33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Testing the Asymptotic Statistic for the Assessment of the Significance of Partial Directed Coherence Connectivity Patterns

J. Toppi*, F. Babiloni, G. Vecchiato, F. Cincotti, F. De Vico Fallani, D. Mattia, S. Salinari and L. Astolfi

Abstract—**Partial Directed Coherence (PDC) is a powerful tool to estimate a frequency domain description of Granger causality between multivariate time series. One of the main limitation of this estimator, however, has been so far the criteria used to assess the statistical significance, which have been obtained through surrogate data approach or arbitrarily imposed thresholds. The aim of this work is to test the performances of a validation approach based on the rigorous asymptotic distributions of PDC, recently proposed in literature. The performances of this method, defined in terms of percentages of false positives and false negatives, were evaluated by means of a simulation study taking into account factors like the Signal to Noise Ratio (SNR) and the amount of data available for the estimation and the use of different methods for the statistical corrections for multiple comparisons. Results of the Analysis Of Variance (ANOVA) performed on false positives and false negatives revealed a strong dependency of the performances from all the factors investigated. In particular, results indicate an amount of Type I errors below 7% for all conditions, while Type II errors are below 10% when the SNR is at least 1, the data length of at least 50 seconds and the appropriate correction for multiple comparisons is applied.**

I. Introduction

IN the latest years, the estimation of functional connectivity has become more and more central in neuroscience to understand brain networks at the basis of motor and cognitive processes. Among different estimators of connectivity between time series, Partial Directed Coherence (PDC) [1] is a spectral estimator of Granger causality between multivariate time series [2]. Different estimators generalizing PDC were developed during the year [3]-[4]-[5] to improve the accuracy and the stability of the connectivity estimation performed. However, all these generalized quantities need a method to assess their statistical significance. Different methods used so far for the validation of connectivity patterns are based on the imposition of an arbitrary threshold, like the Spectral Causality Criterion (SCC) [6], or on a surrogate data approach like in the Shuffling procedure [7], based on an empirical distribution of the null hypothesis obtained with a time consuming procedure.

Recently, a validation approach based on the rigorous asymptotic distributions of PDC was proposed [8] . This new method is based on the assumption that the PDC estimator tends to a χ^2 distribution in the null case (lack of transmission) and has been introduced to overcome the limits of existing methods.

All the assessing methods required the definition of a significance level to be applied for the evaluation of the statistical threshold for every possible connection between signals in the multivariate dataset. Due to the high number of statistical assessments performed, it is necessary to apply corrections to the significance level imposed in the validation process, in order to prevent the occurrence of type I errors [9].

This work aims at testing the performances of the new asymptotic statistic in terms of percentages of type I and type II errors by means of a simulation study. Results will provide an estimation of the levels of error to be attended under specific conditions and a statistical analysis (Analysis Of Variance, ANOVA) of the dependency of the performances from factors like the Signal to Noise Ratio (SNR) and the amount of data used for the estimation. The application of different methods for the corrections for multiple comparisons was also investigated, to suggest the best approach resulting from the analysis of the performances.

Manuscript received April 15, 2011. This work was supported by the European ICT Program FP7-ICT-2009-4 Grant Agreement 247919 DECODER and by the Compagnia San Paolo within the Project “Social and Emotive Hyperbrain”.

J.Toppi is with the Dept. of Computer Science and Systems of the Univ. of Rome “Sapienza” and with IRCCS “Fondazione Santa Lucia”, Rome, Italy. (phone: +39-06-51501510; e-mail: jlenia.toppi@uniroma1.it).

F. Cincotti and D. Mattia, Fondazione Santa Lucia, Rome, Italy.

S. Salinari Dept. of Computer Science and Systems, Univ. of Rome “Sapienza, Rome, Italy.

F. De Vico Fallani, G. Vecchiato and F. Babiloni, Dept. of Physiology and Pharmacology of the University of Rome “Sapienza” and IRCCS “Fondazione Santa Lucia”, Rome, Italy.

L. Astolfi, Dept. of Computer Science and Systems, Dept. of Physiology and Pharmacology, Univ. of Rome “Sapienza”, and IRCCS “Fondazione Santa Lucia”, Rome, Italy.

II. Methods

A. Multivariate Methods for the Estimation of Connectivity

Supposing that the following multivariate autoregressive (MVAR) model is an adequate description of the dataset Y:

$$\sum_{k=0}^{p} \Lambda(k)Y(t-k) = E(t) \quad (1)$$

where **Y(t)** is the data vector in time, **E(t)** is a vector of multivariate zero-mean uncorrelated white noise processes, **Λ(k)** is the matrix of model coefficients at lag **k** and **p** is the model order. In the present study, **p** was chosen by means of the Akaike Information Criteria (AIC) for MVAR processes [10].

To investigate the spectral properties of the examined process, (1) is transformed to the frequency domain:

$$\Lambda(f)Y(f) = E(f)\,,\ \Lambda(f) = \sum_{k=0}^{p} \Lambda(k) e^{-j2\pi f \Delta t k} \qquad (2)$$

where **Δt** is the temporal interval between two samples.

B. Partial Directed Coherence

The PDC [1] is a full multivariate spectral measure, used to determine the directed influences between any given pair of signals in a multivariate data set. This estimator was demonstrated to be a frequency version of the concept of Granger causality [2].

It is possible to define PDC as:

$$\pi_{ij}(f) = \frac{\Lambda_{ij}(f)}{\sqrt{\sum_{k=1}^{N} \Lambda_{ki}(f)\Lambda_{ki}(f)}}\,,\ \sum_{n=1}^{N} \left|\pi_{ni}(f)\right|^2 = 1 \qquad (3)$$

Different estimators generalizing PDC were developed during the years [3]-[4]-[5]. However, different normalization do not affect the asymptotic statistic procedure since they affect the values of the estimator and the threshold in the same way, with a linear dependency.

Squared values of PDC were shown to provide higher accuracy and stability [5].

C. Statistical Assessment of Connectivity Estimate: Asymptotic Statistic

In order to assess the significance of the estimated causal links, the value of functional connectivity for a given pair of signals, obtained by computing PDC, must be compared with a threshold level which is computed for the null case (lack of causality between the considered signals).

Threshold values were estimated using asymptotic statistic method [8] which is based on the assumption that PDC in the null case follows a χ^2 distribution [11]. The statistical threshold is achieved by means of a χ^2 distribution obtained by applying a Monte Carlo method. The percentile related to the significance level imposed is then computed.

D. Statistical Correction in Multiple Comparisons

Due to the high number of comparisons between PDC values and statistical thresholds, a correction for multiple comparisons issue is needed to avoid the occurrence of type I errors (false positives). The statistical theory provides different correction algorithms. We considered here the two most used ones, the traditional Bonferroni adjustments [12] and the more recently introduced False Discovery Rate (FDR) [13].

E. Simulation Study

The simulation study involved the following steps:

1) Generation of different simulated datasets fitting a predefined model, composed by 4 cortical areas and achieved imposing different levels of Signal to Noise Ratio (factor SNR: 0.1, 1, 3, 5, 10) and data length (factor LENGTH: 3000, 10000, 20000, 30000 data samples, corresponding to a signal length of 15 50 100 150 s, at a sampling rate of 200 Hz). The imposed model is reported in Fig.1. $x_1(t)$ is a real signal acquired at the scalp level during an high resolution EEG recording session (64 channels) from a healthy subject in a rest condition with opened eyes. The other signals $x_2(t),\ldots, x_4(t)$ were iteratively achieved according to the predefined scheme. In particular, the signal $x_j(t)$ is obtained adding uncorrelated Gaussian white noise to all the contributions of other signals $x_i(t)$ (with $i\neq j$), each of which amplified of a_{ij} and delayed of τ_{ij}. Connection strengths imposed to the simulated signals are $a_{12}=0.5$, $a_{13}=0.4$ $a_{14}=0.2$, $a_{23}=0.08$. Such values are chosen in a range typical for EEG signals. The values used for the delay in transmission are $\tau_{12}=2$, $\tau_{13}=2$, $\tau_{14}=1$, $\tau_{23}=4$ data samples. To improve the robustness of the successive statistical analysis, the generation of datasets under each combination of factors was repeated 100 times.

2) Evaluation, for each dataset, of MVAR coefficients and estimation of PDC under different conditions.

3) Application of the asymptotic statistic procedures in order to assess the significance of estimated connectivity patterns by imposing a significance level of 0.05 in three different cases: no correction, FDR and Bonferroni adjustments for multiple comparisons (factor CORRECTION).

4) Computation of the total percentage of false positives and false negatives occurred in the assessment of significance of connectivity patterns for all the considered factors.

5) Analysis Of Variance (ANOVA) for repeated measures of the percentage of false positives and false negatives, in order to evaluate the effects of some factors (SNR, LENGTH, CORRECTION) on the performances of the analyzed method.

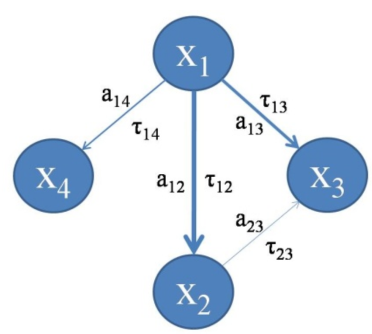

Fig. 1 – Connectivity model imposed in the generation of testing dataset. x1,…, x4 represent the signals of four cortical regions of interest. a_{ij} represent the strength of the imposed connection between nodes i and j, while τ_{ij} represents the delay in transmission applied between the two

signals xi and xj in the generation of the dataset.

F. Statistical Analysis

The statistical analysis consisted in two three-way ANOVAs aiming at studying the effect of factors like the SNR, the amount of data (LENGTH) and the type of adjustments for multiple comparisons on the percentages of false positives and false negatives returned after the statistical assessment of PDC. The within main factors of the ANOVA were SNR (with five levels: [0.1, 1, 3, 5, 10]), LENGTH (with four levels: [3000, 10000, 20000, 30000] data samples) and CORRECTION (with three levels: no corrections, FDR correction and Bonferroni adjustment).

The dependent variables were the percentages of false positives and false negatives returned after the statistical assessment of PDC. The post-hoc analysis with the Duncan test at a statistical significance level p=0.05 was then performed.

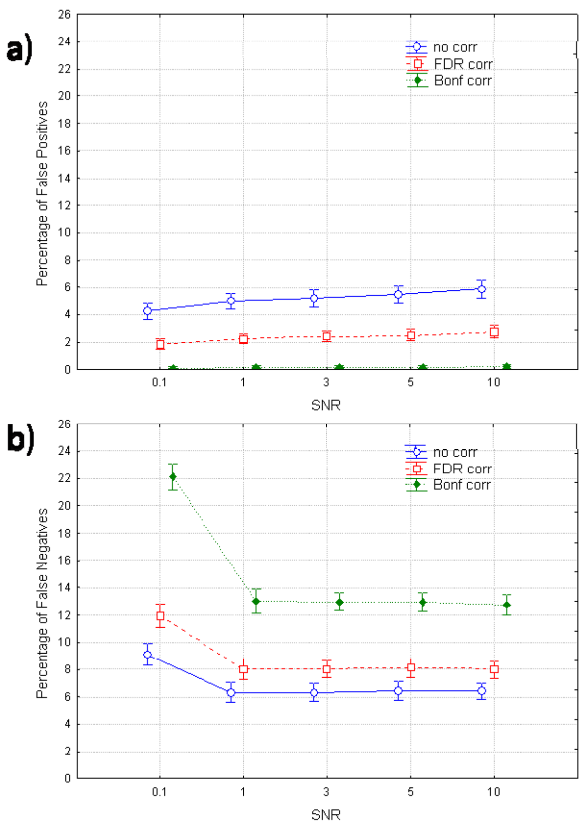

Fig. 2. Results of ANOVA performed on the Percentage of False Positives (a) and False Negatives (b): plot of means with respect to the interaction between the signal to noise ratio (SNR) and statistical corrections for multiple comparisons (CORR). ANOVA shows a high statistical significance (F=25.4, P<0.0001) for the case a) and (F=372, P<0.0001) for the case b), respectively.

III. Results

A MVAR model of order 16 was fitted to each set of simulated data. The procedure of signal generation and PDC estimation was repeated 100 times for each level of factors SNR and LENGTH in order to increase the robustness of the statistical analysis. The percentages of false positives and negatives were computed for each iteration and then subjected to the three way ANOVA.

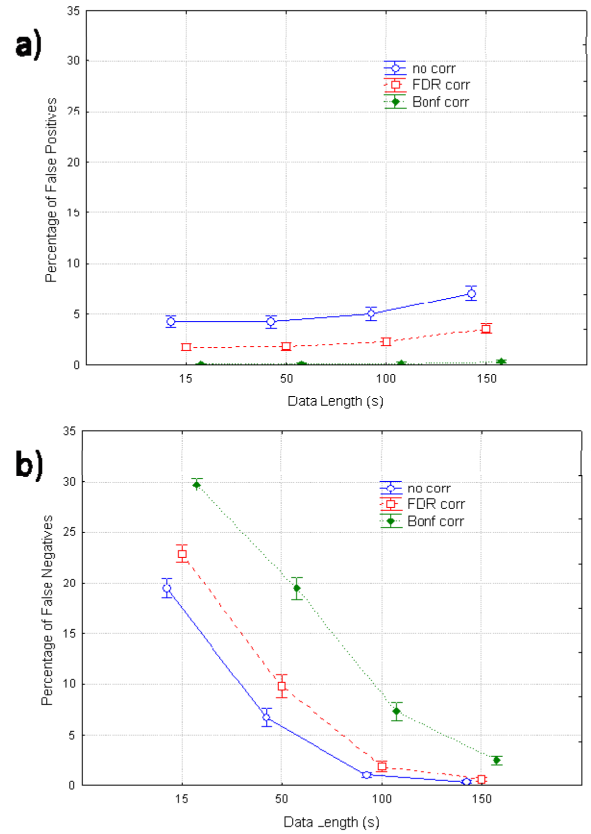

Fig. 3. Results of ANOVA performed on the Percentage of False Positives (a) and False Negatives (b): plot of means with respect to the interaction between signal LENGTH (in seconds) and statistical corrections for multiple comparisons (CORR). ANOVA shows a high statistical significance (F=141.11, P<0.0001) for the case a) and (F=1145.3, P< 0.0001) for the case b), respectively.

Results of three way ANOVA computed by setting as dependent variable the percentage of false positives revealed a strong statistical influence of the main factors SNR ($F = 27.76$, $p < 0.0001$), LENGTH ($F = 184.75$, $p < 0.0001$), and CORRECTION ($F=9308.8$, $p<0.0001$), as well as their interactions SNR x LENGTH ($F=7.36$, $p<0.0001$), SNR x CORRECTION ($F=25.43$, $p<0.0001$), LENGTH x CORRECTION ($F=141.11$, $p<0.0001$).

Results of three way ANOVA computed by setting as dependent variable the percentage of false negatives revealed a strong statistical influence of the main factors SNR ($F=460.62$, $p<0.0001$), LENGTH ($F=9520$, $p<0.0001$), and CORRECTION ($F=16661$, $p<0.0001$), as well as their interactions SNR x LENGTH ($F=213.45$, $p<0.0001$), SNR x CORRECTION ($F=372{,}01$, $p<0.0001$), LENGTH x CORRECTION ($F=1145.3$, $p<0.0001$).

Fig. 2 shows the influence of the different levels of the main factors CORRECTION and SNR on the percentage of false positives (panel a) and false negatives (panel b). The bar on each point represents the 95% confidence interval of the mean errors computed across the simulations. The plot in panel a) indicates a low increase of the percentage of false positives for increased values of SNR of the generated signals, with the mean values remaining under 6% for all the SNR levels and all the corrections.

The plot in panel b) shows a high decrease of the percentage of false negatives for increased values of SNR of the generated signals, but the mean percentage of false negatives overcame 6% threshold for all the SNR levels and all the corrections. Post hoc analysis revealed statistical differences between SNR 0.1 and all the other SNR levels, for each correction.

Fig. 3 shows the influence of the different levels of the main factors CORRECTION and LENGTH on the percentage of false positives (panel a) and false negatives (panel b). The bar on each point represents the 95% confidence interval of the mean errors computed across the simulations. The plot in panel a) indicates a low increase of the percentage of false positives for increased values of signals LENGTH, even if the mean values remained under 8% for all the LENGTH levels and all the CORRECTIONS.

The plot in panel b showed a high decrease of the percentage of false negatives for increased values of data LENGTH, reaching values below 5% for LENGTH 100 and 150s for FDR and Bonferroni CORRECTIONS. Post hoc analysis revealed statistical differences between all the LENGTH levels for all the considered corrections.

IV. DISCUSSION

The results provided by the simulation study suggest that the new method for the assessment of PDC significance is a valid tool for the validation of connectivity patterns. In fact, considering SNR and LENGTH values largely met, for instance, in EEG recordings, the occurrence of type I errors is below 6% for all the three CORRECTIONS levels. Higher percentages of type II errors resulted for both SNR and LENGTH levels. However, it must be noted that the presence of a very weak connection in the model imposed to simulated data (2->3) could be responsible of the increase of the number of false negatives. As expected, as the severity of correction method increased (from no corrections to FDR and Bonferroni) the percentage of false positives is reduced and the percentage of false negatives is increased. In particular, the FDR method seems to provide the best compromise in preventing both type I and type II errors.

In conclusion, the estimation of connectivity patterns on high quality data (good SNR or huge amount of samples) can assure low percentages of both type I and type II errors even without considering severe statistical corrections such as Bonferroni. If the data are characterized by low SNR or signals length, statistical corrections are requested for controlling the estimation performances in terms of type I errors. However, it should be taken into account that this could lead to a loss of weaker connections.

V. CONCLUSION

Results achieved by the simulation study highlighted the possibility to use the new Asymptotic Statistic as a valid alternative to the already existent procedures for the validation of the connectivity patterns estimated by PDC.

In fact, such method overcomes the limits of SCC, by providing a statistical threshold for each connection in the network and for each frequency sample and it takes into account the risk of type I errors. Moreover, the procedures adopted in this approach are faster and less computationally demanding than those used in the Shuffling method, which is currently the standard in the field. Future studies will be focused on a systematic and statistical comparison with the performances (in terms of false positives and false negatives) achieved by the previously available validation methods.

REFERENCES

[1] L. A. Baccalá and K. Sameshima, “Partial directed coherence: a new concept in neural structure determination”, *Biological Cybernetics*, vol. 84, pp. 463-474, 2001.

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

[3] L. A. Baccalá, D. Y. Takahashi, K. Sameshima, “Generalized Partial Directed Coherence (Published Conference Proceedings)”, in *Proc. 15th International Conference on Digital Signal Processing*, Cardiff, 2007, pp. 162-166.

[4] L. Faes, G. Nollo, “Extended causal modeling to assess Partial Directed Coherence”, *Biological Cybernetics*, vol. 103, pp. 387–400, 2010.

[5] L. Astolfi, F. Cincotti, D. Mattia, M. G. Marciani, L. A. Baccalà, F. De Vico Fallani, S. Salinari, M. Ursino, M. Zavaglia, F. Babiloni, “Assessing cortical functional connectivity by partial directed coherence: simulations and application to real data”, *IEEE Transactions on Bio-Medical Engineering*, vol. 53, pp. 1802-1812, Sept. 2006.

[6] K. Sameshima e L. A. Baccalá, “Using partial directed coherence to describe neuronal ensemble interactions”, *Journal of Neuroscience Methods*, vol. 94, pp. 93-103, Dec. 1999.

[7] M. Kaminski, M. Ding, W. A. Truccolo, e S. L. Bressler, “Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance”, *Biological Cybernetics*, vol. 85, pp. 145-157, 2001.

[8] D. Y. Takahashi, L. A. Baccal, K. Sameshima, “Connectivity Inference between Neural Structures via Partial Directed Coherence”, *Journal of Applied Statistics,* vol. 34, pp. 1259-1273, 2007.

[9] T. Nichols e S. Hayasaka, “Controlling the familywise error rate in functional neuroimaging: a comparative review”, *Statistical Methods in Medical Research*, vol. 12, pp. 419-446, Oct. 2003.

[10] H. Akaike, “A new look at statistical model identification”, *IEEE Trans Automat Control,* vol. 19, pp. 716–723, 1974.

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

[12] C. Bonferroni, “Teoria statistica delle classi e calcolo delle probabilità”, *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze*, 1936.

[13] Y. Benjamini and Y. Hochberg, “Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing”, *Journal of the Royal Statistical Society*, vol. 57, pp. 289-300, Jan. 1995.