Computational Neuroscience

Constructing fMRI connectivity networks: A whole brain functional parcellation method for node definition

Eleonora Maggioni[a,b], Maria Gabriella Tana[a,c,d], Filippo Arrigoni[b], Claudio Zucca[b], Anna Maria Bianchi[a,*]

[a] *Department of Electronics Information and Bioengineering (DEIB), Politecnico di Milano, P.za Leonardo da Vinci, 32, 20133 Milan, Italy*
[b] *Scientific Institute IRCCS E.Medea, Via Don Luigi Monza 20, 23842 Bosisio Parini, Lecco, Italy*
[c] *BIND - Behavioral Imaging and Neural Dynamics Center, University "G. d'Annunzio", Chieti, Italy*
[d] *Department of Medicine and Aging Science, University "G. d'Annunzio", Chieti, Italy*

Article history:
Received 26 July 2013
Received in revised form 12 March 2014
Accepted 13 March 2014

1. Introduction

In the last years it came to light that cognitive functions such as vision, attention, memory and language are the result of interactions among highly specialized regions which are widely distributed in the brain (Edelman, 1993; Tononi et al., 1992, 1994). Currently, studies of brain region activations are often integrated with brain connectivity analysis, in order to achieve information on

* Corresponding author at: Department of Electronics Information and Bioengineering (DEIB), Politecnico di Milano, P.za Leonardo da Vinci, 32, 20133, Milan (MI), Italy. Tel.: +39 0223993342; fax: +39 0223993360.
E-mail address: annamaria.bianchi@polimi.it (A.M. Bianchi).

how these regions communicate. Connectivity patterns can refer to a network of anatomical links (structural connectivity), of functional interactions measured in terms of statistical dependencies (functional connectivity) or finally to a network of causal relationships (effective connectivity). While Functional Connectivity (FC) is defined as the temporal correlation between spatially remote neurophysiological or vascular events, the concept of Effective Connectivity (EC) refers to the influence that one system exerts on another, either directly or indirectly (Friston et al., 1993).

As the Blood Oxygen Level Dependent (BOLD) signal has been confirmed to be correlated to neural activity (Logothetis et al., 2001), functional Magnetic Resonance Imaging (fMRI) has become one of the most commonly used noninvasive neuroimaging techniques for the investigation of cerebral connectivity.

One of the critical features of fMRI connectivity analysis regards the spatial definition of the network nodes on which connectivity is measured. When a General Linear Model (GLM) analysis is used to identify active areas, it is common practice to take directly the active clusters as nodes (Deshpande et al., 2008, 2009; Sato et al., 2009) or to manually design spheres (or box shapes) centered on the local maxima of active clusters (Chen et al., 2009; Gao et al., 2011). These approaches do not consider any kind of inhomogeneity inside the active clusters, either anatomical or functional. Instead, nodes should ideally represent brain regions with coherent patterns of extrinsic anatomical or functional connections (Rubinov and Sporns, 2010). Thus, nodes should be defined by grouping voxels with both similar anatomical characteristics and a homogeneous functional behavior. In addition, nodes should not spatially overlap and should be derived from a parcellation scheme which completely covers the surface of the cortex or of the entire brain (Rubinov and Sporns, 2010).

In order to obtain anatomically homogeneous nodes, brain parcellations using probabilistic atlases (Wang et al., 2009) or prior knowledge about the anatomy (e.g. sulcus geometry) (Mangin et al., 2004) can be used. The anatomical parcellation procedure covers all the brain volume without spatial overlapping, but does not take into account any functional dissimilarities within the resulting parcels. By contrast, parcellation based on functional information has its specific challenges, since purely functional clustering algorithms work well only locally and specific methods have to be designed for the parcellation of the entire brain (Grill-Spector et al., 2004).

To enhance the functional homogeneity of the parcels while keeping them connected, anatomo functional parcellation techniques were introduced mainly based on k means clustering algorithms (Thirion et al., 2006; Michel et al., 2012; Kim et al., 2010; Gonzalez-Castillo et al., 2012; Tana et al., 2012). In some of these techniques, the functional distance measure is calculated directly from the time series using correlation metrics (Gonzalez-Castillo et al., 2012); in others, the functional information is derived from a set of features previously extracted from time series, e.g. the beta coefficient or T or F maps estimated during the GLM analysis (Thirion et al., 2006; Michel et al., 2012; Tana et al., 2012). The approaches based on k means have the limitation to require the initial definition of the number of parcels, since the latter was determined by adopting an explorative approach consisting in trying several different values (Gonzalez-Castillo et al., 2012; Makni et al., 2008) or by using the Bayesian Information Criterion (BIC) and cross validation techniques on multiple subjects (Thyreau et al., 2006).

In summary, the desired characteristics of a parcellation algorithm are the capability to produce parcels functionally homogenous and spatially connected, and the possibility to automatically determine the number of parcels.

To cope with these requirements, we developed an intra subject parcellation method that accounts for both spatial information and functional content in the fMRI time series. The method starts from an atlas based parcellation that constrains parcels to lie in one anatomical region (Zalesky et al., 2010) and then splits each anatomical region into functional groups. Voxels with close functional profiles are aggregated in parcels using a statistical measure known as Tononi's cluster index (*TCI*) (Tononi et al., 1998), which is calculated directly on the temporal dynamics of fMRI data. *TCI* measures the instantaneous correlation between voxels by evaluating their mutual statistical dependence in relation to the statistical dependence between them and the rest of the system. The algorithm makes use of the *TCI* values to compare clusters with incremental dimension and to select the ones with the highest functional homogeneity. During the aggregative procedure, which starts from duplets of voxels and gradually adds new voxels to them, a spatial constraint is introduced to guarantee the spatial continuity within each cluster. A crisp functional partition of the system is thus created and the spatial connectedness is meanwhile preserved. The proposed method does not require to define a priori the number of parcels in the system, since the number of resulting parcels depends on the amount of synchronization between the voxels time series.

First, we tested our functional parcellation algorithm on two synthetic datasets, simulating fMRI data acquired both in absence and in presence of stimuli. The simulation results proved the efficiency of the proposed algorithm in grouping time series with a similar functional content. We then tested the algorithm on a set of healthy subjects, during a visual stimulation paradigm. We selected a region in primary visual cortex, the cuneus, and discussed the parcellation results: the clusters distribution was similar within the group, with the main cluster largely overlapping from subject to subject, thus providing confirm of the algorithm reliability. The method was finally applied to a second real dataset, in which fMRI real data were simultaneously recorded with ElectroEncephaloGram (EEG) from two epileptic patients, with the final purpose of studying the EC pattern during ictal activity. In this application, the whole brain parcellation was used to define the nodes of the active network, on which we subsequently applied a connectivity analysis using Granger Causality (GC) measures (Roebroeck et al., 2005). The use of a parcellation scheme for the separation of inhomogeneous signals into different nodes allows a more meaningful connectivity analysis, especially when directionality has to be inferred. We verified that connectivity measures obtained with the novel topological network were congruent with the semiology of the patients' pathologies. Therefore, results on real data confirmed the usefulness of the new parcellation as a preparatory step of node definition for connectivity analysis.

2. Materials and methods

2.1. The Tononi's cluster index

The method here proposed is based on a cluster index proposed by Tononi in (Tononi et al., 1998). The *TCI* of a subset in a system measures the statistical dependence of the elements within the subset in relation to the statistical dependence between this subset and the rest of the system (Tononi et al., 1998). Consider a system S composed of N time series $\{s_i\}$, partitioned into the jth subset S_j^k of $k < N$ elements and its complementary subset $(S - S_j^k)$. The *TCI* is given by:

$$TCI(S_j^k) = \frac{I(S_j^k)}{MI(S_j^k; S - S_j^k)} \tag{1}$$

where $I(S_j^k)$ is the integration of the subset S_j^k and $MI(S_j^k; S - S_j^k)$ is the mutual information between S_j^k and $(S - S_j^k)$ (Tononi et al., 1998). The *TCI* is high for subsets showing both high internal

statistical dependence and high independency (low *MI*) from the rest of the system. Further details on how the *I* and *MI* are computed can be found in Appendix A.

2.2. Brain functional parcellation algorithm

We developed a functional parcellation method which segments fMRI data by grouping BOLD time series into subsets on the basis of their corresponding *TCI* values. The algorithm was implemented in Matlab R2010a (http://www.mathworks.it). Before applying the proposed functional clustering algorithm, a pre processing step was performed, in which the whole brain fMRI volume was first anatomically parceled by using the Harvard Oxford (HO) cortical and subcortical atlas (http://www.cma.mgh.harvard.edu) (Desikan et al., 2006). In this way, each parcel is constrained to lie within one and only one anatomical region. This condition precludes the formation of meaningless parcels, such as those that span both hemispheres or multiple, anatomically distinct regions (Zalesky et al., 2010). The *TCI* algorithm was thus applied on each pre defined anatomical region by performing the steps summarized in Fig. 1, that will be fully explained in the following paragraphs.

2.2.1. System initialization

Hereinafter, the system *S* indicates an anatomical region composed by *N* voxels. As in other applications of the *TCI*, the system Probability Density Function (PDF) was approximated with a multidimensional Gaussian: this allowed to completely characterize the system by its covariance matrix (Jones, 1979). The entropy, integration and mutual information of either the system or its subsets could thus be calculated from the determinant of the corresponding covariance matrices, as performed in (Tononi et al., 1998). These values can vary depending on the system size and on the time series length. Typically, in systems having hundreds of elements the covariance matrix is singular, the entropy *H* and the integration *I* are thus infinite and *TCI* cannot be calculated. Although the remarkable similarity between BOLD time series certainly contributed to the covariance matrix singularity, its occurrence could also be related to the balance between the number of time samples and elements in the system, which influenced the calculation of the statistical parameters of interest. Therefore, the first step of the algorithm is to calculate the determinant of the covariance matrix. A specific step of system size reduction was thus designed to reduce the dimension of large systems with singular covariance matrix, until *H* and *I* assume finite values. The system size reduction step will be described in detail in Section 2.2.2. To compare different subsets, the Tononi's cluster index of each subset $TCI(S_j^k)$ has to be normalized with respect to the subset size and to the null hypothesis of absence of functional clusters (Tononi et al., 1998). The normalized Tononi's cluster index TCI_n is:

$$TCI_n(S_j^k) = \frac{TCI(S_j^k)}{TCI_{homo}^k} \tag{2}$$

where TCI_{homo}^k is the average of the *TCIs* of *k* dimensional subsets belonging to homogeneous systems equivalent to the original.

In the initialization step, a set of homogeneous systems for *TCI* normalization is created. An iterative procedure is performed in order to create a set of homogeneous systems equivalent to the original one. Two systems are defined as equivalent if they have similar integration (i.e. the integration error, defined as the difference between their integration values, is below a given threshold). The steps performed to obtain such set of homogeneous systems are the following.

1. Let *S* be the original system and $S_{ideal,homo}$ an ideal homogeneous system, (i.e. a system where all the pairwise correlation coefficients have the same value σ). Creation of $S_{ideal,homo}$ by initially setting σ to 10^{-3}.
2. Calculation of the covariance error *d* as the absolute difference between the determinants of the covariance matrix of the original system and the covariance matrix of the ideal homogeneous system, $d = abs(\left|COV(S)\right| - \left|COV(S_{ideal,homo})\right|)$, where $|\cdot|$ is the determinant and $abs(\cdot)$ the absolute value. If the condition
$$d < 10^{-\alpha} \quad \text{where}\, \alpha = 1 \tag{3}$$
is accomplished, the procedure is continued with the step 3; otherwise, a new $S_{ideal,homo}$ is iteratively created by increasing the value of σ of a factor 10^{-3}, until Eq. (3) is satisfied.
3. Simulation of twenty homogeneous systems $S_{sim,homo}$, using the previously estimated σ value, using the procedure described in (Tononi et al., 1998). Calculation of the integration error ε as the absolute difference between the average integration over the simulated systems and the integration of the original system, $\varepsilon = abs(I(\bar{S}_{sim,homo}) - I(S))$, where $\bar{\cdot}$ is the average. If the condition
$$\varepsilon < 0.01^* I(S) \tag{4}$$
is satisfied, the set of $S_{sim,homo}$ is saved in order to be used for the *TCI* normalization (Eq. (2)), otherwise a new set of homogeneous systems $S_{sim,homo}$ is created, until a maximum of five simulations trials is reached. If Eq. (4) is still not satisfied, the procedure is replicated from step 1 by increasing of 1 the value of the α parameter of Eq. (3).

2.2.2. System size reduction

The system size reduction procedure was specifically designed to deal with those systems with a nonfinite value of integration. It consists in the following steps.

1. Extraction of all the duplets of voxels in which each voxel is coupled to the six nearest neighbors. The distance between two voxels s_i and s_q of coordinates (x_i, y_i, z_i) and (x_q, y_q, z_q) is calculated as
$$d = (|x_i - x_q| + |y_i - y_q| + |z_i - z_q|)/3 \tag{5}$$
2. Calculation of the duplets integration.
3. Ordering of the duplets with descending integration values, so that the duplets with the higher integration are selected.
4. Exclusion of the duplets having at least one element already selected, moving from the top to the bottom of the ordered list.
5. For each selected duplet, generation of a "virtual" voxel whose signal is the averaging of the two original time series and whose spatial coordinates are at the midpoint between the two original coordinates.

If this new system is still characterized by an infinite integration, this procedure is replicated starting from the system of duplets previously created. A further system is thus created, in which each element represents a quadruplet of the original voxels. The aggregation procedure is repeated until the system is characterized by a finite value of integration. Once obtained a reduced system characterized by a non singular covariance matrix, the analysis can proceed as shown in Fig. 1.

2.2.3. Classification of couples of elements

The duplets which satisfy the spatial constraint described in the previous paragraph are extracted (i.e. each voxel is associated only to its six nearest neighbors). This constraint guarantees that the resulting functional parcels are formed by spatially connected voxels (Thirion et al., 2006). Each duplet $\{s_i, s_q\}$ belonging to this group is then analyzed through the following steps:

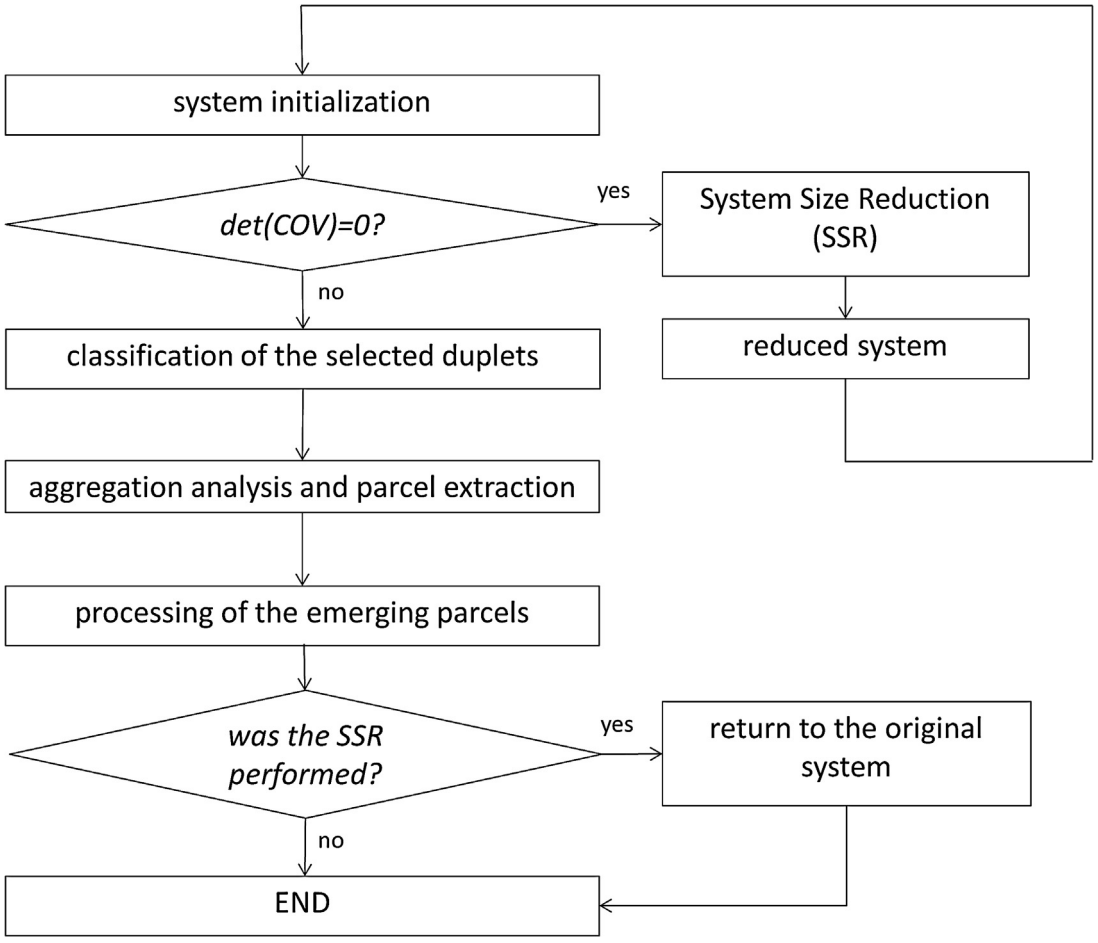

Fig. 1. Parcellation diagram. Main steps of the parcellation algorithm.

1. calculation of the duplet integration $I(\{s_i,s_q\})$.
2. estimation of the mutual information between the duplet and its complementary subset, $MI(\{s_i, s_q\} ; (S - \{s_i, s_q\}))$.
3. calculation of the duplet $TCI(\{s_i,s_q\})$.
4. extraction of all the possible combinations of duplets $\{s_{i,homo},s_{q,homo}\}$ from each of the twenty homogeneous systems saved at the end of the initialization step, calculation of the mean values (over all the duplets and all the homogeneous systems) of integration $I_{homo,2}$ and mutual information $MI_{homo,2}$.
5. calculation of the mean homogeneous *TCI*, $TCI_{homo,2} = I_{homo,2}/MI_{homo,2}$.
6. normalization of the *TCI* of the duplet, $TCI_n(\{s_i,s_q\}) = TCI(\{s_i,s_q\})/TCI_{homo,2}$.

The duplets are then organized in a descending ordered list on the basis of their TCI_n values.

2.2.4. Analysis of subsets with an aggregative procedure

Our parcellation algorithm subdivides the system *S* (i.e. the single anatomical region) in parcels by using an aggregative procedure that gradually adds new voxels to the duplets of voxels of the list computed in the last paragraph.

Starting from the top of the list (i.e. from the duplet with the highest TCI_n), the aggregative procedure on a duplet $\{s_i,s_q\}$ is performed through the following steps:

1. extraction of all the possible triplets $\{s_i,s_q,s_u\}$, where the third voxel s_u is chosen within the two groups of voxels neighbors of s_i and s_q.
2. calculation of the $TCI_n\{s_i,s_q,s_u\}$ of each of the triplets and extraction of the triplet $\{s_i,s_q, s_u\}$ with the highest TCI_n value.
3. starting from the extracted triplet, calculation of the $TCI_n\{s_i,s_q,s_u,s_w\}$ for each quadruplet $\{s_i,s_q,s_u,s_w\}$, where the voxel s_w is chosen in the three groups of voxels neighbors of the voxels s_i, s_q and s_u; extraction of the quadruplet with the highest TCI_n.
4. comparison between the TCI_n values and aggregation of new elements until the TCI_n is minor than the previous one. The subset corresponding to the maximum of TCI_n is saved as functional parcel.

The parcels having elements in common are finally merged in a unique cluster, resulting in a crisp partition of the system. If the system size reduction was performed, at the end of the analysis the results on the reduced system are brought back to the original system, replacing each element of the reduced system with the corresponding original voxels.

The aggregative procedure continues until all the elements of the system are grouped in functional clusters. Therefore, the number of final parcels will depend on the amount of synchronization between the voxels time series and not on a priori user's specifications. Nevertheless, the user can potentially influence the number of final clusters by deciding to stop the aggregative procedure before that all the system is divided into parcels; in this case, each remaining voxel is not used as seed for the aggregation procedure, but is simply inserted in the cluster with which it shares the highest *TCI*. In the latter case, the number of resulting parcels, although

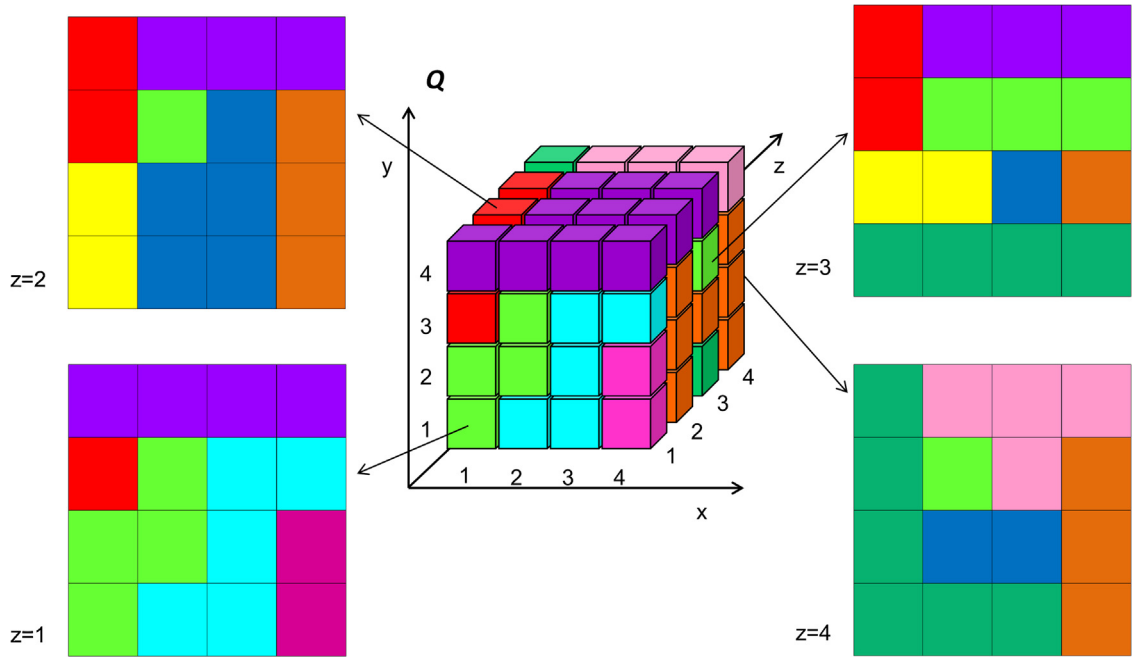

Fig. 2. Simulated system. Functional clusters in a simulated system *Q*, where each element *i* is associated to three spatial coordinates (x_i, y_i, z_i). Different colors are used to represent the 10 different clusters. The three dimensional representation of the system is shown in the center of the figure, while the four sections along the *z*-axis are shown at the corners.

affected by the user's choice, will be strongly dependent on the amount of synchronization of the system.

3. Simulated data

3.1. Generation of synthetic datasets

In order to test the efficiency of our functional parcellation with fMRI data acquired in different experimental conditions, we performed two sets of simulations. The first one simulated ideal resting state (rs) fMRI data (i.e. fMRI data in absence of stimuli), while the second set was used to test the behavior of our method with block and event related fMRI data. To simulate BOLD signals, we first generated time series mimicking the underlying neuronal activity and then we convolved them with the canonical Hemodynamic Response Function (HRF) with variable parameters. We simulated the fMRI data corresponding to an anatomical brain region *Q* with a cubic structure of 64 elements, within which we created 10 spatially connected clusters (Fig. 2).

In both the simulations, each cluster within the anatomical region *Q* was created starting from the simulation of a hub element, to which the other elements of the cluster were connected. In particular, their time series were estimated by adding noise to the hub element time series. In the simulation of ideal rs fMRI data, the time course of the hub element in each cluster was created by convolving a second order AutoRegressive (AR) time series with the canonical HRF (Friston, 2003); we then simulated the signals of the other elements by adding Gaussian noise to the hub of the cluster. The difference between the different clusters was only relative to the AR series parameters. To assess the behavior of the algorithm in different conditions, we changed the Signal to Noise Ratio (SNR) from 10^{-6} to 1 dB, the time series length L from 100 to 1000 samples and the time of repetition (TR) from 2 to 4 s (Table 1).

The second set of simulations was used to check the behavior of our method with block and event related fMRI data. We designed the stimuli with boxcar functions, consisting of alternation between stimuli and resting blocks, both lasting 20 s. To inspect the event related fMRI response, we designed a set of instantaneous events using a binary sequence of impulses. Within the system *Q*, four clusters were simulated as responsive to the stimulus blocks, while the other six clusters were unresponsive to them. Among the latters, two clusters were in resting state, the other four were instead responsive to the binary sequence of events. For all the clusters within the system *Q*, the signals of the elements other than the central one were simulated by adding Gaussian noise to the central time series.

The resting state BOLD data of the unresponsive clusters were simulated as in the first simulations set (i.e. the time series simulating neural activity were convolved with canonical HRFs). In the responsive clusters, the fMRI signals of the hub elements were estimated by convolving the corresponding stimulus function with an HRF with variable parameters, depending on the cluster. The four clusters responsive to block shaped stimuli sequence and the four clusters responsive to the event related stimulation were distinguished one from each other only by differences in the HRFs properties; the four HRFs used at this purpose are plotted in Fig. 3. In summary, in this set of simulations the differences between clusters were encoded by three features: the responsiveness/unresponsiveness to the stimuli sequence (three types of clusters were available: clusters in resting state, clusters responsive to block design stimulation or to event related stimulation), the AR parameters (only for the two unresponsive clusters) and the HRF shape (only for the four block design and the four event related responsive clusters).

We tested the algorithm capability to discriminate between the different clusters by varying the SNR, number of samples and TR values, as already performed in the first set of simulations.

Table 1
Ideal resting state fMRI data simulation results (*N*: number of voxels in the system, *L*: time series length, SNR: signal to noise ratio (dB), TR: repetition time).

N	*L*	SNR	TR	Correct clusters %
64	From 1000 to 100	From 1 to 10^{-6}	From 2 to 4	100%
125	From 1000 to 100	From 1 to 10^{-6}	From 2 to 4	100%

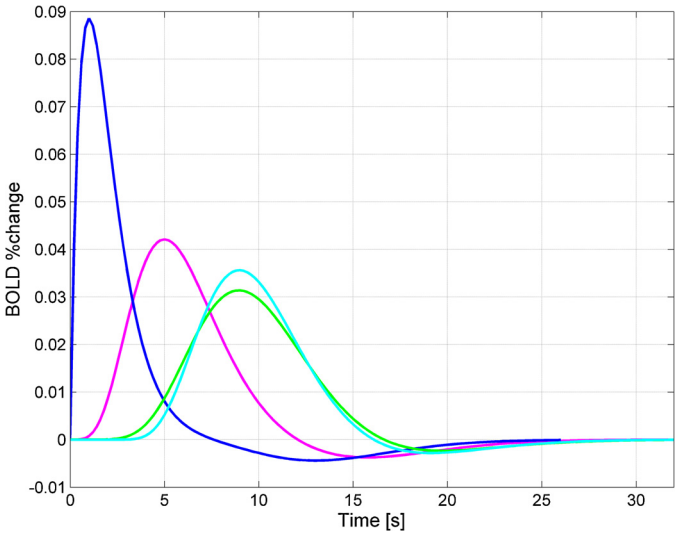

Fig. 3. Hemodynamic Response Functions. The four HRFs used to construct the BOLD signal of the eight (four block-design and four event-related) stimulus-responsive clusters.

To test the algorithm behavior with increasing size and different clusters, an additional run of simulations was performed on a 125 voxels system. In the resting state simulation, the system was composed of 7 clusters, while in the simulation of fMRI data during block design and event related stimulation the clusters were 8, half of them sensitive to the block design and the others to the event related stimulation. The four clusters responsive to block shaped stimuli sequence and the four cluster responsive to the event related stimulation were distinguished one from each other only by differences in the HRFs properties (Fig. 3). The system was tested on the same range of parameters (TRs, SNRs and length) used for testing the system of 64 elements.

3.2. Simulation results

The results of the ideal rs fMRI simulation are summarized in Table 1. The percent of correctly detected clusters was used as metric to measure the algorithm efficiency. In both the systems, having 64 and 125 voxels, with time series having 100 samples or more, the functional parcellation was effective within the entire considered ranges of SNRs and TRs. It is worth to note that with the system with 125 voxels and time series of 100 samples, the algorithm performed well also when the preliminary procedure was applied. The second set of simulations led to the results summarized in Table 2. Our algorithm correctly discriminated clusters with different HRFs properties within a wide range of SNRs values and time series lengths. Regarding the system with 64 voxels, only in one case the algorithm efficiency was minor than 100%: with 100 samples and SNR = 1, the algorithm grouped two clusters responding to block stimulation, having HRF peaks respectively at 6 and 9 s (Fig. 3), with an 80% of identified clusters. In the 125 voxels system, with time series of 300 samples or more, the algorithm was able to detect correctly all the clusters, for the entire range of TRs and SNRs values. With time series having less samples, the percentage of correctly detected clusters dropped to 75%, since the algorithm tended to merge the clusters responsive to the stimulation blocks. Therefore, the simulations results relative to system of 125 voxels are not dissimilar from the simulation results previously obtained with the system of 64 voxels.

The findings of the two sets of simulations proved that the algorithm results are reliable within the usual range of parameters characterizing real fMRI experiments, both in resting state and in presence of stimulations. Across the different simulation runs, the algorithm was able to perform a correct parcellation of the system, independently of the initial number of clusters and of the clusters extension. Moreover, the proposed method appeared to be sensitive to interregional hemodynamic variability.

Table 2
Block-design and event-related fMRI data simulations results (*N*: number of voxels in the system, *L*: time series length, SNR: signal to noise ratio (dB), TR: repetition time).

N	*L*	SNR	TR	Correct clusters %
64	From 1000 to 200	From 1 to 10^{-6}	From 2 to 4	100%
64	100	1	2	80%
125	From 1000 to 300	From 1 to 10^{-6}	From 2 to 4	100%
125	200	1	2	75%

4. Real fMRI data

4.1. Test on healthy subjects – experimental setup

The parcellation method was also applied on real fMRI data recorded from a group of healthy subjects during a visual stimulation protocol, in order to test the inter subject repeatability of the proposed algorithm. The acquisitions took place at the Scientific Institute IRCCS E. Medea of Bosisio Parini (LC, Italy) and were part of the Spider@Lecco project. We used fMRI data relative to eight subjects (3 males, mean age: 27.6 ± 2.67 years) with normal vision and no history of neurological or psychiatric disorders. Each subject signed a written informed consent to the study, whose protocol was approved by the local Ethic Committee.

MR scans were performed with a 3T scanner (Philips Achieva, Best, The Netherlands), equipped with a 32 channels head coil. fMRI data were acquired with a T2*-weighted Gradient-Echo planar sequence (TR = 2 s, TE = 35 ms, flip angle = 85°, 30 axial slices without gap, FOV = 240 mm × 105 mm × 240 mm, voxel size = 1.875 mm × 1.875 mm × 3.5 mm) covering cerebral hemispheres, excluding cerebellum and brainstem. A structural MR image was acquired with a T1-weighted 3D Turbo Field Echo sequence (1 mm isotropic resolution, FOV = 240 mm × 240 mm × 175 mm, TR = 8.19 ms, TE = 3.74 ms, flip angle = 8°) to provide a morphological reference for fMRI data. Visual stimuli were delivered in the MR scanner through MRI-compatible goggles (Resonance Technology Inc.). The experimental procedure consisted in a block-designed Intermittent Photic Stimulation (IPS), created by reversing black and white screens at four different frequencies (6, 8, 10 and 12 Hz). fMRI images were collected in one single run of interleaved blocks of active and rest conditions. Rest epochs lasting 14 s were alternated with active epochs of the same duration and five blocks for each stimulation frequency were included. The subjects were asked to keep their eyes open during all the experiment.

For each subject, we first performed an atlas based parcellation: the brain volume of each subject was subdivided into anatomical regions based on Harvard Oxford atlas using a four steps label propagation procedure (Tana et al., 2012). For sake of simplicity, we decided to focus our analysis on a singular region within the primary visual cortex, which is known to be involved in the response to IPS (Maggioni et al., 2013). We used as example the left cuneus and extracted the relative fMRI time series. Before extraction, the fMRI time series were subjected to realignment, normalization to the MNI template, smoothing with a Gaussian kernel filter (FHWM = 6 mm^3) and removal of the sources of spurious variances (Wang et al., 2009). The results of the parcellation were compared between subjects in terms of number, volume and position (i.e. centroids coordinates) of the detected clusters.

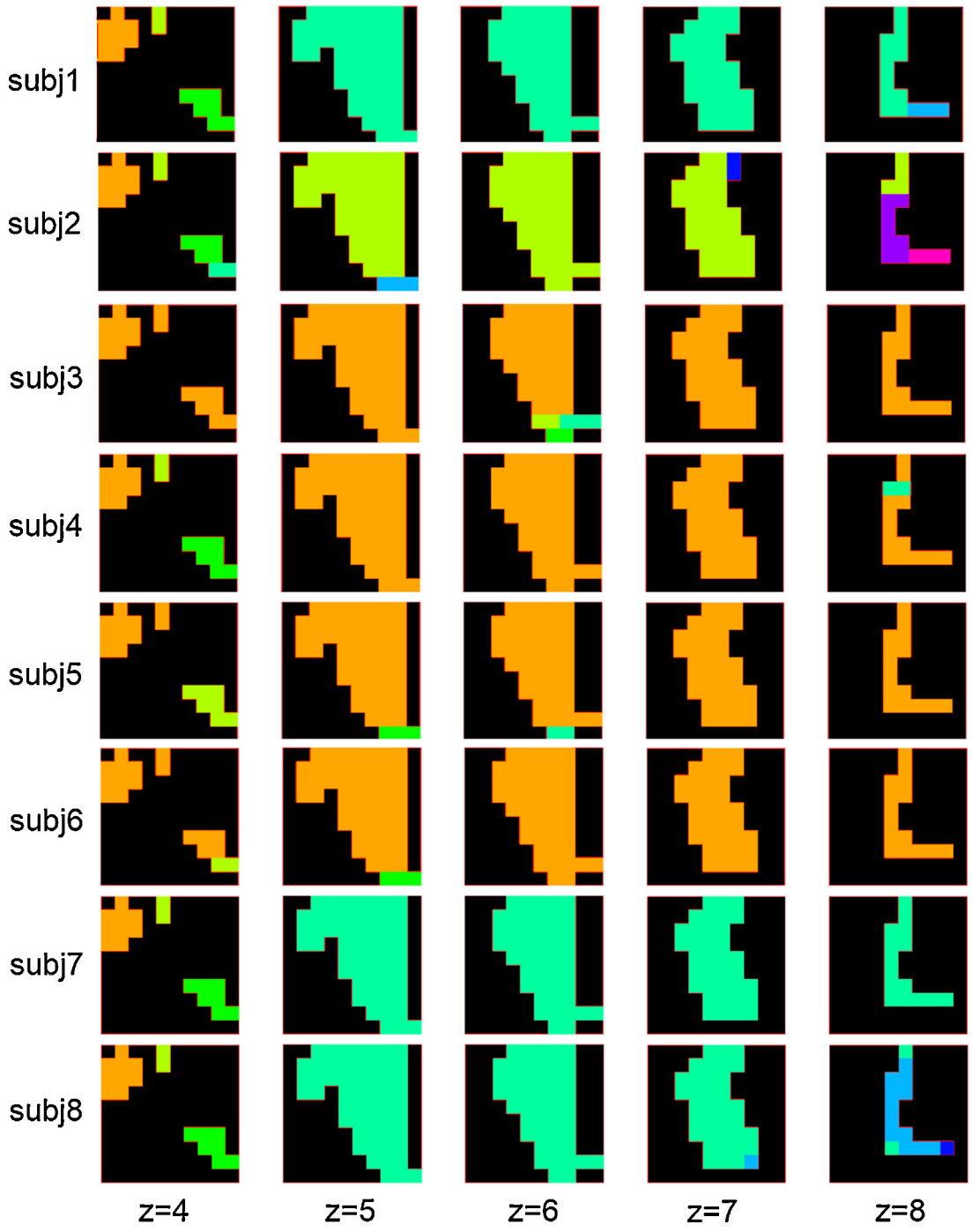

Fig. 4. Parcellation results on healthy subjects. The clusters of each subject are represented with different colors across the five slices of cuneus. The background is colored in black.

4.2. Test on healthy subjects – results

The clusters resulting from the parcellation of the left cuneus in each of the eight subjects are shown in Fig. 4. The clusters distribution was very similar across subjects: the number of clusters within the group varied from 4 to 9 (with a mean value of 5.25 ± 1.75) and the volume of the main (i.e. the larger) cluster ranged from 137 to 164 voxels (with a mean value of 152.75 ± 11.08). In all subjects, such cluster spread from the second to the fourth slice and in the majority of them (6 subjects out of 8) occupied also the fifth slice. In Fig. 5, the coordinates of the main cluster centroids are plotted for each subject. The proximity of the eight centroids, together with the repeatability of the clusters distribution across subjects, confirmed the reliability of the new parcellation algorithm.

4.3. Patients – experimental setup

Our method was then applied to patients' fMRI data, with the aim of testing its capacity as technique for the definition of the nodes of a cerebral network, on which effective connectivity can subsequently be evaluated. The performance of the proposed method was evaluated by assessing its capability to recognize the source and the sink of causal networks (Tana et al., 2012). For this purpose, we used fMRI data recorded from two epileptic patients

Table 3
Clinical characteristics of patients and EEG-fMRI findings (Tana et al., 2012). The fMRI active regions were found using the GLM analysis implemented in Statistical Parametric Mapping (SPM5) software toolbox (http://www.fil.ion.ucl.ac.uk/spm/).

Patient	Anatomical MRI	Epilepsy type	Seizure semiology	Seizure frequency and duration	Ictal fMRI findings	Ictal EEG findings
A (2 years)	Agenesis of Corpus Callosum	Epileptic spasms	Rhythmic spasms, with extension of both upper limbs and upward tonic eye deviation	>20 daily 1–3 min	*Left:* anterior ParaHyppocampal Gyrus (LaPHG), Cerebellum (LCb), anterior and posterior Temporal Fusiform Cortex (LaTFC, LpTFC), Thalamus (LTh). *Right:* Frontal Pole (RFP), posterior Temporal Fusiform Cortex (RpTFC).	Background desynchronized rhythm with bursts of fast activity over right fronto-temporal electrodes spreading to left fronto-temporal ones
B (7 years)	No anatomical MR findings	Parietal Lobe Epilepsy	Seizures with "odd sensation in right arm" followed by choreic/ballistic movements. Progressive right hemiparesis.	10–20 daily 1–2 min	*Left:* PostCentral Gyrus (LPostCG), Lingual Gyrus (LLG). *Right:* superior Parietal Lobule (RsPL)	Rhythmic spike activity in Cz

who underwent simultaneous EEG-fMRI acquisition, whose clinical characteristics and EEG-fMRI findings are summarized in Table 3. The novel whole-brain parcellation scheme was used to define nodes within a set of regions activated during epileptic seizures. Granger Causality Analysis (GCA) was then applied to identify the starting (i.e. the network source) and ending point (i.e. the network sink) of the seizure propagation.

In order to provide a comparison between the proposed approach and other standard parcellation schemes, we chose to analyze already published data and to compare the results of the proposed method with the previous ones. In particular, we used data from the work of (Tana et al., 2012), where a standard anatomical atlas based scheme and a functional hierarchical clustering algorithm were already applied and compared. We selected two datasets from (Tana et al., 2012), showing exactly the same performances for both the atlas based and the hierarchical clustering parcellation, and compared the outcome of the novel algorithm with that of the previous ones.

EEG was recorded with a 37 channels system (Maglink, Neuroscan, Charlotte, NC, USA) inside a 1.5 T MR scanner (GE Cvi/NVi, Milwaukee, WI, USA). fMRI data were acquired with an echo planar imaging (EPI) sequence using axial orientation with the parameters listed in Table 4. A volumetric T1 Spoiled Gradient Echo Recovery (SPGR) 3D sequence was also acquired (slice thickness = 0.6 mm, in plane resolution = 0.94 mm × 0.94 mm). The acquisitions were performed at the Hospital Center of West Lisbon.

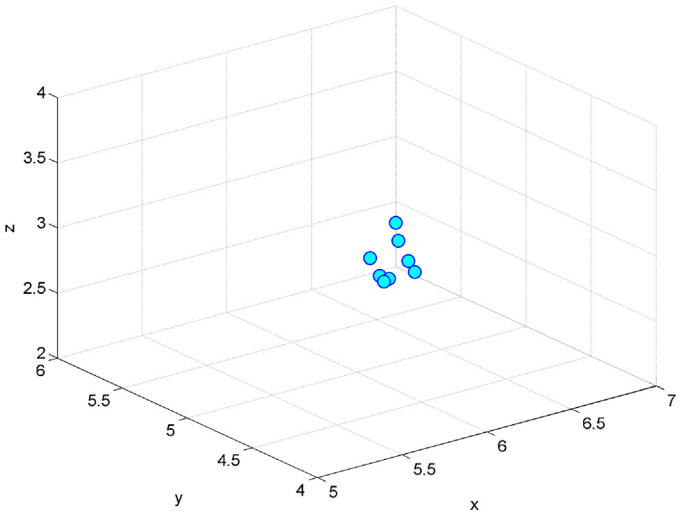

Fig. 5. Main clusters centroids across healthy subjects. Coordinates of centroids relative to the main cluster of each subject.

After brain extraction from structural T1 images, the brain volume of each patient was subdivided into anatomical regions using a four steps label propagation procedure (Tana et al., 2012). Sources of spurious variances, arising from estimated head motion profiles, were removed by linear regression analysis and the obtained residuals were used to replace the original fMRI time series (Wang et al., 2009). Our functional parcellation was then applied to each anatomical region, obtaining a whole brain parcellation scheme.

The onsets of epileptic activity were identified on the EEG signal by an expert neurophysiologist, and the brain regions activated during seizures were found using GLM analysis implemented in Statistical Parametric Mapping (SPM5), as described in Tana et al. (2012). We decided to perform a GCA analysis on the network of brain regions involved in ictal activity, with the aim of identifying the seizure propagation pattern. To properly define the connectivity network, the GLM active brain regions (Table 3) were subdivided into parcels on the basis of the parcellation scheme just described. Each parcel was associated to a node of the network, whose time series was calculated as the spatial average of the time courses of the active voxels within that parcel (Deshpande et al., 2008). The covariance stationarity of the time series was verified by using the Augmented Dickey Fuller (ADF) test (Said and Dickey, 1984), implemented in the GCCA Matlab toolbox (www.anilseth.com). After the definition of the network and before the GCA, in order to account for the possible confounding effects due to the HRF variations across brain areas, we estimated the HRF of each node using a region based non parametric HRF estimation (Makni et al., 2008) and deconvolved BOLD time series with the corresponding HRF (Glover, 1999). For the HRF estimation and deconvolution, an interesting alternative could be the blind deconvolution approach proposed by (Wu et al., 2013). The most important advantage of this approach is the ability to deal with resting state data, since it

Table 4
Parameters of fMRI acquisition.

Patient	Time of Repetition (TR)	Voxel size (mm × mm × mm)	Number of volumes per run
A	2.1	4.38 × 4.38 × 5	170
B	2.1	4.38 × 4.38 × 5	170

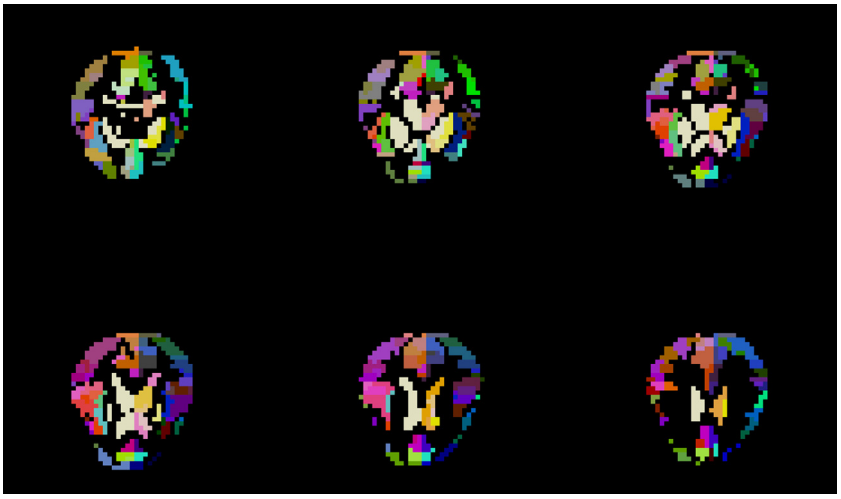

Fig. 6. Patient A's parcels. Six central slices of the patient A's brain fMRI volume divided according to the anatomical/functional parcellation. Each cluster is represented with a different color.

considers them as "spontaneous event related" signals, from which the regional HRF is reconstructed by fitting them with a double gamma function. In our particular dataset of epileptic patients, where the hemodynamic properties can importantly differ from the canonical ones, we adopted the approach from (Makni et al., 2008) to avoid any constraint on the HRF shape, including also those introduced by the use of double gamma function. However, the use of the method by (Wu et al., 2013) could be a valid alternative when resting state data from healthy subjects are analyzed and could be an interesting extension of this work.

The resulting time series of the network nodes were used as channels in an MVAR model. Basing on the MVAR model estimated parameters, the Partial Directed Coherence (PDC) among couples of nodes was calculated (Baccalà and Sameshima, 2001). Since no specific frequency band was observed, we considered a unique connectivity measure by summing the contributions over all the frequencies (Tana et al., 2012). The significance of the connectivity results was assessed by comparing real time series with surrogate data using a phase randomization approach (Korzeniewska et al., 2003). The connectivity values that resulted to be not significant were set to zero and the final EC network was represented in the form of a directed graph. The nodes were then ranked by netflow, which measures the difference of strength between outgoing and incoming connections. Therefore a node i is ranked higher than another node j if and only if the difference between outflow and inflow of i overlays the one of j. The measure of netflow of a node i is calculated as (van den Brink and Gilles, 2009).

$$\sigma_i^{net}(\omega) = \sigma_i^{out}(\omega) - \sigma_i^{in}(\omega) = \sum_{j \in N} \omega(i,j) - \sum_{j \in N} \omega(j,i) \qquad (6)$$

The ranking allows to determine whether a node acts as a driver or it is mainly driven by other nodes, and to make hypothesis about the source and the sink of the network.

4.4. Patients – results

Hereinafter, we will show and describe the outcome of the whole brain parcellation algorithm applied to the two patients' fMRI data and the connectivity results of GCA applied to the nodes obtained from parcellation.

The GCA results of the two patients will be described in separate paragraphs because of the different characteristics of the seizures semiology. Results will be evaluated on the basis of the expected seizure propagation pattern. The GC network will be illustrated in the form of a directed graph, in which the nodes colors code the netflow values and the arrows colors and widths code the interaction strength.

4.4.1. Patient A

Fig. 6 shows the anatomical/functional parcels over the six central slices of patient A's brain.

Using the whole brain parcellation scheme showed in Fig. 7, the active GLM areas were subdivided into 9 parcels belonging to 7 different anatomical regions. In particular, it was possible to distinguish two different functional parcels in the right frontal pole (RFP1, RFP2) and in the right temporal cortex (RpTFC1, RpTFC2).

In Fig. 7 the results of the GCA are represented in the form of a weighted directed graph. One of the two parcels in the right posterior temporal fusiform cortex (RpTFC1) can be reasonably considered as the source of the network, since it is characterized by the highest netflow (i.e. the difference between outgoing and incoming connections). The possibility that the source of

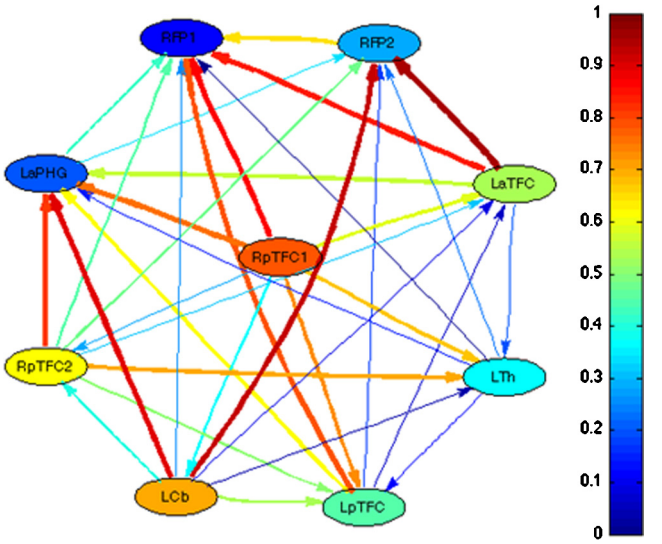

Fig. 7. Connectivity weighted directed graph of Patient A. The colors of nodes code the netflow values, the colors and widths of the arrows the connection intensity. The nodes are labeled with the convention in Table 3.

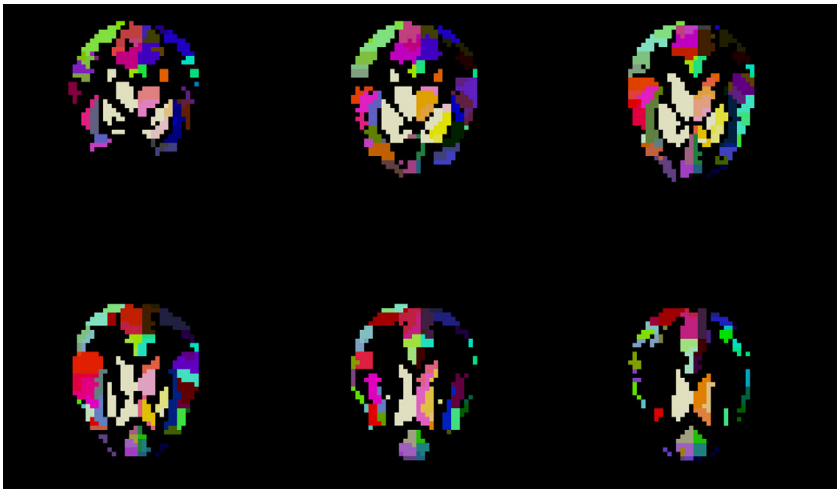

Fig. 8. Patient B's parcels. Six central slices of the patient B's brain fMRI volume divided according to the anatomical/functional parcellation Each cluster is represented with a different color.

epileptic spasms can be located in the right temporal cortex is confirmed by several previous works documenting epileptic spams originated from temporal areas (Akiyama et al., 2005; Mizuno et al., 2011; RamachandranNair et al., 2005).

The pattern of seizure propagation can be also reconstructed from the connectivity graph. A predominant information flow from the right hemisphere to the contralateral one is clearly identifiable (significant connections go from the right temporal region to the left temporal one (LaTFC, LpTFC), the left parahyppocampal gyrus (LaPHG) and the thalamus (LTh)). These results agree with the EEG recordings listed in Table 3, which show a first activity in right temporofrontal electrodes followed by a delayed ictal activity in the contralateral cortex. Finally, the sink of the network is identifiable in the right frontal pole (in particular in RFP1), which appears as driven by the rest of the network, without sending information to any other node.

The present results can be compared with those obtained by (Tana et al., 2012) (patient A corresponds to patient A in that paper) using two other parcellation techniques: anatomical atlas based and functional hierarchical clustering. At this regard, we can observe that the three parcellation schemes produced different number of nodes: 8 for anatomical standard (Tana et al., 2012), 4 for hierarchical clustering (Tana et al., 2012) and 9 for our Tononi's Index based method (Fig. 7). In all the three cases, the source and the sink could be clearly localized in the RpTFC and in the RFP respectively and a predominant information flow from the right hemisphere to the contralateral one was clearly identifiable. The novel algorithm did not produce contradictory results if compared with the others, moreover it allowed to localize the source and the sink with higher precision. Indeed, both RpTFC and RFP were further divided into two functional clusters by our scheme and within each couple of clusters, (RpTFC1-RpTFC2) and (RFP1-RFP2), only one cluster was identified as source (RpTFC1) and sink (RFP1) by the GCA.

4.4.2. Patient B

The parcels resulting from the parcellation over the six central slices of patient B's brain are shown in Fig. 8.

The active GLM areas were subdivided using the whole brain parcellation scheme into 4 parcels belonging to 3 different anatomical regions. The algorithm identified separate functional contributions in the postcentral gyrus of the left hemisphere, which was divided into two nodes (LPostCG1, LPostCG2).

The pattern of propagation of the epileptic seizure is clearly identifiable from the weighted directed graph shown in Fig. 9. The source is found in one of the two functional parcels localized in the left postcentral gyrus (LPostCG2) since it is characterized by the highest netflow value. This finding is in agreement with the clinical diagnosis of Parietal Lobe Epilepsy (PLE). The propagation involves the superior parietal lobule of the right hemisphere (RsPL). The sink is located in the left lingual gyrus (LLG), which is characterized by the lowest netflow value. The overall pattern is supported and confirmed by literature evidences, which show the posterior propagation of seizure activity from the parietal lobe to the occipital regions (Tana et al., 2012; Williamson and Engel, 2008).

Also the results of the second dataset can be compared with those obtained by (Tana et al., 2012) (patient B corresponds to patient D in that paper). At this regard, we can observe that the two functional parcellation schemes divided the GLM clusters into the same number of nodes (i.e. 4 nodes) (Fig. 9), (Tana et al., 2012), while only 3 nodes were the outcome of the standard anatomical scheme (Tana et al., 2012). All the three methods allowed to localize the source in the LPostCG and the sink in the LLG without providing contradictory results. In this case, both functional algorithms (hierarchical clustering and Tononi's Index based)

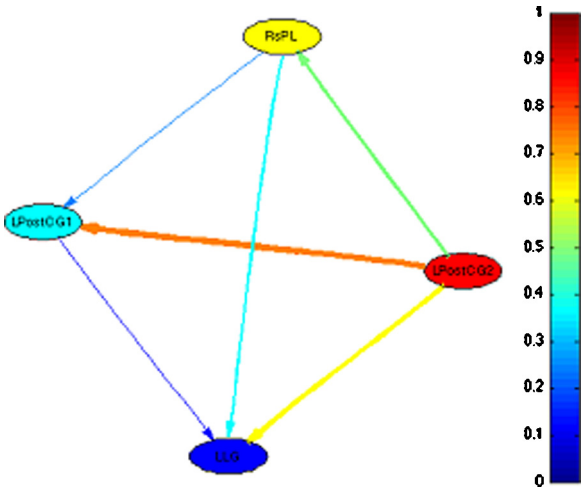

Fig. 9. Connectivity weighted directed graph of Patient B. The colors of nodes code the netflow values, the colors and widths of the arrows the connection intensity. The nodes are labeled with the convention in Table 3.

identified separate functional contributions in the postcentral gyrus of the left hemisphere, which was divided into two nodes (LPostCG1, LPostCG2) having very different netflow ranks. Differently from standard anatomical parcellation, both the functional algorithms allowed to localize the source with a higher precision.

5. Discussion

One of the main challenges in fMRI connectivity studies regards the definition of network topology and, in particular, the construction of network nodes. Indeed, meaningful connectivity results can be obtained only by avoiding that heterogeneously connected brain regions are lumped in a single node. fMRI data should be segmented into clusters representing brain regions having high internal homogeneity and showing a coherent pattern of external connections (Rubinov and Sporns, 2010). In this paper, we propose an algorithm for identifying fMRI clusters in which the structural and functional homogeneity are both preserved.

Our method starts with an atlas based parcellation step that confines each node to the boundaries of one and only one HO region of interest. Since different anatomical regions are highly heterogeneous in structure (Fischl et al., 2002), we avoided aggregating in a unique parcel voxels belonging to different anatomical regions. This constraint precludes the formation of nonsensical parcels spanning well distinct anatomical areas (even belonging to different hemispheres), as it might happen with a pure functional algorithm (Hagmann et al., 2007; Zalesky et al., 2010).

Each anatomical area is then segmented into functional parcels. The algorithm uses the *TCI* as metric for identifying the functional connections within fMRI data. This functional clustering measure allows to identify groups of voxels which are much more strongly interactive among themselves than with the rest of the system (Tononi et al., 1998). The interactivity or statistical dependence is measured in terms of instantaneous functional correlations.

The proposed functional parcellation algorithm is entirely data driven, because it works directly on fMRI time series, without projecting them on a feature space. Several other clustering algorithms are based on a pattern representation of the characteristics of interest, often involving the selection and/or extraction of features. The functional properties of each voxel are often characterized using the beta coefficients or the *T* or *F* maps estimated during the GLM analysis (Michel et al., 2012; Tana et al., 2012; Thirion et al., 2006; Makni et al., 2008). The similarity between elements is quantified by calculating a distance in the feature space, on which fuzzy c means (Yoon et al., 2003) and k means clustering (Mezer et al., 2009) are then applied. Such a type of distance measure can be influenced by the a priori assumptions required by GLM. The feature (beta, *T* or *F*) maps are obtained by regressing a modeled BOLD signal out of the acquired time series. The modeled fMRI signal is constructed on a priori fixed HRF shape and, due to the well known interregional variability of HRF (David et al., 2008; Seth, 2010), can introduce a bias in the regression coefficient estimation and therefore in the construction of the feature maps. On the contrary, our algorithm prevents from making assumptions on fMRI data by calculating the inter voxels distance on the basis of *TCI*, which is an index directly related to the functional content of fMRI time series.

It is worth mentioning that our method is not the first one using distance metrics directly measured on signals (Achard et al., 2006; Kim et al., 2010; Gonzalez-Castillo et al., 2012), but, to the best of our knowledge, the majority of the other existing parcellation techniques are not able to guarantee the spatial connectedness and the structural homogeneity of the resulting parcels. These methods would require post processing steps to guarantee spatial connectedness; for instance, the latter can be ensured by increasing the number of clusters originally estimated until their results are spatially connected (Smith et al., 2009) or by splitting parcels that are not spatially contiguous. On the other hand, Thirion et al. (2006) proposed a functional parcellation algorithm directly incorporating a spatial constraint, but used a subset of features to measure distance instead of working directly on time series. Our approach, instead, is able both to provide a direct measure of the functional information and to ensure the spatial connectivity. This is accomplished by applying a spatial constraint during the aggregative procedure producing the functional clusters. It has to be mentioned that, recently, a spatially constrained spectral clustering has been proposed (Craddock et al., 2012), which subdivides the entire brain into spatially coherent and functionally homogeneous regions. This method is able as ours to account for spatial contiguity and does not employ any projection of functional content on a feature space. Nevertheless, as the other currently available clustering techniques, it requires the a priori specifications of the desired number of parcels in the system.

Conversely, in our method the number of clusters is not imposed, since it depends only on the amount of synchronization between the elements of the system. However, the user could influence the number of the resulting clusters by choosing when stopping the aggregative procedure. This procedure is usually performed on a number of duplets sufficient to parcel the entire system. If the procedure is stopped before the entire system is parceled, only the duplets with higher *TCI* are used as seeds for the aggregation, whereas if a huger number of duplets is chosen, the aggregation is performed also from duplets with a lower *TCI*. This might be a limitation of the method, because clusters constructed from duplets with a lower functional similarity degree could be less meaningful. It is apparent that the final number of parcels is not deterministically predictable from the number of duplets on which the aggregation step is performed, since the former also directly depends on the synchronization amount of the single data set. Further studies are needed to develop methods that determine if and when the aggregative procedure has to be stopped.

The performances of our functional clustering method were checked on both simulated and real fMRI data. The method was tested on synthetic data to prove its capability in identifying the functional parcels of two systems with 64 and 125 voxels, simulating either task based (block and event related) or rs fMRI data. Although the simulations could be further extended, from the tests performed the algorithm results are satisfactory. We verified that the new functional clustering is usually sensitive to differences in the HRFs. The performances of the algorithm get worse when the fMRI signal length decreases, while seem to be invariant with respect to TR and SNR changes. The sensitivity to the signals length is due to the fact that all the statistical parameters are measured from the system covariance matrix. The use of the covariance matrix also restricts the algorithm applications to small or medium systems; in case of a huge number of elements, the matrix determinant is infinitesimal, and this compromises the entire analysis. To overcome this challenge, we developed a pre processing step that performs a system reduction through a preliminary aggregation. This procedure should not affect the correctness of parcellation results, as proved by the 125 voxels simulation; anyway, this is a current limit of the method, which could be overcome in future applications by using different measures of entropy. The use of the singular value decomposition (SVD) for entropy estimation is an interesting option that could overcome the problems of covariance matrix singularity characterizing large systems. Nevertheless, both the here proposed approach and the SVD assume that the system is Gaussian distributed, thereby accounting for statistic relations up to the second order. For this reason, a further interesting extension of the proposed algorithm should better imply entropy measures sensitive also to nonlinear dependencies. It has to be remarked that, although the simulations highlighted some important features of

the algorithm, they are still far from providing a full characterization of the algorithm performance. The purpose of our simulations was simply to highlight the points of strength and of weakness of the parcellation algorithm, without claiming of completeness. The parcellation method was further tested on an actual fMRI dataset relative to healthy volunteers during an IPS paradigm. We focused the comparison on the clusters identified in one anatomical region in primary visual cortex, the left cuneus: despite the expected inter subject variability in the number of clusters, the overall similarity of the clusters distribution across subjects confirmed the reliability of the new functional parcellation scheme.

The method was then applied to real fMRI data recorded from epileptic patients; the resulting whole brain parcellation schemes were shown to be a valid basis for further effective connectivity analysis. In this application on real data, the interest was the study of the interactions among regions activated during the ictal activity; the network for the EC analysis was therefore defined taking from the whole brain parcellation scheme the nodes belonging to the GLM active clusters. The GCA analysis performed on the resulting network led to connectivity patterns that were in line with the clinical evidences about the patients' pathologies. The sources and patterns of seizure propagation estimated with the new network agreed with EEG findings and literature knowledge on such pathologies.

The GCA results on the new network can be discussed with respect to the ones obtained on the same patients in (Tana et al., 2012), using two alternative parcellation schemes. In this paper, the authors compared a pure anatomical parcellation based on the HO atlas with a hierarchical clustering functional parcellation based on the Ward's algorithm (Ward, 1963), where the GLM coefficients were used as distance measure. In both the patients, our GCA results on the new network were in agreement with the results found in (Tana et al., 2012) in terms of source, sink and main flow of seizure propagation. Moreover, the novel parcellation seems to be able to provide a localization of the network source and sink more detailed than a standard anatomical scheme.

Although a more extensive comparative study, comprising a larger group of different parcellation schemes and a larger group of subjects, would be desirable and could be object of further investigation, the accordance between the present and previously published results proves the effectiveness of our parcellation as preparatory step for causal connectivity analysis. It is worth mentioning that the definition of nodes with coherent patterns of anatomical and functional connections is particularly important in directed connectivity analysis. The two just mentioned studies show the novelty to investigate and discuss the impact of different parcellation schemes on the resulting causality pattern.

However, the proposed algorithm potentialities could be fully exploited in complex network analysis investigating the whole brain connectivity.

6. Conclusions

We developed a method for the topological definition of the constitutive nodes of fMRI brain networks. The method defines an intra subject whole brain parcellation scheme based both on anatomical and functional properties. Each parcel groups spatially connected voxels showing instantaneous correlation. Networks can be defined according to the parcellation, associating each parcel of interest to a node, whose homogeneity is therefore guaranteed.

The novel algorithm can be used as a step of node definition preceding further connectivity analysis. The method was first tested on simulated data and on healthy subjects' real data. Both the tests confirmed its capability to identify meaningful parcels. The new method was then applied on real fMRI data recorded from epileptic patients: the algorithm performed well in identifying nodes for subsequent EC analysis, and its accurate partition of the brain domain led up to a detailed definition of the connectivity pattern.

7. Role of the funding source

Grant support was partially provided from Fondazione Cariplo (project #2010-0344, "Spider@Lecco").

Acknowledgments

The authors wish to thank Dr. Alberto Leal from the Centro Hospitalar Psiquiátrico de Lisboa (Lisbon, Portugal) for providing the experimental data set.

Appendix A.

It is known that entropy and mutual information can be thought of as the multivariate generalizations of variance and covariance in univariate statistics (Tononi et al., 1998).

Therefore, the joint Probability Density Function (PDF) of a system S composed of N time series $\{s_i\}$ can be described in terms of entropy and mutual information (Foucher et al., 2005; Papoulis and Pillai, 2002), under the assumption that the system time series conform to a multidimensional stationary stochastic process.

If $H(S_j^k)$ is the entropy of S_j^k, the jth subset with k < N elements, the mutual information MI between S_j^k and its complementary subsystem $(S - S_j^k)$ will be:

$$MI(S_j^k; S - S_j^k) = H(S_j^k) + H(S - S_j^k) - H(S)$$

The mutual information is zero if S_j and its complementary subset $(S - S_j^k)$ are independent variables, otherwise it is greater than zero.

In order to generalize the concept of mutual information, another measure, called integration I, can be introduced. The integration of a system S is defined as the deviation from independence among the N components of S and can be expressed as:

$$I(S) = \sum_{i=1}^{N} H(S_i) - H(S)$$

The integration $I(S)$ represents a multivariate measure of the total amount of deviation from statistical independence within the system. Indeed, $I(S)$ is equal to the sum of values of the mutual information between parts resulting from the recursive bipartition of S down to its elementary components, that is $I(S) = \sum_{i=1}^{N-1} MI\,(s_i; [s_{i+1}, s_{i+2}, \ldots, s_N])$.

The Tononi's cluster index (TCI) of a subset in a system measures the statistical dependence of the elements within the subset in relation to the statistical dependence between this subset and the rest of the system (Tononi et al., 1998). Hence, the TCI relative to S_j^k is found by dividing the subset integration by the mutual information between the subset and the rest of the system:

$$TCI(S_j^k) = \frac{I(S_j^k)}{MI(S_j^k; S - S_j^k)}$$

The TCI is high for subsets showing both high internal statistical dependence and high independency from the rest of the system.

If *PDF* of *S* conforms to a multidimensional Gaussian, the system *S* can be completely characterized by its covariance matrix (Jones, 1979), which is the *N* x *N* symmetric matrix of covariance between all pairs of elements, *COV*(*S*).

If we indicate with σ_i^2 the variance of the *ith* element s_i, i.e. the *ith* diagonal element of *COV*(*S*), the entropy of s_i, is given by

$$H(s_i) = (1/2)\ln(2\pi e\sigma_i^2)$$

and the entropy of *S* is

$$H(S) = (1/2)\ln(2\pi e^N |COV(S)|)$$

Analogously, if S_j^k is the *jth k*-dimensional subset of the system *S*, its entropy is given by

$$H(S_j^k) = (1/2)\ln(2\pi e^k \left|COV(S_j^k)\right|).$$

The mutual information between S_j^k and its complementary subsystem $(S - S_j^k)$ is given by:

$$MI(S_j^k; S - S_j^k) = (1/2)\ln\left(2\pi \frac{\left|COV\left(S_j^k\right)\right| \left|COV\left(S - S_j^k\right)\right|}{\left|COV(S)\right|}\right)$$

The integration of S_j^k will be:

$$I(S_j^k) = -(1/2)\ln\left(\left|CORR(S_j^k)\right|\right)$$

where $CORR(S_j^k)$ is the cross-correlation (normalized covariance) matrix of S_j^k (Tononi et al., 1998).

Finally, the *TCI* will be:

$$TCI(S_j^k) = -\ln\left(\left|CORR(S_j^k)\right|\right) / \ln\left(2\pi \frac{\left|COV(S_j^k)\right| \left|COV(S - S_j^k)\right|}{\left|COV(S)\right|}\right)$$

References

Achard S, Salvador R, Whitcher B, Suckling J, Bullmore E. A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs. J Neurosci 2006;26:63–72.

Akiyama T, Otsubo H, Ochi A, Ishiguro T, Kadokura G, RamachandranNair R, et al. Focal cortical high-frequency oscillations trigger epileptic spasms: Confirmation by digital video subdural EEG. Clin Neurophysiol 2005;116:2819–25.

Baccalà LA, Sameshima K. Partial directed coherence: a new concept in neural structure determination. Biol Cybern 2001;84:463–74.

Chen H, Yang Q, Liao W, Gong Q, Shen S. Evaluation of the effective connectivity of supplementary motor areas during motor imagery using Granger causality mapping. Neuroimage 2009;47:1844–53.

Craddock RC, James GA, Holtzheimer PE, Hu XPP, Mayberg HS. A whole brain fMRI atlas generated via spatially constrained spectral clustering. Hum Brain Mapp 2012;33:1914–28.

David O, Guillemain I, Saillet S, Reyt S, Deransart C, Segebarth C, et al. Identifying Neural Drivers with Functional MRI: An Electrophysiological Validation. PLoS Biol 2008;6:2683–97.

Deshpande G, Hu X, Stilla R, Sathian K. Effective connectivity during haptic perception: a study using Granger causality analysis of functional magnetic resonance imaging data. Neuroimage 2008;40:1807–14.

Deshpande G, LaConte S, James GA, Peltier S, Hu X. Multivariate Granger causality analysis of fMRI data. Hum Brain Mapp 2009;30:1361–73.

Desikan RS, Segonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage 2006;31:968–80.

Edelman GM. Neural Darwinism – selection and reentrant signaling in higher brain-function. Neuron 1993;10:115–25.

Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, et al. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron 2002;33:341–55.

Foucher JR, Vidailhet P, Chanraud S, Gounot D, Grucker D, Pins D, et al. Functional integration in schizophrenia: too little or too much? Preliminary results on fMRI data. Neuroimage 2005;26:374–88.

Friston KJ, Frith CD, Liddle PF, Frackowiak RS. Functional connectivity: the principal-component analysis of large (PET) data sets. J Cereb Blood Flow Metab 1993;13:5–14.

Friston KJ. Statistical parametric mapping. In: Neuroscience databases. Springer; 2003. p. 237–50.

Gao Q, Duan X, Chen H. Evaluation of effective connectivity of motor areas during motor imagery and execution using conditional Granger causality. Neuroimage 2011;54:1280–8.

Glover GH. Deconvolution of impulse response in event-related BOLD fMRI. Neuroimage 1999;9:416–29.

Gonzalez-Castillo J, Saad ZS, Handwerker DA, Inati SJ, Brenowitz N, Bandettini PA. Whole-brain, time-locked activation with simple tasks revealed using massive averaging and model-free analysis. Proc Natl Acad Sci U S A 2012;109:5487–92.

Grill-Spector K, Knouf N, Kanwisher N. The fusiform face area subserves face perception, not generic within-category identification. Nat Neurosci 2004;7:555–62.

Hagmann P, Kurant M, Gigandet X, Thiran P, Wedeen VJ, Meuli R, et al. Mapping Human Whole-Brain Structural Networks with Diffusion MRI. PLoS One 2007:2.

Jones DS. Elementary information theory. Oxford and New York: Clarendon Press; 1979.

Kim JH, Lee JM, Jo HJ, Kim SH, Lee JH, Kim ST, et al. Defining functional SMA and pre-SMA subregions in human MFC using resting state fMRI: functional connectivity-based parcellation method. Neuroimage 2010;49:2375–86.

Korzeniewska A, Manczak M, Kaminski M, Blinowska KJ, Kasicki S. Determination of information flow direction among brain structures by a modified directed transfer function (dDTF) method. J Neurosci Methods 2003;125:195–207.

Logothetis NK, Pauls J, Augath M, Trinath T, Oeltermann A. Neurophysiological investigation of the basis of the fMRI signal. Nature 2001;412:150–7.

Maggioni E, Molteni E, Arrigoni F, Zucca C, Reni G, Triulzi FM, et al. Coupling of fMRI and NIRS measurements in the study of negative BOLD response to intermittent photic stimulation. In: Engineering in Medicine and Biology Society (EMBC), 35th Annual International Conference of the IEEE; 2013. p. 1378–81.

Makni S, Idier J, Vincent T, Thirion B, Dehaene-Lambertz G, Ciuciu P. A fully Bayesian approach to the parcel-based detection-estimation of brain activity in fMRI. Neuroimage 2008;41:941–69.

Mangin JF, Riviere D, Cachia A, Duchesnay E, Cointepas Y, Papadopoulos-Orfanos D, et al. A framework to study the cortical folding patterns. Neuroimage 2004;23:S129–38.

Mezer A, Yovel Y, Pasternak O, Gorfine T, Assaf Y. Cluster analysis of resting-state fMRI time series. Neuroimage 2009;45:1117–25.

Michel V, Gramfort A, Varoquaux G, Eger E, Keribin C, Thirion B. A supervised clustering approach for fMRI-based inference of brain states. Pattern Recognition 2012;45:2041–9.

Mizuno T, Nakagawa E, Sakuma H, Saito Y, Komaki H, Sugai K, et al. Multiple band frequency analysis in a child of medial temporal lobe ganglioglioma. Childs Nerv Syst 2011;27:479–83.

Papoulis A, Pillai SU. Probability, random variables and stochastic processes. fourth ed. New York; 2002.

RamachandranNair R, Ochi A, Akiyama T, Buckley DJ, Soman TB, Weiss SK, et al. Partial seizures triggering infantile spasms in the presence of a basal ganglia glioma. Epileptic Disord 2005;7:378–82.

Roebroeck A, Formisano E, Goebel R. Mapping directed influence over the brain using Granger causality and fMRI. Neuroimage 2005;25:230–42.

Rubinov M, Sporns O. Complex network measures of brain connectivity: uses and interpretations. Neuroimage 2010;52:1059–69.

Said S, Dickey DA. Testing for unit roots in autoregressive-moving average models of unknown order. Biometrika 1984;71:599–607.

Sato JR, Takahashi DY, Arcuri SM, Sameshima K, Morettin PA, Baccala LA. Frequency domain connectivity identification: an application of partial directed coherence in fMRI. Hum Brain Mapp 2009;30:452–61.

Seth AK. A MATLAB toolbox for Granger causal connectivity analysis. J Neurosci Methods 2010;186:262–73.

Smith SM, Fox PT, Miller KL, Glahn DC, Fox PM, Mackay CE, et al. Correspondence of the brain's functional architecture during activation and rest. Proc Natl Acad Sci U S A 2009;106:13040–5.

Tana MG, Bianchi AM, Sclocco R, Franchin T, Cerutti S, Leal A. Parcel-based connectivity analysis of fMRI data for the study of epileptic seizure propagation. Brain Topogr 2012.

Thirion B, Flandin G, Pinel P, Roche A, Ciuciu P, Poline JB. Dealing with the shortcomings of spatial normalization: multi-subject parcellation of fMRI datasets. Hum Brain Mapp 2006;27:678–93.

Thyreau B, Thirion B, Flandin G, Poline JB. Anatomo-Functional Description of the Brain: A Probabilistic Approach. In: ICASSP 2006 Proceedings. 2006 IEEE International Conference on Acoustics, Speech and Signal Processing 2006; 2006.

Tononi G, Sporns O, Edelman GM. Reentry and the problem of integrating multiple cortical areas: simulation of dynamic integration in the visual system. Cereb Cortex 1992;2:310–35.

Tononi G, Sporns O, Edelman GM. A measure for brain complexity: relating functional segregation and integration in the nervous system. Proc Natl Acad Sci U S A 1994;91:5033–7.

Tononi G, McIntosh AR, Russell DP, Edelman GM. Functional clustering: identifying strongly interactive brain regions in neuroimaging data. Neuroimage 1998;7:133–49.

van den Brink R, Gilles RP. The outflow ranking method for weighted directed graphs. Eur J Operational Res 2009;193:484–91.

Wang JH, Wang L, Zang YF, Yang H, Tang HH, Gong QY, et al. Parcellation-dependent small-world brain functional networks: A Resting-State fMRI Study. Hum Brain Mapp 2009;30:1511–23.
Williamson PD, Engel J. Anatomic classification of focal epilepsies. In: Engel J, Pedley TA, editors. Epilepsy: a comprehensive textbook. Philadelphia: Lippincott Williams & Wilkins; 2008. p. 2645–78.
Wu G, Liao W, Stramaglia S, Ding J, Chen H, Marinazzo D. A blind deconvolution approach to recover effective connectivity brain networks from resting state fMRI data. Med Image Anal 2013;17(3):365–74.
Yoon U, Lee JM, Kim JJ, Lee SM, Kim IY, Kwon JS, et al. Modified magnetic resonance image based parcellation method for cerebral cortex using successive fuzzy clustering and boundary detection. Ann Biomed Eng 2003;31:441–7.
Zalesky A, Fornito A, Harding IH, Cocchi L, Yucel M, Pantelis C, et al. Whole-brain anatomical networks: does the choice of nodes matter. Neuroimage 2010;50:970–83.