

Mini-Project (ML for Time Series) - MVA 2023/2024

Mathis Reymond mathis.reymond74@gmail.com

Inès Vati ines.vati@eleves.enpc.fr

February 1, 2024

1 Introduction and contributions

The article at hand focuses on reviewing the current literature regarding the use of Graph Signal Processing (GSP) methods applied to brain graphs obtained from functional imaging techniques. Within this framework, graph representations illustrate brain regions as nodes and depict connections—be they functional or structural—as edges. This approach, viewing the brain through graphs, holds significance in connectomics and the broader field of network neuroscience. It facilitates the correlation between structural neural connections and the functional connectivity, even extending to behavioral observations.

Authors introduce a Fourier paradigm adapted to graph, including a Graph Fourier Transform (GFT) and graph filters. It allows the decomposition of brain signals based on spatial variability relative to the structure of the brain network. Additionally, they introduce signal surrogate generation, notably to assess significance of their results. They lead two experiments. In the first one, subjects perform a classic Navon switching task : they visualize big symbols made of smaller symbols (as big cross made of small circles). Depending on the colour of the small symbols, participants are asked to report either the shape of small or big symbols. Naturally, the response time is increased when the task switch, which allows defining what's called a switch cost. Using GSP tools, authors have shown that having isolated brain regions highly activated is positively correlated with high switch cost.

Authors provide no code. We wrote our own code for the experiments we conducted. Inès focused on graph surrogates and performed the experiment described in 4.1. Mathis performed on the experiment described in 4.2. We state an equal contribution for the rest of the work carried out. Our code is available on [GitHub](#).

Along with qualitative remarks, our contribution lies in the three experiments we conducted. As neither code nor data used in the paper were available, we found another dataset, BOLD5000 [1], on which we performed experiments using the tools introduced in the article.

2 Method

We consider a weighted graph $\mathcal{G} = (\mathcal{V}, S)$ where \mathcal{V} is the set of N vertices associated with specific brain regions and $S \in \mathbb{R}^{N \times N}$ is a matrix connectivity representation (usually the adjacency matrix or Laplacian). We also consider a graph signal $X \in \mathbb{R}^{N \times T}$ where T is the length of the time series.

2.1 Graph Fourier Transform

In order to define a Fourier transform on \mathcal{G} , one must diagonalize this S and express it in the form

$$S = V\Lambda V^{-1} \quad (1)$$

where Λ is a diagonal matrix containing the eigenvalues of S , and V is a matrix of associated eigenvectors. This allows the definition, given a signal x on \mathcal{G} , of the graph Fourier transform of x on \mathcal{G} as

$$\tilde{x} := V^T x \quad (2)$$

Please refer to the appendix A for the computations that draw parallels with the discrete Fourier transform on temporal data.

2.2 Graph filtering

In the context of graph signal processing, filters are commonly defined within the spectral domain. Let h be a transfer function, and we denote H as the diagonal matrix where $H_{ii} = h(\lambda_i)$, and λ_i denotes the eigenvalues of S . The filtered signal $Y_h \in \mathbb{R}^{N \times T}$ is defined as follows

$$Y_h := VHV^T X \quad (3)$$

In this formulation, filtering is applied independently to each time sample $x[t]$. However, unlike traditional time-domain filtering, graph filtering leverages the graph structure to smooth the signal across its edges rather than along the time axis. This distinction becomes obvious when visualizing the impact of graph filtering on a signal. Figure 8 illustrates the differences in the same signal, either filtered with a low-pass filter or a high-pass filter. Colour values on the low-pass filtered signal exhibit greater homogeneity across nodes than on the high-pass filter one.

2.3 Generation of graph surrogate signals

Statistical testing is crucial when it comes to make inferences about the properties of signals. For example, we may want to test whether the observed signal is significantly different from a random signal. Phase-randomization in the temporal Fourier domain [2] offers to preserve the spectral properties of the time series while randomizing the signal. It consists in generating surrogate data from an original time series by randomizing the phases of the original time series while keeping the magnitudes of the Fourier components unchanged. The surrogate signal is given by $Y = XF^H \phi_{time} F$, where ϕ_{time} contains random phase factors, F and F^T indicate the temporal Fourier Transform and Inverse Fourier Transform, respectively.

We can extend this procedure to the graph framework: $Y = V\phi_{graph}V^T X$ where ϕ_{graph} is a diagonal matrix containing random phase factors. ϕ_{graph} can be defined by random sign flips of the graph spectral coefficients. An explanatory scheme is provided in Figure 10.

In [3], one of the objectives is to discern significant signal excursions, which are distinct moments in time when brain activity enters a regime of strong alignment or liberality.

In the context of the authors' study, alignment and liberality are concepts used to characterize different aspects of functional brain activity. Alignment is associated with regions of the brain that activate simultaneously. On the other hand, liberality pertains to areas that exhibit high signal variability, suggesting a more diverse and dynamic pattern of activity. Alignment is computed using a low-pass filter on the graph because low-pass filters allow the preservation of lower-frequency components while attenuating higher-frequency variations. Liberality, conversely, is

computed using a high-pass filter on the graph that emphasizes higher-frequency variations. The authors employed the following procedure to compute excursions in those different regimes :

1. capture alignment and liberality with low-pass and high-pass filters on the original signal
2. generate 1,000 null surrogate signals for each filtered signal, as performed in [4]
3. employ the generated null data to threshold the filtered signals at a α -level of 5%

3 Data

3.1 BOLD5000 dataset

We used the BOLD5000 database [1] in our experiments. It is a large-scale fMRI dataset that captures brain scans from four patients as they view over 5,000 images. The dataset covers a wide range of visual features, categories, and semantics, and can be used to test various hypotheses and models related to visual cognition.

fMRI produces 4D images, relying on the principle that localized neural activity induces changes in metabolism and blood flow. To record cerebral activity during functional sessions, the scanner is tuned to detect this "Blood Oxygen Level Dependent" (BOLD) signal. Brain activity is measured in sessions that span several minutes, during which participants are presented with a variety of images. Simultaneously, participants engage in a valence judgment task for each stimulus, expressing their preference using the descriptors "like", "neutral", and "dislike", which are encoded with labels 1, 2, and 3.

As the mapping between descriptors and labels was not provided, we inferred it ourselves.

3.2 Time series extraction and diagnosis

For each region of interests (ROIs) in the dataset, we extracted the time series, which represents the temporal activity in the brain. We used Nilearn -a Python library that provides tools for neuroimaging data analysis- to extract the signals. Signal extraction is usually achieved by averaging the fMRI time series across the voxels in a region [5].

In this study, time series were z-scored, i.e. shifted to zero mean and scaled to unit variance. We also took into account the confounds in the extraction, as there were provided in the database. Confounds are variables that can affect the brain signal and are not of interest in the study. They can include head motion, physiological noise, and scanner artefacts. In order to extract clean signals from brain regions, it is important to remove the effects of confounds from the data.

The time series obtained from the regions of interest (ROIs) do not appear to be noisy, as shown in Figure 3a. One can assume that signal are weak stationary as their mean and autocorrelation does not appear to vary over time. The spectrogram plotted in figure 5b shows that their spectral properties does noes not change over time. They all exhibit the same variability and mean as they were z-scored. Their values range between -2.5 and 2.5 , except for some regions that have extreme values or outliers (see Fig 3b). They seem to have the same spectral properties except for the retrosplenial complex (RSC) region in left and right hemisphere which has a more chaotic spectrum. Those results are shown in figure 5.

3.3 Brain Graph Construction

In order to represent the strength of axonal connection (structural connectivity), authors of [3] defined A_{ij} as a simple count of the number of streamlines - i.e. estimated individual fibers that connect the regions - using diffusion spectrum imaging. However, since we did not have access to this information, we defined the adjacency matrix \mathbf{A} in another manner. In our setting, it corresponds to the connectivity matrix between brain regions (functional connectivity). Indeed, the extracted signals can be used to compute a correlation matrix between the regions [5]. It is a common metric for computing the edges between the nodes, and is depicted in figure 6a.

4 Results

4.1 Extract excursions in alignment and liberality regimes

In this section, we investigate the following questions : (i) In our setting with only 10 regions, can we exhibit regions with frequent moments of strong alignment and liberality ? (ii) Does the choice of low-pass or high-pass filter influence the percentage of significant excursions ? (iii) What do we obtain if we apply the conventional phase Fourier randomization ?

Mirroring the approach of the authors [3], we propose a non-parametric, model-free statistical test for identifying unexpected fluctuations in time series magnitude over time. We compute the average of the (temporal) 95% percentile on the surrogate signals, as illustrated in Figure 9. The obtained value is used to threshold the "true" time series. The time points that exceed this threshold are said to be significant excursions. The protocol is repeated for all the 20 runs and statistical results are plotted in Figure 11.

The authors chose the 12% smallest and largest eigenvalues for the alignment and liberality filters, respectively. It comes down to almost 2 eigenvalues for our graph. We implement the equivalent version of the filters and two other continuous filters that are displayed in Figure 7. The continuous low-pass filter is defined by $h_{low, continuous}(x) = \frac{1}{1+\tau x}$ with $\tau = 3$ and the continuous high-pass filter is given by $h_{high, continuous}(x) = \frac{2}{1+\exp(-\tau(x-r))}$ with $\tau = 1$ and $r = 5$.

The study by Huang *et al.* [3] focused on 87 brain regions, in contrast to our dataset with only 10 regions. Consequently, it is unsurprising that we do not replicate their reported percentage of excursions. Our findings, depicted in Figure 11, reveal two key observations: (i) the inability to identify regions with a high percentage of excursions, attributed to the limited number of studied regions; and (ii) the similarity in results between continuous and non-continuous low-pass filters. However, for high-pass filters, certain regions, such as the left hemisphere lateral occipital cortex (LHLOC) and the retrosplenial complex (RSC), do not have the same percentage of excursions and the same variability. This discrepancy may be attributed to the high-pass continuous filter accepting more eigenvalues than its non-continuous counterpart, as illustrated in Figure 7. Notably, the majority of eigenvalues surpass 3. Thus, the continuous low-pass filter select a sufficient number of low eigenvalues compare to the non-continuous low-pass filter, which is not replicated by the continuous high-pass filter.

Moreover, (iii) Figure 12 indicates that performing a statistical test with the conventional randomization do not exhibit differences between the brain regions. This outcome is anticipated, given that the phase randomization matrix ϕ is uniformly applied to the nodes of the graph in the same manner. However, the article [3] reported that alignment excursions persist when employing conventional randomization. We can argue that our inability to replicate these results stems from the limited availability of labelled brain regions in our dataset.

4.2 Utilizing Neural Correlates to Predict Subjective Experience of Visual Stimuli

In this experiment, we aim to employ GSP tools to predict the subjective preference of pictures. Specifically, using the BOLD5000 brain images, our objective is to predict categorical labels : 1, 2, or 3. In order to avoid subtleties caused by inter-subject studies, we confine this experiment to data acquired solely during the initial 6 runs of the first session of subject CSI1.

We compute 6 matrices sized $(10, 194)$ corresponding to the activation levels of the 10 brain regions across the 194 time points. This process enables the computation of a connectivity matrix for each of the 6 runs, where connectivity is derived as the average connectivity across the 194 time points. To create a unified representation, we compute the average connectivity matrix from the set of 6 connectivity matrices, which allows defining a brain graph based on this final averaged matrix. We perform diagonalization of the Laplacian matrix associated to this graph, enabling for graph Fourier transform. Given the graph’s 10 vertices, for each of the 6×194 time points, we compute the graph Fourier transform of the activation across the 10 brain regions projected on the graph. This process yields 6×194 samples, each comprising 10 features: the Fourier coefficients.

Using the Sklearn, we train both a simple Multi-Layer Perceptron and a Gradient Boosting Tree. Training involves 80% of the vectors made of the squared Fourier coefficients. Then, we evaluate the models’ performance on the remaining 20% of the data. Regrettably, we do not observe any improvements in balanced accuracy as compared to random classification. However, as this approach breaks down the temporal continuity by considering individual vectors without accounting for the fact that the complete response pattern for each image spans 4 to 5 time points, it was expected to obtain poor results.

That’s why, ultimately, we tried to perform dictionary learning, with the aim to extract specific patterns for each appreciation. Unfortunately, as the patterns we are looking for are 4 to 5 time points long, dictionary learning is not very suited to extract them.

4.3 Subjective Experience Prediction with K-Nearest Neighbors and Distance Time Warping

We refer to a "sub-sequence" as the segment within the time series when the patient is observing a single image. Those sub-sequences last approximately 10 seconds and have a temporal length of 4 to 6 time points.

In this section, our objective is to classify these sub-sequences using a k-Nearest Neighbors (K-NN) algorithm. Let’s index the sub-sequences with I and for any $i \in I$, let’s denote the temporal lengths of the sub-sequence i with T_i . Consider $X_i \in \mathbb{R}^{N \times T_i}$ and $X_j \in \mathbb{R}^{N \times T_j}$ as two sub-sequences. They may have different temporal lengths. Recall that N is the number of brain regions. We define the distance between X and Y as

$$d(X_i, X_j) = \frac{1}{N} \sum_{n=1}^N DTW(X_i[n], X_j[n])$$

where DTW is the Distance Time Warping (DTW) measure.

Using Scikit-learn, we implemented a cross-validation to compare results across various number of neighbors. The balanced accuracies are depicted in Figure 13. The only model surpassing the chance level is the one with $k = 3$ neighbors, displaying a median accuracy of 40%. However, the variability remains substantial. This may be explained by the relatively short length of the sub-sequences. Furthermore, caution is advised when independently comparing sub-sequences from different runs.

References

- [1] N. Chang, J. A. Pyles, A. Gupta, M. J. Tarr, and E. M. Aminoff, “BOLD5000: A public fMRI dataset of 5000 images,” *Scientific Data*, vol. 6, no. 1, p. 49, May 6, 2019, ISSN: 2052-4463. DOI: [10.1038/s41597-019-0052-3](https://doi.org/10.1038/s41597-019-0052-3). arXiv: [1809.01281](https://arxiv.org/abs/1809.01281)[cs, q-bio]. [Online]. Available: <http://arxiv.org/abs/1809.01281> (visited on 11/14/2023).
- [2] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, and J. Doynne Farmer, “Testing for non-linearity in time series: The method of surrogate data,” *Physica D: Nonlinear Phenomena*, vol. 58, no. 1, pp. 77–94, Sep. 15, 1992, ISSN: 0167-2789. DOI: [10.1016/0167-2789\(92\)90102-S](https://doi.org/10.1016/0167-2789(92)90102-S). [Online]. Available: <https://www.sciencedirect.com/science/article/pii/016727899290102S> (visited on 12/17/2023).
- [3] W. Huang, T. A. W. Bolton, J. D. Medaglia, D. S. Bassett, A. Ribeiro, and D. Van De Ville, “A graph signal processing perspective on functional brain imaging,” *Proceedings of the IEEE*, vol. 106, no. 5, pp. 868–885, May 2018, Conference Name: Proceedings of the IEEE, ISSN: 1558-2256. DOI: [10.1109/JPROC.2018.2798928](https://doi.org/10.1109/JPROC.2018.2798928). [Online]. Available: <https://ieeexplore.ieee.org/document/8307490> (visited on 11/03/2023).
- [4] E. Pirondini, A. Vybornova, M. Coscia, and D. Van De Ville, “A spectral method for generating surrogate graph signals,” *IEEE Signal Processing Letters*, vol. 23, no. 9, pp. 1275–1278, Sep. 2016, ISSN: 1070-9908, 1558-2361. DOI: [10.1109/LSP.2016.2594072](https://doi.org/10.1109/LSP.2016.2594072). [Online]. Available: <http://ieeexplore.ieee.org/document/7522633/> (visited on 11/15/2023).
- [5] G. Varoquaux and R. C. Craddock, “Learning and comparing functional connectomes across subjects,” *NeuroImage*, vol. 80, pp. 405–415, Oct. 2013, ISSN: 10538119. DOI: [10.1016/j.neuroimage.2013.04.007](https://doi.org/10.1016/j.neuroimage.2013.04.007). arXiv: [1304.3880](https://arxiv.org/abs/1304.3880)[q-bio]. [Online]. Available: <http://arxiv.org/abs/1304.3880> (visited on 12/14/2023).

Appendices

A Parallel with Discrete Fourier Transform

Pour développer l’intuition, et voir plus clairement le parallèle avec la transformée de Fourier discrète, il peut être intéressant de se demander si un graphe permet de retrouver l’interprétation habituelle de la transformée de Fourier discrète d’un signal temporel. Les graphes cycliques permettent cela. En effet, soit un signal x sur le graphe cycle à N sommets C_N . La matrice laplacienne L_N de ce graphe s’écrit, pour $N = 6$

$$L_6 := \begin{bmatrix} 2 & -1 & 0 & 0 & 0 & -1 \\ -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 \\ 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & -1 & 2 & -1 \\ -1 & 0 & 0 & 0 & -1 & 2 \end{bmatrix}$$

En notant

$$J_N := \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & & & & 1 \\ 1 & 0 & 0 & \dots & 0 \end{bmatrix}$$

on observe que

$$L_N = 2I_N - J_N - J_N^{N-1} \quad (4)$$

Or les valeurs propres de J sont les racines N -ièmes de l'unité $\omega^0, \dots, \omega^{N-1}$ où $\omega := e^{\frac{2i\pi}{N}}$ et pour tout $k \in \llbracket 0, N-1 \rrbracket$,

$$X_k := \begin{bmatrix} 1 \\ \omega^k \\ \omega^{2k} \\ \vdots \\ \omega^{(N-1)k} \end{bmatrix}$$

est vecteur propre de J associé à la valeur propre ω^k . Cela permet donc, en notant W la matrice de transformée de Fourier discrète

$$W := \begin{bmatrix} 1 & 1 & 1 & 1 & \dots & 1 \\ 1 & \omega & \omega^2 & \omega^3 & \dots & \omega^{N-1} \\ 1 & \omega^2 & \omega^4 & \omega^6 & \dots & \omega^{2(N-1)} \\ 1 & \omega^3 & \omega^6 & \omega^9 & \dots & \omega^{3(N-1)} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{N-1} & \omega^{2(N-1)} & \omega^{3(N-1)} & \dots & \omega^{(N-1)(N-1)} \end{bmatrix}$$

d'écrire

$$L_N = W(2I_N - \Omega - \Omega^{N-1})W^{-1} \quad (5)$$

où $\Omega := \text{Diag}(\omega^0, \dots, \omega^{N-1})$

Ainsi, les valeurs propres de L_N sont les $\lambda_k = 2 - \omega^k - \omega^{-k}$ pour $k \in \llbracket 0, N-1 \rrbracket$.

En utilisant la définition de la transformée de Fourier sur les graphes [2](#), on explicite le k -ième coefficient de Fourier

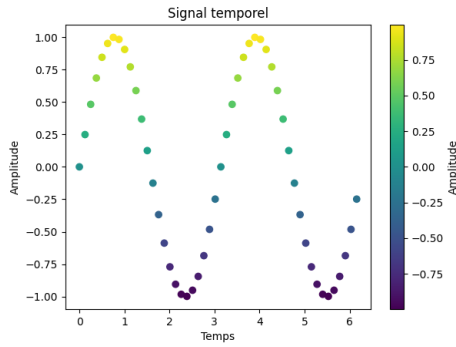
$$\tilde{x}_k = W_k^* x \quad (6)$$

$$= \overline{W_k} x \quad (7)$$

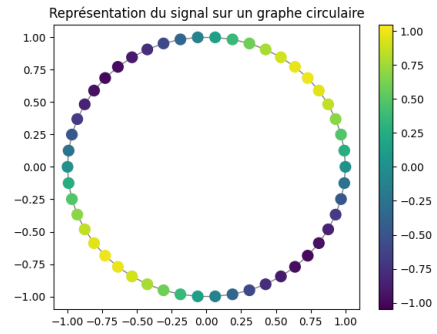
$$= \sum_{n=0}^{N-1} x_n \omega^{-nk} \quad (8)$$

$$= \sum_{n=0}^{N-1} x_n e^{-2i\pi n \frac{k}{N}} \quad (9)$$

On retrouve la définition de la transformée de Fourier discrète.



(a) Représentation temporelle d'un signal échantillonné



(b) Représentation du même signal sur un graphe circulaire

Figure 1: Deux représentations d'un signal échantillonné dont on calcule la transformée de Fourier par les deux définitions dont on dispose : la transformée de Fourier discrète, et la transformée de Fourier d'un graphe

B Data visualization

B.1 MRI Image

Subject 1 at time 10

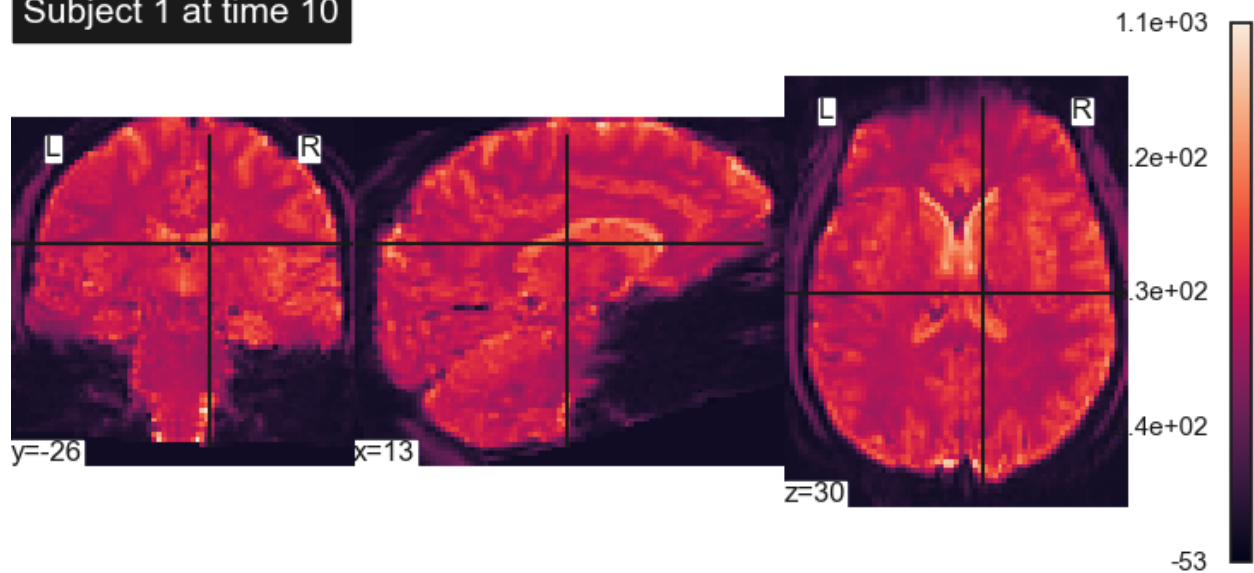
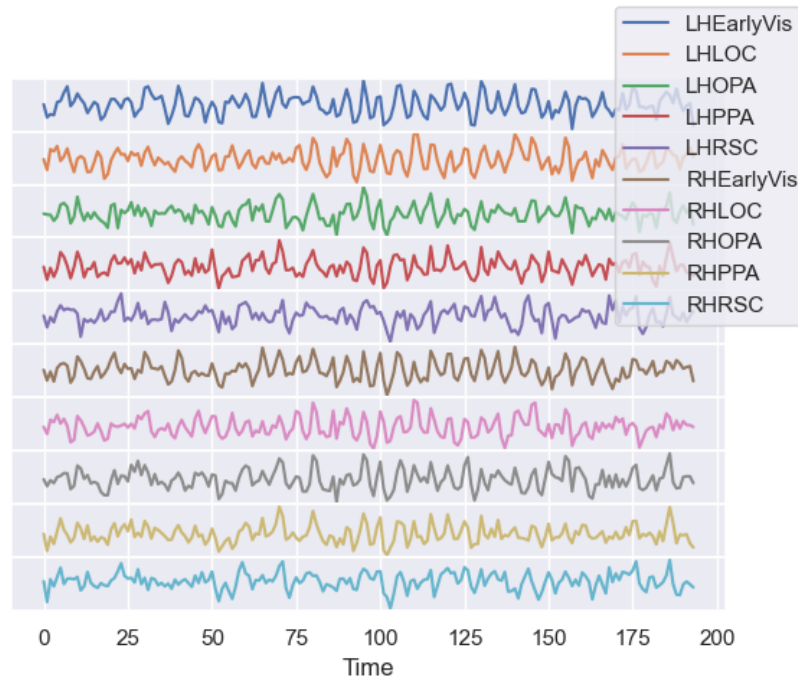
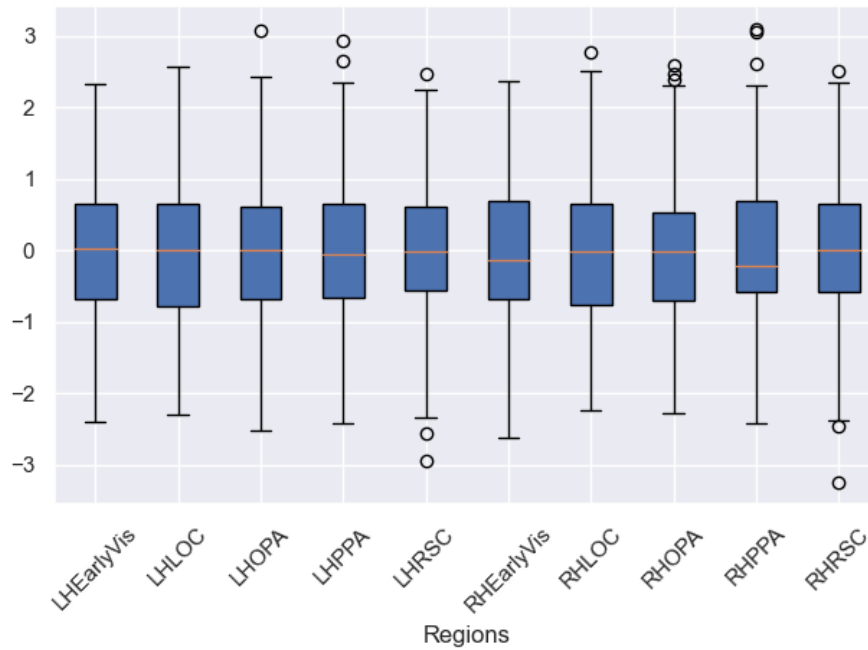


Figure 2: Example of a 3D MRI plotted for the first patient

B.2 Time series diagnosis



(a) Time series in each region of interests



(b) Box plot of the time series

Figure 3: Analysis of the extracted time series. There were 10 ROIs: early visual (EarlyVis), lateral occipital cortex (LOC), occipital place area (OPA), parahippocampal place area (PPA), retrosplenial complex (RSC) for the left hemisphere (LH) and right hemisphere (RH).

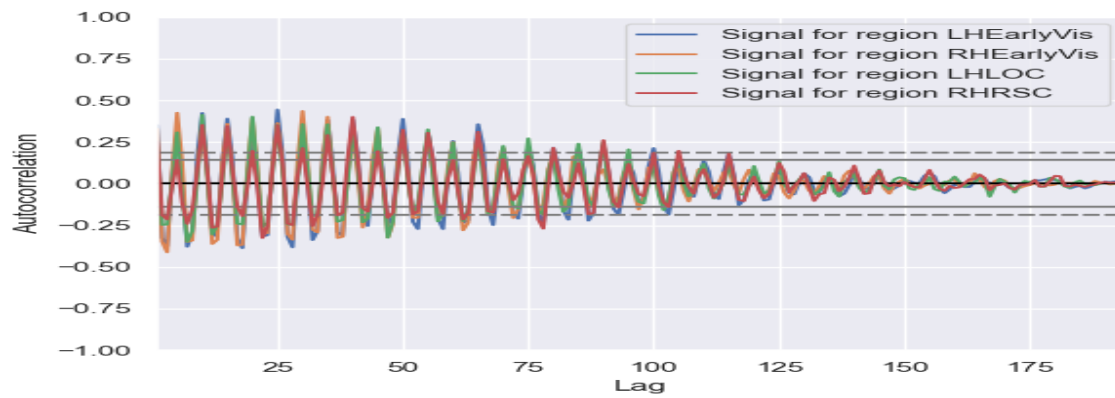
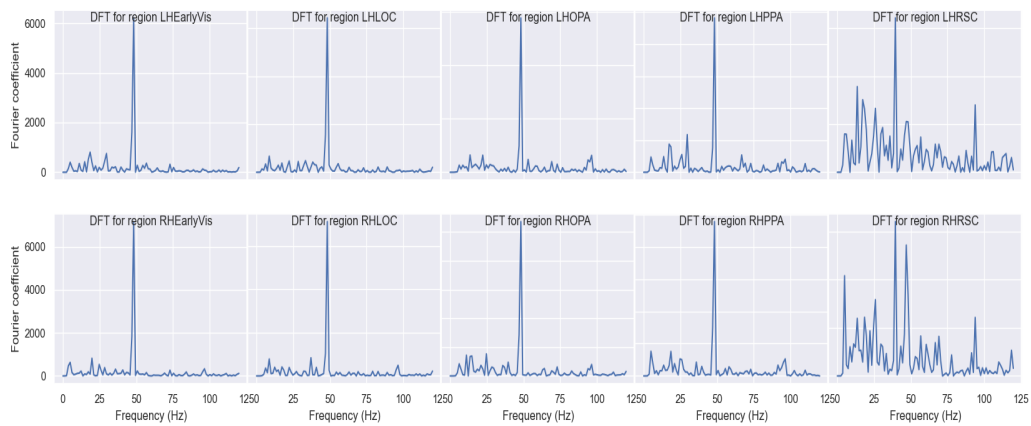
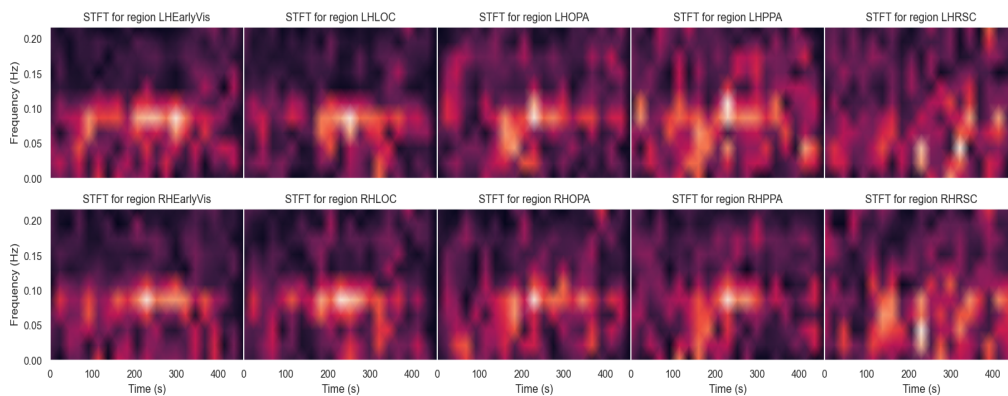


Figure 4: Autocorrelation analysis across various brain regions. Not all results were plotted to enhance readability.



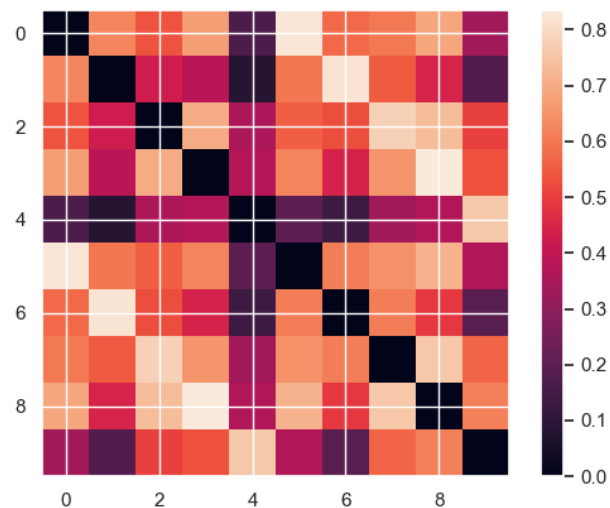
(a) Discrete Fourier Transform (DFT)



(b) Spectrogram

Figure 5: Spectral analysis of the time series for each brain region. (b) The spectrogram was computed with window length of 20 and an overlapping of 16.

C Graph Construction



(a) Connectivity matrix

Brain graph



(b) Functional-connectome graph

Figure 6: Representation of the brain graph. (a) Connectivity matrix for a given patient, which is also the adjacency matrix \mathbf{A} of the graph \mathcal{G} . The entry $A_{i,j}$ of the connectivity matrix is obtained by computing the correlation coefficient between the time series extracted from regions i and j . (b) displays the corresponding *connectome*. This representation offers a unique insight into the patient's neural architecture, highlighting the potential correlates of cognitive processes

D Filters

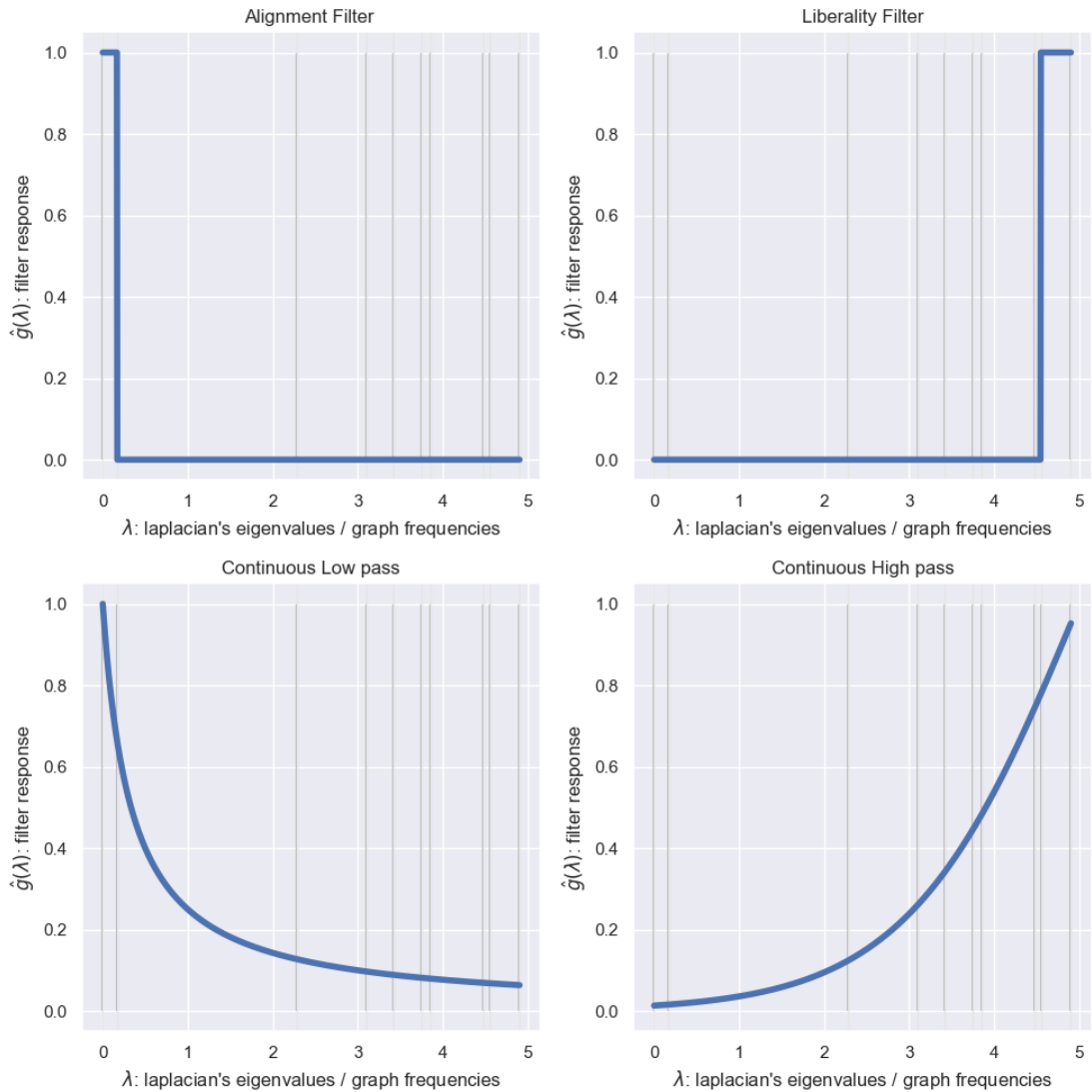
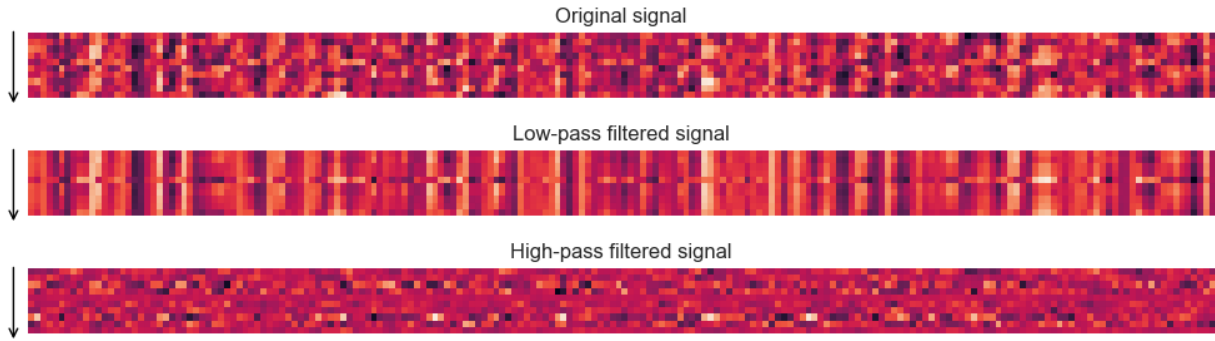
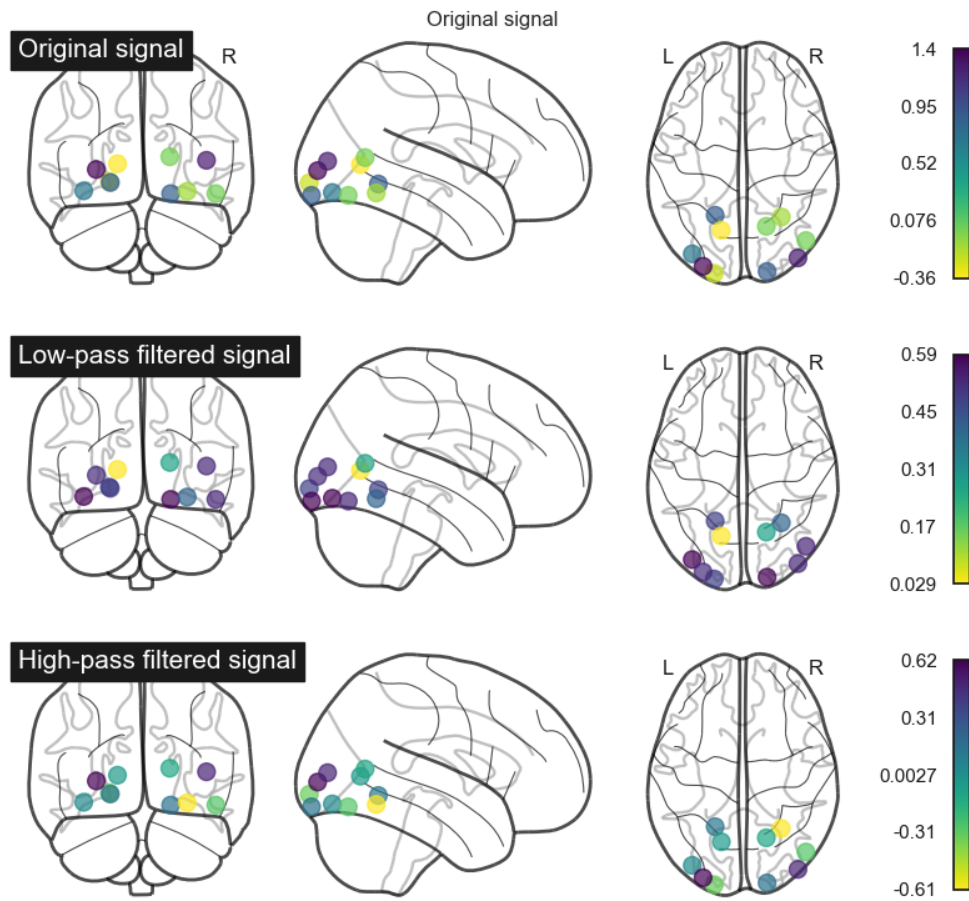


Figure 7: Comparison of different filters for graph signal analysis. The *alignment filter* and the *liberality filter* are the low-pass filter and high-pass filter, respectively, employed in the article[3]. The authors defined their filters by considering the 10% lowest and largest eigenvalues, respectively. In our context, this comes down to filtering approximately 2 eigenvalues.



(a) Filtered signal Y_H



(b) Filtered signal at time $t = 0$ in the brain graph.

Figure 8: Visualization of the filtered signals at each time point in (a) and on the graph in (b). (a) The down arrow in y -axis indicates the brain regions. The x -axis corresponds to the time axis. (b)

E Test statistic

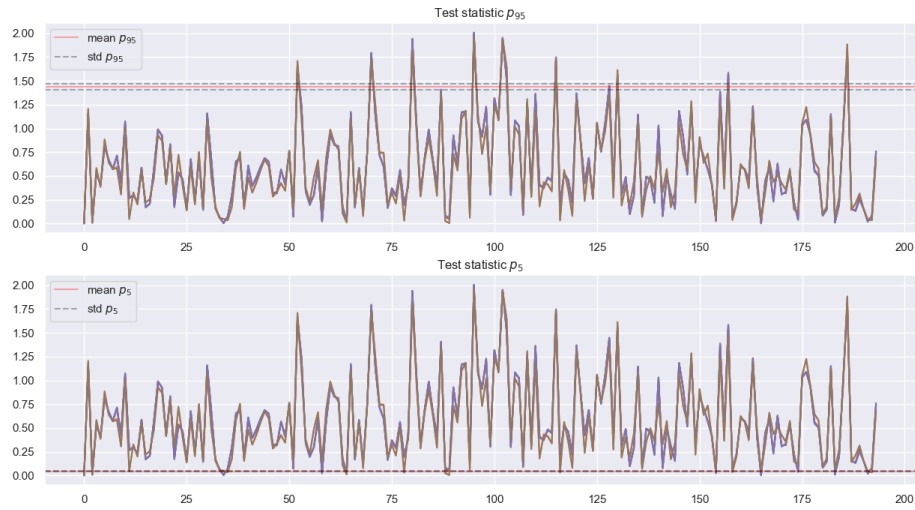


Figure 9: Threshold computed on the surrogate signals. p_{95} and p_5 are respectively the average of the 95% and 5% (for significant low amplitude) percentile on the surrogate signals.

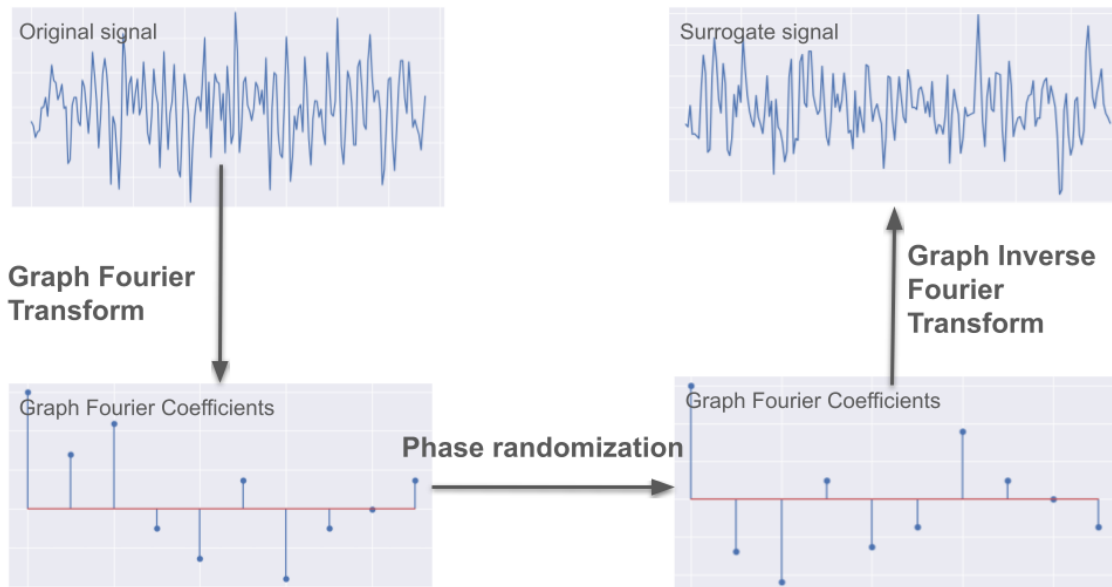
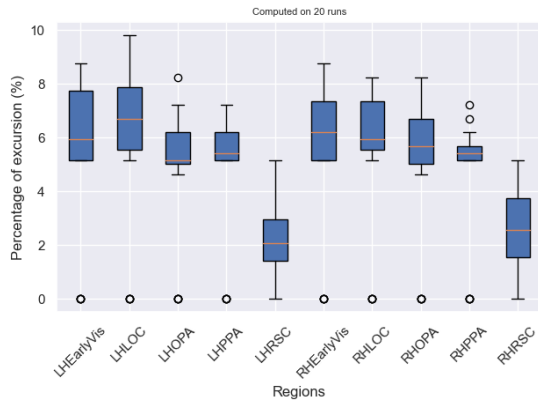
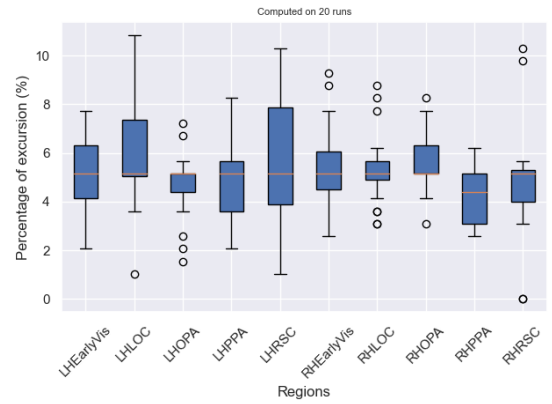


Figure 10: Surrogate signal generation scheme.

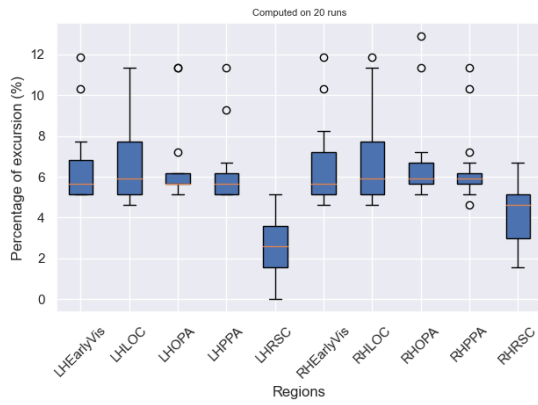
F Results on computing excursions



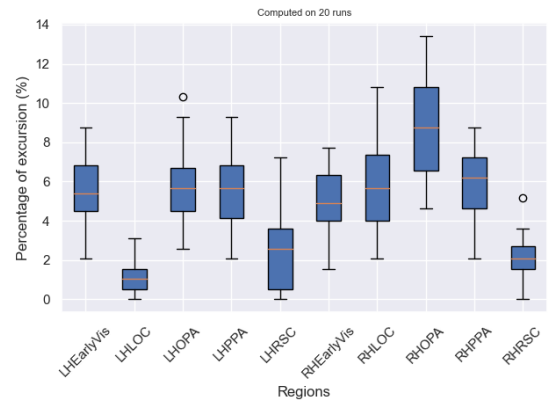
(a) Article-inspired low-pass filter



(b) Article-inspired high-pass filter

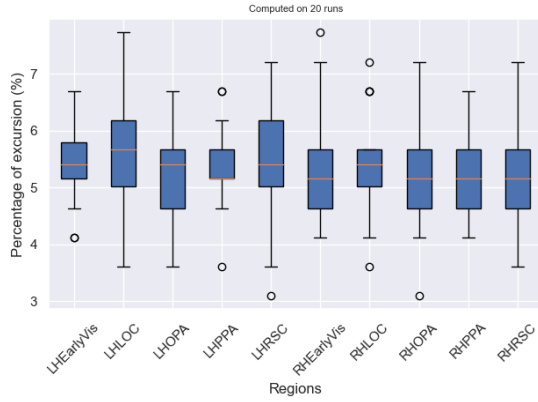


(c) Continuous low-pass filter

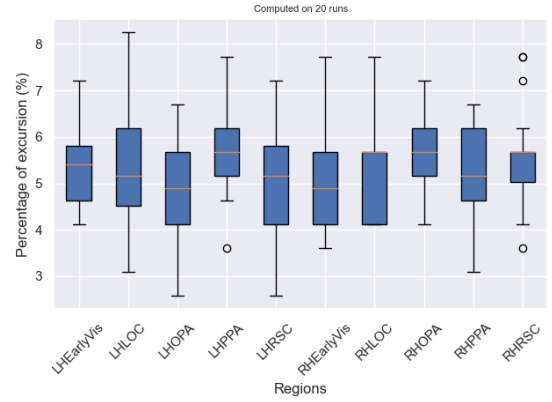


(d) Continuous high-pass filter

Figure 11: Comparison of different filters for computing excursions in alignment and liberality regimes. The method outlined in section 4.1 were repeated on 20 different fMRI of the same patient. For each run, we generate 1,000 null surrogate graph signal.



(a) Continuous low-pass filter



(b) Continuous high-pass filter

Figure 12: Results using the conventional phase Fourier randomization in the temporal domain.

G Accuracy of the k-NN with DTW

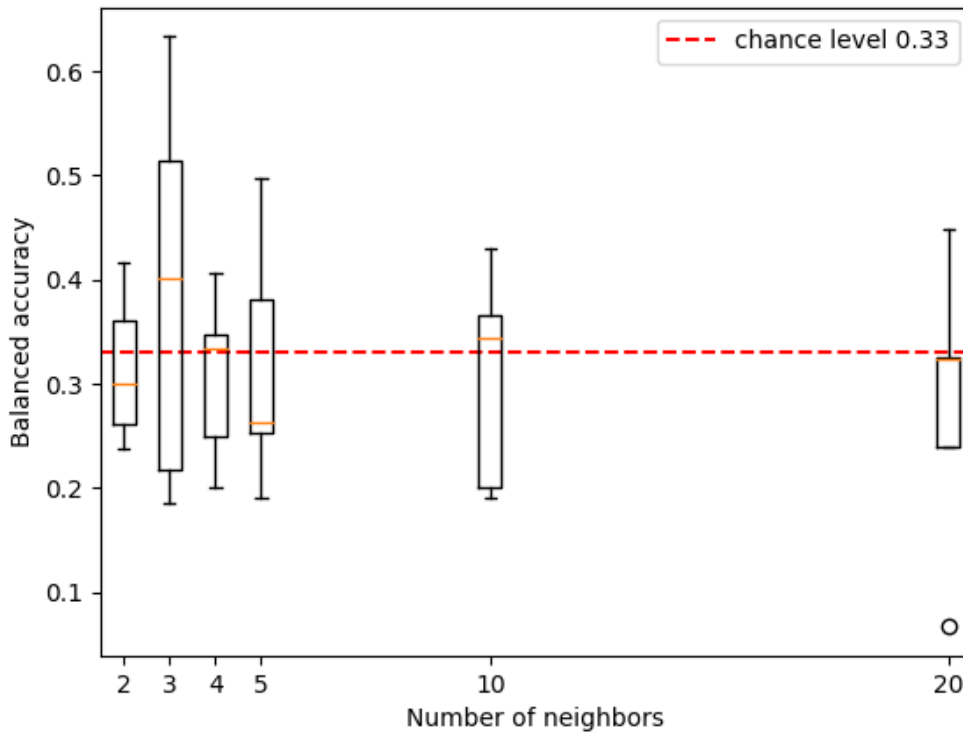


Figure 13: 5-fold cross-validation of the model combining DTW and k-neighbours classifier.