Individual Cortical Connectivity changes after stroke: a resampling approach to enable statistical assessment at single-subject level*

Petti M, Pichiorri F, Toppi J, Cincotti F, Salinari S, Babiloni F, Mattia D and Astolfi L.

***Abstract*— One of the main limitations commonly encountered when dealing with the estimation of brain connectivity is the difficulty to perform a statistical assessment of significant changes in brain networks at a single-subject level. This is mainly due to the lack of information about the distribution of the connectivity estimators at different conditions. While group analysis is commonly adopted to perform a statistical comparison between conditions, it may impose major limitations when dealing with the heterogeneity expressed by a given clinical condition in patients. This holds true particularly for stroke when seeking for quantitative measurements of the efficacy of any rehabilitative intervention promoting recovery of function. The need is then evident of an assessment which may account for individual pathological network configuration associated with different level of patients' response to treatment; such network configuration is highly related to the effect that a given brain lesion has on neural networks. In this study we propose a resampling-based approach to the assessment of statistically significant changes in cortical connectivity networks at a single subject level. First, we provide the results of a simulation study testing the performances of the proposed approach under different conditions. Then, to show the sensitivity of the method, we describe its application to electroencephalographic (EEG) data recorded from two post-stroke patients who showed different clinical recovery after a rehabilitative intervention.**

I. Introduction

Brain connectivity studies play a predominant role in the comprehension of neuroplasticity which occurs after a brain lesion and during the recovery of impaired functions (e.g. as a consequence of a treatment such as specific stimulations, rehabilitation or training based on motor/cognitive tasks [1]). From EEG signals recorded on the scalp, effective connectivity between different brain regions can be estimated by means of Partial Directed Coherence (PDC) [2]. The networks properties can be described by quantifiable indices, like those derived from classical graph theory [3], which were widely used in different applications [4,5], as well as by indices defined *ad hoc* to investigate specific physiological and topological properties of a brain network [6]. This approach is based on the use of advanced techniques for signals processing, which use all the information available from data to build a unique model, thus not providing a distribution for a single experimental condition: this prevents from performing a statistical comparison between two different conditions at the single subject level.

Group analysis is a powerful tool to overcome such limitation and to obtain generalizable results. However, this approach may be hindered by the difficulty to get homogeneous groups, especially in the case of patients, that usually present very heterogeneous characteristics. Moreover, assessing brain organization changes in a specific patient (for instance, as a consequence of a treatment) is of particular interest for clinical applications.

In this study, we propose an approach based on resampling procedures, applied to multi-trial EEG data in order to obtain different datasets representing a distribution of observations for each condition. The working hypothesis is that connectivity patterns computed on such datasets can provide a distribution of patterns allowing to statistically compare brain networks, or their related connectivity indices, at different conditions, for a single subject. Two resampling approaches are here evaluated, based on Jackknife and bootstrap procedures, respectively [7]. The variability introduced in the connectivity estimates by the two approaches was tested by means of a simulation study, providing results that were subjected to an analysis of variance (ANOVA) to evaluate the effects of different factors adopted during the procedure. Then, we applied the procedure to eyes-closed resting state EEG data recorded on subacute stroke patients preceding and following a rehabilitative intervention. In particular, to show the sensitivity of the proposed approach, we considered two patients who showed different clinical recovery after the rehabilitation and we linked the significant variations in brain connectivity to the clinical outcome of the intervention.

II. Methods

A. Partial Directed Coherence

The PDC [2] is a full multivariate spectral measure, used to determine the directed influences between any given pair of signals in a multivariate data set. As a frequency-domain version of Granger causality [8], PDC reveals the existence, the direction and the strength of a functional relationship

*Research partially supported by the Italian Ministry of Health (grant:RF-2010-2319611).

M. Petti (Corresponding Author. Phone: +39 06 51501165; e-mail: manuela.petti@uniroma1.it), L. Astolfi, J. Toppi, F. Cincotti, are with the Department of Computer, Control and Management Engineering, Univ. of Rome "Sapienza" and with Fondazione Santa Lucia Hospital, Rome, Italy.

D. Mattia, F. Pichiorri are with Fondazione Santa Lucia, Rome, Italy.

S. Salinari is with the Department of Computer, Control and Management Engineering, Univ of Rome "Sapienza", Rome, Italy.

F. Babiloni is with the Dept. of Physiology and Pharmacology of the University of Rome "Sapienza" and with Fondazione Santa Lucia, Rome, Italy.

between any given pair of signals in a multivariate data set. It is possible to define PDC as:

$$\pi_{ij}(f) = \frac{\Lambda_{ij}(f)}{\sqrt{\sum_{k=1}^{N} \Lambda_{kj}(f)\Lambda^{*}_{kj}(f)}}, \sum_{n=1}^{N} \left|\pi_{nj}(f)\right|^2 = 1 \quad (1)$$

where *Λ(f)* is a matrix containing the coefficients of associated Multivariate Autoregressive (MVAR) model.

In this study we used the squared formulation of PDC due to its higher accuracy and stability [9]. The patterns significance was assessed by means of an asymptotic statistic method [10].

B. Resampling implementation

To achieve a distribution of connectivity estimations during a single experimental condition, in this study we defined and implemented two methods to perform resampling on EEG data. Given an EEG dataset, characterized by a certain number of trials N_T, we explored two approaches: the Jackknife and the bootstrap [7]. Each method can be applied for *K* iterations, allowing to obtain *K* EEG dataset to be subjected to the connectivity estimation.

The Jackknife method performs a leave-*N*-out approach to trials, where *N* is the percentage of trials to be randomly excluded from the estimation: in this way, each *k* (*k=1,…,K*) dataset is characterized by a number of trials N_E which is smaller than the whole initial dataset.

$$N_E = N_T - N < N_T \quad (2)$$

The second approach is based on data resampling: it keeps constant the number of trials through the exclusion of some of them and the repetition of others, as shown in equation 3:

$$N_E = N_T - N + r \times t = N_T \quad (3)$$

where N_E is the number of trials of each *k* (*k=1,…,K*) dataset to be subjected to the PDC estimation, *N* is the percentage of trials to be randomly excluded and *t* is the number of trials to be repeated *r* times (*r* and *t* are not independent, as their product must be equal to *N)*.

C. Simulation study

In order to test the distribution of connectivity patterns obtained by the proposed approach, we performed a simulation study, in which, starting from a set of EEG data, obtained from a resting-state recording in a healthy subject, we applied both approaches, systematically varying the following factors:

- number of replications of the resampling procedure (factor *K*; four levels: 30, 100, 250, 500);
- percentage of excluded trials in the Jackknife procedure (factor *N;* three levels: 1%, 5%, 10%);
- percentage of excluded trials (factor *N;* 5 levels: 2%, 5%, 10%, 20%, 50%) and number of repetitions of the *t* trials (factor *r;* 3 levels: 1, 2, 4) in the bootstrap procedure.

To evaluate the effect of such factors on the variability of the obtained connectivity distribution, we computed the Frobenius norm of standard deviation, along the resampling replications, of connectivity estimation. In equation 4, the formulation of Frobenius norm of the *N* x *N* matrix SD of standard deviation (*N* is the number of channels) is shown:

$$\left\|SD\right\|_F = \left(\sum_{i=1}^{N} \sum_{j=1}^{N} a_{ij}^{2} \right)^{1/2} \quad (4)$$

where a_{ij} is the entry *ij* of matrix SD. This dependent variable was subjected to two ANOVAs for repeated measures, with 2 within factors (*K*, *N*) and 3 within factors (*K*, *N*, *r*) for Jackknife and bootstrap, respectively.

D. Application to real data

We tested the proposed approach on EEG data acquired from two subacute stroke patients enrolled in a rehabilitative treatment protocol to enhance motor recovery of the upper limb (males, aged 58 and 62, both suffered a left hemispheric stroke 2 months prior to enrollment). Both patients underwent a Brain Computer Interface (BCI)–assisted upper limb Motor Imagery (MI) training (12 sessions, total duration of the treatment 1 month).

To evaluate the sensitivity of the proposed approach, we selected two patients who showed a different clinical recovery, based on the upper limb section of the Fugl-Meyer Assessment (FMA) scale. FMA is a stroke-specific, performance-based impairment index designed to assess motor functioning in patients with post-stroke hemiplegia [11]. A Minimal Clinically Important Difference (MCID) for the upper limb FMA scale was set to 7 points [12]. With respect to this definition, one of the patients reached the MCID after the training (14 points), while the other did not (4 points).

Two minutes of eyes closed resting state EEG (standard 10-20 montage, 20 electrodes) were acquired in two sessions: one preceding (PRE) and the other following (POST) the intervention. Preprocessing included down-sampling at 100 Hz, band pass filtering (1-45 Hz), artifact rejection and 1s-epochs segmentation: with this procedure we obtained for each patient and each condition, an EEG dataset consisting of approximately 100 artifact-free trials. The resampling method was applied setting parameters as follows:

- Bootstrap replications *K* = 20;
- number of trials excluded *N* = 50%;
- repetition of trials *r* = 4.

Brain connectivity achieved between all possible pairs among the 19 electrodes for each frequency in the range [1:45] Hz were averaged within 5 frequency bands, defined according to Individual Alpha Frequency [13]: theta [IAF-6;IAF-2], alpha [IAF-2;IAF+2], beta1 [IAF+2;IAF+11], beta2 [IAF-11;IAF+20] and gamma [IAF+20;IAF+35].

To describe and quantify a networks property known to

be related to stroke effects we computed the Inter-Hemispheric Connectivity (IHC) from PDC patterns estimated in the conditions PRE- and POST-training, for each frequency band. IHC is defined as the mean value of the weights of all the inter-hemispheric connections (*IHC weight*):

$$IHCweight = \frac{\sum_{i=1}^{N_{IHC}} w_i}{N_{IHC}} \quad (5)$$

were N_{IHC} is the number of estimated inter-hemispheric connections and w_i is the weight of *ith* IHC. We performed an ANOVA with *IHC weight* as dependent variable, between factor: subject, within factor: PRE- vs POST-training.

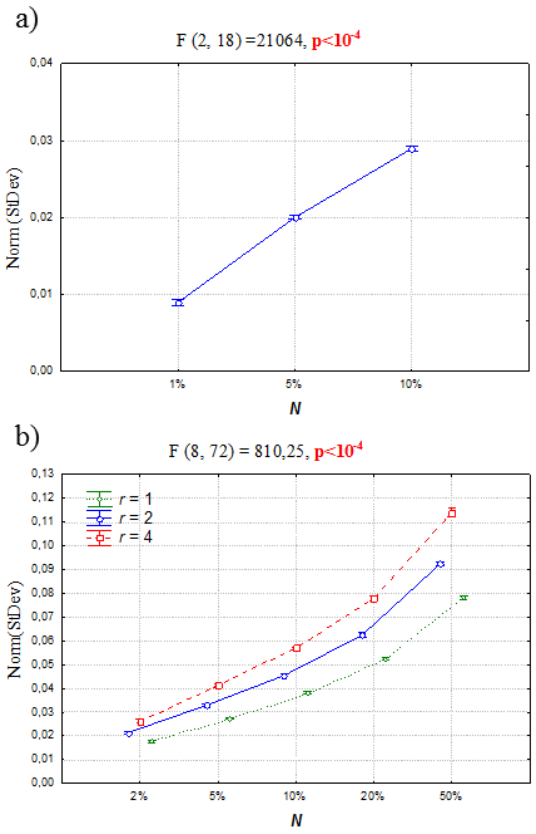

Figure 1. ANOVA performed on the norm of standard deviation of the connectivity distribution. a) plot of means with respect to the factor *N* (number of trials excluded) for Jackknife approach; b) plot of means with respect to the factors *N* (number of trials excluded) and *r* (number of repetitions of *t* trials) for bootstrap approach.

III. Results

A. Simulation study

Results of the ANOVAs performed on the norm of standard deviation revealed that the choice of the factor *K* (resampling replications number), in the proposed range, has no significant effect on any of the methods, whereas the parameters *N* and *r* associated to the two approaches significantly influence the variability of the distributions. Fig. 1.a shows how the standard deviation increases significantly with the number of excluded trials (*N*) in Jackknife approach. The norm of standard deviations obtained is in the order of 10^{-2}. Similarly, in the case of bootstrap (Fig. 1.b) the norm significantly increases with parameter *N* but also with parameter *r*. In this case, standard deviation range is wider ($10^{-2} \div 10^{-1}$), thus allowing to select parameters in order to be more conservative, i.e. to increase variability and to prevent false positives which may result from a distribution with a low standard deviation.

B. Application to real data

Results of ANOVA performed on IHC weight as dependent variable are showed in Figure 2 for Alpha and Beta1 bands. Post-hoc analysis revealed a significant difference in Alpha band (Fig. 2.a) between the PRE and POST conditions for the patient who reached the MCID, indicating a significant increase of inter-hemispheric connectivity after the training with respect to the PRE condition. No significant differences were revealed by the test for the patient who did not reach the MCID. In Beta1 (Fig. 2.b) both patients showed a significant increase of the IHC weight, with a significantly higher value for the patient who achieved a good clinical recovery.

IV. Discussion

The two methods here evaluated for resampling the EEG data have proved to be able to produce a distribution of connectivity in a single condition and for a single subject, allowing statistical analysis. As might be expected, the standard deviation of such distribution increases with the number of excluded trials, for both approaches. However, this affects differently the two methods. In fact, for the Jackknife approach, the increase of standard deviation is related to a loss of accuracy, due to the reduction of the amount of data used for the estimation, which is more and more considerable as N increases. The bootstrap approach, on the contrary, keeps the data amount fixed, and the increase of N, while improving variability, is not expected to significantly compromise the quality of the estimation. Furthermore, Jackknife, while not allowing to increase N to the same extent allowed by bootstrap (to avoid reducing the dataset under the level needed for the connectivity estimation) can lead to a lower variability of the data with respect to the bootstrap, which may enhance the differences when contrasting different conditions and result in false positives. For these reasons, the bootstrap method was chosen to be applied to real data, and we set the parameters to values providing higher standard deviations, to ensure a conservative approach.

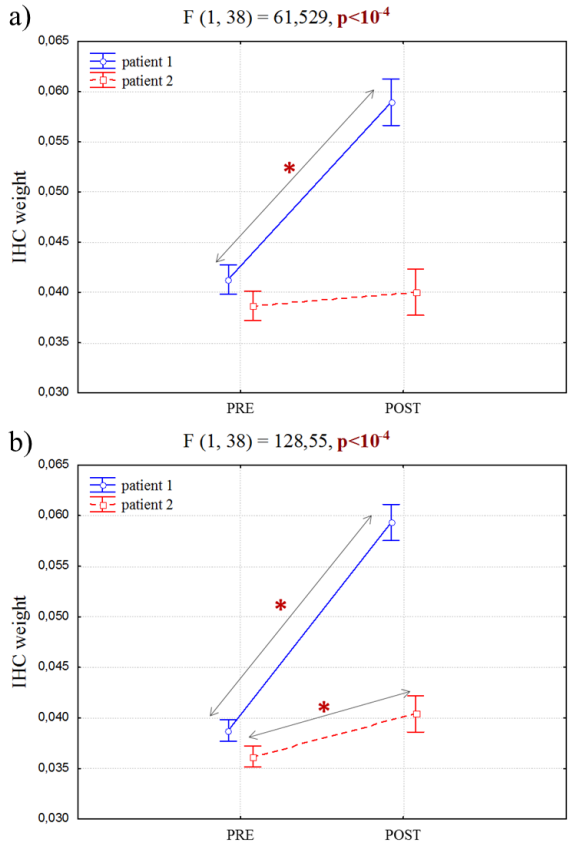

Figure 2. ANOVA performed with *IHC weight* as dependent variable, between factor: subject (patient 1 reached the MCID after the training, while patient 2 did not), within factor: PRE- vs POST-training. a) Alpha band; b) Beta band. The symbol (*) indicates a significant difference resulting from post-hoc tests ($p<0.05$).

The application to patients' data had the aim to show the application of the proposed methodology at a single subject level. We selected two patients that showed a different clinical recovery after a BCI-assisted MI rehabilitative training, thus providing a good test for the sensitivity of the method proposed to changes resulting in the brain networks organization after a rehabilitative intervention. There are evidences of the reduction of inter-hemispheric connectivity as a consequence of stroke, which may be detected even at resting condition [14]. For this reason, we adopted the IHC weight as a descriptor of the network properties to be investigated. Results suggested a significant increase of inter-hemispheric connectivity at rest, in Alpha band, for the patient with a good clinical recovery. This is in line with previous studies indicating significant resting state Alpha rhythm connectivity properties in post stroke patients [15]. Focusing on the rhythm specifically trained by the BCI-MI intervention (i.e. the lower Beta), both patients showed an increase of inter-hemispheric connections, indicating an effect of the training in both cases. However, this increase was significantly higher for the patient with a better clinical recovery, which may suggest that the index is able to quantify the effects of the intervention at the level of cortical organization, in agreement with the clinical outcome.

V. Conclusion

The results of this study suggest that a resampling approach aiming at building a distribution for each experimental condition may be a useful tool to overcome the limitations in the single-subject connectivity evaluation. The application to resting state EEG data in two post stroke patients pre- and post- rehabilitative intervention showed that the proposed method is sensitive to changes related to the clinical outcome, and suggests that it might be adopted, in support of the clinical scales, for the evaluation of the efficacy of a treatment.

References

[1] De Vico Fallani F, Astolfi L, Cincotti F, Mattia D, Marciani MG, Tocci A, Salinari S, Witte H, Hesse W, Gao S, Colosimo A and Babiloni F 2008, "Cortical network dynamics during foot movements". Neuroinformatics, DOI 10.1007/s12021-007-9006-6.

[2] Baccalá LA and Sameshima K, "Partial directed coherence: a new concept in neural structure determination", Biological Cybernetics, vol. 84, pp. 463-474, 2001.

[3] M. Rubinov and O. Sporns, "Complex network measures of brain connectivity: uses and interpretations," *Neuroimage*, vol. 52, no. 3, pp. 1059–1069, Sep. 2010.

[4] M. Petti , J. Toppi, F. Pichiorri, F. Cincotti, S. Salinari, F. Babiloni, L. Astolfi, and D. Mattia "Aged-related changes in brain activity classification with respect to age by means of graph indexes," Conf Proc IEEE Eng Med Biol Soc, 2013.

[5] Wang L, Yu C, Chen H, Qin W, He Y, Fan F, Zhang Y, Wang M, Li K, Zang Y, Woodward TS and Zhu C. "Dynamic functional reorganization of the motor execution network after stroke." Brain 2010: 133; 1224–1238.

[6] J. Toppi, M. Petti, F. De Vico Fallani, G. Vecchiato, A. G. Maglione, F. Cincotti, S. Salinari, D. Mattia, F. Babiloni, and L. Astolfi, "Describing Relevant Indices from the Resting State Electrophysiological Networks," Conf Proc IEEE Eng Med Biol Soc, 2012.

[7] B. Efron, "Bootstrap methods: another look at the Jackknife," vol. 7, pp. 1–26, 1979.

[8] Granger CWJ, "Investigating causal relations by econometric models and cross-spectral methods," Econometrica, vol. 37, pp. 424–438, 1969.

[9] Astolfi L, Cincotti F, Mattia D, Marciani MG, Baccalà 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 Transactions on Bio-Medical Engineering, vol. 53, pp. 1802-1812, Sept. 2006.

[10] Takahashi DY, Baccalà LA, Sameshima K, 2007. "Connectivity Inference between Neural Structures via Partial Directed Coherence". Journal of Applied Statistics 34, 1259–1273.

[11] Fugl-Meyer, Jaasko, Leyman, Olsson, & Steglind, The post stroke hemiplegic patient. I. A method for evaluation of physical performance. Scand J Rehabil Med 1975;7:13-31.

[12] Sivan M, O'Connor RJ, Makower S, Levesley M, Bhakta B, "Systematic review of outcome measures used in the evaluation of robot-assisted upper limb exercise in stroke." J Rehabil Med 2011; 43: 181–189

[13] Klimesch W, 1999. "EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis". Brain Research Reviews 29, 169–195.

[14] Wu W, Sun J, Jin Z, Guo X, Qiu Y, Zhu Y, Tong S, "Impaired neuronal synchrony after focal ischemic stroke in elderly patients". Clin Neurophysiol. 2011; 122(1): 21-6.

[15] Dubovik S, Ptak R, Aboulafia T, Magnin C, Gillabert N, Allet L, Pignat JM, Schnider A, Guggisberg AG, EEG alpha band synchrony predicts cognitive and motor performance in patients with ischemic stroke. Behav Neurol. 2013;26(3):187-9.