Modelling cognitive loads in schizophrenia by means of new functional dynamic indexes

Angela Lombardi [a,b], Cataldo Guaragnella [c], Nicola Amoroso [b,a], Alfonso Monaco [a], Leonardo Fazio [f], Paolo Taurisano [d], Giulio Pergola [d], Giuseppe Blasi [d,e], Alessandro Bertolino [d,e], Roberto Bellotti [b,a], Sabina Tangaro [a,*]

[a] Istituto Nazionale di Fisica Nucleare, Sezione di Bari, Via E. Orabona 4, 70125, Bari, Italy

[b] Dipartimento Interateneo di Fisica "M. Merlin", Università degli Studi di Bari "A. Moro", Via E. Orabona 4, 70125, Bari, Italy

[c] Dipartimento di Ingegneria Elettrica e dell'Informazione, Politecnico di Bari, Via E. Orabona 4, 70125, Bari, Italy

[d] Dipartimento di Scienze Mediche di Base, Neuroscienze e Organi di Senso, Università degli Studi di Bari "A. Moro", Piazza Giulio Cesare 11, 70124, Bari, Italy

[e] Azienda Ospedaliero-Universitaria Consorziale Policlinico, 70124, Bari, Italy

[f] Sezione di Neuroradiologia, IRCSS Casa Sollievo della Sofferenza, San Giovanni Rotondo, Foggia, Italy

ARTICLE INFO

Keywords:
Functional connectivity
fMRI
Cross recurrence plots
Dynamic connectivity
Schizophrenia

ABSTRACT

Functional connectivity analysis techniques have broadly applied to capture phenomenological aspects of the brain, e.g., by identifying characteristic network topologies for healthy and disease-affected populations, by highlighting several areas important for the global efficiency of the brain during some cognitive processing and at rest. However, most of the known methods for quantifying functional coupling between fMRI time series are focused on linear correlation metrics. In this work, we propose a multidimensional framework to extract multiple descriptors of the dynamic interaction among BOLD signals in their phase space. A set of metrics is extracted from the cross recurrence plots of each couple of signals to form a multilayer connectivity matrix in which each layer is related to a specific complex dynamic phenomenon. The proposed framework is used to characterize functional abnormalities during a working memory task in patients with schizophrenia. Some topological descriptors are then extracted from both multilayer connectivity matrices and the most used Pearson-based connectivity networks to perform a binary classification task of normal controls and patients. The results show that the proposed connectivity model outperforms the statistical correlation-based connectivity in accuracy, sensitivity and specificity. Moreover, the statistical analysis of the selected features highlights that several dynamic metrics could better identify disease-related dynamic states in brain activity than the statistical correlation among physiological signals.

1. Introduction

Physiological and biological systems involve complex processes whose dynamics exhibit nonlinear interactions (Deisboeck and Kresh, 2007). In particular, the complexity of the brain arises from different aspects.

In recent decades, scientists have tried to incorporate the dense network of relationships and mechanisms of large-scale synchronization that are the basis of the functioning of neuronal networks by using the complex network framework (Sporns, 2010). The network formulation describes the brain as a graph composed by nodes (i.e., brain regions) linked by edges (their functional connectivity) (Bullmore and Sporns, 2009; Sporns, 2010, 2011). Neuroimaging techniques have been extensively applied to investigate the macroscale functional organization of the human brain (Fornito et al., 2016). Many functional connectivity (FC) analysis techniques have been proposed to uncover different dynamic neural mechanisms such as communication, information processing and neural integration from functional magnetic resonance imaging (fMRI) scans (Sporns et al., 2005).

The functional connectome indicates the complete map of the functional links among all the regions of interest in which the total brain is partitioned. Such connections usually quantify the statistical similarity between the time series at each pair of regions, i.e., the functional connectivity (Bullmore and Bassett, 2011). Although the high spatial resolution of the fMRI data allows a detailed mapping of the connections, the

* Corresponding author.
E-mail address: sonia.tangaro@ba.infn.it (S. Tangaro).

low temporal resolution limits the number of methods than can be applied to assess the statistical dependences between the time series. Typically, linear correlation metrics such as Pearson correlation and partial correlation are used to estimate the FC (Van Den Heuvel and Pol, 2010). Dynamic FC analysis aims at describing simultaneously the dynamic relations in time and space of brain activity by observing the evolution of statistical similarity among the fMRI time series into a specific domain (time or frequency) (Calhoun et al., 2014; Bassett et al., 2011; Betzel et al., 2015; De Domenico, 2017).

For example, dynamic FC analysis of fMRI data has been used to detect changes in FC measurements both in healthy populations and across diagnoses and relate such changes to behavioral and cognitive outcomes (Cohen, 2018). Temporal fluctuations have been examined both in rest conditions and in the presence of specific tasks (Gonzalez-Castillo and Bandettini, 2018), to such an extent that the term "chronnectome" has been specifically defined to indicate a connectivity analysis model to investigate nodal activity and time-varying connectivity patterns (Calhoun et al., 2014). Commonly, time series data collected from multiple brain regions are windowed in time of in frequency and statistical similarity through correlation metrics is inferred between the ROIs for each window. A multilayer network is finally built where each layer represents the connectivity pattern for a specific window. Connectivity matrices can be then compared across windows to assess dynamical changes of FC.

The multilayer formulation of dynamic connectivity enables the application of multilayer community detection algorithms which can track a given community over time allowing the definition of new metrics to quantify the flexibility and stability of community formation across layers (Betzel and Bassett, 2017). Such new metrics have been successfully used to study the reconfiguration of functional organization during learning (Bassett et al., 2011), aging (Betzel et al., 2015) and increased cognitive loads (Braun et al., 2015).

These findings suggest some important aspects concerning FC: (i) the mathematical framework of multilayer networks extends some concepts of complex networks (De Domenico et al., 2013) and offers a series of metrics to comprehensively analyze the different interactions between brain components at multiple scales (De Domenico, 2017; De Domenico et al., 2016); (ii) both temporal- and frequency-based multilayer FC analysis could highlight salient topological properties of the brain, which are not detectable by univariate approaches.

However, most of the FC techniques adopt linear correlation metrics, including Pearsons correlation and partial correlation, to assess the statistical interactions between the physiological time series, assuming linear functioning. This hypothesis could represent a limitation for the FC analysis as neurodynamics are actually characterized by complex and nonlinear phenomena that arise from feedback responses to environmental stimuli and physiological interactions among various subsystems that comprise the brain (Marmarelis, 2004). Nonlinear dynamic principles and nonlinear time series tools have been exploited for in-depth investigation of neuronal transient responses and coupling mechanisms with the goal of enhancing the understanding of human cognition and dynamic processes underlying normal and pathological brain states (Rabinovich et al., 2006). A key aspect of dynamical analysis of a system is the reconstruction of its phase space, i.e. the identification of all the dynamic trajectories of the system from which the time series is sampled.

Dynamical analysis methods have been extended for bivariate characterization of coupled interactive systems. Specifically, generalized synchronization measures have been suggested to assess the dynamic interaction of neurophysiological signals in their reconstructed phase space (Pereda et al., 2005; Stam, 2005; Sauer et al., 1991). In detail, generalized synchronization between pairs of signals is measured by mapping the trajectories of the signals into a common phase space and then comparing their local neighboring states. Recently, an efficient way to explore the dynamic behaviour of the trajectories of interacting systems in phase space has been proposed (Lombardi et al., 2017). The proposed method quantifies the functional coupling between fMRI time series by embedding pairs of ROI time series into a common phase space and by using cross recurrence plots (CRPs) (Marwan and Kurths, 2002) to define a new synchronization index. Besides the proposed synchronization metric, several complex metrics exist to perform a formal quantification of graphical patterns of recurrence plots (RPs) and CRPs that have been proven to be useful in a great variety of context (Webber and Marwan, 2015; Lombardi et al., 2016). Recurrence Quantification Analysis (RQA) is a model-free analysis that includes a set of generalized indices, each related to specific complex phenomena occurring in phase space (Webber and Zbilut, 2005). Hence, its application to brain signal analysis could reveal changing dynamics, critical events and distinctive features of cognitive states and diseases.

In this work, we present a framework based on both CRPs and graph analysis aimed at capturing dynamic changes of functional connectivity and providing straightforward markers of the dynamic states in brain activity to characterize pathological conditions. The principal idea underlying this analysis is that if each RQA metric is representative of different dynamic phenomena in the phase space, a multivariate model that uses simultaneously all the measurements extracted from the CRP, could describe more completely the dynamic behaviour of the interacting systems. We used the fMRI data of a cohort of subjects including controls and schizophrenic patients acquired during a working memory task, to verify whether the new markers could be associated with different cognitive loads and detect differences between the groups of subjects. In brief, we applied the multilayer cross-recurrence framework to fMRI data of the two groups of subjects performing both a 2-back and 0-back experiment to define task-evoked networks. Thus, some topological descriptors of the multilayer networks were extracted and a supervised classification pipeline was applied to identify the set of features that achieve the best classification accuracy of the two groups of subjects. The reasons for the N-back task fMRI data selection are twofold: (i) schizophrenia-affected areas related to working memory are documented in literature, so a comparison between the set of identified features and the known areas is simplified; (ii) by using the two task-evoked datasets it is possible to verify if group differences in connectivity change as functions of task conditions in order to detect dynamic descriptors that reflect different cognitive loads in the two populations.

FC was also assessed by using the Pearson correlation coefficient between couple of time series. The connectivity matrices obtained with the Pearson coefficient are used as a benchmark to which the other complex dynamic indices are compared. Indeed, these matrices represent connectivity patterns obtained with simple statistical similarity between pairs of time series. Similarly, topological descriptors were extracted from these networks, and the predictive power of the features to classify the two groups of subjects was evaluated.

2. Materials

2.1. Subjects

49 participants and 42 patients with schizophrenia were included in this study. Control subjects were evaluated using the Non-Patient Structured Clinical Interview for DSM-IV (DATA, 1997) to exclude any psychiatric condition. Other exclusion criteria were: a significant history of drug or alcohol abuse, active drug abuse in the previous year, experience of a head trauma with loss of consciousness and any other significant medical condition. We also reported handedness (Edinburgh Inventory) (Oldfield, 1971), total IQ (TIB), PANSS symptoms scores and equivalent medication scores (Gardner et al., 2010) (see Table 1). The present study was approved by the local ethics committee.[1] Written informed consent was obtained by all participants after a complete description of the procedures, in accordance with the Helsinki Declaration.

[1] Comitato Etico Locale Indipendente Azienda Ospedaliera Ospedale Policlinico Consorziale Bari.

Table 1
Demographic data and clinical information of patients and healthy controls.

	Controls	Patients	P value[a]
Demographics			
Mean age (years) (range)	30(23 – 62)	34(23 – 50)	0.07
Gender (male/female)	26/23	32/10	0.02
Handedness (Median ± SD)	0.80 ± 0.2	0.80 ± 0.3	0.78
IQ (Median ± SD)	107.8 ± 13.8	83.7 ± 20.10	< 0.001
Head motion			
mFD 0-back (Median ± SD)	0.16 ± 0.09	0.17 ± 0.10	0.23
mFD 2-back (Median ± SD)	0.13 ± 0.07	0.14 ± 0.09	0.13
Task performance			
0-back (Median ± SD)	100 ± 6.67	98 ± 7.12	0.02
2-back (Median ± SD)	85 ± 20.59	47 ± 23.35	< 0.0001
Medication			
Equivalent (Median ± SD)	NA	550 ± 258	
Symptoms			
Total PANSS (Median ± SD)	NA	36 ± 7.5	

[a] Resulting from the Wilcoxon rank sum test except for Gender, for which the Chi-square test was adopted.

2.2. Task

Participants performed the N-Back working memory task, in which a sequence of stimuli is presented and the subject has to remember the stimulus from "N" steps earlier (Bertolino et al., 2004). The stimuli consisted of numbers (1–4) presented in random sequence and displayed at the points of a diamond-shaped box. The control condition (0-back) simply required the subjects to identify the current stimulus. In the working memory condition, the task required the collection of a stimulus seen two stimuli earlier (2-Back). The task was organized in a block design, consisting of eight alternating 0-back and 2-back conditions, each lasting 30 s. Each 30 s block includes 14 n-back trials with an inter-stimuli interval of 2000 ms. Each run lasted 4 min and 8 s, from which dummy scans were acquired and discarded, obtaining 120 volumes. Information about task performance accuracy for both cohorts are reported in Table 1.

2.3. fMRI data

Blood oxygen level-dependent (BOLD) signal was recorded by a GE Signa 3T scanner (General Electric, Milwaukee, WI), using a gradient-echo planar imaging sequence (repetition time, 2000 ms; echo time, 28 ms; 20 interleaved axial slices; thickness, 4 mm; gap, 1 mm; voxel size, 3.75 × 3.75 × 5 mm; flip angle, 90°; field of view, 24 cm; matrix, 64 × 64). The first four scans were discarded to allow for equilibration effect.

We performed standard fMRI preprocessing steps using SPM12.[2] The images were realigned to correct for motion artifacts. Movement parameters were extracted to exclude data affected by excessive head motion (2.5 mm of translation or 2.5° of rotation). Realigned images were resliced to a 3.75 mm isotropic voxel size, co-registered to high-resolution T1-weighted structural images, spatially normalized into a standard space (Montreal Neurological Institute), smoothed with a 10 mm full-width at half-maximum isotropic kernel and temporal band-pass filtered (0.01 – 0.25 Hz). Additionally, nuisance signal correction was done on the data by regressing out (1) linear trends in the time series; (2) mean time-series from the white matter (WM) and the cerebrospinal fluid (CSF); (3) 24 motion parameters obtained by motion correction (as specified in the Friston24 model (Friston et al., 1996)); and (4) signals extracted using the CompCor algorithm (Behzadi et al., 2007). We used five components for CompCor-based nuisance regression. To extract mean time-series from the WM, gray matter, and CSF the anatomical MR data were automatically segmented. Mean frame-wise displacement (mFD) (Power et al., 2012) is also reported in Table 1 for both conditions in order to compare head motion between the two groups.

[2] (http://www.fil.ion.ucl.ac.uk/spm).

3. Methods

3.1. General framework

The general framework is shown in Fig. 1. The brain volume of each subject was divided in 246 non-overlapping anatomical regions of interest according to the Brainnetome Atlas (Fan et al., 2016). Thirty regions from the most ventral part of the brain not acquired during scans were discarded and are not included in the following analysis. For each of the 216 remaining ROIs, a single time series was extracted by averaging the fMRI time series over all the voxels within the ROI.

As in (Lombardi et al., 2017), we firstly reconstructed the phase space of each time series by using its m-dimensional time delay embedded vector:

$$\overrightarrow{x}_i = \left(u_i, u_{i+\tau}, \ldots, u_{i+(m-1)\tau}\right) \tag{1}$$

where $\{u_i\}_{i=1}^{N}$ is the BOLD time series, $m = 6$ is the embedding dimension end $\tau = 1$ is the time delay.

For each participant, functional connectivity between all pairwise combinations of ROI time series was assessed by computing their CRP and then by calculating the 17 RQA metrics as described in the following subsection resulting in 17 connectivity matrices.

A tensor structure with 17 undirected weighted networks as layers is finally composed of all the 17 connectivity matrices.

In order to assess the importance of the regions with respect the rest of the network, the following graph metrics have been extracted from each weighted graph layer $W^{[\alpha]} = \{w_{ij}^{[\alpha]}\}, \alpha = 1, \ldots, 17$ and for each node of the network $i = 1, \ldots, 216$:

- strength s_i^α:

$$s_i^\alpha = \sum_{j=1}^{N} w_{ij}^\alpha \tag{2}$$

- betweenness centrality b_i^α:

$$b_i^\alpha = \sum_{i \neq j \neq t} \frac{\sigma_{jt}^\alpha(i)}{\sigma_{jt}^\alpha} \tag{3}$$

where σ_{jt}^α denotes the number of shortest paths from j to t and $\sigma_{jt}^\alpha(i)$ denotes the number of shortest paths from j to t that pass through i (Freeman, 1977);

- clustering coefficient:

$$c_i^\alpha = \frac{2}{k_i^\alpha\left(k_i^\alpha - 1\right)} \sum_{j,k} \left(w_{ij}^\alpha w_{jk}^\alpha w_{ki}^\alpha\right)^{1/3} \tag{4}$$

where $\sum_{j,k} w_{ij}^\alpha w_{jk}^\alpha w_{ki}^\alpha$ is sum of triangle intensities attached to the node i (Onnela et al., 2005);

- pagerank centrality p_i^α, an iterative centrality metric that assigns the importance to each node by random walking on the network. At first, each node has $p = 1$; then each node spreads the centrality value to its neighbours according to the link weights along the output connections. Finally, the importance of a node is determined by both criteria of quality and quantity of the pages linked to it. The p_i value of the node i at the step t is defined as:

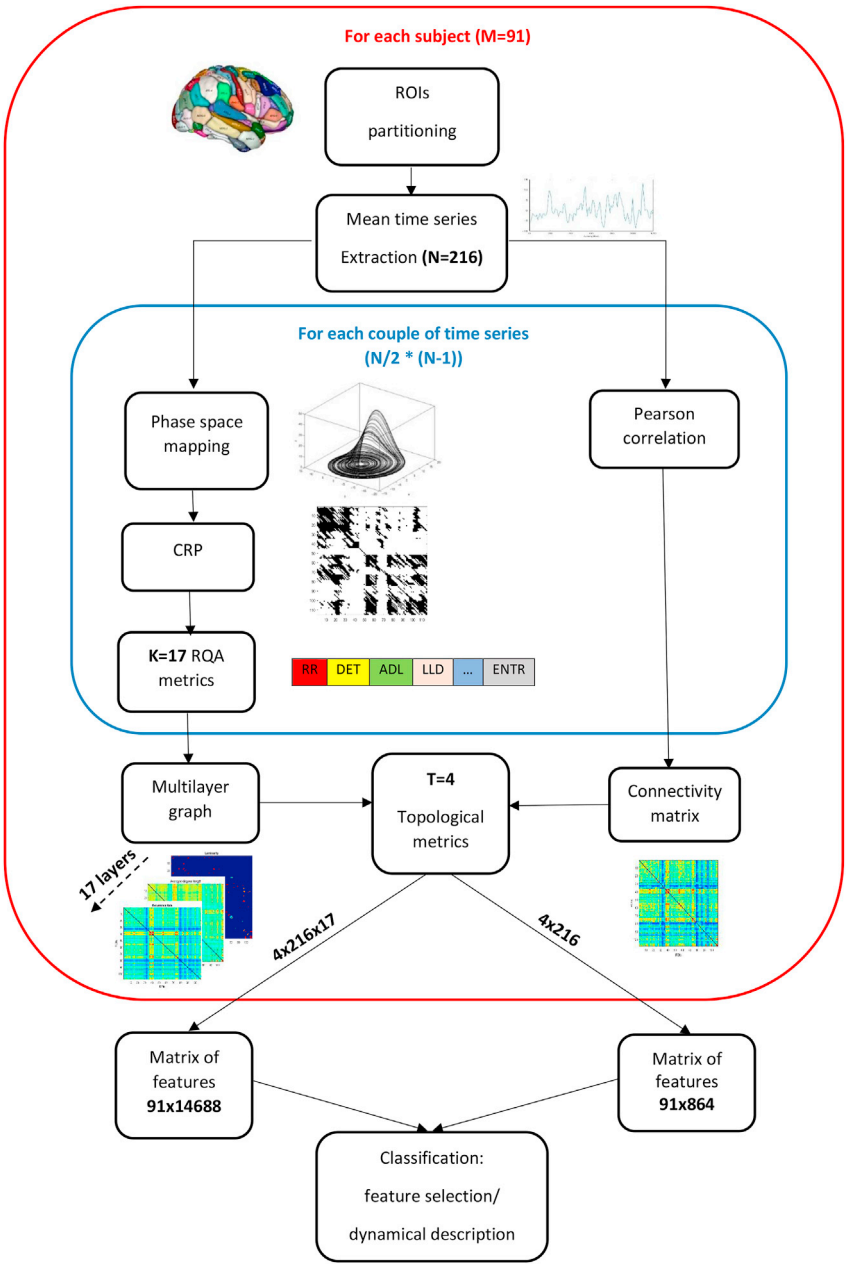

Fig. 1. Flow-chart showing the multi-recurrence framework and the comparison with the Pearson-based functional connectivity.

$$p_i^{\alpha}(t) = \sum_{j=1}^{N} \frac{w_{ij}^{\alpha} p_j^{\alpha}(t-1)}{s_j^{\alpha}} \tag{5}$$

where the iterations will stop if the steady state of the p values of all the nodes is reached.

At the end, a 91 × 14688 matrix is constructed, in which each feature is labelled as: *rqa layer - graph metric - roi* according to the layer from which it belongs, the graph metric and the node (ROI) of the network.

In order to compare the phase space approach to the temporal analysis, functional connectivity for each subject was also assessed by computing the Pearson correlation coefficient between all pairwise combinations of ROI time series:

$$r = \frac{\sum_{n=1}^{N}(x_n - \overline{x})(y_n - \overline{y})}{\sqrt{\sum_{n=1}^{N}(x_n - \overline{x})^2}\sqrt{\sum_{n=1}^{N}(y_n - \overline{y})^2}} \tag{6}$$

where $\{x_i\}_{i=1}^{N}$ and $\{y_i\}_{i=1}^{N}$ are the two time series and $\overline{x}$ and $\overline{y}$ denote their sample means. Then, the same local graph metrics were extracted from the connectivity matrices resulting in a 91 × 864 matrix, in which each feature is labelled as: *graph metric - roi*.

3.2. RQA metrics

As already mentioned, a CRP enables the comparison of the trajectories of two distinct systems in the same embedding space (Marwan and Kurths, 2002). A CRP provides information on the degree of similarity of each state of a system with any other state of the second system. For two systems with trajectories respectively $\vec{x}_i$ $(i = 1, \ldots, N)$ and $\vec{y}_j$ $(j = 1, \ldots, N)$, the CRP is defined as:

$$CR_{i,j}(\varepsilon) = \Theta\left(\varepsilon - \left|\left|\vec{x}_i - \vec{y}_j\right|\right|\right) \tag{7}$$

where Θ is the Heaviside function, ε is a threshold for closeness, N is the number of considered states for each system and $||\cdot||$ a norm function. A

generic entry $CR_{i,j}$ in the resulting $N \times N$ array is set to one if the distance between the points $\vec{x}_i$ and $\vec{y}_j$ is smaller than the threshold ε or to zero elsewhere.

The graphical patterns of a CRP are related to similarities and differences between the dynamic evolutions of the couple of systems under investigation. Single, isolated recurrence points indicate rare states that can appear due to noise. Diagonal lines occur when the evolution of the states is similar at different times. Vertical and horizontal black lines mark time periods in which the sequence of states changes very slowly or do not change at all, while vertical and horizontal white bands result from states which occur rarely.

The analytical evaluation of the geometric structures of the CRPs involves a set of complex measures, mostly based on the distribution of the lengths of the diagonal and vertical/horizontal lines in the plot (Zbilut and Webber, 1992; Marwan et al., 2007).

In this work, five major classes of measures defined as follows are considered:

1. Recurrence density - based measures, i.e. metrics based on the density of the points in a CRP.
2. Measures based on the distribution $P(l)$ of lengths l of the diagonal lines. These metrics mostly quantify the period during which the two systems visit close orbits in the phase space.
3. Measures based on the distribution $P(v)$ of vertical line lengths v. This distribution is used to quantify laminar phases during which the states of the systems change very slowly.
4. Recurrence network - based measures. A CRP can be also interpreted as the adjacency matrix of a graph $G(V,E)$, whose links E connect the set of vertices V that represent neighbor points in phase space. Topological properties of recurrence networks have been related to statistical properties of the phase space density, revealing further complex dynamical aspects of time series. In particular, some network-based metrics have been associated with invariant characteristics of phase space, i.e., independent from a particular embedding (Donner et al., 2010).
5. Line of synchronization (LOS) - based measures. A CRP can also exhibit the main diagonal known as LOS. The presence of LOS implies the identity of the states of the two systems in the same time intervals, so its structure can be analyzed to extract information about the synchronization of the two time series.

More details and a qualitative description of the measures can be found in Appendix.

3.3. The statistical framework

A statistical framework was developed in order to select only the most significant features among the 14688 of the multilayer recurrence connectivity and among the 864 features from the Pearson's matrices for the binary classification problem (i.e., healthy versus patients) and for each of the two conditions. Indeed, mass univariate hypothesis tests detect the cross-group differences by comparing each feature individually across the two populations. This technique is effective in describing the most important differences between two classes but has some limitations (Arbabshirani et al., 2017).

We stress that in this context the classification of subjects for the 0-back and 2-back conditions and for the two analyses (multi-recurrence framework and Pearson correlation) is used to verify the presence of particular dynamic phenomena in the phase space that characterize the interactions among the brain regions in the two populations beyond the statistical correlation among their activity. In this context, the selection of the features showing the greatest predictive power is of paramount importance: indeed, these features indicate which dynamic metric and which area is particularly relevant to explain the differences between the two groups as the cognitive load increases.

Hence, a multivariate machine learning approach has been integrated with statistical techniques to ensure the stability of the performance values achieved by the model and the set of the most predictive features. A Support Vector Machine (SVM) classifier was used to identify multivariate patterns of multi-recurrence connectivity related to the diagnostic groups. A k-folds validation procedure was repeated to perform a robust feature selection. The best subset in terms of classification performances was selected to evaluate the final set of most predictive features to identify the dynamic phenomena which characterize the two populations during both cognitive conditions. The main steps of the framework are shown in Fig. 2 and are described in the following subsections more in details.

3.3.1. Feature selection

Feature selection techniques are powerful tools that attempts to:

- removing irrelevant, noisy and redundant features, with the aim to avoid overfitting and improving classification performance;
- reducing the computational complexity of the learning algorithm;
- providing a deeper insight into the data, highlighting which features are most informative for classification.

Feature selection algorithms are divided into three categories: filters, wrappers and embedded methods (Guyon and Elisseeff, 2003). Filters evaluate each feature without interaction with classifiers by using several criteria related to correlations among features or amount of shared information. Wrappers find a feature subset that has the minimum cross-validation error on the training data, treating the classifier as a black box. Searching methods such as sequential forward selection and simulate annealing are examples of wrappers. Embedded methods incorporate variable selection as part of the training process.

In this work, Support Vector Machine-Recursive Feature Elimination (SVM-RFE), which integrates in a single consistent framework both feature selection and pattern classification, was selected to perform feature selection. SVM-RFE is an embedded method, introduced by Guyon et al. (2002) in the context of gene selection for cancer classification.

The main intuition of SVMs is to find a separating hyperplane with the largest possible margin on either side (Cortes and Vapnik, 1995). A particularity of the algorithm is that the weights w_i of the decision function $D(x)$ are a function of only a small subset of the training examples, called support vectors.

Recursive feature elimination, in its broad sense:

1. Trains the classifier;
2. Computes a ranking criterion for all features;

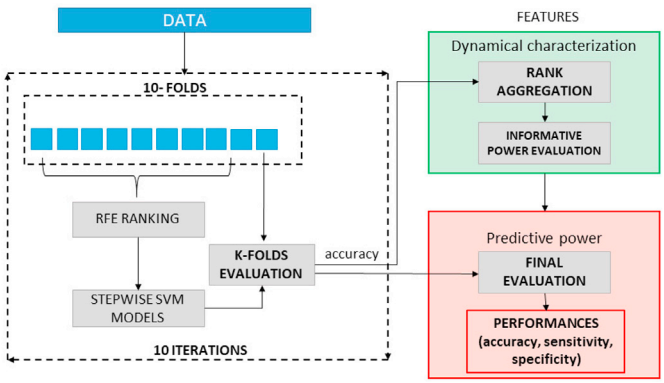

Fig. 2. Statistical framework to select the most discriminative features for both the experiments. A nested feature selection was performed on the training set in each round of the k-fold validation. Then 100 stepwise SVM models were trained by progressive increasing the training set size. A consensus ranking procedure was used to select the most stable features with the highest accuracy.

3. Removes the feature with smallest ranking criterion, as it has the least effect on classification.

This process is iteratively computed until all the features have been removed. SVM-RFE is intimately related to the SVM model, as the ranking criterion for feature k is the square of the k-th element of w: $J(k) = w_k^2$. In particular, in this work, the SVM-RFE with correlation bias reduction, as proposed in (Yan and Zhang, 2015), has been adopted. In the original formulation of SVM-RFE, one problem is not addressed: when some of the features are highly correlated, the assessing criteria of these features will be influenced, so their importance will be underestimated. The authors called this phenomenon correlation bias. When a subset of features is removed in one iteration of RFE, a group of correlated features may be removed entirely. This may happen either because the features are truly irrelevant, or because their ranking criteria have been incorrectly underestimated. In both conditions, the method proposed in (Yan and Zhang, 2015) moves representative feature of the group back to the surviving feature list. In this way, the group can be evaluated again in the next iteration without the influence of the correlation bias. The group representative can be chosen as the feature with the highest criterion in this iteration. Therefore, this strategy does not change the candidate feature set or the ranking criterion, but monitors and corrects the potentially wrong decisions due to the correlation bias.

The output of SVM-RFE is a list of ranked features. Feature selection can be achieved by choosing a group of k_{opt} top-ranked features. Since k_{opt} is not known a priori, the RFE algorithm was performed in each fold of a cross-validation procedure. In this way, a nested feature selection was obtained since the ranking of the features is blind to the test set. This approach was chosen in order to avoid the "double dipping" problem. Indeed, performing the feature selection on the whole dataset could introduce bias in the final classification model. The dangerous effects of the double dipping have been widely described and result essentially in overestimating the values of accuracy and area under the ROC curve (AUC) (Singhi and Liu, 2006; Kriegeskorte et al., 2009). In some neuroimaging studies different recommendations have been provided to avoid the bias introduced by this procedure (Pereira et al., 2009; Olivetti et al., 2010; Eskildsen et al., 2013).

In particular, here 10 re-sampling of a 10-fold cross-validation were executed producing 100 bootstraps of the train set. In each iteration, 9-folds of the original training set were input to RFE and then stepwise SVM models were trained for ranked subsets of increasing size (e.g., the top 10, 20, 30, and so on up to 14688 ranked features for the multilayer cross-recurrence connectivity). Each stepwise model was tested on the left fold and the performances of each model were stored for successive evaluations. For the classification analysis the LIBSVM software version 3.23[3] used with the default parameters (cost $c = 1$). The purpose of stepwise analysis is to identify the particular subset of features that maximizes classification performance. Consequently, the output of the analysis is a performance curve from which the number of features k_{opt} can be assessed.

3.3.2. Stability of the features

Some measures are usually employed to assess the predictive power of a classifier. Given the output statistics of a classifier in terms of true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN), we used:

1. Accuracy (ACC):

$$ACC = \frac{TP + TN}{TP + FP + TN + FN} \tag{8}$$

2. Sensitivity or true positive rate (TPR):

$$TPR = \frac{TP}{TP + FN} \tag{9}$$

3. Specificity or true negative rate (TNR):

$$TNR = \frac{TN}{TN + FP} \tag{10}$$

Here, the number of features k_{opt} to be selected was chosen as the one yielding the highest accuracy of all the stepwise models. Since, in principle, the features of each subset corresponding to the top k_{opt} elements of each model can be different from one cross-validation round to another, a consensus ranking algorithm was used to select the most common features across all the 100 models. In particular, we adopted a rank aggregation procedure to combine the multiple ranked lists, (i.e., the base rankers), into a final aggregated ranked list, which is considered more reliable than the base rankers (Deng et al., 2014). This problem is of a critical relevance in genomic applications, where it is often required to combine large amounts of genetic data collected by multiple laboratories with different sources of variability and batch effects (Li et al., Xiao). To determine the stability of the selected subset of features, we adopted the robust rank aggregation algorithm (Kolde et al., 2012). Briefly, this approach computes the position of each item in the final ranking by comparing its position in all the ranked lists to a non-informative null model of random permutations of the items. A numerical score for each item is then assigned according to the reference beta distributions of order statistics and the Bonferroni correction is applied to compute P-values and find statistically significant items. The P-values are then sorted to obtain the final ranking. The algorithm was also extended for accounting partial rankings, i.e., lists where only k top elements are available.

3.4. Prediction of task performance from connectivity

Table 1 shows the in-scanner task accuracy of the participants. For both 0-back and 2-back, schizophrenic subjects had lower accuracy compared to controls. In order to explore relation between the selected connectivity features and the subjects' task performance, we predicted task accuracy scores from the most significant connectivity features of each condition as well as from both methods (Pearson and multi-recurrence framework). In particular, a general linear model was used to predict the accuracy score of each subject during each task condition with a leave-one-out validation scheme.

3.5. Statistical analysis of phenotypic information

We performed a Wilcoxon rank-sum test to compare age and handedness between the two cohorts and Chi-square test for the gender. As shown in Table 1, only a significant difference in gender between the two groups of subjects was found, suggesting that this factor might affect the discrimination power of the classification models. So we performed the classification by including the gender variable as an additional feature. Results are displayed in SI-Fig. 2 and reveal that the performance did not significantly improve compared with the classification accuracy achieved by neglecting this information.

4. Results

4.1. Classification performance

Fig. 3 shows the mean accuracy achieved by all the stepwise models with the standard error for both conditions respectively for the multi-recurrence framework (Fig. 3a) and Pearson correlation (Fig. 3b). If on the one hand the trend of curves shows a decay in performance after a significant peak and therefore the presence of poorly informative and redundant features, on the other hand it is interesting to note that the 0-back multi-recurrence features provide always a better discrimination

[3] https://www.csie.ntu.edu.tw/~cjlin/libsvm/.

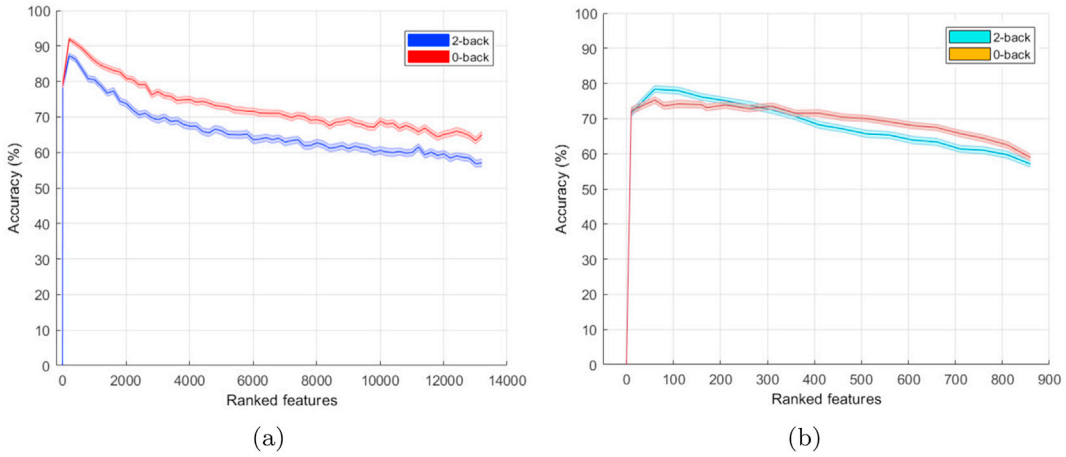

Fig. 3. Shadow performance curves with the mean accuracy achieved by all the stepwise models with the standard errors for both conditions for (a) the multi-recurrence framework and (b) Pearson correlation.

between normal controls and patients. The accuracy curves for Pearson connectivity reported in Fig. 3b show similar performances for the two conditions. The analysis of the averaged performances revealed $k_{opt} = 200$ for the multilayer cross-recurrence framework and $k_{opt} = 60$ for the Pearson connectivity. Hence the top 200 features of the 100 SVM models trained with the cross-recurrence features and the top 60 of the 100 SVM models trained with the graph metrics extracted from the Pearson correlation matrices were selected for the stability analysis. Fig. 4 shows the mean values with standard errors of the accuracy, sensitivity and specificity of all the 100 stepwise SVMs models for the selected k_{opt} for each experiment. Numerical values of the same performances are also reported in Table 2. In both the experiments (i.e., 0-back and 2-back) the multi-recurrence connectivity features have been shown to be significantly more effective than those extracted from the Pearson's matrices in discriminating patients against controls. Average values resulting from 5000 permutations of the labels of subjects for each experiment are also reported for comparison to the chance level. The performance values of all the experiments resulted significantly different from the empirical chance level. The confusion matrices are listed in SI-Section 1 of Supplementary Information file.

Table 2
Performances of the classifications for the most discriminative features for each experiment.

Experiment	ACC (%)	TPR (%)	TNR (%)
2-back multi-recurrence	87.25 ± 0.67	94.16 ± 0.99	78.76 ± 1.23
2-back Pearson	78.38 ± 0.95	87.01 ± 1.40	67.87 ± 1.68
0-back multi-recurrence	91.91 ± 0.52	97.24 ± 0.58	85.29 ± 1.12
0-back Pearson	75.26 ± 0.93	80.82 ± 1.38	68.47 ± 1.89

4.2. Ranked features

In order to obtain the final set of ranked features, the stable k_{opt} features were selected through the robust rank aggregation algorithm. Since the multilayer recurrence model implicitly defines hierarchical features, it is possible to calculate global statistics, i.e., defined on a single hierarchical level of the features independently of the others, or local statistics, i.e., by selecting a sequence of the hierarchical levels. Here, the global statistics concerning the levels of the multi-recurrence graphs and the most important ROIs are shown.

Specifically, the frequency of occurrence of the layers across the features' labels are used to show importance ranking of the layers at global level. Similarly, the occurrence frequency of each ROI of the partition across the k_{opt} features' labels are employed to rank the most important regions.

The ranking of the layers are shown in Fig. 5. Synchronization metrics (HARD SYNC, SYNC) are among the top ranked layers for the 2-back condition (rates: 12% and 10%, respectively); only the HARD SYNC

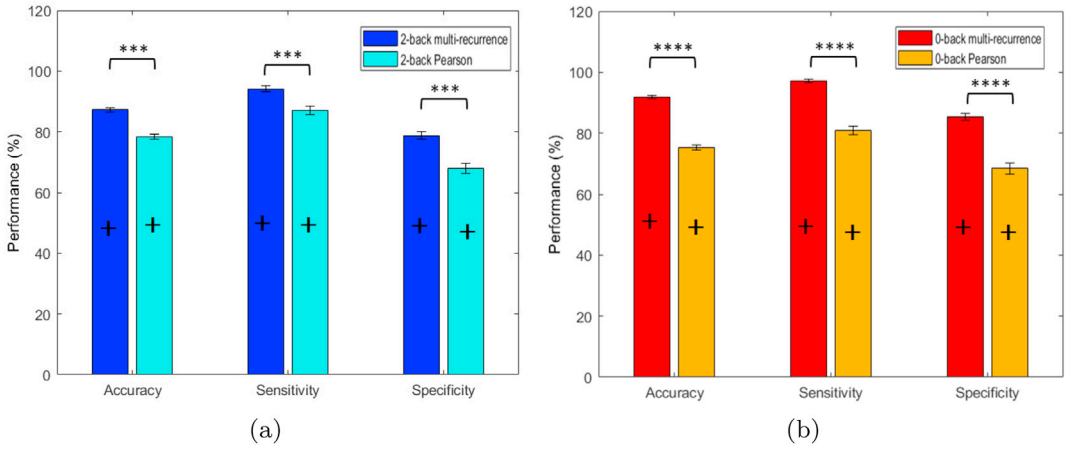

Fig 4. Mean values and standard errors of the accuracy, sensitivity and specificity of the stepwise SVMs obtained for the first k_{opt} ranked features for (a) the 2-back condition and (b) the 0-back condition. The asterisks denote statistically significant differences ($****p < 0.00001$, $***p < 0.0001$, $**p < 0.001$, $*p < 0.01$) for the indicated groups (Wilcoxon's rank sum test). The black crosses denote chance level resulted from 5000 permutation of subjects' labels.

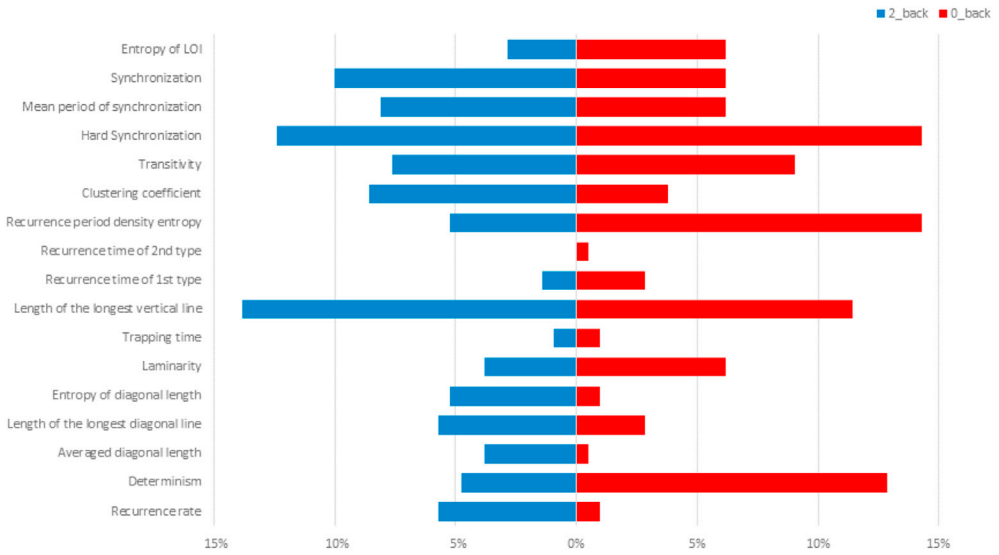

Fig. 5. Ranking of the multi-recurrence layers derived from the frequency of occurrence of the stable features according to the layer from which they belong for both the experiments.

index among the diagonal-based RQA metrics prevails in the 0-back experiment (14%). The LLV layer exhibits high frequencies of occurrence across the features for both experiments (14% for 2-back and 11% for 0-back). It is worth noting that determinism is prevalent among the most important features only for the visual-motor task (13%), while the occurrence percentage is negligible in the working memory task (5%).

In Fig. 6 are shown the most significant ROIs for the 2-back condition. Table 3 also reports the same regions with MNI coordinates. Fig. 7 and Table 4 show the most significant ROIs and relative coordinates for the 0-back condition.

4.3. *Prediction of task performance from connectivity*

We performed a correlation analysis between the true task accuracy

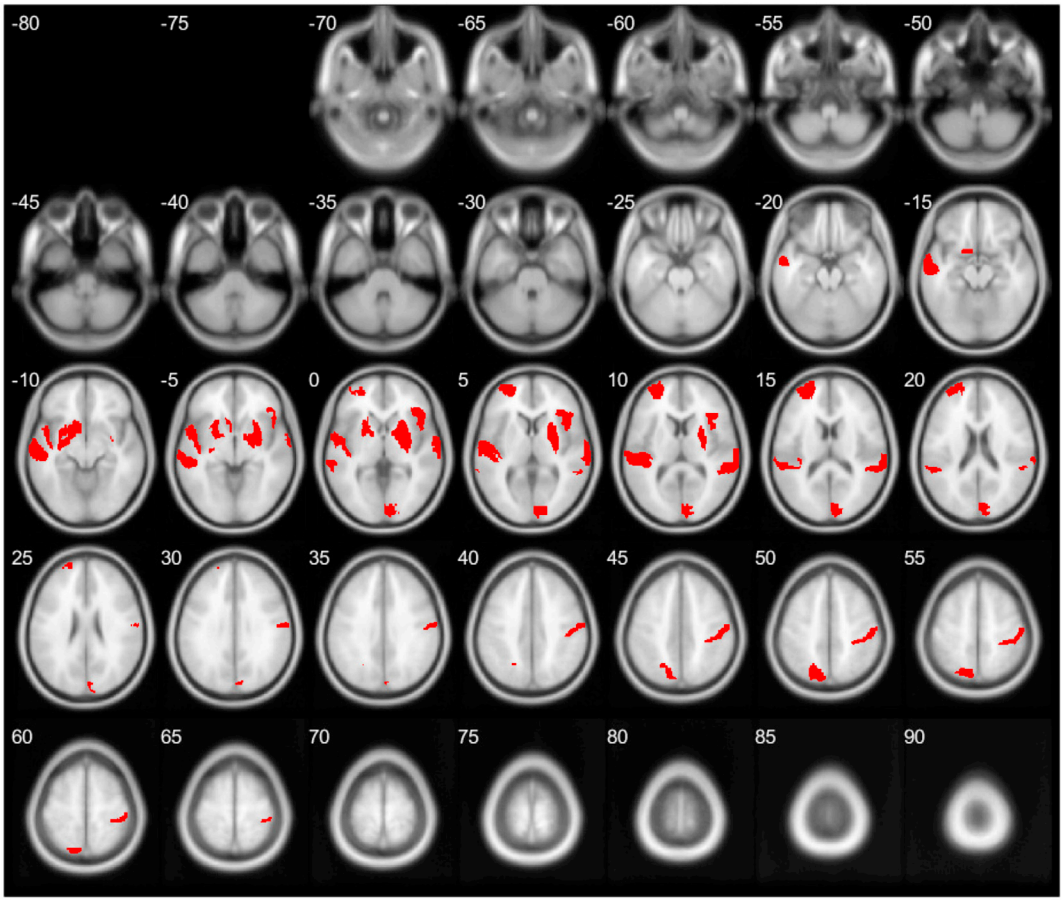

Fig. 6. Slice view of the most significant ROIs for the classification of the control/schizophrenic subjects during the 2-back condition.

Table 3
Significant ROIs with MNI coordinates for the classification of the control/schizophrenic subjects during the 2-back condition.

ROI	MNI coordinates		
	x	y	z
A46, area 46 L	− 28	56	12
A44op, opercular area 44 R	42	22	3
A41/42, area 41/42 L	− 54	− 32	12
TE1.0 and TE1.2 L	− 50	− 11	1
A22c, caudal area 22 R	66	− 20	6
aSTS, anterior superior temporal sulcus L	− 58	− 20	− 9
cpSTS, caudoposterior superior temporal sulcus R	57	− 40	12
A7c, caudal area 7 L	− 15	− 71	52
A2, area 2 R	48	− 24	48
cCunG, caudal cuneus gyrus R	8	− 90	12
GP, globus pallidus R	22	− 2	3
NAC, nucleus accumbens L	− 17	3	− 9
vmPu, ventromedial putamen L	− 23	7	− 4
vmPu, ventromedial putamen R	22	8	− 1
dlPu, dorsolateral putamen R	29	− 3	1

Table 4
Significant ROIs with MNI coordinates for the classification of the control/schizophrenic subjects during the 0-back condition.

ROI	MNI coordinates		
	x	y	z
A9/46v, ventral area 9/46 R	42	44	14
A1/2/3ll, area1/2/3 (lower limb region) L	− 8	− 38	58
A35/36c, caudal area 35/36 L	− 25	− 25	− 26
dmPOS, dorsomedial parietooccipital sulcus(PEr) L	− 12	− 67	25
A1/2/3ulhf, area 1/2/3 L	− 50	− 16	43
A1/2/3tonIa, area 1/2/3 L	− 56	− 14	16
dIa, dorsal agranular insula L	− 34	18	1
dId, dorsal dysgranular insula R	38	5	5
rCunG, rostral cuneus gyrus L	− 5	− 81	10
cCunG, caudal cuneus gyrus L	− 6	− 94	1
vmPOS, ventromedial parietooccipital sulcus L	− 13	− 68	12

scores during the two conditions and the predicted scores for both Pearson and multi-recurrence framework in order to evaluate the prediction performance. As shown in Table 5, none of the prediction performance was statistically significant. Further details are provided in SI-Fig. 3.

5. Discussion

5.1. Significant RQA layers

In this work a novel approach to investigate the interacting behaviour of neural dynamics is shown. In particular, a multidimensional framework to map the coupling interaction of fMRI time series into a common phase space and extract dynamic descriptors by means of cross recurrence quantification analysis is proposed. The aim of this study was to demonstrate that generalized indexes of synchronization computed from CRP together with a complex network-based approach could be effectively used to reveal further insights into different cognitive mechanisms in the presence of psychiatric disorders such schizophrenia.

Many studies have explored the brain at various levels of details, confirming its multi-scale nature (Betzel and Bassett, 2017). In particular, the brain has shown multiple temporal scales with behaviour ranging from milliseconds to the entire lifespan and can be explored at different topological scales ranging from individual nodes to the whole network.

Recently, the phase space reconstruction has been used to perform a mutual connectivity analysis (MCA) to estimate directed links among

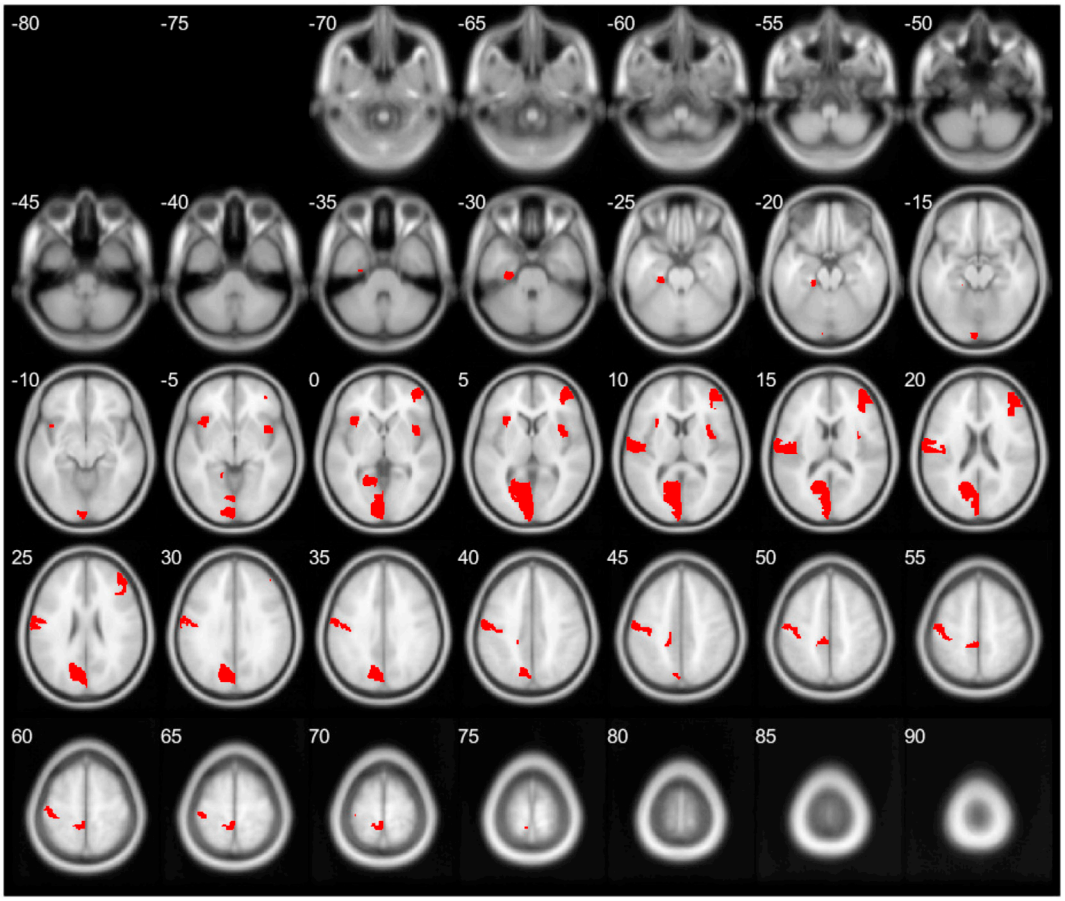

Fig. 7. Slice view of the most significant ROIs for the classification of the control/schizophrenic subjects during the 0-back condition.

Table 5
Correlation between the true accuracy scores of the participants during the two conditions and the predicted scores obtained by using a general linear model with the most discriminating features for both Pearson and multi-recurrence framework.

Experiment	R	p-value
2-back multi-recurrence	0.15	0.14
2-back Pearson	0.14	0.18
0-back multi-recurrence	0.03	0.76
0-back Pearson	− 0.15	0.13

fMRI time series by quantifying the cross-prediction between the dynamic trajectories of each pair of signals (DSouza et al., 2018). The analysis showed that the nonlinear modelling of the phase space reconstruction could offer an useful hypotheses-free method to assess effective connectivity. Here, the multidimensional phase space approach has been compared with the linear temporal correlation analysis. Indeed, by means of the cross recurrence analysis it is possible to compare a specific state of the dynamic system, i.e., the values of the variables that describe it at a given time, with any other state of another system. With the Pearson correlation between pairs of signals, a linear temporal analysis is performed by comparing the two systems in the same time intervals. The multilayer cross-recurrence analysis represents a generalization of the temporal analysis since trajectory segments of the two systems are compared not necessarily for successive temporal instants. In fact, RQA metrics quantify "recurrent" dynamic behaviors of two systems, or rather identify reciprocal patterns in the entire phase space. In addition, recurrence metrics are linked to some invariant measurements in the phase space but has been found to be useful for analyzing also biological short, noisy and non-stationary data (Webber and Marwan, 2015; Zbilut et al., 2002; Marwan et al., 2007). For example, RQA have been used to detect N400 ERP components in EEG data (Schinkel et al., 2009), extract nonlinear information from EEG signals strongly related to severity of autism spectrum disorder (ASD) and predict early stages of ASD (Bosl et al., 2018). RQA has been also adopted to find activation patterns in fMRI data in an univariate data-driven fashion by detecting autocorrelated and recurrent time series (Bianciardi et al., 2007).

The classification results show that the multivariate approach is significantly more effective in classifying control subjects and patients. It should be noted that the classification performances are used to select the optimal subset of features for each experiment (0-back and 2-back with both Pearson correlation and multi-recurrence approach) and that the greater predictive power of the multi-recurrence features may indicate that other dynamic phenomena in the phase space could characterize differently the two populations besides the simple temporal synchronization. This is also confirmed by the analysis of occurrence of the layers of the multi-recurrence structure resulting most significant for the classification between schizophrenic subjects and controls. In details, the two synchronization-related metrics (HARD SYNC and SYNC) exhibits high percentage of occurrence in the 2-back condition. The HARD SYNC is one of the most frequent layer for both the experiments. The metric HARD SYNC emphasizes the concept of synchronization expressed by the SYNC as it is a binary metric that indicates the connectivity between a pair of ROIs only in case of perfect synchronization between the two areas (i.e., in case of the presence of the LOS) and it assigns a zero value otherwise. It is worth noting that metrics that assess averaged quantities such as TT, ADL, LAM and RR, are the least frequent among all the levels, so the two populations do not show any significant differences with respect to these metrics. In contrast, among the most frequent levels, there is the LLV for both conditions. These findings suggest that the two groups of subjects have similar dynamic behaviors over the entire observation time interval, but differ over the maximum time intervals in which the couple of regions are synchronized (or desynchronized). The layer DET resulted also important in classify the two groups during the 0-back condition. This metric is related to the "predictability" of a system, as it indicates the percentage of points of the RP that form diagonal lines. In the bivariate case, a high level of determinism could imply a large amount of time intervals during which two systems are mutually synchronized. Higher determinism in electro-physiological signals was found to be associated with absence of some diseases such as epilepsy and autism (Bosl et al., 2017).

5.2. Significant ROIs

Functional connectivity studies on whole-brain activity or specific seed-based analysis have shown alterations in brain systems such as fronto-temporal networks, cingulo-opercular circuits (implicated in salience processing), default mode network and fronto-parietal networks (involved in high cognitive functions) (Mamah et al., 2013). In addition, several studies have also associated the schizophrenia disease with abnormal configurations of hub regions (van den Heuvel and Fornito, 2014). Most connectivity studies refer to resting state data, aimed at assessing intrinsic functional brain connectivity (Lynall et al., 2010; Alexander-Bloch et al., 2010; Damaraju et al., 2014). Task-based studies have mainly addressed the detection of communities during several tasks and their functional reorganization over time (Ma et al., 2012). Here, the classification analysis revealed the involvement of different brain regions according to the working memory task condition used for classification purposes. Indeed, ROIs better discriminating between controls and patients during the higher working memory load condition are represented by the dorsolateral prefrontal cortex (DLPFC) and by parietal and striatal nodes (see Fig. 6). This brain pattern is coherent with previous investigations highlighting key areas for working memory processing (Linden, 2007) and with previous reports of brain phenotypes associated with schizophrenia diagnosis, both at the structural and functional level (Callicott et al., 1999; Bertolino et al., 2000, 2003). Specifically, DLPFC function appears impaired in schizophrenia. In patients with this brain disorder, DLPFC efficacy in generating and maintaining the neural network underlying working memory function is compromised. Patients show abnormal DLPFC activation while performing working memory tasks (Arnsten, 2011). It is therefore relevant that we found a high information content related with control-patient discrimination in DLPFC regions, such as BA46, BA9, or BA44. Specifically, BA46 is part of the DLPFC, which together with the hippocampus has been reported as a central hub in several models of schizophrenia pathophysiology (Dean and Murray, 2005). In addition, DLPFC activity is strongly related with basal ganglia signaling during cognitive engagement. Striato-prefrontal connectivity patterns have also been associated with schizophrenia (Bertolino et al., 2008). On the other hand, the pattern of brain regions showing higher significance levels in classifying patients from controls based on the 0-back condition reported in Fig. 7 is mainly constituted by the BA9 and by several brain regions of the attentional network involved in attentional control processing. This is coherent with other investigations supporting both the centrality of these regions within this specific task condition (Callicott et al., 1999, 2003; Bertolino et al., 2000, 2003) and differential activity and connectivity in attentional networks between psychosis patients and controls (Blasi et al., 2009; Sambataro et al., 2010). Taken together, we believe that our results are consistent with previous investigation and reveal the brain functional dysregulation of patients with schizophrenia during working memory performance, thus corroborating the hypothesis that the proposed method can capture also nonlinear connectivity patterns reflecting a coherent clinical meaning.

5.3. Prediction of task performance from connectivity

As reported in literature (Barch and Sheffield, 2014), cognitive dysfunction in schizophrenia is reflected in decreased performance in several cognitive domains. Since the task performance wouldn't be fully comparable across cohorts (healthy/patients), a classification bias could be introduced when task-based functional connectivity features are

involved to discriminate between the two cohorts. The correlation analysis between the accuracy scores and the predicted ones, reveals that there are no significant relationships between the accuracy scores and the predicted values using the connectivity features. It is worth noting that the lack of significant predictions of cognitive domain scores from the selected connectivity features could reflect the independence of these features from the performance of the subjects during each task.

6. Limitations

Our study presents some limitations that may influence our findings. First, this study was designed to assess both linear and non-linear dynamics of the neural activity during different cognitive loads. Nevertheless, the BOLD signals result from mixed contributions of different sources like vascular and metabolic activities due to the change of the relative levels of oxyhemoglobin and deoxyhemoglobin, so fMRI is only related to global correlates of neural activity. These properties should be considered in order to interpret the results: even if the proposed framework can effectively classify the two groups of subjects, the most discriminating features cannot be ascribed to a differentiation of local neural properties. Other techniques such as EEG and MEG could result more sensitive to local neural variations, so further analyses may be needed to extend the validity of the framework to other functional connectivity methods. Second, in this study we adopted a cross-validation framework to classify 49 healthy individuals and 42 patients with schizophrenia and extract the most significant connectivity features. Although cross-validation is commonly used to test the models ability to generalize to an independent dataset while avoiding problems like overfitting or selection bias (Cawley and Talbot, 2010), independent tests on an external validation set should be required in order to make the results more robust and to ensure a better generalization of the classification models. Future work will include collecting more data to further strengthen the effectiveness of our results. Third, the current study is conducted using BOLD data during a working memory task. Given the cognitive dysfunction in schizophrenia, the patients' task performance wouldn't be fully comparable to that of healthy individuals, and this could potentially introduce classification bias. In fact, although we have not identified significant predictions of cognitive domain scores from the selected connectivity features, it is possible that several factors, such as awareness or subjects' anxiety levels, may have had an impact on the final performance. The antipsychotic treatment of patients may also influence neuroimaging-related findings during working memory tasks (Schlagenhauf et al., 2008). Our additional correlation analysis between each connectivity feature and PANSS/medication scores reported in SI-Section 4 of Supplementary Information do not show significant connectivity features-symptom scores and connectivity features-medication scores correlations, however a more accurate analysis could highlight specific effects of medication on task-based connectivity. Finally, in this work we compared the classification performance of the proposed multi-recurrence framework and the statistical correlation-based connectivity. The rationale underlying this choice is that here we explored the effectiveness of a phase-space based approach to compare recurrent behaviour of BOLD time series for describing the level of interaction between the physiological activity of the different regions of interest so the simple statistical correlation index represents a “baseline” against which to compare the set of dynamic descriptors. Other methods, based on dynamic connectivity analysis (Allen et al., 2014) have been proven to outperform the static connectivity approach in classifying schizophrenia (Rashid et al., 2016). They mainly aim at capturing the time variability of connectivity, by using several indexes such as those related to the reconfiguration of communities over time. Future work will address a more complete comparison between the proposed framework and dFC methods through the identification of specific dynamic metrics.

7. Conclusion

In this work, we present a framework based on both CRPs and graph analysis to fully describe dynamic functional connectivity. The framework was used to describe dynamic states in fMRI data of two clinical populations: a group of controls and a pathological group of subjects affected by schizophrenia performing a visual-motor and a working memory task. The effectiveness of the method in identifying dynamic states characterizing the two populations during the two conditions was validated with a supervised classification procedure. The results highlight that the multivariate framework can actually reveal brain dynamics that are significantly related to schizophrenia. Such dynamic behaviors are also different from the simple statistical correlation between time series. This study represents an exploratory step in which nonlinear metrics have been evaluated as possible indexes of functional connectivity for prediction of schizophrenia. Further research will be required to fully explore the neurophysiological meaning of fMRI signal complexity through RQA metrics and their possible application in broader clinical contexts.

Conflicts of interest

AB is a stockholder of Hoffmann-La Roche Ltd. He has also received consulting fees from Biogen and lecture fees from Otsuka, Janssen, Lundbeck. GB has received honorary by Lundbeck and Otsuka for lectures. All other authors declare no biomedical financial interests and no potential conflicts of interest.

Ethics statement

All experiments were performed with the informed consent of each participant or caregiver, in line with the Code of Ethics of the World Medical Association (Declaration of Helsinki). Local institutional ethics committees approved the study.

Research data for this article

Data not available/Data will be made available on request. Code available upon request from the corresponding author.

Acknowledgments

This work was partially supported by the Italian Ministry of Health (grant PE-2011-02347951) and by Regione Puglia. This project has received funding from the European Union Seventh Framework Programme for research, technological development and demonstration under grant agreement no. 602450 (IMAGEMEND). GPs position is funded by the European Unions Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement no. 798181 (FLOURISH). This paper reflects only the authors' views and the European Union and Research Executive Agency are not liable for any use that may be made of the information contained therein.

The authors gratefully acknowledge Dr. Tiziana Quarto and Roberta Passiatore for their help with data collection.

Appendix A. Supplementary data

Appendix

1. Recurrence density - based measures. The simplest measure is the *recurrence rate* (RR), a generalization of the cross correlation sum:

$$RR = \frac{1}{N^2}\sum_{i,j=1}^{N} CR_{i,j} \tag{11}$$

2. Measures based on the distribution $P(l)$ of lengths l of the diagonal lines. Among these:

- the *determinism* (DET) is the ratio of the recurrence points that form diagonal structures to all points:

$$DET = \frac{\sum_{l=l_{min}}^{N} lP(l)}{\sum_{l=1}^{N} lP(l)} \tag{12}$$

- The *average diagonal line length* (ADL) is the average time in which two segments of the trajectory move close together:

$$ADL = \frac{\sum_{l=l_{min}}^{N} lP(l)}{\sum_{l=1}^{N} P(l)} \tag{13}$$

- The *length of the longest diagonal line* (LLD) found in the CRP is related to maximal time period in which the two systems are synchronized:

$$LLD = \left(\{l_i\}_{i=1}^{N_l}\right), \tag{14}$$

where N_l is the total number of diagonal lines.

- The *entropy of diagonal length* (EDL) shows the complexity of the diagonal lines in a CRP. It is the Shannon entropy of the probability $p(l)$ to find a diagonal line of length l in the plot:

$$EDL = -\sum_{l=l_{min}}^{N} p(l)\ln p(l) \tag{15}$$

3. Measures based on the distribution $P(v)$ of vertical line lengths v.

- The ratio of recurrence points forming vertical structures to all recurrence points of the CRP is called *laminarity* (LAM):

$$LAM = \frac{\sum_{v=v_{min}}^{N} vP(v)}{\sum_{l=1}^{N} vP(v)} \tag{16}$$

where N_v is the total number of diagonal lines.

- The *average length of vertical lines* (TT) is the trapping time and represents the average time in which the systems are trapped into a specific state:

$$TT = \frac{\sum_{v=v_{min}}^{N} vP(v)}{\sum_{v=1}^{N} P(v)} \tag{17}$$

- The *length of the longest vertical line* (LLV) is analogous to LLD for the vertical lines:

$$LLV = \left(\{v_l\}_{l=1}^{N_v}\right) \tag{18}$$

- From a CRP it is possible to extrapolate the *recurrence times*. Let's consider the recurrence points of the ith row $\{CR_{ij}\}_{j=1}^{N}$ which correspond to the set of points of the trajectory which fall into the ε-neighborhood of an arbitrary chosen point at i. The recurrence times between these recurrence points (*recurrence times of first type*) are:

$$\{RT1_k = j_{k+1} - j_k\}_{k\in\mathbb{N}} \tag{19}$$

Removing all consecutive recurrence points with $RT1_k = 1$ to avoid tangential motion the *recurrence times of second type* are:

$$\left\{RT2_k = j'_{k+1} - j'_k\right\}_{k\in\mathbb{N}} \tag{20}$$

where the set of the remaining recurrence points is used. It turns out that $RT2$ measures the time distance between the beginning of subsequent recurrence structures in the CRP along the vertical direction and it can be considered as an estimate of the average of the lengths of white vertical lines in a column of the plot.

- The *normalized entropy of the recurrence time distribution (RPDE)* of the time series $P(t)$ is defined as:

$$RPDE = -(lnT_{max})^{-1} \sum_{t=1}^{T_{max}} P(t) lnP(t) \tag{21}$$

where T_{max} is the largest recurrence value.

4. Recurrence network - based measures.

- The *clustering coefficient (CLUST)* is the average of the local clustering coefficient which gives the probability that two neighbours (i.e. recurrences) of any state are also neighbours (Marwan et al., 2009):

$$CLUST = \sum_{v=1}^{N} \frac{1}{k_v(k_v-1)} \frac{\sum_{i,j=1}^{N} CR_{v,i} CR_{i,j} CR_{j,v}}{N} \tag{22}$$

where k_v is the number of neighbours of node v.

- The *transitivity (TRANS)* of a CRP is defined as the ratio of the number of close triangles λ_G to the number of subgraphs with 2 edges and 3 vertices τ_G in the network:

$$TRANS = \frac{3\lambda_G}{\tau_G} \tag{23}$$

It can be seen as a quantitative measure of the geometric structural complexity of the trajectories in phase space (Donner et al., 2011).

5. LOS - based measures. In addition to the SYNC metric, other two metrics were defined to quantify the temporal synchronization behaviour of the two systems. In detail:

The *mean period of synchronization (MEAN SYNC)* simply quantifies the mean period during which the two systems are synchronized:

$$MEAN\ SYNC = \frac{1}{N} \sum_{j=1}^{N_d} l_j \tag{24}$$

- *SYNC* is proportional to the ratio of the sum of the lengths of the subsegments l_j along the LOS to the total number of samples N:

$$SYNC = \frac{1}{N} \frac{\sum_{j=1}^{N_d} l_j}{N_d} \tag{25}$$

where N_d is the total number of subsegments.

- *Hard Synchronization (HARD SYNC)*, a binary response variable that gives information about the presence/absence of the LOS:

$$HARD\ SYNC = \begin{cases} 1, & \text{if } CR_{i,i} = 1, \forall i = 1, \ldots, N \\ 0, & \text{otherwise} \end{cases} \tag{26}$$

- The *entropy of LOS (ENTR LOS)* is the Shannon entropy of the probability $p(d)$ to find a subsegment of length d along the LOS and is an index of the complexity of the synchronization periods of the two time series:

$$ENTR\ LOS = -\sum_{d=1}^{N} p(d) \ln p(d) \tag{27}$$

References

Alexander-Bloch, A.F., Gogtay, N., Meunier, D., Birn, R., Clasen, L., Lalonde, F., Lenroot, R., Giedd, J., Bullmore, E.T., 2010. Disrupted modularity and local connectivity of brain functional networks in childhood-onset schizophrenia. Front. Syst. Neurosci. 4, 147.

Allen, E.A., Damaraju, E., Plis, S.M., Erhardt, E.B., Eichele, T., Calhoun, V.D., 2014. Tracking whole-brain connectivity dynamics in the resting state. Cerebr. Cortex 24 (3), 663–676.

Arbabshirani, M.R., Plis, S., Sui, J., Calhoun, V.D., 2017. Single subject prediction of brain disorders in neuroimaging: promises and pitfalls. Neuroimage 145, 137–165.

Arnsten, A.F., 2011. Prefrontal cortical network connections: key site of vulnerability in stress and schizophrenia. Int. J. Dev. Neurosci. 29 (3), 215–223.

Barch, D.M., Sheffield, J.M., 2014. Cognitive impairments in psychotic disorders: common mechanisms and measurement. World Psychiatr. 13 (3), 224–232.

Bassett, D.S., Wymbs, N.F., Porter, M.A., Mucha, P.J., Carlson, J.M., Grafton, S.T., 2011. Dynamic reconfiguration of human brain networks during learning. Proc. Natl. Acad. Sci. Unit. States Am. 108 (18), 7641–7646.

Behzadi, Y., Restom, K., Liau, J., Liu, T.T., 2007. A component based noise correction method (compcor) for bold and perfusion based fmri. Neuroimage 37 (1), 90–101.

Bertolino, A., Esposito, G., Callicott, J.H., Mattay, V.S., Van Horn, J.D., Frank, J.A., Berman, K.F., Weinberger, D.R., 2000. Specific relationship between prefrontal neuronal n-acetylaspartate and activation of the working memory cortical network in schizophrenia. Am. J. Psychiatry 157 (1), 26–33.

Bertolino, A., Sciota, D., Brudaglio, F., Altamura, M., Blasi, G., Bellomo, A., Antonucci, N., Callicott, J.H., Goldberg, T.E., Scarabino, T., et al., 2003. Working memory deficits and levels of n-acetylaspartate in patients with schizophreniform disorder. Am. J. Psychiatry 160 (3), 483–489.

Bertolino, A., Caforio, G., Blasi, G., De Candia, M., Latorre, V., Petruzzella, V., Altamura, M., Nappi, G., Papa, S., Callicott, J.H., et al., 2004. Interaction of comt val108/158 met genotype and olanzapine treatment on prefrontal cortical function in patients with schizophrenia. Am. J. Psychiatry 161 (10), 1798–1805.

Bertolino, A., Fazio, L., Caforio, G., Blasi, G., Rampino, A., Romano, R., Di Giorgio, A., Taurisano, P., Papp, A., Pinsonneault, J., et al., 2008. Functional variants of the dopamine receptor d2 gene modulate prefronto-striatal phenotypes in schizophrenia. Brain 132 (2), 417–425.

Betzel, R.F., Bassett, D.S., 2017. Multi-scale brain networks. Neuroimage.

Betzel, R.F., Mišić, B., He, Y., Rumschlag, J., Zuo, X.-N., Sporns, O., 2015. Functional Brain Modules Reconfigure at Multiple Scales across the Human Lifespan. 1510.08045.

Bianciardi, M., Sirabella, P., Hagberg, G.E., Giuliani, A., Zbilut, J.P., Colosimo, A., 2007. Model-free analysis of brain fmri data by recurrence quantification. Neuroimage 37 (2), 489–503.

Blasi, G., Taurisano, P., Papazacharias, A., Caforio, G., Romano, R., Lobianco, L., Fazio, L., Di Giorgio, A., Latorre, V., Sambataro, F., et al., 2009. Nonlinear response of the anterior cingulate and prefrontal cortex in schizophrenia as a function of variable attentional control. Cerebr. Cortex 20 (4), 837–845.

Bosl, W.J., Loddenkemper, T., Nelson, C.A., 2017. Nonlinear eeg biomarker profiles for autism and absence epilepsy. Neuropsych. Electrophy. 3 (1), 1.

Bosl, W.J., Tager-Flusberg, H., Nelson, C.A., 2018. Eeg analytics for early detection of autism spectrum disorder: a data-driven approach. Sci. Rep. 8 (1), 6828.

Braun, U., Schäfer, A., Walter, H., Erk, S., Romanczuk-Seiferth, N., Haddad, L., Schweiger, J.I., Grimm, O., Heinz, A., Tost, H., et al., 2015. Dynamic reconfiguration of frontal brain networks during executive cognition in humans. Proc. Natl. Acad. Sci. Unit. States Am. 112 (37), 11678–11683.

Bullmore, E.T., Bassett, D.S., 2011. Brain graphs: graphical models of the human brain connectome. Annu. Rev. Clin. Psychol. 7, 113–140.

Bullmore, E., Sporns, O., 2009. Complex brain networks: graph theoretical analysis of structural and functional systems. Nat. Rev. Neurosci. 10 (3), 186.

Calhoun, V.D., Miller, R., Pearlson, G., Adalı, T., 2014. The chronnectome: time-varying connectivity networks as the next frontier in fmri data discovery. Neuron 84 (2), 262–274.

Callicott, J.H., Mattay, V.S., Bertolino, A., Finn, K., Coppola, R., Frank, J.A., Goldberg, T.E., Weinberger, D.R., 1999. Physiological characteristics of capacity constraints in working memory as revealed by functional mri. Cerebr. Cortex 9 (1), 20–26.

Callicott, J.H., Egan, M.F., Mattay, V.S., Bertolino, A., Bone, A.D., Verchinksi, B., Weinberger, D.R., 2003. Abnormal fmri response of the dorsolateral prefrontal cortex in cognitively intact siblings of patients with schizophrenia. Am. J. Psychiatry 160 (4), 709–719.

Cawley, G.C., Talbot, N.L., 2010. On over-fitting in model selection and subsequent selection bias in performance evaluation. J. Mach. Learn. Res. 11 (Jul), 2079–2107.

Cohen, J.R., 2018. The behavioral and cognitive relevance of time-varying, dynamic changes in functional connectivity. Neuroimage 180, 515–525.

Cortes, C., Vapnik, V., 1995. Support vector machine. Mach. Learn. 20 (3), 273–297.

Damaraju, E., Allen, E.A., Belger, A., Ford, J.M., McEwen, S., Mathalon, D., Mueller, B., Pearlson, G., Potkin, S., Preda, A., et al., 2014. Dynamic functional connectivity analysis reveals transient states of dysconnectivity in schizophrenia. Neuroimage: Clinic 5, 298–308.

DATA, D., 1997. Structured Clinical Interview for DSM-IV axis I Disorders. American Psychiatric Press, Washington.

De Domenico, M., 2017. Multilayer modeling and analysis of human brain networks. GigaScience 6 (5), 1–8.

De Domenico, M., Solé-Ribalta, A., Cozzo, E., Kivelä, M., Moreno, Y., Porter, M.A., Gómez, S., Arenas, A., 2013. Mathematical formulation of multilayer networks. Phys. Rev. X 3 (4), 041022.

De Domenico, M., Sasai, S., Arenas, A., 2016. Mapping multiplex hubs in human functional brain networks. Front. Neurosci. 10.

Dean, K., Murray, R.M., 2005. Environmental risk factors for psychosis. Dialogues Clin. Neurosci. 7 (1), 69.

Deisboeck, T., Kresh, J.Y., 2007. Complex Systems Science in Biomedicine. Springer Science & Business Media.

Deng, K., Han, S., Li, K.J., Liu, J.S., 2014. Bayesian aggregation of order-based rank data. J. Am. Stat. Assoc. 109 (507), 1023–1039.

Donner, R.V., Heitzig, J., Donges, J.F., Zou, Y., Marwan, N., Kurths, J., 2011. The geometry of chaotic dynamics–a complex network perspective. 1102.1853.

Donner, R.V., Zou, Y., Donges, J.F., Marwan, N., Kurths, J., 2010. Recurrence networksa novel paradigm for nonlinear time series analysis. New J. Phys. 12 (3), 033025.

DSouza, A.M., Abidin, A.Z., Chockanathan, U., Schifitto, G., Wismüller, A., 2018. Mutual connectivity analysis of resting-state functional mri data with local models. Neuroimage 178, 210–223.

Eskildsen, S.F., Coupé, P., García-Lorenzo, D., Fonov, V., Pruessner, J.C., Collins, D.L., Initiative, A.D.N., et al., 2013. Prediction of alzheimer's disease in subjects with mild cognitive impairment from the adni cohort using patterns of cortical thinning. Neuroimage 65, 511–521.

Fan, L., Li, H., Zhuo, J., Zhang, Y., Wang, J., Chen, L., Yang, Z., Chu, C., Xie, S., Laird, A.R., et al., 2016. The human brainnetome atlas: a new brain atlas based on connectional architecture. Cerebr. Cortex 26 (8), 3508–3526.

Fornito, A., Zalesky, A., Bullmore, E., 2016. Fundamentals of Brain Network Analysis. Academic Press.

Freeman, L.C., 1977. A set of measures of centrality based on betweenness. Sociometry 35–41.

Friston, K.J., Williams, S., Howard, R., Frackowiak, R.S., Turner, R., 1996. Movement-related effects in fmri time-series. Magn. Reson. Med. 35 (3), 346–355.

Gardner, D.M., Murphy, A.L., O'Donnell, H., Centorrino, F., Baldessarini, R.J., 2010. International consensus study of antipsychotic dosing. Am. J. Psychiatry 167 (6), 686–693.

Gonzalez-Castillo, J., Bandettini, P.A., 2018. Task-based dynamic functional connectivity: recent findings and open questions. Neuroimage 180, 526–533.

Guyon, I., Elisseeff, A., 2003. An introduction to variable and feature selection. J. Mach. Learn. Res. 3 (Mar), 1157–1182.

Guyon, I., Weston, J., Barnhill, S., Vapnik, V., 2002. Gene selection for cancer classification using support vector machines. Mach. Learn. 46 (1), 389–422.

Kolde, R., Laur, S., Adler, P., Vilo, J., 2012. Robust rank aggregation for gene list integration and meta-analysis. Bioinformatics 28 (4), 573–580.

Kriegeskorte, N., Simmons, W.K., Bellgowan, P.S., Baker, C.I., 2009. Circular analysis in systems neuroscience: the dangers of double dipping. Nat. Neurosci. 12 (5), 535–540.

Li, X., Wang, X., Xiao, G., 2017. A comparative study of rank aggregation methods for partial and top ranked lists in genomic applications. Briefings Bioinf. 20 (1), 178–189.

Linden, D.E., 2007. The working memory networks of the human brain. Neuroscientist 13 (3), 257–267.

Lombardi, A., Guccione, P., Guaragnella, C., 2016. Exploring recurrence properties of vowels for analysis of emotions in speech. Sensors & Transducers 204 (9), 45.

Lombardi, A., Tangaro, S., Bellotti, R., Bertolino, A., Blasi, G., Pergola, G., Taurisano, P., Guaragnella, C., 2017. A novel synchronization-based approach for functional connectivity analysis. Complexity.

Lynall, M.-E., Bassett, D.S., Kerwin, R., McKenna, P.J., Kitzbichler, M., Muller, U., Bullmore, E., 2010. Functional connectivity and brain networks in schizophrenia. J. Neurosci. 30 (28), 9477–9487.

Ma, S., Calhoun, V.D., Eichele, T., Du, W., Adalı, T., 2012. Modulations of functional connectivity in the healthy and schizophrenia groups during task and rest. Neuroimage 62 (3), 1694–1704.

Mamah, D., Barch, D.M., Repovš, G., 2013. Resting state functional connectivity of five neural networks in bipolar disorder and schizophrenia. J. Affect. Disord. 150 (2), 601–609.

Marmarelis, V.Z., 2004. Nonlinear Dynamic Modeling of Physiological Systems, vol 10. John Wiley & Sons.

Marwan, N., Kurths, J., 2002. Nonlinear analysis of bivariate data with cross recurrence plots. Phys. Lett. 302 (5), 299–307.

Marwan, N., Romano, M.C., Thiel, M., Kurths, J., 2007. Recurrence plots for the analysis of complex systems. Phys. Rep. 438 (5), 237–329.

Marwan, N., Donges, J.F., Zou, Y., Donner, R.V., Kurths, J., 2009. Complex network approach for recurrence analysis of time series. Phys. Lett. 373 (46), 4246–4254.

Oldfield, R.C., 1971. The assessment and analysis of handedness: the edinburgh inventory. Neuropsychologia 9 (1), 97–113.

Olivetti, E., Mognon, A., Greiner, S., Avesani, P., 2010. Brain decoding: biases in error estimation. In: Brain Decoding: Pattern Recognition Challenges in Neuroimaging (WBD), 2010 First Workshop on. IEEE, pp. 40–43.

Onnela, J.-P., Saramäki, J., Kertész, J., Kaski, K., 2005. Intensity and coherence of motifs in weighted complex networks. Phys. Rev. 71 (6), 065103.

Pereda, E., Quiroga, R.Q., Bhattacharya, J., 2005. Nonlinear multivariate analysis of neurophysiological signals. Prog. Neurobiol. 77 (1), 1–37.

Pereira, F., Mitchell, T., Botvinick, M., 2009. Machine learning classifiers and fmri: a tutorial overview. Neuroimage 45 (1), S199–S209.

Power, J.D., Barnes, K.A., Snyder, A.Z., Schlaggar, B.L., Petersen, S.E., 2012. Spurious but systematic correlations in functional connectivity mri networks arise from subject motion. Neuroimage 59 (3), 2142–2154.

Rabinovich, M.I., Varona, P., Selverston, A.I., Abarbanel, H.D., 2006. Dynamical principles in neuroscience. Rev. Mod. Phys. 78 (4), 1213.

Rashid, B., Arbabshirani, M.R., Damaraju, E., Cetin, M.S., Miller, R., Pearlson, G.D., Calhoun, V.D., 2016. Classification of schizophrenia and bipolar patients using static and dynamic resting-state fmri brain connectivity. Neuroimage 134, 645–657.

Sambataro, F., Blasi, G., Fazio, L., Caforio, G., Taurisano, P., Romano, R., Di Giorgio, A., Gelao, B., Bianco, L.L., Papazacharias, A., et al., 2010. Treatment with olanzapine is associated with modulation of the default mode network in patients with schizophrenia. Neuropsychopharmacology 35 (4), 904.
Sauer, T., Yorke, J.A., Casdagli, M., 1991. Embedology. J. Stat. Phys. 65 (3), 579–616.
Schinkel, S., Marwan, N., Kurths, J., 2009. Brain signal analysis based on recurrences. J. Physiol. Paris 103 (6), 315–323.
Schlagenhauf, F., Wüstenberg, T., Schmack, K., Dinges, M., Wrase, J., Koslowski, M., Kienast, T., Bauer, M., Gallinat, J., Juckel, G., et al., 2008. Switching schizophrenia patients from typical neuroleptics to olanzapine: effects on bold response during attention and working memory. Eur. Neuropsychopharmacol. 18 (8), 589–599.
Singhi, S.K., Liu, H., 2006. Feature subset selection bias for classification learning. In: Proceedings of the 23rd International Conference on Machine Learning. ACM, pp. 849–856.
Sporns, O., 2010. Networks of the Brain. MIT press.
Sporns, O., 2011. The human connectome: a complex network. Ann. N. Y. Acad. Sci. 1224 (1), 109–125.
Sporns, O., Tononi, G., Kötter, R., 2005. The human connectome: a structural description of the human brain. PLoS Comput. Biol. 1 (4), e42.
Stam, C.J., 2005. Nonlinear dynamical analysis of eeg and meg: review of an emerging field. Clin. Neurophysiol. 116 (10), 2266–2301.
van den Heuvel, M.P., Fornito, A., 2014. Brain networks in schizophrenia. Neuropsychol. Rev. 24 (1), 32–48.
Van Den Heuvel, M.P., Pol, H.E.H., 2010. Exploring the brain network: a review on resting-state fmri functional connectivity. Eur. Neuropsychopharmacol. 20 (8), 519–534.
Webber Jr., C.L., Marwan, N., 2015. Recurrence Quantification Analysis. Springer.
Webber Jr., C.L., Zbilut, J.P., 2005. Recurrence quantification analysis of nonlinear dynamical systems. In: Tutorials in Contemporary Nonlinear Methods for the Behavioral Sciences, pp. 26–94.
Yan, K., Zhang, D., 2015. Feature selection and analysis on correlated gas sensor data with recursive feature elimination. Sensor. Actuator. B Chem. 212, 353–363.
Zbilut, J.P., Webber, C.L., 1992. Embeddings and delays as derived from quantification of recurrence plots. Phys. Lett. 171 (3–4), 199–203.
Zbilut, J.P., Thomasson, N., Webber, C.L., 2002. Recurrence quantification analysis as a tool for nonlinear exploration of nonstationary cardiac signals. Med. Eng. Phys. 24 (1), 53–60.