# On the Efficient Computation of Partial Coherence from Multivariate Auto-Regressive Models

Luiz A. Baccalá

Telecommunications and Control Engineering Department

Escola Politécnica of the University of São Paulo,

Av. Prof. Luciano Gualberto, Trav. 3, #158,

São Paulo, SP, Brazil CEP 05508-900.

e-mail: baccala@lcs.poli.usp.br

## Abstract

Recent approaches to EEG signal analysis have increasingly employed joint multivariate autoregressive (MAR) models to describe relations between signals collected from different electrodes. Among descriptors of pairwise relations, partial coherences stand out as they exclude the common influences of other simultanouesly observed signals and thereby decompose the analysis into sets of direct mutual influences between signal pairs. Here we describe an efficient method to reduce the complexity of partial coherence computation when MAR models are available.

**keywords:**
partial coherence, spectral analysis, autoregressive models, multiple time series, wedge product

## 1 Introduction

Recent applications of spectral estimation to the problem of signal pathway analysis in the brain [1, 2, 3, 4] have renewed interest in the computation of partial coherences [5] between pairs of time series when many simultaneously time series are recorded and analyzed simultaneously. This interest is justified by the distinctive role this quantity plays in isolating the exclusive mutual relations between two time series through suppressing those effects that can be attributed to other simultaneously measured signals [5] . Even though most published algorithms for its computation rely on estimates of the joint-spectral density matrix $\mathbf{S}(f)$, we show here that whenever Multivariate Auto Regressive (MAR) models are available significant computational savings can be achieved that altogether dispense with the prior computation of $\mathbf{S}(f)$. This matter assumes added relevance in view of the growing popularity of MAR models specially in EEG analysis [1, 6].

## 2 Background

Given a set of time series $x_k(t)$ $,i = 1 \dots N$, the partial coherence $\gamma_{ij}^2(f)$ between the pair of time series $x_i(t)$ and $x_j(t)$ measures the adequacy, in terms of mean squared error prediction power, of a linear model $T_{ij}(f)$ connecting $x_i(t)$ and $x_j(t)$ when the influence of all other time series $x_k(t)$ $k \neq i, j$ are subtracted from $x_i(t)$ and $x_j(t)$. By definition [5],

$$\gamma_{ij}^2(f) = \frac{\left|S_{\bar{x}_i \bar{x}_j}(f)\right|^2}{S_{\bar{x}_i}(f) S_{\bar{x}_j}(f)} \tag{1}$$

where $S_{\bar{x}_i}(f)$, and $S_{\bar{x}_i \bar{x}_j}(f)$ stand for the auto- and cross-power spectral densities of $\bar{x}_m(t) = x_m(t) - \mathfrak{P}\{\mathfrak{x}_\mathfrak{m}(\mathfrak{t}) | \mathfrak{x}_\mathfrak{k}(\mathfrak{t}), \mathfrak{k} \neq \mathfrak{i}, \mathfrak{j}\}$, $m = i, j$ which result from discounting the optimal linear least squares (Wiener) joint prediction using the other series. This leads naturally to the computation of

$$.S_{\bar{x}_i \bar{x}_j}(f) = S_{x_i x_j}(f) - \mathbf{s}_{x_i \mathbf{x}}^T(f) \mathbf{S}_{\mathbf{xx}}^{-1}(f) \mathbf{s}_{\mathbf{x} x_j}(f) \tag{2}$$

where

$$\mathbf{s}_{x_i \mathbf{x}}(f) = \begin{bmatrix} S_{x_j x_1}(f) \\ \vdots \\ S_{x_j x_{j-1}}(f) \\ S_{x_j x_{j+1}}(f) \\ \vdots \\ S_{x_j x_N}(f) \end{bmatrix}$$

and $\mathbf{S}_{\mathbf{xx}}(f)$ is the cross-power spectral density of the reduced joint process describing $x_k(t)$ $k \neq i, j$ [5]. Thus all the information for computation can be obtained by adequate choice of submatrices in the joint power spectral density matrix $\mathbf{S}(f)$. Thus each of the $N(N+1)/2$ auto- and cross-spectra $S_{\bar{x}_i \bar{x}_j}(f)$ requires inverting an $N-1 \times N-1$ matrix leading to $O(N^5)$ multiplications to obtain all the needed quantities.

An ever more popular means to estimate $\mathbf{S}(f)$ involves the use of linear prediction, specially through Multichannel Auto-Regressive models

$$\begin{bmatrix} x_1(t) \\ \vdots \\ x_N(t) \end{bmatrix} = \sum_{r=1}^{p} \mathbf{A}_i \begin{bmatrix} x_1(t-r) \\ \vdots \\ x_N(t-r) \end{bmatrix} \tag{3}$$

$$+ \begin{bmatrix} e_1(t) \\ \vdots \\ e_N(t) \end{bmatrix} \tag{4}$$

where the resulting joint spectral density estimate is given by the factors

$$\mathbf{S}(f) = \mathbf{H}(f) \mathbf{\Sigma} \mathbf{H}(f)^H \tag{5}$$

with $\mathbf{\Sigma}$ standing for the prediction error covariance
matrix and $\mathbf{H}(f) = \bar{\mathbf{A}}^{-1}(f) = (\mathbf{I} - \mathbf{A}(f))^{-1}$, where $\mathbf{A}(f) = \sum_{i=1}^{p} \mathbf{A}_i z^{-i} \Big|_{z=e^{-j2\pi f}}$.

We show here that whenever a fitted model in Eq. (3) is available, a simpler and more convenient means of computing Eq. (1) is possible without prior need to estimate Eq. (5) for use in Eq. (2).

# 3 The Computationally Improved Method

In the general case, proved in the Appendix, we have:

**Method** For each frequency $f$, consider the partition of

$$\bar{\mathbf{A}} = \mathbf{I} - \mathbf{A}(f) = \begin{bmatrix} \bar{\mathbf{a}}_1(f) & \bar{\mathbf{a}}_2(f) & \dots & \bar{\mathbf{a}}_N(f) \end{bmatrix}$$

into columns. Then the partial coherence between $x_i(t)$ and $x_j(t)$ at each such frequency is given by:

$$\gamma_{ij}^2(f) = \frac{|\langle \bar{\mathbf{a}}_i(f), \bar{\mathbf{a}}_j(f) \rangle|^2}{\langle \bar{\mathbf{a}}_i(f), \bar{\mathbf{a}}_i(f) \rangle \langle \bar{\mathbf{a}}_j(f), \bar{\mathbf{a}}_j(f) \rangle} \tag{6}$$

where $\langle \bar{\mathbf{a}}_i(f), \bar{\mathbf{a}}_j(f) \rangle = \bar{\mathbf{a}}_i^H(f) \mathbf{\Sigma}^{-1} \bar{\mathbf{a}}_j(f)$ and $\mathbf{\Sigma}$ is the prediction error covariance matrix.

**Remark 1** *A simple interpretation of Eq. (6) is that it measures the square of the cosine of the angle between the column vectors $\bar{\mathbf{a}}_i(f)$ and $\bar{\mathbf{a}}_j(f)$ with $\mathbf{\Sigma}^{-1}$ for metric.*

**Remark 2**
*For each frequency, one may compute all pairwise coherences as squares of the upper diagonal elements of $\bar{\mathbf{A}}^H \mathbf{\Sigma}^{-1} \bar{\mathbf{A}} = \mathbf{S}^{-1}(f)$ followed*

*by a suitable normalization of its rows and columns by the main diagonal of* $\mathbf{S}^{-1}(f)$*. This observation is consistent with the expression for* $\gamma_{ij}^2(f)$ *used by [7] as ratios of certain minors of* $\mathbf{S}(f)$*.which also leads to* $O(N^5)$ *multiplications.*

**Remark 3** *The complexity of* $\bar{\mathbf{A}}^H\mathbf{\Sigma}^{-1}\bar{\mathbf{A}}$ *is* $N^3$ *(inversion of* $\mathbf{\Sigma}$**)** $+2N^3$ *(2 matrix multiplications), where only the latter term need be accounted for each new desired frequency.*

To illustrate the meaning of the present method we consider the case where $N = 3$ and $\Sigma$ is the identity matrix. Then, in evaluating $\gamma_{12}^2(f)$ we may rewrite Eq. (2) as:

$$.S_{\bar{x}_i\bar{x}_j}(f) = \frac{1}{S_{x_3x_3}(f)} \times \det \begin{bmatrix} S_{x_ix_j}(f) & S_{x_3x_j}(f) \\ S_{x_ix_3}(f) & S_{x_3x_3}(f) \end{bmatrix} \tag{7}$$

for $i, j \neq 3$.

We may compute

$$S_{x_ix_j}(f) = E[X_i(f)^* X_j(f)] \tag{8}$$

with help of a vector representation of $X_k(f)$ in terms of a vector $\mathbf{X}_k(f) = [H_{k1}(f), H_{k2}(f), H_{k3}(f)]$ (row $k$ of $\mathbf{H}(f)$ in Eq. (5)) which corresponds to a decomposition of $X_k(f)$ in terms of convenient orthonornal (because $\mathbf{\Sigma} = \mathbf{I}$) independent increment processes $\mathbf{dZ}^T$=$[d\mathbf{Z}_1(f)\ d\mathbf{Z}_2(f)\ d\mathbf{Z}_3(f)]^T$ so that

$$X_k(f) = \mathbf{dZ}^T\mathbf{X}_k(f) \tag{9}$$

$$= \sum_{l=1}^{3} H_{kl}(f) d\mathbf{Z}_l(f) \tag{10}$$

In this representation $E[X_i(f)^* X_j(f)] = \langle \mathbf{X}_i(f), \mathbf{X}_j(f) \rangle$ is a scalar product whereupon follows $\langle d\mathbf{Z}_i(f), d\mathbf{Z}_j(f) \rangle = E[dZ_i(f)^* dZ_j(f)] = 1$ from the assumed orthonormality between $dZ_i(f)$. Using vector analysis [8], we can rewrite the numerator of Eq. (7) as $\langle \mathbf{X}_i(f), \mathbf{X}_j(f) \rangle \langle \mathbf{X}_3(f), \mathbf{X}_3(f) \rangle - \langle \mathbf{X}_i(f), \mathbf{X}_3(f) \rangle \langle \mathbf{X}_3(f), \mathbf{X}_j(f) \rangle$ which equals

$$\langle \mathbf{X}_i(f) \wedge \mathbf{X}_3(f), \mathbf{X}_j(f) \wedge \mathbf{X}_3(f) \rangle \tag{11}$$

where in $3D$, $\wedge$ corresponds to the usual vector product. Using the basis $d\mathbf{Z}$, the vector product components for $\mathbf{X}_i(f) \wedge \mathbf{X}_k(f)$ are given by

$$\begin{bmatrix} H_{i2}(f)H_{k3}(f) - H_{i3}(f)H_{k2}(f), \\ H_{i3}(f)H_{k1}(f) - H_{i1}(f)H_{k3}(f) \\ H_{i1}(f)H_{k2}(f) - H_{k1}(f)H_{i2}(f) \end{bmatrix} \tag{12}$$

which is proportional to the $j-th$ column ($j \neq i, k$) of the inverse of $\mathbf{H}(f)$, with $1/\det \mathbf{H}(f)$ for propotionality constant. Upon taking the scalar products in the RHS of Eq. (11) leads to a particular case of Eq. (6).

# 4 Final Remarks and Conclusion

The alternative of obtaining $\mathbf{S}^{-1}(f)$ directly from Eq. (5), though its complexity is roughly the same order of magnitude ($4N^3-$ two matrix inversions and two matrix multplications) as for Eq. (6), should be avoided because of the accumulation of numerical errors due to likely poor conditioning of $\bar{\mathbf{A}}$ at certain frequencies.

Thus when MAR models are available, one should prefer Eq. (6) to compute partial coherences since only one matrix inversion is necessary ($\mathbf{\Sigma}^{-1}$) regardless of how many frequency points are desired, and also because only scalar products appear in the computation of $\gamma_{ij}^2(f)$ for each frequency leading at once to lower complexity and increased accuracy.

Thus when available, a joint multivariate autoregressive model describing $N$ time series permits finding the partial coherence functions between pairs of time series with a reduction in complexity from $O(N^5)$ to $O(N^3)$ multiplications foresaking matrix inversions for each frequency estimate thereby also leading to more accurate partial coherence estimates.

**Remark 4** *This complexity reduction is important given the present availability of EEG data acquisition systems with as many as* $N = 128$ *channels.*

**Remark 5** *The expression of the partial coherence as in Eq. (6) lies at the heart of a recently introduced method [11] for structural inference based on the analysis of multiple neural signals.*

# 5 Appendix - Proof of the Improved Method

The proof of the algorithm in the general case follows from the observation that (1) is the cosine of the dihedral angle [8] between the hyperplanes spanned the vectors $[\mathbf{X}_i(f), \{\mathbf{X}_k(f), k \neq i, j\}]$ and $[\mathbf{X}_j(f), \{\mathbf{X}_k(f), k \neq i, j\}]$. This quantity is proportional to the scalar product between the $(N-1)-$ forms given by the wedge vector products $\mathbf{V}_m(f) = \mathbf{X}_m(f) \wedge (\bigwedge_{k \neq i,j} \mathbf{X}_k(f))$, $m = i, j$, [9] or

$$\gamma_{ij}^2(f) = \frac{|\langle \mathbf{V}_i(f), \mathbf{V}_j(f) \rangle|^2}{\langle \mathbf{V}_i(f), \mathbf{V}_i(f) \rangle \langle \mathbf{V}_j(f), \mathbf{V}_j(f) \rangle} \quad (13)$$

**Remark 6** *Essentially* $V_m(f)$ *is a direct frequency domain measure of the complement* $\bar{x}_m(t) = x_m(t) - \mathfrak{P}\{\mathfrak{x}_\mathfrak{m}(\mathfrak{t}) | \mathfrak{x}_\mathfrak{k}(\mathfrak{t}), \mathfrak{k} \neq \mathfrak{i}, \mathfrak{j}\}$ *orthogonal to the subspace generated by* $x_k(t), k \neq i, j$ *at frequency* $f$.

The wedge product can be computed directly from the components of $\mathbf{X}_p(f) = \sum_{q=1}^{N} \tilde{H}_{pq}(f) d\mathbf{Z}_q(f)$ where without loss of generality, $\tilde{H}_{pq}(f)$ are the elements in a redefinition of (5) as

$$\begin{aligned} \mathbf{S}(f) &= \mathbf{H}(f)\mathbf{\Sigma}\mathbf{H}(f)^H \\ &= \left(\mathbf{H}(f)\mathbf{\Sigma}^{1/2}\right)\left(\mathbf{\Sigma}^{1/2}\mathbf{H}(f)\right)^H \\ &= \left(\tilde{\mathbf{H}}(f)\right)\left(\tilde{\mathbf{H}}(f)\right)^H \end{aligned} \quad (14)$$

In this case, using a result from [10] (sec I.11 p.83), it follows that, $V_{mp}(f)$, the $p-th$ component of $V_m(f)$, up to a sign due to the ordering of factors in the wedge product (that cancels out in $|\langle \mathbf{V}_i(f), \mathbf{V}_j(f) \rangle|^2$), equals $\left[\tilde{\mathbf{H}}(f)\right]_{mp}$, the minor obtained by eliminating row $m$ and column $p$ in $\tilde{\mathbf{H}}(f)$. It is easy to see that the latter value is proportinal to each entry at column $m$, row $p$ of the inverse of $\tilde{\mathbf{H}}(f)$ by a factor $(-1)^{m+p} \times \det \tilde{\mathbf{H}}(f)$. Upon computation of the scalar product $\langle \mathbf{V}_m(f), \mathbf{V}_n(f) \rangle$ as

$$\begin{aligned} &= \sum_{p=1}^{N} \left[\tilde{\mathbf{H}}(f)\right]_{mp} \left[\tilde{\mathbf{H}}(f)\right]_{np} \\ &= \left(\det \tilde{\mathbf{H}}(f)\right)^2 (-1)^{m+n} \\ &\times \left\langle col_m \tilde{\mathbf{H}}^{-1}(f), col_n \tilde{\mathbf{H}}^{-1}(f) \right\rangle \\ &= \left(\det \tilde{\mathbf{H}}(f)\right)^2 (-1)^{m+n} \left\langle \mathbf{\Sigma}^{-1/2}\bar{\mathbf{a}}_m, \mathbf{\Sigma}^{-1/2}\bar{\mathbf{a}}_n \right\rangle \\ &= \left(\det \tilde{\mathbf{H}}(f)\right)^2 (-1)^{m+n}\, \bar{\mathbf{a}}_m^H \mathbf{\Sigma}^{-1} \bar{\mathbf{a}}_n \end{aligned} \quad (15)$$

where $\left(\det \tilde{\mathbf{H}}(f)\right)^2$ cancels out upon substitution into (13) thus completing the proof. □

# 6 Acknowlegdements

I thank Dr. Max Gerken suggestions which improved this manuscript. This work was supported by CNPQ 301273/96-0 and FAPESP 1999/07461-2 grants.

# References

[1] Franaszczuk PJ, Bergey GK, Kaminski MJ. "Analysis of mesial temporal seizure onset and propagation using the directed coherence function method. " Electroencephalogr Clin Neurophysiol 1994, 91:413-27.

[2] Sherman DL, Tsai YC, Rossell LA, Mirski MA, Thakor NV. "Spectral analysis of a thalamus-to-cortex seizure pathway." IEEE Trans Biomed Eng 1997, 44:657-64.

[3] Baccalá LA, Sameshima K. "Directed Coherence: a tool for exploring causality

and feedback among brain structures." In Nicolelis MAL (ed.) Methods for Simultaneous Neuronal Ensemble Recordings. CRC Press, Boca Raton, 1998, pp 177–90.

[4] Baccalá LA, Sameshima K, Ballester G, Valle AC, Timo-Iaria C. "Studying the Interaction between Brain Structures via Directed Coherence and Granger Causality. " Applied Signal Processing 1998, 5:40–8.

[5] Bendat JS, Piersol AG. Random Data: Analysis and Measurement Procedures. John Wiley, New York, 1986.

[6] Anderson CW, Stolz EA, Shamsunder S. "Multivariate Autoregressive Models for Classification of Spontaneous Electroenecephalographic Signals During Mental Tasks." IEEE Trans Biomed Eng 1998, 45:277–88.

[7] Franaszczuk PJ, Blinowska KJ, Kowalczyk M. "The Application of parametric multichannel spectral estimates in the study of electrical brain activity ". Biol Cybern 1985, 51:239–47.

[8] Burke WL. Applied Differential Geometry. Cambridge University Press, Cambridge, 1985.

[9] Hestenes D. New Foundations for Classical Mechanics. D. Reidel, Dordrecht, 1987.

[10] Deheuvels R. Formes Quadratiques et Groupes Classiques. Presses Universitaires de France, Paris, 1981.

[11] Baccala, L.A. and Sameshima, K. "Partial Directed Coherence a New Concept in Neural Structure Determination. Biol. Cybern. 84(6):463–74.