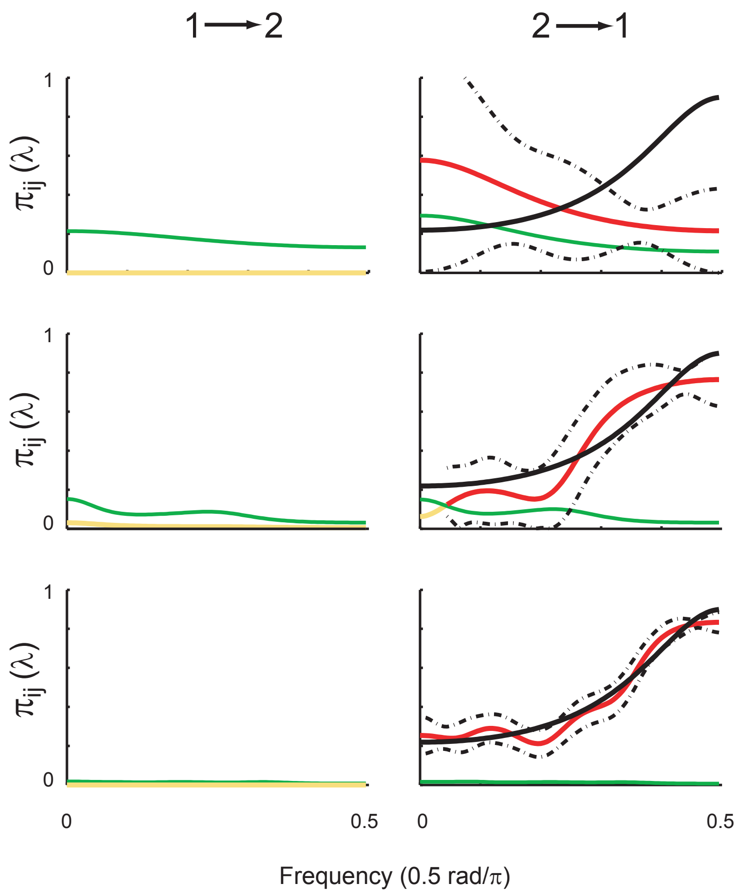

Figure 1. Results of simulations for 20 (upper panels), 200 (middle panels) and 2000 (bottom panels) time points. The solid green line indicates the threshold value under the hypothesis of zero PDC (α = 5%) and the dotted black line indicates the 95% confidence intervals. The solid black line is the true value for the simulated model, the red line is the estimated value when above the null hypothesis threshold and the yellow line stands the estimated values below the thresholds.

5. Conclusion

Rigorous asymptotic distributions for estimated PDC statistics were obtained and illustrated via the simple example of a VMA process.

Acknowledgements

This work was supported in part by CAPES (DYT), FAPESP-CInAPCe Project (LAB, KS), AASDAP, CNPq (LAB, KS) and FAPESP grant 02/06925-1 (LAB).

References

Astolfi L, Cincotti F, Mattia D, Marciani MG, Baccalá LA, de Vico Fallani F, Salinari S, Ursino M, Zavaglia M, Babiloni F. Assessing cortical functional connectivity by partial directed coherence: simulations and application to real data. *IEEE Transaction on Biomedical Engeneering*, 53(9):1802-1812, 2006.

Astolfi L, Bakardjian H, Cincotti F, Mattia D, Marciani MG, De Vico Fallani F, Colosimo A, Salinari S, Miwakeichi F, Yamaguchi Y, Martinez P, Cichocki A, Tocci A, Babiloni F. Estimate of causality between independent cortical spatial patterns during movement volition in spinal cord injured patients. *Brain Topography*, 19(3):107-123, 2007.

Baccalá LA, Sameshima K. Partial directed coherence: a new concept in neural structure determination. *Biological Cybernetics*, 84(6): 463–474, 2001.

Baccalá LA, Alvarenga MY, Sameshima K, Jorge CL, Castro LH. Graph theoretical characterization and tracking of the effective neural connectivity during episodes of mesial temporal epileptic seizure. *Journal of Integrative Neuroscience*, 3(4): 379–395, 2004.

Fanselow EE, Sameshima K, Baccalá, LA, Nicolelis MAL. Thalamic bursting in rats during different awake behavioral states. *Proceedings of the National Academy of Sciences of the United States of America*, 98(26): 15330–15335, 2001.

Kamiński MJ, Blinowska KJ. A new method of the description of the information flow in the brain structures. *Biological Cybernetics*, 65(3):203-210, 1991.

Kamiński MJ, Ding M, Truccolo WA, Bressler SL. Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance. *Biological Cybernetics*, 85(2):145-157, 2001.

Lütkepohl H. Introduction to Multiple Time Series Analysis. Springer-Verlag, Berlin, 1993.

Priestley MB. Spectral Analysis and Time Series, Two-Volume Set, Volume I and II Academic Press, London, 1981.

Sameshima K, Baccalá LA. Using partial directed coherence to describe neuronal ensemble interactions. *Journal of Neuroscience Methods*, 94(1): 93–103, 1999.

Schelter B, Winterhalder M, Eichler M, Peifer M, Hellwig B, Guschlbauer B, Lucking, C, Dahlhaus R, Timmer J. Testing for directed influences among neural signals using partial directed coherence. *Journal of Neuroscience Methods*, 152(1-2): 210–219, 2006.

Supp GG, Schlögl, A, Fiebach CJ, Gunter TC, Vigliocco G, Pfurtscheller G, Petsche H. Semantic memory retrieval: Cortical couplings in object recognition in the n400 window. *European Journal of Neuroscience*, 21(4): 1139–1143, 2005.

Supp GG, Schlögl A, Trujillo-Barreto N, Müller MM, Gruber T. Directed cortical information flow during human object recognition: analyzing induced EEG gamma-band responses in brain's source space. *Public Library of Science ONE*, 2(1):e684, 2007.

Yamashita O, Sadato N, Okada T, Ozaki T. Evaluating frequency-wise directed connectivity of BOLD signals applying relative power contribution with the linear multivariate time-series models. *Neuroimage*, 25(2):478-490, 2005.

Takahashi DY, Baccalá LA, Sameshima K. Connectivity inference via partial directed coherence. *Journal of Applied Statistics*, in press, 2007.