

# Functional connectivity analysis of cerebellum using spatially constrained spectral clustering

Vasileios C. Pezoulas, *Student Member, IEEE*, Kostas Michalopoulos, *Member, IEEE*, Manousos A. Klados, *Member, IEEE*, Sifis Micheloyannis, Nikolaos G. Bourbakis, *Fellow, IEEE*, and Michalis Zervakis, *Senior member, IEEE*

**Abstract**— The human cerebellum contains almost fifty percent of the neurons in the brain, although its volume does not exceed ten percent of the total brain volume. The goal of this study is to derive the functional network of the cerebellum during resting-state and then compare the ensuing group networks between males and females. Towards this direction, a spatially constrained version of the classic spectral clustering algorithm is proposed and then compared against conventional spectral graph theory approaches, such as, spectral clustering, and N-cut, on synthetic data as well as on resting-state fMRI data obtained from the Human Connectome Project (HCP). The extracted atlas was combined with the anatomical atlas of cerebellum resulting to a functional atlas with 46 Regions of Interest (ROIs). As a final step, a gender-based network analysis of cerebellum was performed using the data-driven atlas along with the concept of the Minimum Spanning Trees. The simulation analysis results confirm the dominance of the spatially constrained spectral clustering approach in discriminating activation patterns under noisy conditions. The network analysis results reveal statistically significant differences in the optimal tree organization between males and females. In addition, the dominance of the Left VI lobule in both genders supports the results reported in a previous study of ours. To our knowledge, the extracted atlas comprises the first resting-state atlas of cerebellum based on HCP data.

**Keywords**— cerebellum, resting-state fMRI, spatially constrained spectral clustering, gender, minimum spanning trees

## I. INTRODUCTION

Resting state functional magnetic resonance imaging (rs-fMRI) has emerged as a valuable technique for mapping the functional organization of the brain. The identification of the intrinsic functional networks from fMRI data is important for understanding the connections and interactions of different brain regions. Functional connectivity is usually measured using the Pearson's correlation coefficient between pairs of voxels or regions of interests (ROIs). In particular, a mean Blood-oxygen-level dependent (BOLD) signal is calculated for each Region of Interest (ROI) and correlations are pairwise calculated among ROIs. An open problem with this approach is

the meaningful definition of such ROIs. A common methodology is to define the specific regions for analysis manually using anatomical images, in combination with prior knowledge of the anatomical connectivity of the region. This approach is quite popular although it has several drawbacks. One of these is that a region defined anatomically may not be functionally homogeneous. It has been observed that adjacent regions could have significantly different connectivity patterns, although being part of the same anatomical region [1, 2].

In a voxel-based analysis, the signals from all voxels are used to extract the corresponding pairwise correlations. We can distinguish two major approaches for the analysis of resting-state functional networks. In the first approach, the recorded fMRI time series is correlated to the time course of a seed voxel/ROI, which is usually determined in a separate experiment. Different studies have used this approach to examine the resting state network related to motor, visual, and auditory stimuli. The hypothesis is that during resting state the fMRI oscillations of functionally related regions are correlated [3, 4]. A dataset is collected while the subjects perform a task that is expected to activate the functionally related areas (for example, motor, visual, memory tasks). The data are analyzed to identify the significantly activated areas and the voxels with the highest activity are identified. These voxels are used as seed voxels in the related resting state dataset. The correlation between the time courses of all the voxels and the seed voxel is calculated and used to generate a statistical map. This implies that the studies are limited by the capability to accurately determine the seed regions during the activation study. The effectiveness of this technique also depends on the selected seed voxel and whether it is a good exemplar of the related voxels during the resting state. Finally, the selection of a seed voxel is also influenced by noise and other artifacts of the data collection process (e.g., head motion, physiological oscillations).

The other approach considers region or voxel-wise correlations and attempts to extract the temporal and spatial patterns during the resting state activity. It is based on the analysis of the intrinsic hemodynamic oscillations at rest and

Date submitted: 1/29/2018. The dataset was provided [in part] by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University. The corresponding author (V. C. Pezoulas) has received a scholarship of excellence from the Pancretan Endowment Fund for this research work.

Vasileios C. Pezoulas was with the School of Electrical and Computer Engineering, Technical University of Crete, Chania, 73100 Greece. He is now with the Unit of Medical Technology and Intelligent Information Systems,

Department of Materials Science and Engineering, University of Ioannina, Ioannina 45110 GR (email: bpezoulas@gmail.com).

Kostas Michalopoulos and Nikolaos G. Bourbakis are with the Center of Assistive Research Technologies, Wright State University, Dayton, OH 45435 USA (email: michalopoulos.2@gmail.com, nikolaos.bourbakis@wright.edu).

Michalis Zervakis is with the School of Electrical and Computer Engineering, Technical University of Crete, Chania 73100 Greece (e-mail: michalis@display.tuc.gr).

Manousos A. Klados is with the School of Life & Health Sciences, Aston University, Birmingham, UK (e-mail: m.klados@aston.ac.uk).

Sifis Micheloyannis, deceased, was with the School of Medicine, University of Crete, Heraklion 71003 Greece.

the construction of functional networks between distant brain regions [1, 5]. These networks can be analyzed using graph theoretical approaches to uncover global network characteristics of nodes of importance. The accurate identification of the nodes of the network is very important for the actual interpretation of the graph analysis [1, 6]. Several methods have been proposed to uncover and characterize the resting state networks, including Principal Component Analysis (PCA) [7], Independent Component Analysis (ICA) [8-10], as well as clustering methods, including hierarchical and spectral clustering [11, 12]. These are model free methods that allow us to explore the connectivity patterns without the need of a priori selected seed regions.

There are several inherent difficulties with this type of model-free analysis. Most importantly, it is not straightforward how to transfer information from the subject-level functional parcellations to the group-level. The main difficulty lies on the inherent variability between subjects that produces different solutions and nodes for each subject. Various solutions have been proposed to address this problem. In decomposition techniques like ICA, the bold time-series are concatenated across subjects to derive the group-level [11] decomposition. A common approach for clustering algorithms is to either generate the average affinity matrix for the group and apply the clustering algorithm on the average matrix [11] or extract the individual clusters and perform a secondary group-wise clustering over the individual maps [12]. In this paper, we are using the group-average affinity matrix approach to form our group results. Averaging over the subject level affinity matrices removes part of the subject variability and enhances the common connectivity patterns. Nevertheless, selecting the correct number of clusters plays a crucial role and the stability of the solution is not always guaranteed. In our study we use a cross-validation approach to determine the number of clusters based on the reliability of the produced solutions.

Another difficulty of the voxel-wise clustering approach is that we have to incorporate spatial constraints to the estimated solution. Without such constraints, there is no guarantee that the resulting clusters will be spatially coherent. In most cases, extensive post-processing or additional constraints are required to extract the final set of functional nodes [11]. To alleviate this problem, either the number of clusters is adjusted until a spatially contiguous result is achieved or spatially disconnected clusters are split into more coherent groups [11].

Few methods directly incorporate a spatially coherent methodology. Recently, a spatially constrained version of the Normalized cut (N-CUT) algorithm has been proposed [11]. The N-CUT algorithm and the spectral-clustering family of algorithms have shown to outperform other algorithms in terms of clustering quality and robustness to outliers, especially compared to region growing techniques [11]. Spectral graph theoretic approaches have been employed in several resting-state fMRI studies for the investigation of brain's connectivity networks [12, 13]. More specifically, the spectral clustering concept was applied in [13] in order to investigate the topological organization of brain's subnetworks modules in Attention Deficit Hyperactivity Disorder (ADHD) subjects, combined with the network's entropy where the silhouette method was employed for determining the appropriate number of clusters (e.g., Regions of Interest).

The methodology presented in [11] is a spatially constrained version of the N-cut algorithm and it was used to generate a whole brain fMRI atlas. The proposed methodology constraints the voxel-wise affinity matrix based on the spatial distance between the voxels, by setting the correlation of distant voxels to zero. In this paper, we proceed one step further and examine the effect of the spatial constraint on the spatial coherence of the resulting clusters in the group affinity matrix. Towards this direction, we apply the constrained spectral clustering algorithm to a simulated dataset to demonstrate the limitations of the standard spectral clustering algorithm in the application of functional parcellation and we studied the spatial coherence of the resulting solutions.

One other parameter of interest in brain organization studies is the selected number of clusters. We study the effect of the spatial constraint to the reliability of the group solution by expanding on the procedure in [13]. We construct the group affinity matrix by taking the average over the subject level connectivity matrices. To test the effect of the spatial constraint on the reliability of the solution, we are using a Leave-One-Out cross-validation scheme and we calculate the optimum number of clusters for each fold using different spatial constraints.

After the exploratory stages, we apply this methodology on an fMRI dataset for the functional parcellation of the cerebellum. The human cerebellum is a structure located underneath the cerebral hemispheres. It contains almost fifty percent of the neurons in the brain, although its volume does not exceed ten percent of the total brain volume. The cerebellum has been associated mainly with the planning and execution of movement. Recent studies though have linked many areas of the cerebellum to language, emotion, and attention networks. Most studies have examined the functional connectivity of the cerebellum in relation to other cortical areas, although the cerebellum networks present a significantly different organization from the cortical ones [14-17].

Although most studies have focused on the functional association of cerebellum and cerebral regions, few studies have explored the resting state functional networks within the cerebellum. For example, the authors in [18] investigated the associations of gender and IQ (Intelligent Quotient) with cerebellum using methods from Graph Theory in order to construct functional connectivity graphs and further compute graph theoretical measures. The implication of schizophrenia in cerebellum's functional connectivity has been also examined using a voxel-wise approach for constructing graphs [19, 20]. In addition, several resting-state cortico-cerebellar studies in the literature take into account the relationship of cerebellum with cortical and subcortical brain regions using anatomical and self-organizing map methods [21].

The major goal of our study is to derive the functional network of the cerebellum using a spatially constrained spectral clustering approach and compare the derived group networks between genders. This approach is different from a previous study of ours [18] where cerebellum's functional connectivity was assessed based on a pre-defined anatomical atlas. Our results reveal significant differences in the optimal tree configurations between males and females, as well as, a left hemispheric dominance in cerebellum's functional connectivity, highlighting the importance of lobule Left VI during information transfer within the cerebellum.

## II. SUBJECTS

Our dataset was collected from the Human Connectome Project (HCP), an open-source database aiming to provide a deep examination of the human brain connectome [22]. The HCP is the result of efforts of co-investigators from the University of California, Los Angeles, Martinos Center for Biomedical Imaging at Massachusetts General Hospital (MGH), Washington University, and the University of Minnesota. Resting-state fMRI datasets were selected from 100 healthy subjects (50 males; 50 females) after the first large HCP S500 and MEG2 data release and prepared for further analysis. The HCP investigators collected 91 MRI volumes/slices per subject, with the following scanner parameters: TR = 720 ms, TE = 33.1 ms, FA = 52, FOV = 208x180 mm, slice thickness = 2.0 mm; 72 slices; 2.0 mm isotropic voxels, multiband factor = 8, echo spacing = 0.58 ms, bandwidth = 2290 Hz/Px.

## III. HCP ACQUISITION PROTOCOL

HCP investigators have already performed fundamental preprocessing pipelines on the rsfMRI data [22]. More specifically, two MR functional pipelines were applied and are briefly mentioned in the sequel. The first one, the fMRI volume pipeline, removes spatial distortions, realigns volumes using FSL's FLIRT-based motion correction, normalizes the intensity of 4D images to a global mean, registers data into MNI space and finally masks the data with the final brain mask derived from FreeSurfer segmentation. The second pipeline, namely the fMRI surface pipeline, aims to transform the time-series from volume space to CIFTI grayordinates standard space with 2mm average surface vertex and subcortical volume spacing. Surface data were smoothed using a 2mm full width at half maximum (FWHM) kernel. Independent Component Analysis (ICA) was also applied using FSL's MELODIC tool with automatic dimensionality estimation (please see [22] for further details).

## IV. CEREBELLUM'S PARCELLATION AND FURTHER PREPROCESSING

Whole cerebellum's parcellation procedure was performed individually per subject through a standard cerebellum anatomical atlas template which was already registered in the MNI152 space and was originally developed by Diedrichsen and colleagues [23]. The anatomical atlas contains 28 Regions of Interest (ROIs) which were combined in order to examine the whole atlas. BOLD rs-fMRI time-series were extracted from 21522 cerebellar voxels in total and each voxel's time-course was bandpass filtered with a 50<sup>th</sup> order finite impulse response (FIR) bandpass filter that has been designed to meet the specifications set by the HCP signals ( $f_{c,1} = 0.01$  Hz;  $f_{c,2} = 0.1$  Hz; stopband attenuation 1 and 2 = 60 dB; passband ripple = 1 dB; fs = 1.38 Hz, which is equal to 1/TR, where TR = 720 ms) applied in zero phase mode so as to eliminate low frequency noise not implicated in resting-state functional connectivity (i.e. slow scanner drifts and influences of higher frequencies reflecting possible cardiac or respiratory oscillations) [12].

## V. VOXEL-WISE FUNCTIONAL GRAPH CONSTRUCTION

Pearson's zero-lag temporal correlation coefficient was employed as a functional connectivity measure and was

computed for every possible pair of voxels time-series. The result is a 21522x21522 correlation matrix (adjacency matrix) per subject, where the element  $(i, j)$  represents the correlation between the filtered BOLD time-series of voxels  $i$  and  $j$ . In fact, each correlation matrix can be modeled as a graph  $G = (V, E)$  where  $V$  is the set of nodes and  $E$  is the set of edges. The nodes are defined as the voxel locations and the edges reflect the correlation values (weights) between each pair of nodes. In this study, weighted and undirected graphs were constructed, where the negative correlations were discarded. Finally, the average correlation matrix across all 100 subjects was computed in order to apply the proposed clustering procedures.

## VI. CLUSTERING APPROACHES

### A. Spectral Clustering

Our main goal here is to develop an appropriate method to identify the cerebellar resting-state networks (RSNs), which will be subsequently used for extracting the functional ROIs the cerebellum. We are using Spectral clustering to organize the voxels in functional clusters based on their fMRI signal. In this context, cerebellar voxels are organized in clusters based on the similarity of their correlation patterns. Specifically, given an  $m \times m$  similarity matrix  $X$ , our goal is to partition  $X$  into  $k$ -subsets. First, we define the Laplacian matrix, say  $L$ , as follows:

$$L = D - X \quad (1)$$

where  $D$  is an  $m \times m$  diagonal matrix with the degree of each node (i.e. the number of neighbors connected to that node) on the main diagonal. In fact, it is mathematically proven [24] that the optimal partitioning of  $X$  can be found by solving the generalized eigen-problem of its Laplacian matrix:

$$L = UAU^{-1} \quad (2)$$

where  $A$  is an  $k \times k$  diagonal matrix with the eigenvalues on the main diagonal (sorted in descending order),  $U$  is an  $m \times k$  matrix with the largest  $k$ -eigenvectors stacked in columns and  $U^{-1}$  is the Moore-Penrose pseudoinverse matrix of  $U$  which can be computed by applying Singular Value Decomposition (SVD) on  $U$  [25]. After the computation of the eigenvectors matrix  $U$ , a  $k$ -means approach is applied in order to cluster the first (largest)  $k$ -eigenvectors  $u_1, u_2, \dots, u_k$  of  $U$  into clusters  $C_1, C_2, \dots, C_k$ . Then, each voxel (point in 3D space), say  $x_i, i = 1, 2, \dots, m$ , is assigned to cluster  $Y_j, j = 1, 2, \dots, k$ , if and only if the corresponding row of  $U$ , say  $y_i$ , belongs to cluster  $C_j$ :

$$Y_j = \{x_i \mid y_i \in C_j\} \quad (3)$$

### B. The N-cut algorithm

A counterpart approach for graph partitioning, which is also based on spectral graph theory, is the Normalized cut (N-cut) algorithm which was originally introduced by Shi and Malik [24]. The goal of the N-cut algorithm is to partition a graph  $G$  into a pre-specified number of N-subsets (clusters) by cutting the edges that connect these subsets so that the similarity between the voxels of each cluster is smaller than the similarity within the voxels of different subsets. This method utilizes the

similarity matrix associated with the graph to derive a cost function for each partition. Assume that we want to partition  $\mathbf{G}$  into two disjoint sets  $\mathbf{A}$  and  $\mathbf{B}$ , where  $\mathbf{A} \cup \mathbf{B} = \mathbf{V}$ . To achieve this, the algorithm minimizes the following cut cost function:

$$\text{cut}(\mathbf{A}, \mathbf{B}) = \sum_{x_i \in \mathbf{A}, x_j \in \mathbf{B}} w_{ij} \quad (4)$$

where  $x_i, x_j$  are voxels in clusters  $\mathbf{A}, \mathbf{B}$ , respectively and  $w_{ij}$  denotes the similarity between voxels with indices  $i$  and  $j$  (i.e.  $(i, j)$ -th correlation coefficient in the similarity matrix  $\mathbf{X}$ ). A common problem related with this cost function however is that it may result to single voxel clusters. To address this problem, the Normalized cut cost is defined as:

$$Ncut(\mathbf{A}, \mathbf{B}) = \frac{\text{cut}(\mathbf{A}, \mathbf{B})}{\sum_{x_i \in \mathbf{A}, x_k \in \mathbf{V}} w_{ik}} + \frac{\text{cut}(\mathbf{A}, \mathbf{B})}{\sum_{x_j \in \mathbf{B}, x_k \in \mathbf{V}} w_{jk}} \quad (5)$$

The normalization factor introduced by the N-cut algorithm forces the cut in a direction to minimize the similarity between clusters and at the same time maximize the similarity within the clusters. The optimal partitioning of  $\mathbf{X}$  can be found by solving the generalized eigen-problem of its Laplacian matrix  $\mathbf{L}$  [24].

### C. Spatially Constrained Spectral Clustering

A common problem with the spectral clustering methods is the fact that they may create non-contiguous clusters, i.e. clusters with its voxels being spatially distributed. A solution to this problem is offered by introducing a spatial constraint in the algorithmic procedure [11]. More specifically, instead of applying the spectral clustering procedure on the original correlation matrix, the idea is to multiply it with a distance matrix, say  $\mathbf{S} = (s_{ij})_{i,j=1,2,\dots,m}$ , which can be computed using a distance function (i.e. the Euclidean distance). In order to examine the effect of the spatial constraint in the clustering procedure, different thresholds can be applied on  $\mathbf{S}$  to derive the binary thresholded matrix  $\mathbf{S}_{rc}$ :

$$\mathbf{S}_{rc} = \begin{cases} 1, & s_{ij} \leq rc \\ 0, & s_{ij} > rc \end{cases} \quad (6)$$

where  $rc$  is the threshold parameter. Finally, we define:

$$\mathbf{W}_{rc} = \mathbf{S}_{rc} \circ \mathbf{X} \quad (7)$$

where  $\circ$  denotes element-wise multiplication with  $\mathbf{X}$ , which is a non-negative similarity matrix (i.e.  $x_{ij} \geq 0, \forall i, j = 1, 2, \dots, m$ ). By applying the spectral clustering concept on  $\mathbf{W}_{rc}$  instead of  $\mathbf{X}$ , we can control the number of neighboring voxels per cluster and therefore perform a spatially constrained version of the original spectral clustering algorithm. The appropriate  $rc$  value is selected based on a proposed cross-validation procedure which is described on the Results section.

### D. Clustering Evaluation

To evaluate the clustering solutions produced by the algorithms we used the Silhouette and the Davies-Bouldin indices to measure the compactness and separation of the

clusters. The silhouette index compares the similarity of a voxel with all other voxels in its own cluster to the voxels in all the other clusters [26].  $SI(i)$  ranges between -1 and 1, with high silhouette values denoting that the voxel is well matched to the voxels in  $\mathbf{A}$  and poorly matched to the voxels in the rest of clusters. We compute the average silhouette index across the voxels of a cluster and use this measure as a clustering evaluation metric.

The Davies-Bouldin index [27] is another well-known clustering evaluation metric. A good clustering scheme is characterized by a large separation between clusters and a small within-cluster scatter. The optimal number of clusters is the one with the smallest Davies-Bouldin index.

To evaluate the homogeneity of the ROIs in each segmentation, we use a sliding 3-D sphere in the resulting cluster maps from both clustering approaches to compute the Shannon entropy of the voxels' labels within the sphere, until the entire number of cerebellum voxels is examined. More specifically, for each voxel of interest, we first compute a vector of its neighboring voxels' clustering indices, say  $\mathbf{y}$ , and afterwards estimate the probability distribution of  $\mathbf{y}$ , say  $\mathbf{p}_y$ . Then, Shannon's entropy is computed as follows:

$$E(\mathbf{p}_y) = - \sum_{j=1}^n p_y(j) \cdot \log(p_y(j)) \quad (8)$$

where  $n$  is the number of different clustering indices in  $\mathbf{y}$ . Finally, the average entropy is used as a ROI homogeneity quantifier. Small entropy values indicate better ROI cohesion.

We quantify the consistency between each pair of cluster maps (i.e. 3D matrices with clustering labels or atlases) among all runs using the Dice similarity coefficient. The Dice coefficient measures the similarity between a pair of cluster maps. For two cluster maps,  $\mathbf{A}$  and  $\mathbf{B}$ , the Dice similarity coefficient is defined as the ratio of twice the number of voxels common to both cluster maps divided by the total number of voxels in both cluster maps [28]:

$$\text{Dice} = \frac{2 \cdot |\mathbf{A} \cap \mathbf{B}|}{|\mathbf{A}| + |\mathbf{B}|} \quad (9)$$

leading to values between 0 and 1, where the latter corresponds to a perfect correspondence between two cluster maps.

## VII. GENDER-BASED NETWORK ANALYSIS

A gender-based analysis of cerebellum's network was also performed for each subject separately, based on the functional atlas of cerebellum that was obtained from the application of the SCSC procedure on the average correlation matrix with the appropriate  $k$  and  $rc$  values (which are determined through the proposed cross-validation procedure and the average is over the entire population).

### A. Graph construction

Recall that a graph  $\mathbf{G}$  can be defined as a pair  $(\mathbf{V}, \mathbf{E})$  where  $\mathbf{V}$  is a set of vertices and  $\mathbf{E}$  is a set of edges. Cerebellum's functional connectivity was assessed by computing Pearson's correlation coefficients between each pair of the functional

ROIs, leading to an  $n \times n$  correlation (adjacency) matrix per subject, where  $n$  is now the number of functional ROIs. Negative correlations were excluded from further analysis, with the elimination of the corresponding edge.

### B. Minimum Spanning Trees

An alternative way of modeling a weighted graph  $G$  is based on the notion of the Minimum Spanning Tree (MST), which is an acyclic (loopless) subgraph connecting all nodes of  $G$  [29]. A spanning tree is a connected subgraph of the original graph with  $n$ -nodes and exactly  $n-1$  edges [30, 31]. An MST is a spanning tree that manages to preserve only the edges that minimize the total cost defined as the sum of weights of its edges. The MST is a straightforward method that overcomes biases introduced by comparing networks with a different number of edges and eliminates disconnected syndromes within the network [32, 33]. Here, MSTs were constructed using Kruskal's scheme [34]. Nevertheless, since we are only interested in evaluating the strongest connections, e.g. the edges with the largest weights [31, 35], the edge weights were defined as the inverse of functional connectivity estimates (or  $1/\text{correlation coefficient}$ ). The outcome forms an acyclic subgraph that maximizes the total weight and, thus includes the strongest connections of MST. In total, 100 MSTs were computed based on the above procedure. These MSTs were finally categorized into two equally populated gender-based groups (50/50 males/females) with the aim of detecting and characterizing hubs within all ROIs as well as testing for gender differences.

### C. Local and global MST measures

To assess the functionality of each MST, three nodal and six global metrics were recruited. For a given node  $i \in V$ , where  $V$  is the set of  $n$ -nodes, the local MST metrics are estimated as features of hubs [30]. Degree (DEG),  $k_i$ , is the number of neighbors connected to the  $i$ -th node. Betweenness centrality (BC),  $BC_i$ , defines the number of shortest paths between any two nodes  $k$  and  $l$  in the network passing through  $i$ , divided by the total number of shortest paths between these two nodes [36]. A BC value of zero indicates a leaf node, whereas a value of one indicates a central node in a star network. The node with the highest BC value plays an important role in the overall network's communication [37], since it has the highest load and reflects the highest number of shortest paths between any two nodes that run through this node [35, 38]. The eccentricity of node  $i$ , is the longest shortest path from node  $i$  to any other node in the MST. A node with low eccentricity is more central in a tree and therefore eccentricity acts as a measure of the central topological organization of a network [37, 39].

Six global measures are used in order to evaluate the global structure and utility of each MST. Leaf number,  $N_{leaf}$ , is the number of nodes with degree equal to one. It has a lower bound of 2 and an upper bound of  $n - 1$ . However, it is often useful to compute the fraction of leaf nodes,  $L_f = N_{leaf} / (n - 1)$  in the MST [31, 38]. Degree correlation,  $r$ , is an index of whether the degree of a node is influenced by the degree of its neighboring connected vertices [35] and is equal to the Pearson correlation coefficient of the MST's degree sequence (known as Pearson degree correlation). Diameter,  $d$ , is the maximum eccentricity (longest shortest path) of the nodes and has an

upper bound of  $n - N_{leaf} + 1$ . A small diameter leads to better global communication between remote nodes [37]. Radius,  $r$ , is defined as the smallest eccentricity in the tree, with small radius reflecting more central topology. Kappa or degree divergence,  $\kappa$ , is a measure of the broadness of the degree distribution [30, 38] that is related to resilience against attacks, epidemic spreading and the synchronizability (i.e. information flow) of tree nodes [31, 39]. High  $\kappa$  values suggest the existence of high degree nodes in the tree, mostly related to scale-free networks [39]. Tree hierarchy,  $T_h$ , is a hierarchical metric that quantifies the balance between diameter reduction and overload prevention [35].  $T_h$  approaches zero for a line-like topology and for a star-like topology  $T_h$  approaches 0.5. It varies in-between for intermediate networks. A desirable tree configuration is characterized by a combination of small diameter (i.e., short distances as in a star-like topology) and prevention of the overload of the central tree nodes (hubs). The latter corresponds to  $T_h$  values around 0.5.

### D. Hub analysis

Nodes with high BC and DEG values are characterized as critical nodes (hubs) and are used to determine the information flow within the network. In order to detect the most important nodes (hubs), the total number of males and females exhibiting the maximum DEG, BC values, was calculated and then divided by the total number of subjects.

### E. Statistical analysis

As described above, three local and six global metrics were computed in order to examine the topological and functional characteristics of every MST. These feature datasets are, in general, non-normal so that natural log-transformation was applied in order to approximate normal distribution properties, with the addition of a very small constant ( $1 \cdot e^{-24}$ ) for avoiding zero-value transforms. Statistical analysis was performed using 1-way balanced ANOVA in order to test for differences between males and females. All p-values were corrected based on False Discovery Rate (FDR) using the Benjamini-Hochberg procedure [40] with the significance level set to 0.05.

## VIII. RESULTS

### A. Application on synthetic time-series

A  $128 \times 128$  block containing  $k$ -randomly distributed patterns was generated. Each pattern has a representative or seed signal which is a sinus wave with random superimposed noise and a unique random frequency varying on the interval  $(0, f_N)$ , where  $f_N$  is the Nyquist frequency equal to  $f_s/2$ . The rest of the signals in a pattern have the same frequency as the pattern's seed signal with random superimposed noise. For realistic purposes,  $f_s$  was set to 1.388 Hz, which is equal to the fMRI scanner's ratio  $1/TR$  (i.e.,  $1/0.72$ ). The number of square blocks was set to  $k = 6$  and the number of samples was set to 100 per time-course. After generating the synthetic BOLD time-series, Pearson's correlation coefficient was computed between each pair of time-series, leading to a  $16384 \times 16384$  adjacency matrix, where the element  $(i, j)$  is the correlation between the time-series of points  $i$  and  $j$  in the  $128 \times 128$  block.

Then, the SC, N-cut and SCSC clustering approaches were applied to the corresponding correlation matrices. An extra distance matrix is necessary for the application of the SCSC with  $rc = 2$  (covering 12 points), where its  $(i, j)$ -th element is the Euclidean distance between points  $i$  and  $j$ . Finally, Shannon's entropy was computed on each clustermap using various sphere's radius values, i.e.,  $r = 2$  (12 points);  $r = 3$  (28 points);  $r = 4$  (48 points);  $r = 5$  (80 points);  $r = 6$  (112 points), to evaluate each clustermap's homogeneity.

According to Fig. 1, it is obvious that the SC and N-cut algorithms are not able to distinguish the patterns from the noisy background across all three runs. This is also supported by their corresponding entropy distributions, which tend to exhibit high entropy values (i.e., large uncertainty with non-uniform distributions) implying that the homogeneity of the resulted clustermaps is too small. On the other hand, the SCSC algorithm correctly discriminates the patterns from the background noise, facilitated by the spatial constraint that restricts the number of neighbors and reduces the noise levels to a large extent. As a matter of fact, the resulting clusters are non-spatially distributed (i.e., contiguous). This outcome can be also verified from the corresponding entropy distributions, which tend to be smooth and uniform, with the entropy values approximating zero (i.e., very small uncertainty) implying high clustermap homogeneity.

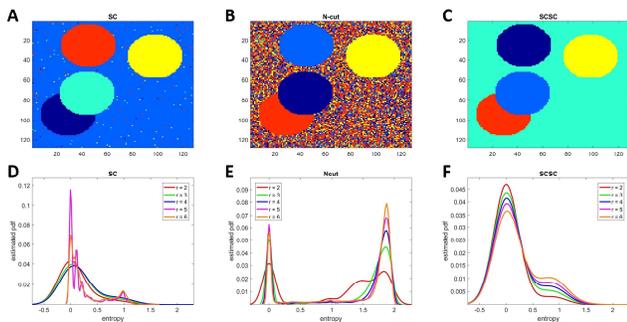


Fig. 1. Simulation results after the application of the (A) Spectral, (B) N-cut and (C) Spatially Constrained Spectral clustering procedures, respectively, alongside with the corresponding clustermap's entropy distributions (D), (E), (F), using various radius values.

### B. Cross-validation for performance evaluation

In order to validate the consistency of the number of clusters across the total population, a repeated cross-validation procedure is executed by first computing the average correlation matrices across 10 randomly selected subjects for 10 times (runs) in total; i.e., 10 average correlation matrices across 10 random subjects per run and re-applying for 10 random runs.

According to Fig. 2, the SCSC procedure for  $rc = 8$  on  $S$  was stable at  $k = 4$  clusters across all runs. The same comment can be done for  $rc = 12$  but at  $k = 3$  clusters. In addition, the case for  $rc = 5$  is interesting, since it is stable at  $k = 4$  clusters across 9 out of 10 runs. The rest of thresholding values do not exhibit any consistency at all. The fact that  $rc = 5$  has the second best performance at  $k = 4$  further supports to fix the number of clusters to 4. In a similar manner, the  $rc$  value of 8 is preferred over  $rc = 12$ , since the former covers 1947 voxels instead of the large number of 5763 voxels.

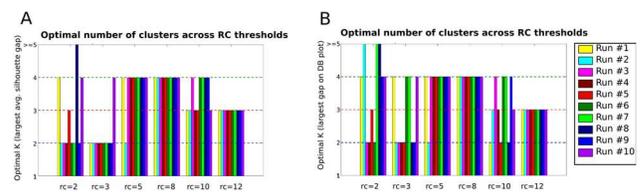


Fig. 2. The optimal number of clusters across each  $rc$  value (horizontal axis) for all validation runs, as indicated by the largest gap on each run's (A) avg. silhouette and (B) Davies-Bouldin indices evaluated across 30 clusters (vertical axis).

We have also evaluated the consistency of the resulting cluster maps. For this purpose, the Dice similarity coefficient was evaluated individually for each pair of 10 cluster maps obtained through the SCSC procedure.

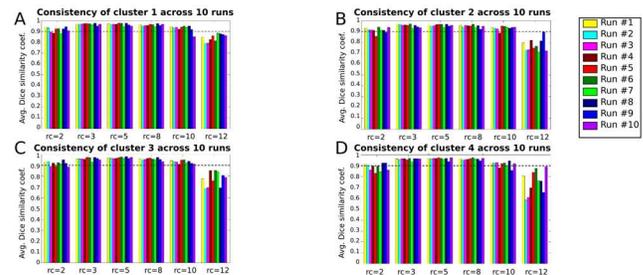


Fig. 3. Consistency evaluation results for each  $rc$  value across all runs for cluster (A) 1, (B) 2, (C) 3 and (D) 4. The black dashed line indicates a clustering consistency of 0.9 (90%).

The selection of  $rc = 8$  with  $k = 4$  clusters gave the most stable threshold across all 10 runs of the proposed cross-validation procedure (Fig. 2) and produces cluster maps with each cluster having at least 95% consistency across 10 runs (Fig. 3). According to these tests, the SCSC will be applied on the original data using  $rc = 8$  on  $S$  ( $k = 4$  clusters).

### C. Resting-state network atlas

The SCSC algorithm was applied to the average correlation matrix computed across all 100 subjects, with the number of clusters set to  $k = 4$  with an  $rc = 8$  on  $S$ . The result is a clustermap or atlas (a 3D matrix with clustering labels assigned per voxel), which has been already registered in the MNI coordinate space (template) and represents the four identified cerebellar resting-state networks (Fig. 4(A)). These four clusters are well-separated and consistent. For the sake of completeness and for comparison purposes only, the SC and N-cut algorithms have been also applied to the same matrix for  $k = 4$ . Fig. 4(B), (C) illustrate the corresponding 3D cluster maps with the Shannon's entropy distributions across all voxels, for various sphere's radius values (i.e.,  $r = 2$  (33 voxels);  $r = 3$  (123 voxels);  $r = 4$  (257 voxels);  $r = 5$  (515 voxels);  $r = 6$  (924 voxels)).

Note that the clusters derived from the SCSC algorithm (Fig. 4(A)) are spatially contiguous and homogeneous whereas the clusters obtained from the SC (Fig. 4(B)) and N-cut (Fig. 4(C)) algorithms are spatially distributed. It is also obvious that the SCSC algorithm achieves small entropy values (Fig. 4(D)) and consistent ROIs with better cohesion than the SC (Fig. 4(E)) and N-cut (Fig. 4(F)) algorithms.

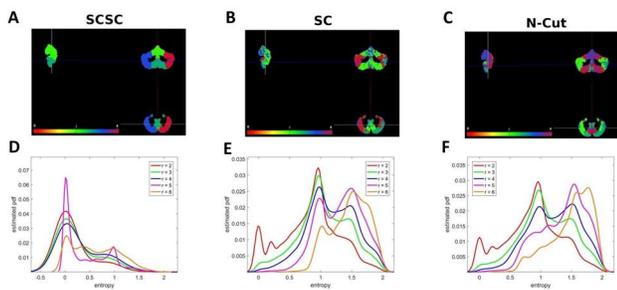


Fig. 4. The resulting cluster maps for (A) SCSC, (B) SC and (C) N-cut along with the corresponding voxel-wise entropy distributions (D, E, F).

#### D. The final cerebellar atlas

In order to extract a much more detailed atlas reflecting both anatomic and functional characteristics, the RSN atlas obtained from the SCSC approach is combined with the anatomical atlas of the cerebellum, which was already obtained from the SUIT toolbox [23] presented in Fig. 5(A) using BrainNet Viewer [41]. After projecting the RSN atlas (Fig. 5(B)) onto the anatomical atlas, the resulting overlaid regions can be treated as functional ROIs. Regions that contain less than 10 voxels are ignored from further analysis. As a result, the final number of extracted ROIs is equal to 46 (Fig. 5(C)). In fact, this procedure aims to extract a large number of functional ROIs with anatomic coherence, in order to be used in data-driven network analysis of cerebellum. The projection is valid since the voxel coordinates of the anatomical atlas lie on the same space (i.e., MNI space) with those of the RSN atlas.

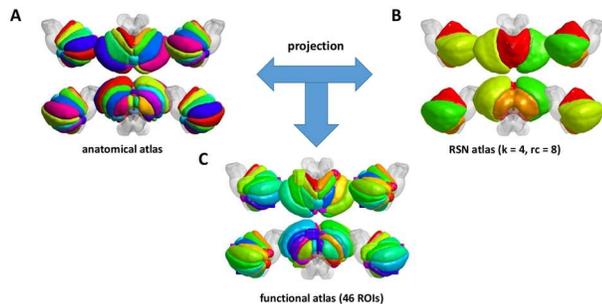


Fig. 5. The construction steps of cerebellum's functional atlas. (A) The anatomical atlas is projected on (B) the RSN atlas leading to a new (C) functional atlas with 46 functional ROIs.

Table I summarizes useful information concerning the functional ROIs locations as well as the number of voxels per ROI. Regions with less than 10 voxels are considered insufficient and have been discarded.

#### E. Network analysis

The average correlation networks across males and females are depicted on Fig 6 (A), (B). Afterwards, the corresponding MSTs were computed for each population and various local and global metrics were computed for each MST. The average MSTs for both males and females are presented on Fig. 6 (C), (D), respectively.

From Table II we notice a significant difference in the tree organization between males and females ( $p = 0.0386$ ,  $F = 4.3938$ ). Furthermore, despite the small differences, tree-hierarchy is close to 0.4 for both populations, indicating that

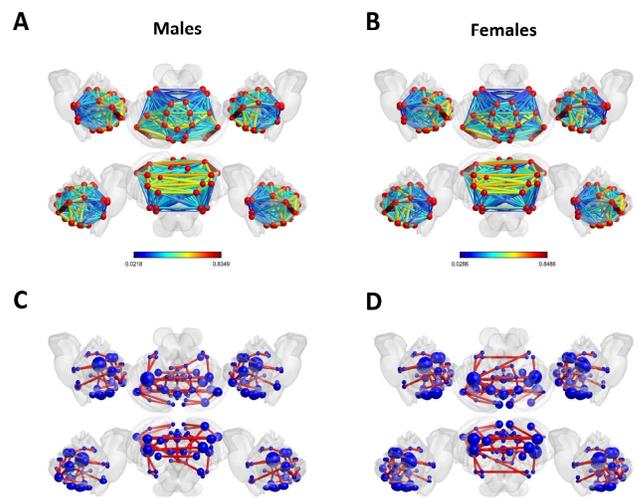


Fig. 6. The average correlation networks of both (A) males and (B) females along with the corresponding MSTs for (C) males and (D) females, projected on cerebellum's surface using BrainNet Viewer [41].

tree topologies of both genders are close to a star network topology (e.g., an optimal topology) with many central hubs.

TABLE II  
STATISTICAL ANALYSIS RESULTS BASED ON GENDER FOR THE MAIN NETWORK METRICS

Metric	Males	Females	F	p-values
	Mean±SD	Mean±SD		
$d$	0.3298±0.1323	0.3096±0.0782	0.2803	0.5977
$r$	0.2042±0.1269	0.1834±0.0471	0.5030	0.4799
$L_f$	0.5298±0.0572	0.5333±0.0500	0.1592	0.6908
$T_h$	0.3807±0.0463	0.3999±0.0466	4.3938	<b>0.0386</b>
$\kappa$	15.3239±1.8603	15.5915±1.0772	1.1942	0.2772
$r_{deg}$	-0.2163±0.1045	-0.2472±0.1008	2.5079	0.1165

with **bold** highlight: statistical significant result(s) ( $p < 0.05$ ).

According to Fig. 7, functional node 10 located in lobule Left VI, is a critical node for both genders with an observed male dominance; it exhibits the highest DEG in 32% of males and 22% of females and the highest BC in 32% of males and 20% of females. This is in concordance with the left cerebellar hemispheric dominance, as well as with Left VI's importance during information transfer within cerebellum's network [18].

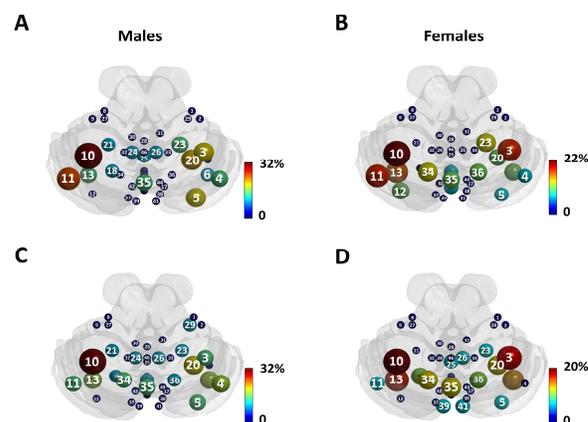


Fig. 7. Nodes with the highest (A) DEG and (B) BC values across males and females.

Moreover, there are additional nodes that can be characterized as hubs (but to a lower extent), such as functional node 3 located in lobule Right VI and is mostly a hub for the female population; it exhibits the highest DEG in 8% of males and 12% of females and the highest BC in 4% of males and 14% of females. The same stands for functional node 13 (located in lobule Left VIIb), which appears to have the highest DEG in 10% of females and the highest BC in 12% of females. Finally, functional node 11 (located in lobule Left Crus I) exhibits the highest DEG in 14% of males and 12% of females, without however any influence to BC.

TABLE I  
FUNCTIONAL ROIS INFORMATION

ROI	% of voxels	Cluster	Lobule	MNI coordinates		
				X	Y	Z
1	0.195	4	Right I-IV	24	-32	-32
2	0.437	4	Right V	28	-36	-30
3	4.535	4	Right VI	30	-54	-28
4	11.514	4	Right Crus I	38	-68	-32
5	7.704	4	Right Crus II	26	-78	-42
6	3.633	4	Right VIIb	32	-66	-52
7	1.166	4	Right VIIIa	32	-58	-54
8	0.098	3	Left I-IV	-20	-32	-30
9	0.646	3	Left V	-26	-36	-30
10	4.874	3	Left VI	-28	-56	-26
11	10.966	3	Left Crus I	-38	-68	-32
12	8.182	3	Left Crus II	-26	-76	-42
13	3.689	3	Left VIIb	-28	-66	-52
14	2.704	3	Left VIIIa	-28	-56	-54
15	0.121	3	Left X	-26	-36	-42
16	0.107	2	Left VIIb	-6	-72	-46
17	0.302	2	Right VIIb	10	-72	-48
18	1.059	2	Left VIIIa	-16	-64	-52
19	0.483	2	Vermis VIIIa	0	-68	-42
20	2.779	2	Right VIIIa	24	-58	-54
21	3.048	2	Left VIIIb	-18	-50	-54
22	0.506	2	Vermis VIIIb	0	-64	-42
23	3.341	2	Right VIIIb	18	-50	-54
24	2.574	2	Left IX	-6	-54	-48
25	0.525	2	Vermis IX	0	-56	-38
26	3.020	2	Right IX	6	-54	-48
27	0.497	2	Left X	-20	-36	-46
28	0.209	2	Vermis X	0	-48	-36
29	0.650	2	Right X	22	-36	-46
30	2.792	1	Left I-IV	-6	-46	-16
31	2.806	1	Right I-IV	8	-44	-18
32	2.718	1	Left V	-10	-54	-16
33	1.766	1	Right V	12	-54	-16
34	2.477	1	Left VI	-12	-66	-20
35	1.766	1	Vermis VI	0	-70	-20
36	2.402	1	Right VI	14	-66	-20
37	0.149	1	Left Crus I	-8	-78	-26
38	0.107	1	Right Crus I	8	-76	-26
39	0.488	1	Left Crus II	-4	-80	-34
40	0.376	1	Vermis Crus II	0	-74	-32
41	0.497	1	Right Crus II	6	-80	-34
42	0.144	1	Left VIIb	-6	-72	-38
43	0.125	1	Vermis VIIb	0	-68	-32
44	0.144	1	Right VIIb	8	-70	-36
45	0.544	1	Vermis VIIIa	0	-66	-34
46	0.079	1	Vermis IX	0	-54	-30

## IX. DISCUSSION

The cerebellum participates in various functional processes as a direct implication of its strong connectivity with multiple cerebellar regions. Separate regions of the cerebellum are

connected to different cerebral areas forming a complex topography. Various studies have mapped the functional association of the cerebellum regions and the corresponding cerebellar structures. More specifically, the cerebellum has been associated with sensory, motor, and cognitive tasks as well as with the integration of motor and sensory information. The cerebellum contains multiple somatomotor representations topographically organized; an inverted somatomotor map exists in the anterior lobe and a secondary posterior map inversely organized to the anterior [42]. The inverted representation extends from lobule V to lobule VI. Lobules Crus I and II are not associated with motor areas but present mainly projections to the prefrontal cortex [42].

To determine the most important nodes that act as hubs in the network we used the BC of each node and normalized it by the mean BC of the node across subjects. The nodes that their mean BC was 1.5 times higher than the average centrality were identified as hubs. The results indicate a symmetry in the distribution of hubs between the left and the right cerebellum with only a couple of notable differences. The nodes identified as hubs include regions in lobules VI, VII, and VIII, representing the anterior sensorimotor map. The lobules Crus I and II were also hubs for both genders.

Comparing the strength of the hubs between the two hemispheres of the cerebellum the nodes in the left hemisphere presented higher BC than the nodes in the right hemisphere (Mann–Whitney U test [43] males:  $p < 0.01$ , females:  $p < 0.01$ , Fig. 8(A), (B)). Although the number of nodes between the left and the right are the same, the nodes of the left lobe seem to be more integrated than the right ones. It is also known that the projections between the cerebellum and the corresponding cerebral areas are mostly contralateral [44]. This property implies that the right hemisphere of the cerebellum is mostly associated with linguistic processing while the left is associated with the processing of spatial information [42]. It seems that this difference is mapped to the resting state functional network as well. The left hemisphere contains more hubs associated with somatomotor lobules (Left VI, VIIb, VIIIa) and the nodes of the network present higher centrality than the right lobe. Moreover, it has been found that the more efficient network organization in women reflects the different hemispheric organization between genders [18]. This is a fact that supports the difference in the optimal organization between both genders since the MSTs reflect the network’s topological structure.

The spatially constrained algorithm divided the Left VI lobule into two hubs. The first Left VI hub is located in the main anterior lobe and is associated with all kinds of movement, marked with green in Fig. 8(C), (D) [14]. The second hub is located closer to the Vermis and has been implicated in affective processing especially for the emotions of sadness and disgust [45]. In the right hemisphere, only the corresponding anterior VI lobule has significantly increased BC in both genders. The second Right VI node, associated with emotional processing [44] is only identified as a hub in the female population. This difference can be attributed to the different strategies employed for processing emotional stimuli, between males and females. It has been shown that females are more likely to activate midline limbic structures, including the cerebellum, than males. In our case, the lobules implicated with emotional processing play a more central role in females.

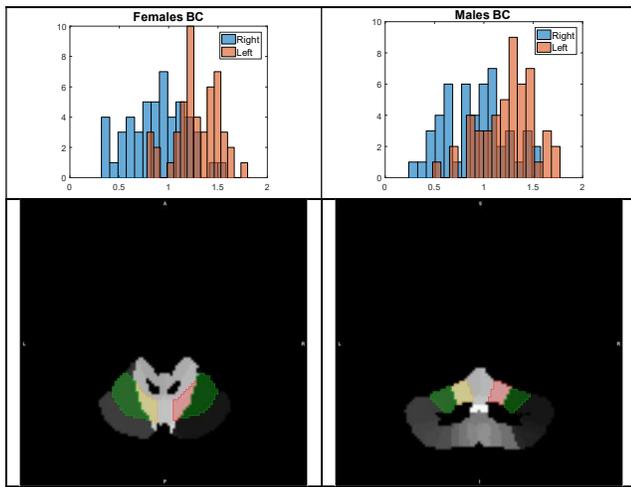


Fig. 8. Top Row: A. The distribution of the BC values (averaged across subjects) for the nodes in the Left and the Right lobe for the female group. B. The corresponding distribution for the male group. Bottom Row: The Left/Right VI sensory hubs are marked green. With yellow is marked the second Left VI hub common to both groups. With Red is marked the Right VI hub specific to the female group.

## X. CONCLUSIONS

In this work, we studied the effects of spatial constraints on the derivation of the functional atlas of the human cerebellum during resting-state. We demonstrated the effectiveness of the spatially constrained spectral clustering to derive spatial coherent clusters. The cross-validation study revealed that the effect of the spatial constraint is significant to the clustering results and it should be taken into consideration in studies that are using techniques that are using spatial constraints. It appears that the radius of the spatial constraint has a significant impact on the number of clusters selection.

Using the cross-validation procedure to select parameters, we extracted the group functional atlas and used it to perform a gender-based functional connectivity analysis of cerebellum's network. Our hub analysis findings denote a left hemispheric dominance in cerebellum's connectivity, highlighting the importance of lobule Left VI during information transfer. This finding is in line with other studies that have characterized lobule VI as a transition region bridging the motor-related networks with those related to cognitive regions [21]. Our methodology split lobule VI into two nodes that correspond to distinct functional specializations of the Lobule. The node associated with emotional processing was identified as a hub in both lobes only in female subjects reflecting the differences in affective processing between the two genders.

## ACKNOWLEDGMENT

The dataset was provided [in part] by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University. The author Pezoulas VC has received

a scholarship of excellence from the Pancretan Endowment Fund for this research work.

## REFERENCES

- [1] M. P. van den Heuvel and H. E. Hulshoff Pol, "Exploring the brain network: a review on resting-state fMRI functional connectivity," *Eur. Neuropsychopharmacol.*, vol. 20, pp. 519-534, Aug. 2010.
- [2] M. H. Lee, C. D. Smyser, and J. S. Shimony, "Resting-state fMRI: a review of methods and clinical applications," *AJNR Am. J. Neuroradiol.*, vol. 34, pp. 1866-1872, Oct. 2013.
- [3] B. Biswal, F. Z. Yetkin, V. M. Haughton, and J. S. Hyde, "Functional connectivity in the motor cortex of resting human brain using echo-planar MRI," *Magn. Reson. Med.*, vol. 34, no. 4, pp. 537-541, Oct. 1995.
- [4] M. J. Lowe, B. J. Mock, and J. A. Sorenson, "Functional connectivity in single and multislice echoplanar imaging using resting-state fluctuations," *Neuroimage*, vol. 7, no. 2, pp. 119-132, Oct. 1998.
- [5] E. Bullmore and O. Sporns, "Complex brain networks: graph theoretical analysis of structural and functional systems," *Nat. Rev. Neurosci.*, vol. 10, no. 3, pp. 186-198, Mar. 2009.
- [6] J. D. Power, A. L. Cohen, S. M. Nelson, G. S. Wig, K. A. Barnes, J. A. Church, A. C. Vogel, T. O. Laumann, F. M. Mielzlin, and B. L. Schlaggar, "Functional network organization of the human brain," *Neuron*, vol. 72, no. 4, pp. 665-678, Nov. 2012.
- [7] K. J. Friston, "The disconnection hypothesis," *Schizophr. Res.*, vol. 30, no. 2, pp. 115-125, Mar. 1998.
- [8] C. F. Beckmann, M. DeLuca, J. T. Devlin, and S. M. Smith, "Investigations into resting-state connectivity using Independent Component Analysis," *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, vol. 360, no. 1457, pp. 1001-1013, May 2005.
- [9] M. De Luca, C. F. Beckmann, N. De Stefano, P. M. Matthews, and S. M. Smith, "fMRI resting state networks define distinct modes of long-distance interactions in the human brain," *Neuroimage*, vol. 29, no. 4, pp. 1359-1367, Feb. 2006.
- [10] V. D. Calhoun, T. Adali, G. D. Pearlson, and J. J. Pekar, "A method for making group inferences from functional MRI data using Independent Component Analysis," *Hum. Brain Mapp.*, vol. 14, no. 3, pp. 140-151, Nov. 2001.
- [11] R. C. Craddock, G. A. James, P. E. Holtzheimer, 3rd, X. P. Hu, and H. S. Mayberg, "A whole brain fMRI atlas generated via spatially constrained spectral clustering," *Hum. Brain Mapp.*, vol. 33, pp. 1914-1928, Aug. 2012.
- [12] M. van den Heuvel, R. Mandl, and H. Hulshoff Pol, "Normalized cut group clustering of resting-state FMRI data," *PLoS One*, vol. 3, no. 4, pp. e2001, Apr. 2008.
- [13] J. R. Sato, D. Y. Takahashi, M. Q. Hoexter, K. B. Massier, and A. Fujita, "Measuring network's entropy in ADHD: a new approach to investigate neuropsychiatric disorders," *Neuroimage*, vol. 77, pp. 44-51, Aug. 2013.
- [14] M. Manto, J. M. Bower, A. B. Conforto, J. M. Delgado-Garcia, S. N. da Guarda, M. Gerwig, C. Habas, N. Hagura, R. B. Ivry, P. Mariën, M. Molinari, E. Naito, D. A. Nowak, N. Oulad Ben Taib, D. Pelisson, C. D. Tesche, C. Tilikete, D. Timmann, "Consensus paper: roles of the cerebellum in motor control--the diversity of ideas on cerebellar involvement in movement," *Cerebellum*, vol. 11, no. 2, pp. 457-487, Jun. 2012.
- [15] Pernet, Cyril R, Jean Poline, Jean Demonet, and Guillaume A Rousselet. 2009. "Brain classification reveals the right cerebellum as the best biomarker of dyslexia," *BMC Neurosci.*, vol. 10, no. 1, pp. 67, Jun. 2009.
- [16] J. X. O'Reilly, C. F. Beckmann, V. Tomassini, N. Ramnani, and H. Johansen-Berg, "Distinct and overlapping functional zones in the cerebellum defined by resting state functional connectivity," *Cereb. Cortex*, vol. 20, no. 4, pp. 953-965, Apr. 2010.
- [17] R. L. Buckner, F. M. Krienen, A. Castellanos, J. C. Diaz, and B. T. Yeo, "The organization of the human cerebellum estimated by intrinsic functional connectivity," *J. Neurophysiol.*, vol. 106, no. 5, pp. 2322-2345, Jul. 2011.
- [18] V. C. Pezoulas, M. Zervakis, S. Michelogiannis, and M. A. Klados, "Resting-state functional connectivity and network analysis of cerebellum with respect to crystallized IQ and gender," *Front. Hum. Neurosci.*, vol. 11, p. 189, Apr. 2017.

- [19] H. Liu, G. Fan, Ke X., and F. Wang, "Changes in cerebellar functional connectivity and anatomical connectivity in schizophrenia: a combined resting-state functional MRI and diffusion tensor imaging study," *J. Magn. Reson. Imaging*, vol. 34, p. 1430-1438, Dec. 2011.
- [20] C. Zhuo, C. Wang, L. Wang, X. Guo, Q. Xu, Y. Liu, and J. Zhu, "Altered resting-state functional connectivity of the cerebellum in schizophrenia," *Brain Imaging and Behavior*, vol. 12, p. 383-389, Mar. 2017.
- [21] J. A. Bernard, R. D. Seidler, K. M. Hassevoort, B. L. Benson, R. C. Welsh, J. L. Wiggins, et al., "Resting state cortico-cerebellar functional connectivity networks: a comparison of anatomical and self-organizing map approaches," *Front. Neuroanat.*, vol. 6, p. 31, 2012.
- [22] D. C. Van Essen, S. M. Smith, D. M. Barch, T. E. Behrens, E. Yacoub, and K. Ugurbil, "The WU-Minn Human Connectome Project: an overview," *Neuroimage*, vol. 80, pp. 62-79, Oct. 2013.
- [23] J. Diedrichsen and E. Zotow, "Surface-based display of volume-averaged cerebellar imaging data," *PLoS One*, vol. 10, no. 7, pp. e0133402, Jul. 2015.
- [24] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 8, pp. 888-905, Aug. 2000.
- [25] J. C. A. Barata, and M. S. Hussein. The Moore-Penrose pseudoinverse: a tutorial review of the theory. *Braz. J. Phys.*, vol. 42, no. 1-2, pp. 146-165, Apr. 2012.
- [26] P. J. Rousseeuw, "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis," *Journal of Computational and Applied Mathematics*, vol. 20, pp. 53-65, Nov. 1987.
- [27] D. L. Davies and D. W. Bouldin, "A cluster separation measure," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 1, no. 2, pp. 224-227, Apr. 1979.
- [28] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297-302, Jul. 1945.
- [29] N. A. Laskaris and A. A. Ioannides, "Exploratory data analysis of evoked response single trials based on minimal spanning tree," *Clin. Neurophysiol.*, vol. 112, no. 4, pp. 698-712, Apr. 2001.
- [30] C. J. Stam, P. Tewarie, E. Van Dellen, E. C. van Straaten, A. Hillebrand, and P. Van Mieghem, "The trees and the forest: characterization of complex brain networks with minimum spanning trees," *Int. J. Psychophysiol.*, vol. 92, no. 3, pp. 129-38, Jun. 2014.
- [31] P. Tewarie, E. van Dellen, A. Hillebrand, and C. J. Stam, "The minimum spanning tree: an unbiased method for brain network analysis," *Neuroimage*, vol. 104, pp. 177-188, Jan. 2015.
- [32] C. J. Stam and E. C. van Straaten, "The organization of physiological brain networks," *Clin. Neurophysiol.*, vol. 123, no. 6, pp. 1067-1087, Jun. 2012.
- [33] E. van Diessen, T. Numan, E. van Dellen, A. W. van der Kooij, M. Boersma, D. Hofman, R. van Lutterveld, B. W. van Dijk, E. C. van Straaten, A. Hillebrand, and C. J. Stam, "Opportunities and methodological challenges in EEG and MEG resting state functional brain network research," *Clin. Neurophysiol.*, vol. 126, no. 8, pp. 1468-1481, Aug. 2015.
- [34] J. B. Kruskal, "On the shortest spanning subtree of a graph and the traveling salesman problem," *Proc. Amer. Math. Soc.*, vol. 7, no. 1, pp. 48-50, Feb. 1956.
- [35] M. Boersma, D. J. Smit, D. I. Boomsma, E. J. De Geus, H. A. Delemarre-van de Waal, and C. J. Stam, "Growing trees in child brains: graph theoretical analysis of electroencephalography-derived minimum spanning tree in 5- and 7-year-old children reflects brain maturation," *Brain Connect.*, vol. 3, no. 1, pp. 50-60, Jan. 2013.
- [36] M. Rubinov and O. Sporns, "Complex network measures of brain connectivity: uses and interpretations," *Neuroimage*, vol. 52, no. 3, pp. 1059-1069, Sep. 2010.
- [37] E. van Dellen, H. de Waal, W. M. van der Flier, A. W. Lemstra, A. J. C. Slooter, L. L. Smits, E. C. van Straaten, C. J. Stam, and P. Scheltens, "Loss of EEG network efficiency is related to cognitive impairment in dementia with lewy bodies," *Mov. Disord.*, vol. 30, no. 13, pp. 1785-1793, Nov. 2015.
- [38] P. Tewarie, A. Hillebrand, E. van Dellen, M. M. Schoonheim, F. Barkhof, C. H. Polman, C. Beaulieu, G. Gong, B. W. van Dijk, and C. J. Stam, "Structural degree predicts functional network connectivity: a multimodal resting-state fMRI and MEG study," *Neuroimage*, vol. 97, pp. 296-307, Aug. 2014.
- [39] W. M. Otte, E. van Diessen, S. Paul, R. Ramaswamy, V. P. Subramanyam Rallabandi, C. J. Stam, and P. K. Roy, "Aging alterations in whole-brain networks during adulthood mapped with the minimum spanning tree indices: the interplay of density, connectivity cost and life-time trajectory," *Neuroimage*, vol. 109, pp. 171-189, Apr. 2015.
- [40] Y. Benjamini and Y. Hochberg, "Controlling the false discovery rate: a practical and powerful approach to multiple testing," *J. R. Stat. Soc. Series B Stat. Methodol.*, vol. 57, no. 1, pp. 289-300, 1995.
- [41] M. Xia, J. Wang, and Y. He, "BrainNet Viewer: a network visualization tool for human brain connectomics," *PLoS One*, vol. 8, no. 7, pp. e68910, Jul. 2013.
- [42] C. J. Stoodley and J. D. Schmahmann, "Evidence for topographic organization in the cerebellum of motor control versus cognitive and affective processing," *Cortex*, vol. 46, no. 7, pp. 831-844, Jul-Aug. 2010.
- [43] H. B. Mann and D. R. Whitney, "On a test of whether one of two random variables is stochastically larger than the other," *Annals math. stat.*, vol. 18, no. 1, pp. 50-60, Mar. 1947.
- [44] P. Brodal, "The pontocerebellar projection in the rhesus monkey: an experimental study with retrograde axonal transport of horseradish peroxidase," *Neuroscience*, vol. 4, no. 2, pp. 193-208, Feb. 1979.
- [45] O. Baumann and J. B. Mattingley, "Functional topography of primary emotion processing in the human cerebellum," *Neuroimage*, vol. 61, no. 4, pp. 805-811, Jul. 2012.