Brain functional connectivity in schizophrenia revealed by functional magnetic resonance imaging

Miguel Lourenço Farinha
miguel.lourenco.farinha@tecnico.ulisboa.pt
Instituto Superior Técnico, Lisbon, Portugal

November 2021

Abstract

Brain activity during rest has been demonstrated to evolve through a repertoire of functional connectivity (FC) patterns, whose alterations may provide biomarkers of schizophrenia - a psychotic disorder characterised by dysfunctional brain connectivity. In this study, differences between the dynamic exploration of resting-state networks using functional magnetic resonance imaging (fMRI) data from 71 schizophrenia patients and 74 healthy controls were investigated using a method focusing on the dominant fMRI signal phase coherence pattern at each time point. Through the lens of dynamical systems theory, brain activity in the form of temporal FC state trajectories was examined for intergroup differences by calculating the fractional occupancy, dwell time, limiting probability of each state and the transition probabilities between states. Results showed a medium to large effect size for a globally synchronised state which was found to have decreased fractional occupancy in schizophrenia. The fractional occupancy and limiting probability of FC states overlapping with canonical functional subsystems was identified to have increased in schizophrenia, with small to medium effect sizes. Medium to large effect size for the reduced probability of remaining in a global integrative state and for a number of altered state-to-state transition probabilities were revealed in schizophrenia. Finally, this study showed that using K-medoids clustering did not influence the observed intergroup differences - highlighting the utility of dynamical systems theory to better understand brain activity. Combined, these findings expose pronounced differences between schizophrenia patients and healthy controls - supporting and extending current knowledge regarding disrupted brain dynamics in schizophrenia.

Keywords: Resting-state fMRI, Dynamic functional connectivity, Functional networks, LEiDA, Schizophrenia

1. Introduction

Functional magnetic resonance imaging (fMRI) data can be used to characterise brain activity at "rest" as the time-resolved emergence and dissolution of functionally meaningful networks, in other words, the constant reconfiguration of resting-state functional connectivity (FC) patterns or states over time [5]. These resting-state networks (RSNs) have been extensively analysed across neuroimaging studies [23] and their characterisation in the temporal domain has been suggested to provide potential biomarkers of several disorders [17]. Schizophrenia (SZ) is a chronic brain disorder typified by disruption to thought processes, perception, cognition and behaviours, for which there is still a lack of biomarkers. To date, previous studies investigating dynamic functional connectivity (dFC) have suggested that compared to healthy controls (HCs), patients with SZ spend more time in FC states characterised by weak connectivity [18] and less time in FC states which represent strong, large-scale brain connectivity [4, 6, 21]. Furthermore, when SZ patients transition into the FC state of strongest connectivity, they switch states very rapidly [18]. Overall, SZ patients have been found to exhibit fewer changes between connectivity patterns compared to HCs [13].

Most research on dFC in SZ has been carried out using independent component analysis to extract time courses of networks which were subsequently used to estimate dFC through sliding-window analysis (SWA) [17]. However, the choice of the window length affects the temporal resolution of the SWA approach - raising questions over its validity [17]. In this study, to overcome this weakness, the Leading Eigenvector Dynamics Analysis (LEiDA) method, based on Blood Oxygenation Level Dependent (BOLD) phase coherence, is used to investigate dFC at an instantaneous level [3, 8]. There are three primary aims of this study: (1) assess whether intergroup differences exist by estimating and characterising time courses of recurrent FC states using tools from dynamical systems theory; (2) examine the validity of the partitions resulting from the clustering procedure; (3) understand the influence of using the K-medoids algorithm instead of the K-means algorithm on the ability to differentiate SZ patients from HCs. This work hypothesised to find abnormal dFC in SZ patients characterised by: (1) reduced excursions to a FC state possibly involved in the integration of segregated functional connections; (2) increased excursions to a number of FC states which represent functionally segregated networks.

2. Materials & Methods

This Section outlines the materials and methods utilised in this study.

2.1. Neuroimaging data

Neuroimaging data was obtained from the publicly available repository COBRE preprocessed with NIAK 0.17 - lightweight release [2]. The neuroimaging data included preprocessed resting-state fMRI (rs-fMRI) data from 72 SZ patients and 74 HCs. The rs-fMRI data featured 150 echo planar imaging BOLD volumes obtained in 5 minutes, with repetition time (TR) = 2 s, echo time = 29 ms, acquisition matrix = 64×64 mm^2, flip angle = 75° and voxel size = $3 \times 3 \times 4$ mm^3.

Inspection of the BOLD data for each subject resulted in the exclusion of one subject whose data did not include all 150 BOLD volumes. Therefore, the final dataset used in this analysis included 71 SZ patients (80.28% males) and 74 HCs (68.92% males). Both groups had an age range of 18-65 years old. A two-sided Wilcoxon Rank-Sum test with Bonferroni correction did not identify a significant difference between the mean age of the groups ($p = 0.4253$). The framewise displacement (FD) provided a quantitative indication of each subject's head motion during the scanning period. The same statistical test detected a significant intergroup difference in the group mean FD ($p < 0.001$). Specifically, on average, the BOLD signals of SZ patients were characterised by larger amounts of head motion (FD).

The acquisition and preprocessing of the data are fully described in detail in [2]. Here, contrary to the typical rs-fMRI preprocessing procedure, the neuroimaging data was not subject to spatial smoothing, temporal filtering and nuisance regression.

2.2. Parcellation

Following the methodology used by [3, 7, 10, 24], the canonical Anatomic Automatic Labeling (AAL) template was used to parcellate the entire brain of each participant into 90 cortical and sub-cortical non-cerebellar regions. Accordingly, for each region in the brain, at each time point (TR), the BOLD signals were averaged over all voxels belonging to that brain area to compute the regional BOLD time courses. For each subject, this resulted in a $N \times T$ BOLD dataset, where $N = 90$ is the number of brain areas and $T = 150$ is the number of volumes in each scan.

2.3. Computation of dynamic functional connectivity

To compute the phase relationship between each pair of AAL regions, first the instantaneous phase of the BOLD signals across all brain regions $n \in \{1, \ldots, N\}$ for each time $t \in \{2, \ldots, T - 1\}$, $\theta(n, t)$, were estimated by computing the Hilbert transform of their BOLD regional time courses [8]. Here, the first and last TR of each fMRI scan were excluded due to possible signal distortions induced by the Hilbert trans-

form [24]. The Hilbert transform enables the capture of the time-varying phase of a BOLD signal at each time, t, by converting it into its analytical representation (see Figure 1(a) (top left)) [3, 8]. To obtain a whole-brain pattern of BOLD phase synchrony, the phase coherence between areas n and p at each time t, $dFC(n, p, t)$, was estimated using Equation (1):

$$dFC(n, p, t) = cos(\theta(n, t) - \theta(p, t)) \qquad (1)$$

where phase coherence values range between -1 (areas n and p in anti-phase at time t) and 1 (areas n and p have synchronised BOLD signals at time t), as shown in Figure 1(a) (bottom left). This computation was repeated for all pairwise combinations of brain areas (n, p), with $n, p \in \{1, \ldots, 90\}$, at each time point t, with $t \in \{2, \ldots, 149\}$, and for all subjects. For each subject, the resulting dFC was a three-dimensional tensor with dimension $N \times N \times T'$, where $T' = 148$, i.e., 148 $dFC_{90 \times 90}(t)$ matrices were estimated.

2.4. Functional connectivity leading eigenvector

To characterise the evolution of the phase coherence matrix over time with reduced dimensionality, the current study employed the LEiDA method which considers only the leading eigenvector, $V_1(t)$, of each $dFC(t)$ matrix [3]. In detail, as observed in Figure 1(a) (middle), the leading eigenvector, $V_1(t)$, is a $N \times 1$ vector that captures the dominant connectivity pattern of BOLD phase coherence at time t, i.e., $V_1(t)$ represents the main orientation of the BOLD phases over all brain areas [3]. Under this framework, for each time t, the associated leading eigenvector partitions the N brain areas into two communities by separating the elements with different signs in $V_1(t)$ [3, 14]. When all elements of $V_1(t)$ have the same sign, the BOLD phases between brain regions are coherent, which is indicative of a global mode of phase coherence governing all BOLD signals. This implies that all brain regions belong to the same community. Contrarily, if the elements of $V_1(t)$ have different signs (i.e., positive and negative), the connectivity pattern between brain regions is not coherent. As a result, each brain area is assigned to one of the two communities according to their BOLD phase relationship. Additionally, the absolute value of each element in the leading eigenvector weighs the contribution of each brain area to the assigned community [3, 14]. The dominant FC pattern of the dFC matrix at time t can also be reconstructed back into matrix format by computing the $(N \times N)$ outer product $V_1(t)V_1^{\mathsf{T}}(t)$, as shown in Figure 1(a) (top right). Given that if $V_1(t)$ is a leading eigenvector, so is $-V_1(t)$, following the procedure of [7, 10, 24], it was ensured that most of the elements in $V_1(t)$ had negative values. This is because by assigning positive values to the brain areas whose BOLD phases did not follow the global mode, functional brain networks were distinctly detected, as
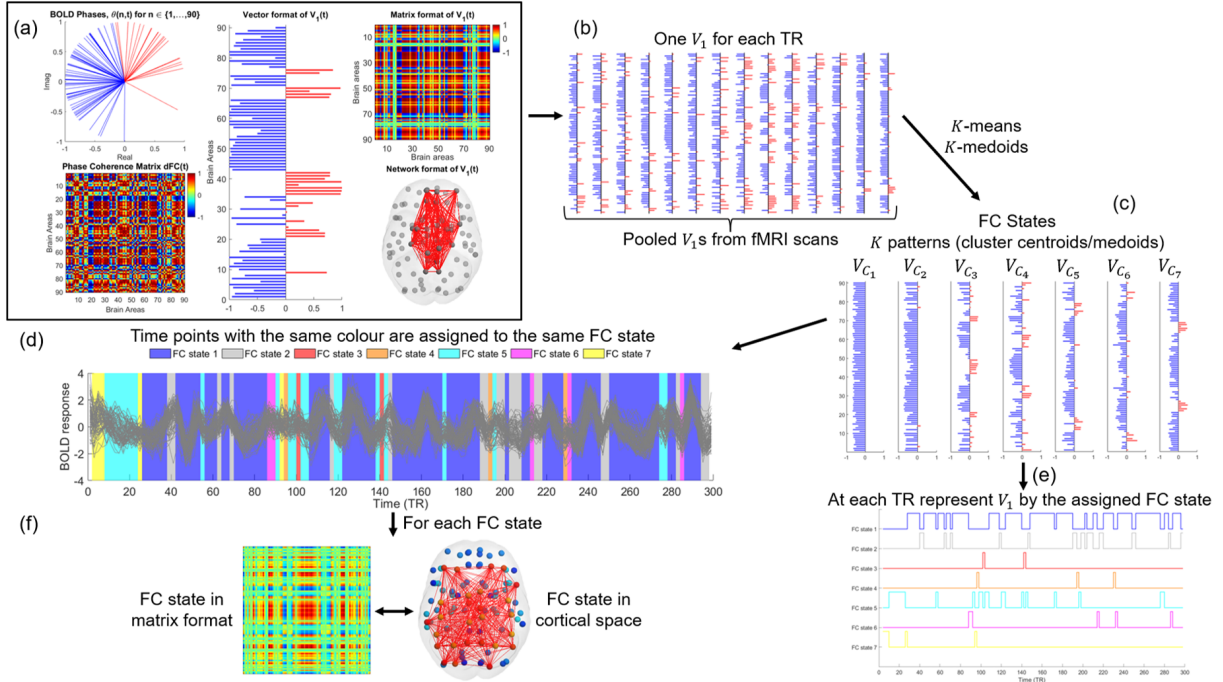
Figure 1: Graphical illustration of the estimation and characterisation of the temporal trajectories of recurrent FC states obtained by using Leading Eigenvector Dynamics Analysis (LEiDA). (a) BOLD phases of all $N = 90$ brain areas in the complex plane at time t (top left); BOLD phase coherence matrix at time t, $dFC(t)$ (bottom left); Vector representation of the leading eigenvector, $V_1(t)$, of $dFC(t)$ (middle); Matrix representation of $V_1(t)$ (top right); Network representation of $V_1(t)$, with links between the areas with positive elements in $V_1(t)$ plotted in red (bottom right). **(b)** The leading eigenvectors are computed for each time point and from all fMRI scans. **(c)** The pooled leading eigenvectors are partitioned into K clusters using a clustering algorithm. The cluster centroids/medoids are assumed to represent recurrent patterns of BOLD phase coherence (FC states). **(d,e)** The leading eigenvector at each TR is represented by the centroid/medoid of the cluster to which it was assigned by the clustering procedure. This originates time courses of FC states for each fMRI session. The time courses are then characterised using tools from dynamical systems theory. **(f)** Each FC state can be represented as a $N \times N$ matrix (outer product) and as a network in cortical space (elements with positive sign linked by red edges).

seen in Figure 1(a) (bottom right). Importantly, this approach was found to explain most of the variance of observed BOLD phase coherence data variation, while substantially reducing its dimensionality. In fact, the leading eigenvector accounted for more than 50% of the variance in phase coherence at all time points and for all subjects.

2.5. Estimation of FC states
Upon computing the leading eigenvector of the phase coherence matrix for each recording frame, the next step in the analysis was to characterise the evolution of the dFC over time by identifying recurrent FC states in the data, as illustrated in Figure 1(c) [3].

Firstly, the conventional LEiDA clustering analysis was performed by applying the K-means algorithm [1] to the dataset of all leading eigenvectors computed at the set of volumes $\{2, \ldots, 149\}$ across all 145 participants (see Figure 1(b)). This corresponded to a combined total of $148 \times 145 = 21460$ leading eigenvectors. Lastly, the impact of using the K-medoids algorithm [1] to conduct a LEiDA analysis was investigated through its application to the combined total of 21460 leading eigenvectors. Here, both algorithms were run with a value of K from 2 to 20, i.e., dividing the set of leading eigenvectors into

$K = \{2, 3, \ldots, 20\}$ clusters. Furthermore, in both clustering analyses, the cosine distance was used as the distance metric for minimisation and the algorithms were run 1000 times to minimise the chances of getting trapped in a local minima [3, 7, 10, 24].

2.6. Characterisation of FC state trajectories
Independently of the algorithm, the LEiDA clustering procedure outputs one distinct clustering solution for each value of K clusters. Specifically, each clustering solution contains K clusters $\mathbf{C} = \{C_1, \ldots, C_K\}$, with $K \in \{2, \ldots, 20\}$ - decomposing the N-dimensional phase space of pooled leading eigenvectors into a K-dimensional state space. Let V_{C_α} be the vector of dimension $N \times 1$ representing the centroid/medoid of cluster C_α, where $\alpha \in \{1, \ldots, K\}$. Then, for a given clustering solution, each V_{C_α} represents a recurrent FC state, as depicted in Figure 1(c) [3].

For each clustering solution, the set of estimated K FC states was used to obtain, for each participant, time courses of FC states (as represented in Figure 1(d,e)). This was accomplished by representing each V_1 at time t by the FC state (centroid/medoid) of the cluster to which it was assigned by the clustering algorithm. Specifically, following the conceptual framework proposed by [24], resting-state BOLD time se-

ries were assumed to temporally evolve through a finite state trajectory of recurrent patterns of BOLD phase coherence. Following this rationale, each clustering solution with K FC states was assumed to define a finite state space $S = \{1, \ldots, K\}$. Furthermore, for a clustering solution with K clusters, the cluster (FC state) to which $V_1(t)$ was assigned at time t, denoted by V_t, was assumed to define a stochastic process, $\{V_t : t \in \{2, \ldots, 149\}\}$, with an associated finite state space given by S. Consequently, considering the Markov property [9] holds, each temporal trajectory of FC states was assumed to define a time-homogeneous Discrete Time Markov Chain (DTMC). Importantly, it must be noted that, although brain activity is an uninterrupted process, the restricted fMRI scanning windows implied the state trajectories were temporally limited - resulting in a number of DTMCs not spanning the entire state space.

A number of descriptive measures were considered to characterise the properties of the temporal trajectories of FC states observed in SZ patients and HCs. Notably, these measures have been shown to provide relevant insights on aberrant dynamic brain activity in previous LEiDA analyses [3, 7, 10].

2.6.1. Fractional occupancy

The fractional occupancy (probability of occurrence) of a FC state α represents the proportion of times V_t is assigned to cluster C_α throughout a scan [24]. The fractional occupancy of FC state α for the fMRI scan of subject s, $P_\alpha^{(s)}$, is calculated (estimation) as follows:

$$P_\alpha^{(s)} = \frac{1}{T'} \sum_{t=1}^{T'} \mathbb{1}_{\left\{V_t^{(s)} \in C_\alpha\right\}} \, , \, \alpha \in \{1, \ldots, K\} \quad (2)$$

where $T' = 148$ is the number of time points (first and last volume of each scan were excluded), $\mathbb{1}$ is the indicator function and $V_t^{(s)}$ is the FC state to which $V_1(t)$ was assigned at time t. For each clustering solution, this property was estimated for each of the K FC states separately for each fMRI scan.

2.6.2. Dwell time

The dwell time (mean duration) of a FC state α represents the mean number of consecutive epochs spent in that state throughout the duration of a scan s [24], $DT_\alpha^{(s)}$, and is defined (estimation) as:

$$DT_\alpha^{(s)} = \frac{1}{k_\alpha^{(s)}} \sum_{d_\alpha=1}^{k_\alpha^{(s)}} R_{d_\alpha} \, , \, \alpha \in \{1, \ldots, K\} \quad (3)$$

where $k_\alpha^{(s)}$ is the number of consecutive periods in which $V_t^{(s)}$ was assigned to cluster C_α and R_{d_α} is the duration of each of the $k_\alpha^{(s)}$ periods. For each clustering solution, the dwell time was estimated for each of the K FC states separately for each fMRI scan.

2.6.3. One-step transition probability matrix

According to [24], considering a clustering solution with state space $S = \{1, \ldots, K\}$, the probability of being in FC state α at time t and transition to FC state β at time $t + 1$ is given by the following expression:

$$P_{\alpha\beta}^{(s)} = \frac{1}{T'-1} \sum_{t=1}^{T'-1} \mathbb{1}_{\left\{V_t^{(s)} \in C_\alpha \, , \, V_{t+1}^{(s)} \in C_\beta\right\}} \quad (4)$$

with $\alpha, \beta \in \{1, \ldots, K\}$. From Equation (4), for a clustering solution with K FC states, it follows that the Transition Probability Matrix (TPM) of the fMRI scan of subject s, $\mathbf{P}^{(s)}$, is defined (estimation) as:

$$\mathbf{P}^{(s)} = P\left[V_{t+1}^{(s)} \in C_\beta \mid V_t^{(s)} \in C_\alpha\right] = \frac{P_{\alpha\beta}^{(s)}}{P_\alpha^{(s)}} \quad (5)$$

with $\alpha, \beta \in \{1, \ldots, K\}$. For the tentative optimal clustering solution, a TPM was estimated separately for the DTMC of each fMRI scan.

2.6.4. Limiting probability

In this study, the limiting distribution was only estimated for irreducible and aperiodic DTMCs [9], with finite state space given by the tentative optimal state trajectories. Therefore, for every subject, s, with a DTMC satisfying the aforementioned criteria, it followed that:

$$\lim_{t \to \infty} \left(P_{\alpha\beta}^{(s)}\right)^t = \pi_\beta^{(s)} > 0 \, , \, \alpha, \beta \in S \quad (6)$$

where the estimate of the row vector denoting the stationary distribution of the DTMC [9], $\boldsymbol{\pi}^{(s)} = \left[\pi_\beta^{(s)}\right]_{\beta \in S}$, with dimension $1 \times |S|$, is given by:

$$\boldsymbol{\pi}^{(s)} = \mathbf{1} \times \left(\mathbf{I} - \mathbf{P}^{(s)} + \mathbf{ONE}\right)^{-1} \quad (7)$$

where $\mathbf{1}$ is a $1 \times |S|$ vector of ones, $\mathbf{I}$ is the identity matrix with rank $|S|$, $\mathbf{P}^{(s)}$ is the TPM of subject s and $\mathbf{ONE}$ is a $|S| \times |S|$ matrix all of whose entries are one. Due to the inclusion criteria imposed on the DTMCs defined by the optimal state trajectories, i.e., irreducibility and aperiodicity, only 37 and 46 DTMCs from the HC and SZ groups, respectively, were analysed. For a given FC state β, π_β (element β of the row vector $\boldsymbol{\pi}$) was the measure to be used to perform intergroup comparisons. Importantly, since only aperiodic DTMCs were considered, π_β can be understood as the limiting probability that the DTMC is in FC state β and as the long-run fraction of time the DTMC spends in FC state β. It must be noted that intergroup comparisons between the estimated stationary distributions were not performed in this study.

2.7. Intergroup comparisons

In this research, hypothesis tests to compare the group mean of the properties (Sections 2.6.1 to 2.6.4)

4

calculated from the temporal state trajectories observed in SZ patients and HCs were performed using Monte Carlo permutation tests [16] by adapting the procedure used by [3, 7, 10]. To produce an accurate approximate estimation of the permutation distribution, these tests were conducted using 100000 permutations. Here, depending on the result of a Levene's test (used to assess the homogeneity between the group variances) the Monte Carlo permutation tests were performed based on one of the following two statistics (under the null hypothesis):

$$
T_0^* = \begin{cases} \dfrac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_{\text{HC}}} + \frac{S_2^2}{n_{\text{SZ}}}}}, & p_{\text{Levene's test}} < 0.05 \\[3ex] \dfrac{\bar{X}_1 - \bar{X}_2}{\sqrt{S_{\text{p}}\left(\frac{1}{n_{\text{HC}}} + \frac{1}{n_{\text{SZ}}}\right)}}, & \text{otherwise} \end{cases} \quad (8)
$$

where $\bar{X}_1$ and $\bar{X}_2$ are the random sample means, S_1 and S_2 are the random sample standard deviations and n_{HC} and n_{SZ} are the sample sizes for the HC and SZ groups, respectively. The pooled random standard deviation, S_{p}, is given by $S_{\text{p}} = \left((n_{\text{HC}} - 1)S_1^2 + (n_{\text{SZ}} - 1)S_2^2\right)/\left(n_{\text{HC}} + n_{\text{SZ}} - 2\right)$. In each hypothesis test, the same statistic was used throughout all data permutations to ensure the consistency of the method.

2.8. Comparison to resting-state functional networks
The functional relevance of the estimated FC states was investigated by assessing whether there was a significant spatial overlap between their centroids/medoids and any of the seven reference RSNs defined by [22]. This was accomplished by employing the procedure used by [10, 24]. Specifically, the seven RSNs were transformed into seven non-overlapping vectors with dimension 1×90, where each entry of the vectors corresponded to the proportion of voxels of each AAL brain area that were assigned to each of the seven RSNs. Finally, the Pearson correlation coefficient was used to assess the spatial overlap between these seven RSNs and the centroids/medoids V_{C_α} with $\alpha \in \{1, \ldots, K\}$ (all negative values of V_{C_α} were set to zero so that only areas thought to define relevant functional networks were considered).

2.9. Validation of clustering solutions
This Section describes the methods used to validate clustering results.

2.9.1. Unsupervised internal cluster validation criteria
The quality of clustering solutions outputted by the clustering algorithms was evaluated using the average Silhouette coefficient and the Dunn's index [1].

2.9.2. External validation clustering agreement measures
Clustering outputs from distinct algorithms were compared using the Adjusted Rand Index (ARI) and the Variation of Information (VI) clustering agreement measures [1].

2.9.3. Clustering stability evaluated by K-fold cross-validation
The stability of clustering solutions was assessed according to a 10-fold cross-validation procedure adapted from [11]. Firstly, the sample of the pooled leading eigenvectors was split into two subsamples, referred to as training and test samples. Secondly, a clustering algorithm was applied to the training sample - yielding partition P_1. Subsequently, a Nearest Centroid classifier assigned each observation of the test sample to the cluster of partition P_1, whose centroid was nearest - resulting in the class set P_2 of the test sample. The same clustering algorithm was then applied to the test sample - producing the cluster set P_3. Finally, partitions P_2 and P_3 were compared based on the ARI, VI and percent agreement (fraction of objects correctly assigned). This procedure was repeated for each of the 10 cross-validation folds.

2.10. Software
This analysis used MATLAB R2019b [12], the Statistics and Machine Learning Toolbox™ and the Econometrics Toolbox™.

3. Results
3.1. Intergroup differences across partition models detected by using the K-means algorithm
The collection of clustering solutions was investigated to search for FC states whose fractional occupancy and dwell time most significantly and consistently differed between SZ patients and HCs. For a partition model with K clusters, K hypothesis tests were performed. Consequently, to account for the increased probability of false positives, the significance threshold was adjusted to $\alpha_2 = 0.05/K$ using a Bonferroni correction. Additionally, a conservative significance threshold of $\alpha_3 = 0.05/\sum_{K=2}^{20} K$ was considered to encompass both dependent and independent null hypotheses across partition models.

Figure 2(b) presents, for each clustering solution, the K two-sided p-values obtained from evaluating whether the group mean fractional occupancy of a FC state differed between SZ patients and HCs. From the inspection of Figure 2(b), it is apparent that, across all partition models, the clustering procedure consistently returned FC states whose mean fractional occupancy differed significantly between groups - falling below the corrected significance thresholds α_2 and α_3.

Closer inspection of Figure 2(b) shows there are significant intergroup differences in the mean fractional occupancy of FC state 1 for a range of clustering solutions ($p < \alpha_3$, two-tailed tests). In fact, the mean fractional occupancy of this state was found to be significantly decreased in SZ patients compared to HCs ($p < \alpha_3$ for $K \in \{2, \ldots, 18\}$, one-tailed tests), as suggested in Figure 2(a). Interestingly, for all partition models, the centroid associated with FC state 1 revealed this recurrent FC pattern represents a global mode of BOLD phase coherence (all elements of the
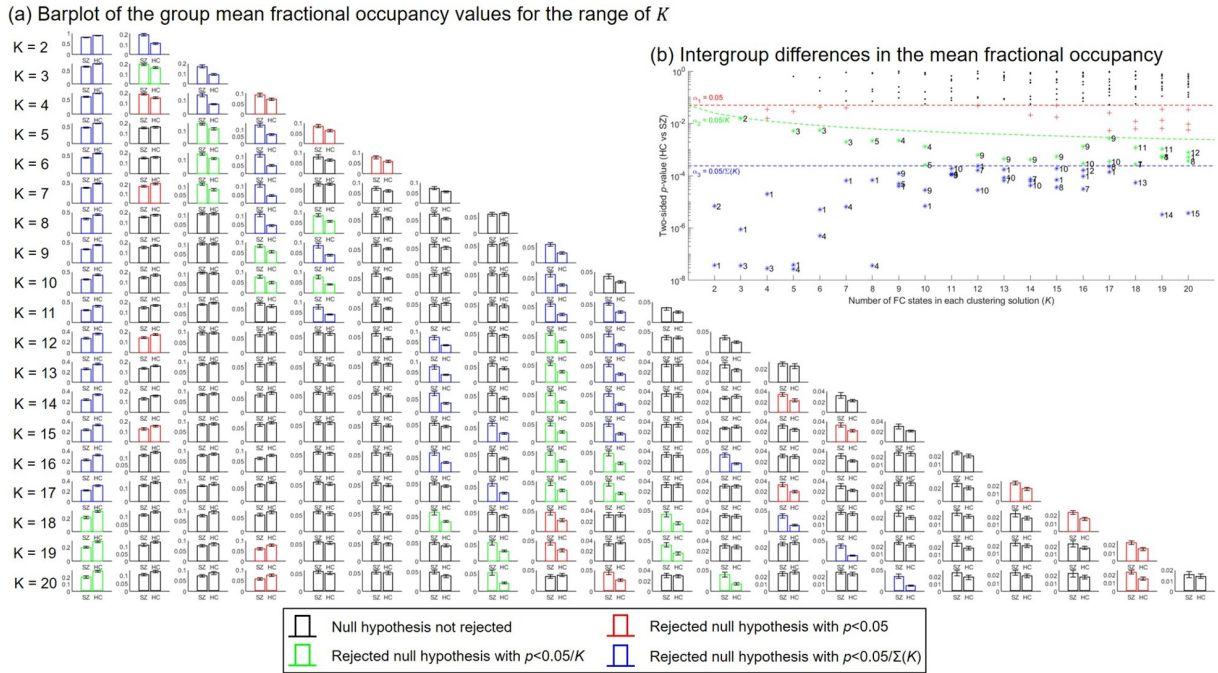
(a) Barplot of the group mean fractional occupancy values for the range of K

(b) Intergroup differences in the mean fractional occupancy

Figure 2: Detection of FC states with mean fractional occupancy significantly different between SZ patients and HCs. (a) Barplot of the estimated mean fractional occupancy with associated standard error of each FC state for each group. For each FC state, the colour of the bars indicates whether the null hypothesis of no intergroup differences in the mean fractional occupancy was rejected (two-tailed tests). The standard error of each bar was calculated as the standard deviation of the sample data divided by the square root of the sample size. **(b)** Two-sided p-values obtained for the intergroup comparisons of the mean fractional occupancy of each FC state for each partition model. FC states (clusters) are ranked according to their estimated fractional occupancy, where cluster 1 consisted of the largest number of objects and cluster K consisted of the least number of objects.

centroid had the same sign). Hence, FC state 1 is referred to as the Global Mode.

As depicted in Figure 2(b), across all clustering solutions, further non-global FC states are characterised by significant intergroup differences in the group mean fractional occupancy ($p < \alpha_3$, two-tailed tests). Interestingly, all these states were typified by a significantly higher mean probability of occurrence in the SZ group compared to the HC group ($p < \alpha_3$, one-tailed tests), as presented in Figure 2(a). Visual inspection of these non-global FC states revealed they represent varying forms of the same underlying connectivity patterns. Specifically, states detected for lower values of K could be obtained by combining the fine-grained FC patterns identified in partition models with larger values of K - evidencing the dependence among the hypothesis tests performed across clustering solutions.

The analysis of mean dwell time estimates of detected FC states suggested that this measure did not allow as much consistent and clear differentiation between groups compared to the estimates of the fractional occupancy of FC states. In fact, the mean dwell time of the Global Mode was reduced significantly in SZ patients compared to HCs in only 7 clustering solutions ($p < \alpha_3$, one-tailed tests). Conversely, across all partition models, the mean dwell time was identified as significantly increased in the SZ group compared to the HC group in only two FC states

($p < \alpha_3$, one-tailed tests). Notably, these non-global states were highly correlated (Pearson's $r = 0.996$) - reinforcing the fact that significant intergroup differences were consistently detected across similar FC patterns.

3.2. Overlap with reference functional networks
Investigation of the overlap between the centroids of the detected FC states and the seven canonical functional networks defined by [22] confirmed that intergroup differences were consistently detected in a number of varying forms of the same FC patterns. Interestingly, FC state 1 did not significantly overlap with any of the seven reference RSNs - indicating this global state did not reveal the activation of any particular subset of functionally coupled brain regions. Additionally, across partition models, the non-global states with a significantly increased mean fractional occupancy (and mean dwell time) in SZ were found to repeatedly overlap with the reference Somatomotor, Dorsal Attention and Limbic networks. Accordingly, FC states with functional activity possibly related to that of the aforementioned canonical RSNs recur more often (and last for larger consecutive periods of time) in SZ patients compared to HCs.

3.3. Internal validation of K-means clustering solutions
As shown in Figure 3, the highest average Silhouette coefficient and Dunn's index were obtained for clustering solutions with a low number of FC states,

6

which are of limited interest for the present study. Contrarily, for clustering solutions with more than 12 clusters, both validation measures remained relatively constant at low values - indicating such partitions are also of limited interest.
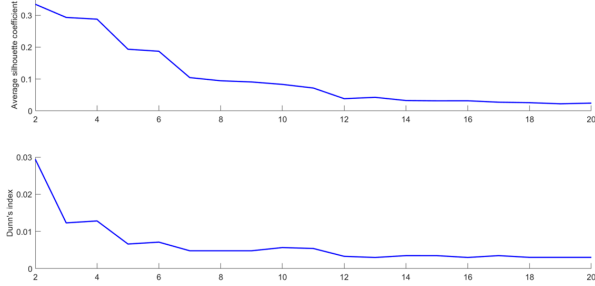

Figure 3: Internal validation of K-means clustering results. Average Silhouette coefficient and Dunn's index used to evaluate the quality of clustering solutions.
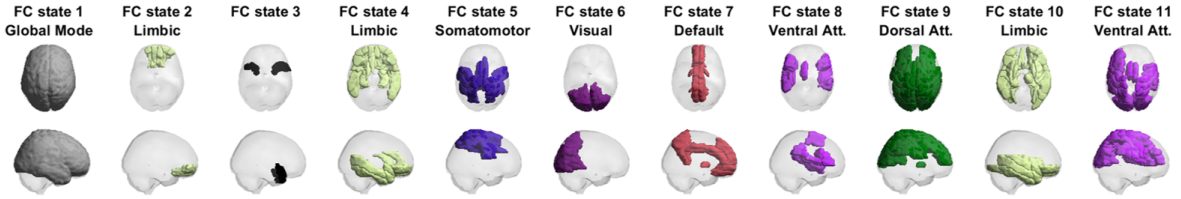
Notably, for clustering solutions with K between 7 and 11, the average Silhouette coefficient decreased smoothly and the Dunn's index remained approximately constant, as observed in Figure 3 - suggesting these partition models are of potential interest for further analysis.

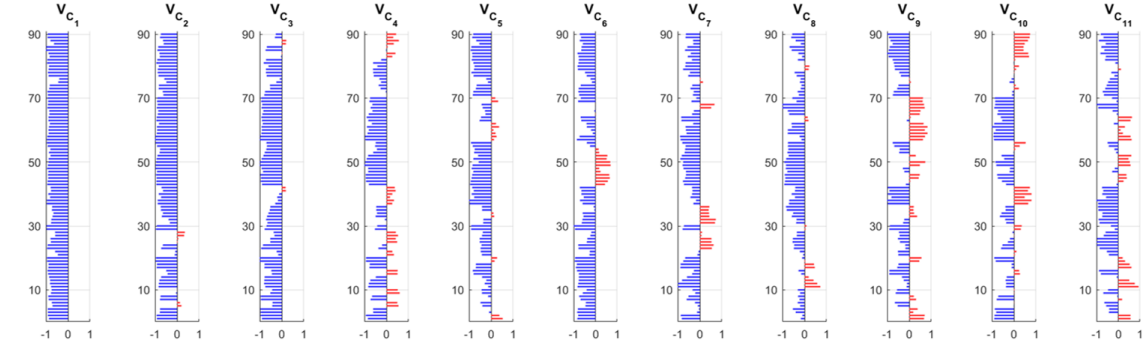## 3.4. Selection of the optimal clustering solution

For the subsequent analysis, the partition model with 11 FC states was selected as the optimal K-means clustering solution. This decision was based upon the ability to identify a collection of FC states with properties that significantly differed between groups and the quality and stability of the actual partition of the data.

The collection of 11 BOLD phase coherence patterns and their fractional occupancy values are presented in Figure 4. The non-global FC states were found to significantly correlate with six of the seven RSNs estimated by [22], as shown in Figure 4(a). From Figure 4(b), it is apparent that the 11 FC states represent BOLD phase coherence between distinct subsets of brain areas. Furthermore, significant intergroup differences were identified in the mean fractional occupancy of 4 FC states, as observed in Figure 4(c). Closer inspection of Figure 4(a) and (b) revealed that these 4 states represent distinct functionally meaningful networks. The mean fractional occupancy of FC state 1 was significantly decreased in SZ patients compared to HCs (Hedge's $g = 0.694$, medium to large effect size), with $29.8 \pm 19.8\%$ and $39.8 \pm 14.9\%$ (estimated mean $\pm$ std) for the SZ
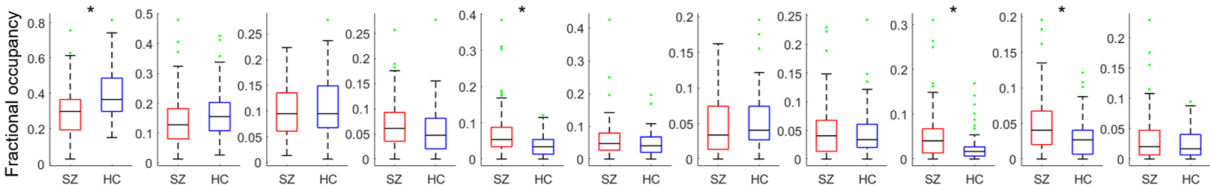

Figure 4: Repertoire of BOLD phase coherence states obtained from the selected optimal clustering solution with $K = 11$ clusters. The FC states are arranged (left-to-right) according to decreasing estimated fractional occupancy. Each FC state is represented by a centroid V_{C_α}, with $\alpha \in \{1, \dots, 11\}$, of size $N \times 1$. **(a)** Cortical rendering of all brain areas with positive values in V_{C_α}. The functional network defined by [22] with which V_{C_α} most significantly overlapped is indicated as subtitle. **(b)** Vector representation showing the N elements in V_{C_α}, representing the contribution of each brain area to FC state α. **(c)** Boxplot of the fractional occupancy values for each FC state for the SZ and HC groups. Asterisks indicate significant intergroup differences ($p < \alpha_3$, one-tailed tests). Green points represent outliers, according to the Tukey criterion.

and HC groups respectively. Furthermore, the mean fractional occupancy of FC states 5, 9 and 10 was significantly increased in SZ patients compared to HCs (Hedge's $g = \{0.611, 0.630, 0.629\}$, respectively, medium to large effect size), with $7.61 \pm 7.78\%$, $5.67 \pm 6.34\%$ and $5.14 \pm 4.09\%$ for the SZ group and $4.02 \pm 3.15\%$, $2.52 \pm 3.24\%$ and $2.94 \pm 2.81\%$ for the HC group, respectively (estimated mean $\pm$ std).

3.5. Stability of the optimal clustering solution

The minimum, maximum and average values of the percent agreement, ARI and VI obtained across the 10 cross-validation folds are provided in Table 1. These values indicate respectively, good levels of association and paired agreement between partitions of the test sample and that the amount of information that was lost in changing from the class set P_2 to the cluster set P_3 of the test sample was relatively low. Therefore, the optimal clustering solution is considered valid and appropriate for further analysis.

Table 1: **Stability analysis of the optimal clustering solution.** Minimum, maximum and average value obtained across the 10 cross-validation folds for each clustering agreement measure.

Indices	Minimum	Maximum	Average
Percent agreement	0.523	0.698	0.591
Adjusted Rand	0.678	0.773	0.720
Variation of information	1.181	1.504	1.336

3.6. Transition probability matrix of the optimal state trajectories

With respect to the optimal clustering solution, for all participants, the individual DTMC defined by the temporal trajectories through the finite state space $S' = \{1, \ldots, 11\}$, was characterised by its estimated TPM. For each probability of transitioning from state α to state β ($\alpha \rightarrow \beta$, $\alpha, \beta \in S'$), a two-sided p-value was obtained by testing whether its group mean differed between groups. The state-to-state transition probabilities that were significantly affected in SZ patients compared to HCs are depicted in Figure 5.

As shown in Figure 5, the mean probability of remaining in FC state 1 was significantly reduced in SZ patients compared to HCs (Hedge's $g = 0.726$, medium to large effect size). Furthermore, the mean probability of remaining in FC state 7 was significantly reduced in the SZ group compared to the HC group (Hedge's $g = 0.515$, medium effect size). Lastly, the mean probability of transitioning from FC state 1 to FC states 5 and 10 (Hedge's $g = \{0.452, 0.513\}$, respectively, small to medium effect size) and from FC state 9 to FC state 10 (Hedge's $g = 0.461$, small to medium effect size) was significantly increased in SZ patients compared to HCs.
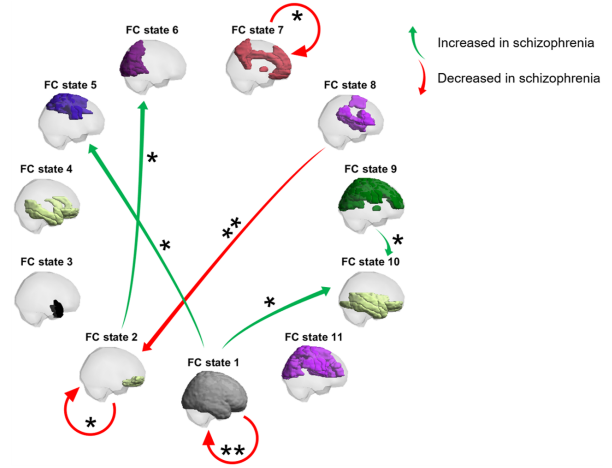

Figure 5: **Transition diagram of the state-to-state transitions significantly altered in SZ patients.** Arrows represent a mean transition probability that was significantly increased (green) or decreased (red) in SZ patients compared to HCs. Single and double asterisks indicate, respectively, significant intergroup differences with $p < 0.05/11$ and $p < 0.05/(11 \times 11)$ (one-tailed tests).

3.7. Limiting probabilities of the optimal FC states

For the subgroup of 37 HCs and 46 SZ patients with irreducible and aperiodic DTMCs, the estimated mean long-run proportion of TRs spent in FC state 1 was, respectively, 0.309 ± 0.104 and 0.272 ± 0.111 (mean $\pm$ std). Surprisingly, no intergroup differences were found in the mean limiting probability of this state (two-tailed test; Hedge's $g = 0.342$, small to medium effect size). Only the mean limiting probability of FC states 5 and 10 was identified as significantly increased in the SZ subgroup compared to the HC subgroup ($p < 0.05$, one-tailed tests; Hedge's $g = \{0.464, 0.449\}$, respectively, small to medium effect size).

3.8. Influence of using the K-medoids algorithm instead of the K-means algorithm

The application of the K-medoids algorithm was found to enable the detection of FC states with a mean fractional occupancy and a mean dwell time that consistently and significantly differed between groups. Similarly with the findings produced by the K-means algorithm, the K-medoids algorithm identified a FC state which represented a globally BOLD synchronised pattern whose mean fractional occupancy was significantly decreased in SZ patients. Additionally, the mean fractional occupancy of a number of non-global FC states related to the reference Somatomotor, Dorsal Attention and Limbic RSNs was found to be significantly increased in SZ patients. Although the algorithms outputted dissimilar clustering solutions, for each partition model, the FC states detected by each of the algorithms with significant intergroup differences in the mean fractional occupancy and mean dwell time ($p < \alpha_2$, two-tailed tests) were found to be highly correlated. Therefore, both clustering algorithms were found to effectively identify sim-

ilar FC states whose properties provide the capacity to differentiate SZ patients from HCs.

4. Discussion

This study investigated differences between the resting-state dFC observed in SZ patients and HCs.

Across partition models, a globally BOLD synchronised state recurs less in SZ patients. This finding is in line with those of previous studies using different approaches to investigate dFC [4, 18, 21]. Conversely, a number of non-global FC states recur more often in SZ patients, as shown in [4, 13, 18]. These non-global states, which have been previously referred to as "ghost" attractor states [5, 24], were found to have functional relevance related to that of the Somatomotor, Dorsal Attention and Limbic reference RSNs defined by [22]. Notably, despite the lack of a full understanding of the relationship between connectivity patterns observed in Electroencephalography (EEG) and fMRI, these findings could be speculated to portray a temporal dynamics related to that observed with EEG microstates measured at a different time resolution. In fact, in line with the aforementioned findings, EEG studies have reported an increased recurrence of a microstate associated with the Limbic RSN in SZ patients compared to HCs [19]. However, the same study showed that the recurrence of a microstate associated with the Attention RSN was decreased in SZ patients [19]. This differed from the findings presented here.

From the analysis of the optimal clustering solution, it was found that when in the Global Mode, SZ patients are more likely not to remain in that state in the next time instant, in concordance with findings from previous studies [18]. Importantly, this more frequently occurring pattern of global BOLD phase coherence has been linked to greater neural flexible switching via integration or segregation of different functional connections [3, 15]. Therefore, the reduced ability of SZ patients to access and remain in this state could be hypothesised to provoke reduced brain flexibility, i.e., the dynamical state trajectories of SZ patients will, to a greater extent, stay restricted to a fixed set of "ghost" attractor states rather than flexibly transitioning through the full collection of FC states enabled by the recurrent transitions into the functionally integrative Global Mode. This hypothesis seemed to be consistent with previous research which found whole-brain integration of higher-order networks was impaired in SZ [6]. Lastly, SZ patients were found to transition more likely than HCs from the Global Mode to FC states 5 and 10. It can therefore be suggested that the decreased recurrence of FC state 1 resulted from the increased propensity of SZ patients to transition to networks related to the Somatomotor (state 5) and Limbic (state 10) RSNs. This could have in turn led to their increased recurrence (and limiting proba-

bility) - supporting the hypothesis of reduced brain integration of segregated functional connections in SZ.

One interesting finding is that, compared to HCs, SZ patients switch more frequently from a FC state related to the Dorsal Attention RSN (state 9) to a FC state related to the Limbic RSN (state 10). This finding is consistent with that of EEG studies which reported unexpectedly more transitions from a microstate associated with the Attention RSN to a microstate associated with the Limbic RSN [20]. Furthermore, SZ patients present a decreased ability to remain in a FC state functionally related to the Default RSN. Notably, this RSN has been linked to core processes of human cognition [23]. Therefore, this finding may be related to the cognitive symptoms observed in patients with schizophrenia. Additional intergroup differences were detected in a number of other state-to-state transition probabilities. It may be the case therefore that the observed abnormal dynamical state transitions provide potential biomarkers of this disorder.

On the question of the influence of using the K-medoids algorithm to conduct a LEiDA analysis, this study found that similar intergroup differences are captured by employing either the K-medoids algorithm or the K-means algorithm. This finding suggests that the choice of an optimal clustering algorithm should rely not only on statistical and cluster validation analyses, but also on concepts and methods from dynamical systems theory [3, 5, 24]. On the one hand, from the definition of the K-means algorithm, the detected FC states are not necessarily observations from the input dataset, but could rather be interpreted as averaged recurrent unobserved FC patterns; hence their designation as "ghost" attractor states [5, 24]. However, the definition of the K-medoids algorithm implied the detected FC states are observed recurrent FC patterns. Research questions pertaining to the functional meaning of the detected FC patterns underline the need to employ tools from dynamical systems theory to provide further insights into the dynamical regime of brain activity [3, 5, 24].

5. Conclusions & Future work

Resting-state dynamic functional connectivity comparisons were conducted between SZ patients and HCs by employing and extending the LEiDA method. Through the characterisation of the temporal expression of different FC patterns, this study found that SZ patients exhibit a reduced capacity to access and remain in a globally BOLD synchronised state. An implication of this is the possibility that, even in the absence of any explicit task, SZ patients transition more frequently to network patterns that are commonly activated during specific tasks.

This study proposed examining the limiting probability of FC states. However, considerably more re-

search will need to be conducted to determine the utility of this measure. Furthermore, one shortcoming of this study which could have affected the measurements of both the state-to-state transition and state limiting probabilities was the low temporal resolution of the data. Future work should utilise data with higher temporal resolution to improve the utility of these properties as possible biomarkers of SZ. Secondly, this study did not take into account the effect of variables such as age, gender and clinical history of patients while assessing intergroup differences - attributing intergroup differences only to the effect of the group. Further research is required to determine whether these variables or their interaction could explain the variability found between groups. Another issue that was not addressed in this work was whether not applying nuisance regression strategies influenced the LEiDA method and therefore, the observed intergroup differences. Further research on this question could contribute with valuable insights into this controversial preprocessing step. In this study, intergroup comparisons were performed by utilising permutation tests. Further work might alternatively consider bootstrap tests to assess intergroup differences. Finally, this study provides unbiased and statistically rigorous evidence for differences between SZ patients and HCs. Nevertheless, the potential to serve as biomarkers of SZ and the clinical implications of the results derived herein remain to be analysed. Future investigations should gather a diverse panel of experts to explore how these findings could be applied to improve our understanding of SZ.

References

[1] C. C. Aggarwal and C. K. Reddy. *Data Clustering: Algorithms and Applications*. Chapman & Hall/CRC, 1st edition, 2013.

[2] P. Bellec. COBRE preprocessed with NIAK 0.17 - lightweight release, November 2016.

[3] J. Cabral, D. Vidaurre, P. Marques, R. Magalhães, P. S. Moreira, J. M. Soares, G. Deco, N. Sousa, and M. L. Kringelbach. Cognitive performance in healthy older adults relates to spontaneous switching between states of functional connectivity during rest. *Scientific reports*, 7(1):1–13, 2017.

[4] E. Damaraju, E. Allen, A. Belger, J. Ford, S. McEwen, D. Mathalon, B. Mueller, G. Pearlson, S. Potkin, A. Preda, J. Turner, J. Vaidya, T. van Erp, and V. Calhoun. Dynamic functional connectivity analysis reveals transient states of dysconnectivity in schizophrenia. *NeuroImage: Clinical*, 5:298–308, 2014.

[5] G. Deco and V. K. Jirsa. Ongoing cortical activity at rest: Criticality, multistability, and ghost attractors. *Journal of Neuroscience*, 32(10):3366–3375, 2012.

[6] D. Dong, M. Duan, Y. Wang, X. Zhang, X. Jia, Y. Li, F. Xin, D. Yao, and C. Luo. Reconfiguration of Dynamic Functional Connectivity in Sensory and Perceptual System in Schizophrenia. *Cerebral Cortex*, 29(8):3577–3589, 10 2018.

[7] C. A. Figueroa, J. Cabral, R. J. T. Mocking, K. M. Rapuano, T. J. van Hartevelt, G. Deco, P. Expert, A. H. Schene, M. L. Kringelbach, and H. G. Ruhé. Altered ability to access a clinically relevant control network in patients remitted from major depressive disorder. *Human Brain Mapping*, 40(9):2771–2786, 2019.

[8] E. Glerean, J. Salmi, J. M. Lahnakoski, I. P. Jääskeläinen, and M. Sams. Functional magnetic resonance imaging phase synchronization as a measure of dynamic functional connectivity. *Brain Connectivity*, 2(2):91–101, 2012. PMID: 22559794.

[9] V. G. Kulkarni. *Introduction to Modeling and Analysis of Stochastic Systems*. Springer Texts in Statistics. Springer-Verlag New York, 2 edition, 2011.

[10] L.-D. Lord, P. Expert, S. Atasoy, L. Roseman, K. Rapuano, R. Lambiotte, D. J. Nutt, G. Deco, R. L. Carhart-Harris, M. L. Kringelbach, and J. Cabral. Dynamical exploration of the repertoire of brain networks at rest is modulated by psilocybin. *NeuroImage*, 199:127–142, 2019.

[11] M. C. Martins and M. G. M. S. Cardoso. Evaluation of clusters of credit cards holders. *Revista de Ciências da Computação*, 3(3):1–11, 2008.

[12] MATLAB. *MATLAB version 9.7.0.1190202 (R2019b))*. The MathWorks Inc., Natick, Massachusetts, 2019.

[13] R. L. Miller, M. Yaesoubi, J. A. Turner, D. Mathalon, A. Preda, G. Pearlson, T. Adali, and V. D. Calhoun. Higher dimensional meta-state analysis reveals reduced resting fmri connectivity dynamism in schizophrenia patients. *PLOS ONE*, 11(3):1–24, 03 2016.

[14] M. E. J. Newman. Finding community structure in networks using the eigenvectors of matrices. *Phys. Rev. E*, 74:036104, Sep 2006.

[15] J. S. Nomi, S. G. Vij, D. R. Dajani, R. Steimke, E. Damaraju, S. Rachakonda, V. D. Calhoun, and L. Q. Uddin. Chronnectomic patterns and neural flexibility underlie executive function. *NeuroImage*, 147:861–871, 2017.

[16] F. Pesarin and L. Salmaso. *Permutation tests for complex data: theory, applications and software*. John Wiley & Sons, 2010.

[17] M. G. Preti, T. A. Bolton, and D. Van De Ville. The dynamic functional connectome: State-of-the-art and perspectives. *NeuroImage*, 160:41–54, 2017. Functional Architecture of the Brain.

[18] L. Rabany, S. Brocke, V. D. Calhoun, B. Pittman, S. Corbera, B. E. Wexler, M. D. Bell, K. Pelphrey, G. D. Pearlson, and M. Assaf. Dynamic functional connectivity in schizophrenia and autism spectrum disorder: Convergence, divergence and classification. *NeuroImage: Clinical*, 24:101966, 2019.

[19] J. N. Ramos da Cruz, O. Favrod, M. Roinishvili, E. Chkonia, A. Brand, C. Mohr, P. Figueiredo, and M. Herzog. Eeg microstates are a candidate endophenotype for schizophrenia. *Nature Communications*, 11(3089):1–11, 2020.

[20] K. Rieger, L. Diaz Hernandez, A. Baenninger, and T. Koenig. 15 years of microstate research in schizophrenia – where are we? a meta-analysis. *Frontiers in Psychiatry*, 7:22, 2016.

[21] L. Sanfratello, J. M. Houck, and V. D. Calhoun. Dynamic functional network connectivity in schizophrenia with magnetoencephalography and functional magnetic resonance imaging: Do different timescales tell a different story? *Brain Connectivity*, 9(3):251–262, 2019. PMID: 30632385.

[22] B. T. Thomas Yeo, F. M. Krienen, J. Sepulcre, M. R. Sabuncu, D. Lashkari, M. Hollinshead, J. L. Roffman, J. W. Smoller, L. Zöllei, J. R. Polimeni, B. Fischl, H. Liu, and R. L. Buckner. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, 106(3):1125–1165, 2011. PMID: 21653723.

[23] M. P. van den Heuvel and H. E. Hulshoff Pol. Exploring the brain network: A review on resting-state fmri functional connectivity. *European Neuropsychopharmacology*, 20(8):519–534, 2010.

[24] J. Vohryzek, G. Deco, B. Cessac, M. L. Kringelbach, and J. Cabral. Ghost attractors in spontaneous brain activity: Recurrent excursions into functionally-relevant bold phase-locking states. *Frontiers in Systems Neuroscience*, 14:20, 2020.