Connectivity estimation of three parametric methods on simulated electroencephalogram signals

Hugo Vélez-Pérez, Valérie Louis-Dorr, Radu Ranta, Michel Dufaut

***Abstract*— The global framework of this paper is the connectivity estimation in multichannel electroencephalogram (EEG) recordings, modeled as multidimensional autoregressive (AR) processes. The coherence, directed transfer function and partial directed coherence functions are evaluated on two simulated EEG signals for their later application on real EEG recordings. The results were evaluated computing the relative error and a second proposed performance criterion (η) based on the entropy of the estimated connectivity matrix.**

I. INTRODUCTION

Some works based on stereo electroencephalography (SEEG) recordings ([1], [2]) have showed a certain electrophysiological synchronization between deep brain anatomical structures during seizures. This deep synchronization can be expressed by correlations on scalp electrodes in electroencephalogram (EEG) recordings, even if this activity transfer is only partial. This hypothesis has been reinforced by Franaszczuk and Bergey [3] and by Caparos [4], who proposed either parametric or non parametric methods to estimate the synchronization between rhythmic waves often observed at seizure onset.

The objective of this work is to evaluate, on two simulated EEG signals, the capacity of three parametric methods, coherence (C), directed transfer function (DTF) [5] and partial directed coherence (PDC) [6], based on the autoregressive model (ARM), to reveal existent interactions between signals. This analysis will guide their possible later application to real scalp EEG signals. In section 2, we remind the theoretical bases of ARM applied to EEG analysis. The methods derived from the ARM and the used evaluation criteria are presented in the third section. In the fourth section we present the simulated EEG signals used in this study, while the results and the discussion are presented in the following section. Finally, in last section, we conclude and we note some perspectives of this work.

II. EEG MODELING

The most widely model used in EEG signal processing is the ARM. EEG signals are non stationary signals; nevertheless, is possible to consider a certain stationarity for short time intervals. For a given time interval, the multidimensional ARM is:

$$\mathbf{x}(t) = \sum_{k=1}^{p} \mathbf{A}(k)\mathbf{x}(t-k) + \mathbf{e}(t) \tag{1}$$

Manuscript recived on July 21, 2008

All authors are with Research Centre for Automatic Control (CRAN-UMR 7039), National Center for Scientific Research (CNRS), Nancy University, 2 Avenue de la Forêt de Haye, 54516, Vandoeuvre lès Nancy, France. hugo.velez-perez@ensem.inpl-nancy.fr

with $\mathbf{A}(k) = \begin{bmatrix}\mathbf{a}_1(k) & \mathbf{a}_2(k) & \dots & \mathbf{a}_n(k)\end{bmatrix}^T$ as the $n \times n$ ARM coefficients (ARMC) matrix, n the number of channels, $\mathbf{x}(t-k)$ the time-delayed values vector, p the model order and $\mathbf{e}(t)$ the error vector. To solve (1) is necessary to find p (smaller than the sequence length) and next to estimate the ARMC.

To find p, the classical approach minimizes the Akaike's Information Criterion (AIC) [7], based on the maximum likelihood estimation of the signal probability density function. To find $\mathbf{A}(k)$, the solution of Yule-Walker's equations [8] must be computed. For this application, the Levinson-Durbin's algorithm [9] is applied.

III. SYNCHRONIZATION IN FREQUENCY DOMAIN

Once the multichannel ARM has been obtained and to investigate its spectral properties, (1) can be written in frequency domain as:

$$\mathbf{x}(f) = \bar{\mathbf{A}}(f)^{-1}\mathbf{e}(f) = \mathbf{H}(f)\mathbf{e}(f) \tag{2}$$

where $\bar{\mathbf{A}}(f) = \mathbf{I} - \mathbf{A}(f) = \begin{bmatrix}\bar{\mathbf{a}}_1(f) & \bar{\mathbf{a}}_2(f) & \dots & \bar{\mathbf{a}}_n(f)\end{bmatrix}$ and $\mathbf{I}$ the identity matrix. $\mathbf{H}(f)$ is called the transfer function matrix of the system. The power spectral matrix $\mathbf{S}(f)$ is obtained by:

$$\mathbf{S}(f) = \mathbf{H}(f)\mathbf{V}\mathbf{H}^*(f) \tag{3}$$

where * denotes transposition and complex conjugate and $\mathbf{V}$ is the noise covariance matrix.

A. Coherence

The first studied estimator is the well known C function between two channels i and j, introduced as a measure of linear association degree between two signals in frequency domain:

$$C_{ij}(f) = \frac{|s_{ij}(f)|}{\sqrt{|s_{ii}(f)||s_{jj}(f)|}} \tag{4}$$

where s_{ij} are elements of $\mathbf{S}(f)$. C is a positive and symmetric function, normalized between 0 and 1.

B. Directed Transfer Function

In 1991, Kaminski and Blinowska introduced the DTF estimator [5], defined as:

$$\gamma_{ij}(f) = \frac{|H_{ij}(f)|}{\sqrt{\sum_{m=1}^{n} |H_{im}(f)|^2}} \tag{5}$$

The DTF describes the ratio of the influence from the j-th channel to i-th channel with respect to the influence of all inflows to i-th channel.

C. Partial Directed Coherence

In 1999, Baccalá and Sameshima, proposed the PDC, defined as [6]:

$$\pi_{ij}(f) = \frac{\left|\bar{a}_{ij}(f)\right|}{\sqrt{|\bar{\mathbf{a}}_j^*(f)||\bar{\mathbf{a}}_j(f)|}} \tag{6}$$

where $\bar{a}_{ij}(f)$ is the i, j-th element and $\bar{\mathbf{a}}_j$ a vector column of $\bar{\mathbf{A}}(f)$. PDC was introduced to provide a clearer connectivity of Granger causality, showing the direct interactions (*feedforward* and *feedback*) between every pair of channels within a multichannel process.

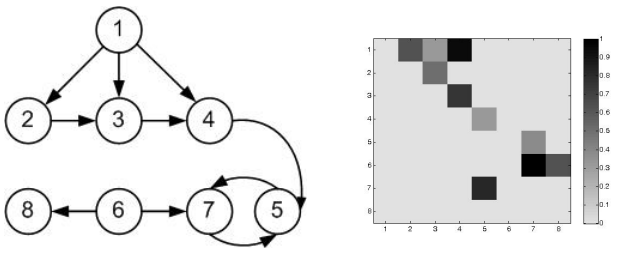
D. Evaluation criteria

1) Relative Error ($E_{relative}$): To evaluate the performance of the previous estimators, an error function was introduced by Astolfi in [10]:

$$E_{relative}(f) = \frac{\sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}(\bar{\zeta}_{ij}(f) - \bar{\hat{\zeta}}_{ij}(f))^2}}{\sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}\bar{\zeta}_{ij}(f)^2}} \tag{7}$$

where $\bar{\zeta}_{ij}(f)$ and $\bar{\hat{\zeta}}_{ij}(f)$ are the mean of C, DTF or PDC function from the theoretical and estimated models respectively. These averages are computed by frequency band in [10]. As in this work we aim to evaluate the global connectivity estimation regardless of the frequency, we consider a $E_{relative}$ computed over the whole frequency domain.

2) Entropic criterion (η): A connectivity graph, as presented in Fig. 1(a), does not take into account neither time delays nor frequency relations. From Parseval's theorem, we might consider a global matrix $\mathbf{A}_g = \sqrt{\sum_{k=1}^{p}\mathbf{A}(k)^2}$, which contains the connectivity information as given by the graph, regardless of the time delay. Therefore, the non-zero coefficients will correspond only to the graph edges and their amplitude to amount of the transfered energy.



(a) (b)

Fig. 1. Connectivity graph (a) and its $\mathbf{M}_A$ image representation (b).

In a similar way, a $\mathbf{Z} = \{\zeta_{ij}\}$ matrix can be obtained by grouping all ζ_{ij} values obtained from C $\left(\zeta_{ij} = \sqrt{\sum_f \zeta_{ij}(f)}\right)$, DTF or PDC $\left(\zeta_{ij} = \sqrt{\sum_f \zeta_{ij}(f)^2}\right)$. $\hat{\mathbf{Z}} = \{\hat{\zeta}_{ij}\}$ will correspond to estimated functions. Assuming that a connection should exist also in frequency domain (*i.e.* a non-zero coefficient in $\mathbf{Z}$), we propose a new criterion to compare the connectivity information given by $\mathbf{A}_g$ and $\mathbf{Z}$. First, as we are only interested in connectivity estimation, two new matrices $\mathbf{M_A}$ and $\mathbf{M_Z}$ ($(p-1)\times(p-1)$ sized) are constructed by eliminating diagonal terms from previous matrices. Next, we normalize the $\mathbf{M_A}$ and $\mathbf{M_Z}$ elements into a range between 0 and 1. To compare the connectivity information contained in $\mathbf{M_Z}$ to the information contained in $\mathbf{M}_A$, we compute the Shannon entropy, redefined for a matrix $\mathbf{M} = \{m_{ij}\}$ like:

$$H(\mathbf{M}) = -\sum_{i=1}^{n}\sum_{j=1}^{n} p(m_{ij})\log_2(p(m_{ij})) \tag{8}$$

The entropy can be estimated from a simple histogram, to avoid the estimation of the probability law $p(m_{ij})$. The connectivity information from $\mathbf{M}_A$ can be compared:

- with $\mathbf{M_Z}$ (based on the theoretical C, DTF, PDC), which evaluates if the synchronisation method can accurately describe the connectivity pattern regardless of the possible modeling errors;
- with $\mathbf{M}_{\hat{\mathbf{Z}}}$ (estimated C, DTF, PDC), which takes into account the possible estimation errors of the ARMC.

We can compute then two version of a new η criterion, a theoretical value (η) and an estimated one ($\hat{\eta}$):

$$\eta = \left|1 - \frac{H(\mathbf{M_Z})}{H(\mathbf{M}_A)}\right| \tag{9}$$

$\hat{\eta}$ is obtained replacing $\mathbf{Z}$ by $\hat{\mathbf{Z}}$. A value close to 0 indicates similar information in $\mathbf{A}_g$ and $\mathbf{Z}$ ($\hat{\mathbf{Z}}$). The lowest theoretical value η indicates the most appropriate method to evaluate connectivity patterns, while an important change in the estimated value $\hat{\eta}$ (comparing to η) indicates a particular sensibility of the method to possible model estimation errors.

IV. Models and Simulations

A. Model I (MI)

The first simulated model was inspired from the model used by Baccalá in [11].

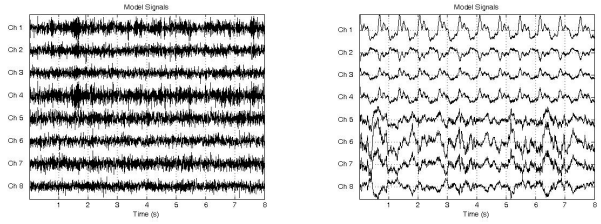
$$\begin{aligned}
x_1(k) &= 0.95x_1(k-1) - 0.75x_1(k-2) + e_1(k)\\
x_2(k) &= -0.5x_1(k-1) + e_2(k)\\
x_3(k) &= 0.25x_1(k-1) - 0.4x_2(k-2) + e_3(k)\\
x_4(k) &= 0.75x_1(k-1) + 0.1x_4(k-1) - 0.6x_3(k-2) + e_4(k)\\
x_5(k) &= 0.2x_5(k-1) + 0.65x_7(k-1) + 0.25x_4(k-2) + e_5(k)\\
x_6(k) &= 0.75x_6(k-1) - 0.4x_6(k-2) + e_6(k)\\
x_7(k) &= 0.8x_6(k-1) - 0.3x_5(k-2) + e_7(k)\\
x_8(k) &= -0.5x_6(k-1) + e_8(k)
\end{aligned}$$

This model behaves as an oscillator driving, directly or indirectly, the other structures according to Fig. 1(a). It contains two independent sources (channels 1 and 6), whereas the other channels are generated by 1 and 6. For this model, $p = 2$, the ARM coefficients are chosen in $[-1;1]$ interval and e is a zero mean Gaussian noise with 0.5 standard deviation.

B. Model II (MII)

Because the objective is to apply these methods to real EEG recordings, it is necessary to verify if it is possible to introduce in the ARM signals generated in a different way. The second proposed model is identical to the first one, for the mixing, but 2 simulated EEG are introduced in channels

1 and 6 instead of the previous ARM signals. The used signals are 2 newborn simulated EEG signals taken from Stevenson and Rankine [12]: a background signal in channel 1 and a seizure signal in channel 6. The signals produced by simulating MI and MII are shown in Fig. 2.



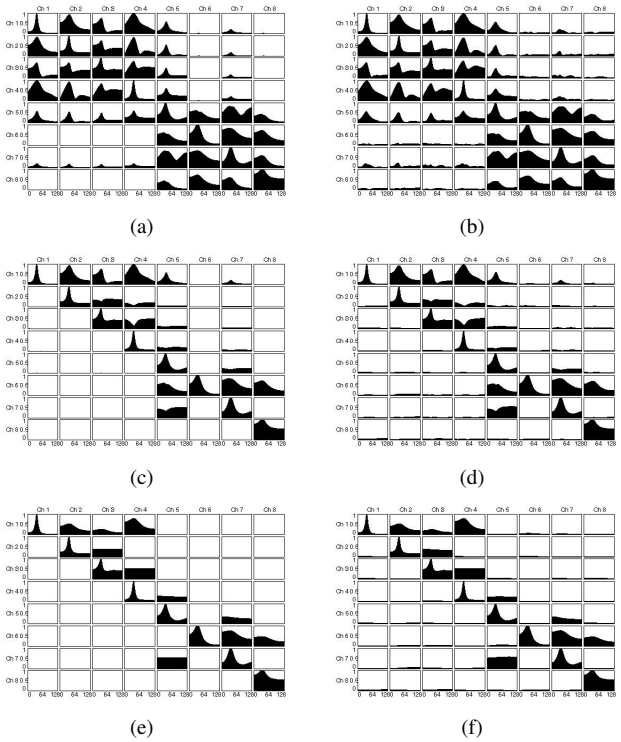
(a) (b)

Fig. 2. Signals obtained by simulating the MI (a) and MII (b).

V. Results and discussion

A hundred simulations were executed in order to obtain mean results over different noise generations. For MI, most of the times, the estimated p did agree with theoretical p; however, for MII, p varied between simulations, which puts in evidence that if signals are not generated in an AR way, the model becomes more difficult to estimate.

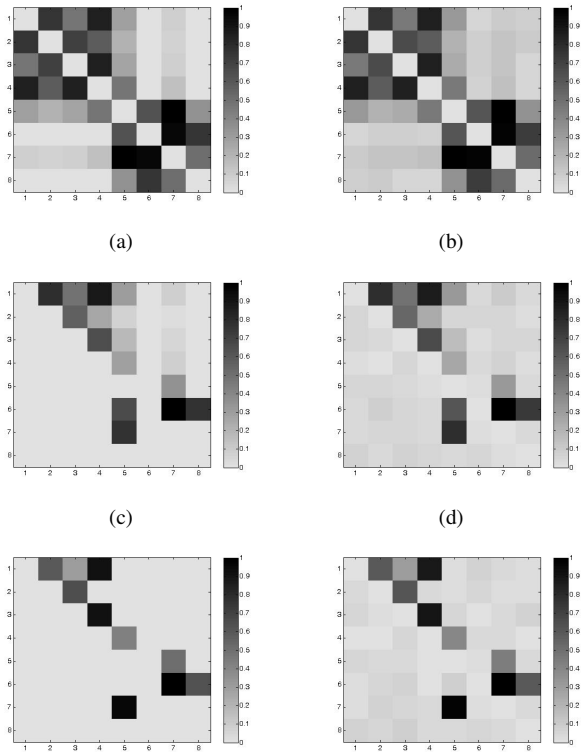
In Fig. 3, graphs of theoretical (left column) and estimated (right column) C, DTF and PDC of MI are presented. For all these estimators, we can compute $\mathbf{M_Z}$ matrices, which can be represented by scaled gray level images (Fig. 4) to be compared with the scaled image of $\mathbf{M}_A$ matrix (Fig. 1(b))(a 0 diagonal was reintroduced for displaying facility).



(a) (b) (c) (d) (e) (f)

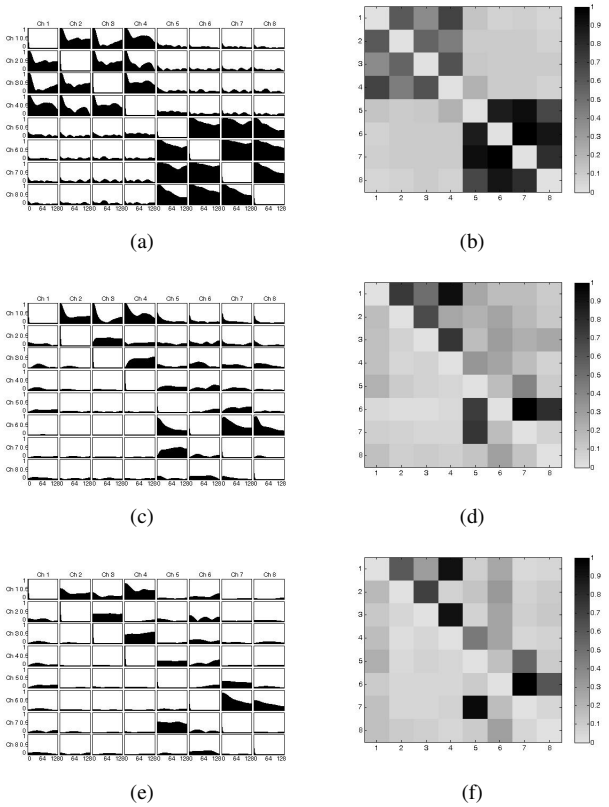
Fig. 3. Theoretical and estimated C (a-b), DTF (c-d) and PDC (e-f) for MI.

Because C is a positive and symmetric function, all channel relations are present, so a symmetric graph is obtained.



(a) (b) (c) (d) (e) (f)

Fig. 4. Gray scale representation for theoretical and estimated C (a-b), DTF (c-d) and PDC (e-f) for MI.



(a) (b) (c) (d) (e) (f)

Fig. 5. Estimated and gray scale representation for C (a-b), DTF (c-d) and PDC (e-f) for MII.

For DTF, direct and indirect relations are expressed, proving its notion of causality. Finally, for PDC, only the direct

connectivities are displayed. We can appreciate in Fig. 3 how the estimated functions are very close to theoretical ones. Comparing the gray scale image of $\mathbf{M}_A$ in Fig. 1(b) with the right column of Fig. 5, we notice that estimated PDC image is very close to the $\mathbf{M}_A$ image, while DTF presents little differences and the C image is quite far.

These observations are confirmed by Table I, which contains the mean values for the computed criteria. Unlike in [10], in our tests (only one type of noise) PDC seems to be more robust than DTF and C (smaller $E_{relative}$). However, these $E_{relative}$ values are not far between them. We also notice that the greatest value of η and $\hat{\eta}$ is given by C, followed by the DTF, whereas the smallest value corresponds to PDC in both cases. That is explained because C highlights all channel relations (symmetric function), DTF estimates the direct and indirect connexions and finally PDC only highlights the direct relations between channels. According to results of $\hat{\eta}$, PDC outperforms DTF and C, as expected.

TABLE I

$E_{relative}$, η FOR MI AND $\hat{\eta}$ FOR MI AND MII.

Model	Criterion	C	DTF	PDC
Model I	$E_{relative}$	0.0636	0.0646	**0.0614**
	η	1.5345	0.3984	**0.0641**
	$\hat{\eta}$	1.4909	0.4181	**0.0384**
Model II	$\hat{\eta}$	1.2601	0.6206	**0.2129**

However, the results for MII are more questionable. As we appreciate in Fig. 5, when we introduce more realistic signals in the mixing, invalid theoretical connections appear between channels, leading to an incorrect connectivity interpretation. We remark that is not possible to obtain theoretical graphs for C, DTF and PDC because we do not have access to theoretical values of signals x_1 and x_6. Therefore, the η and $E_{relative}$ can not be obtained. As for MI, the connectivity of MII is better highlighted by PDC than by DTF or C. In spite of PDC being the best estimator, we can observe that for MII, $\hat{\eta}_{PDC}$ is much closer to $\hat{\eta}_{DTF}$ than MI. In fact, during our simulations, we observed also the inversion of these two values. Also, we can observe an important degradation of the PDC estimator (from 0.04 to 0.21) when MII is used, while the DTF estimator is much more robust (from 0.42 to 0.62) and less influenced by the quality of the AR modeling . Therefore, it is necessary to have in mind that, if ARM is questionable for real or close to real signals, frequency domain parametric connectivity estimators may lead to erroneous interpretations (especially PDC).

VI. Conclusion and future works

In this work, the property of C, DTF and PDC to put in evidence the connectivity in EEG models was evaluated. If we are interested in studying direct and/or indirect connections, DTF and PDC are the best relationship estimators. According to both used criteria, PDC is the best direct connectivity estimator, if the ARM is correctly estimated. Nevertheless, in spite of DTF is not being able to reveal only the direct connections between the channels like PDC, it also seems a good estimator, whose results are not far from PDC results. Moreover, the DTF is less sensible to modeling difficulties. From the results obtained in the MII, we can conclude that, in spite of the sensibility of these estimators to the presence of more realistic signals in the model, it is possible to apply them to real EEG recordings. The proposed entropic criterion η highlights direct interchannel relationships and permits a quantitative evaluation of the connectivity in function. This connectivity estimation becomes an important aspect in the characterization of seizures in scalp EEG recordings and, particularly, in the location of the anatomical origin of seizures.

Future investigations aim to evaluate the noise influence. In order to account for the quasi stationarity of signals in longer simulated and real EEG, a sliding window could be applied. Because our underlying hypothesis here is that the results obtained for the whole frequency domain can be extrapolated by frequency band, we think it is possible to work on electrophysiological frequency bands of real EEG in order to describe or to obtain different connectivity graphs for each band.

References

[1] F. Wendling, F. Bartolomei, F. Bellanger, P. Chauvel. 'Interpretation of interdependencies in epileptic signals using a macroscopic physiological model of the EEG'. *Elsevier Clinical Neurophysiology 112(7)*: 1201-1218. 2001.

[2] M. Chavez. 'Analyse des signaux SEEG intercritiques: Apport des modèles dynamiques non linéaires'. *PhD Thesis*, Université de Rennes 1, France, 2001.

[3] P.J. Franaszczuk, G.K. Bergey. 'Application of the directed transfer function method to mesial and lateral onset temporal lobe seizures'. *Brain Topography 11(1)*: 13-21. 1997.

[4] M. Caparos, V. Louis, F. Wendling, L. Maillard, D. Wolf. 'Automatic lateralization of temporal lobe epilepsy based on scalp EEG'. *Elsevier Clinical Neurophysiology 117(11)*: 2414-2423. 2006.

[5] M.J. Kaminski, K.J. Blinowska. 'A new method of the description of the information flow in the brain structures'. *Biological Cybernetics 65(3)*: 203-210. 1991.

[6] L. Baccalá, K. Sameshima. 'Partial directed coherence: a new concept in neural structure determination'. *Biological Cybernetics 84(6)*: 463-474. 2001.

[7] H. Akaike. 'A new look at statistical model identification'. *IEEE Transactions on Automatic Control 19(6)*: 716-723. 1974.

[8] V. Louis-Dorr, M. Caparos, F. Wendling, J.P. Vignal, D. Wolf 'Extraction of reproducible seizure patterns based on EEG scalp correlations'. *Biomedical Signal Processing and Control 2(3)* 154-162. 2007.

[9] J. Durbin. 'The fitting of time series model'. *Review of the International Statistical Institute 28(3)*: 233-244. 1960.

[10] Astolfi L., Cincotti F., Mattia D., Marciani G., Baccalá L., et al. 'Comparison of different cortical connectivity estimators for high-resolution EEG recordings'. *Human Brian Mapping 28*: 143-157. 2007.

[11] L. Baccalá, K. Sameshima. 'Overcoming the limitations of correlation analysis for many simultaneously processed neural structures'. *Progress in Brain Research 130*. 2001.

[12] N. Stevenson, L. Rankine, M. Mesbah, B. Boashash. 'Newborn EEG seizure using time-frequency signals synthesis'.