Brain Oscillatory and Network Activity during resting states

Siddhant Tandon[1,*]

[1]Data Science Student, La Sapienza, Roma
[*]siddhant.tandon203@gmail.com, Matricola:1771650

ABSTRACT

The project focuses on understanding brain functionality related to the Eye Movement. More Specifically, we analyze Brain Signals recorded by EEG for two different Eye states i.e. open and close respectively. The analysis involves different methodologies including Spectral Analysis, Graph Connectivity, Motif Analysis and community detection. I make a brief comparison of two Eye States for each of these methodologies. Finally, I discuss the overall understanding and learning outcome of the project.

Introduction

Understanding brain functionality has been a major interest among the computational neuroscientists. Over the year, different methodologies have been introduced and implemented to accomplish that. In this project, I study the brain functionality related to Eye Movement. We receive Data Set from the following URL https://physionet.org/physiobank/database/eegmmidb/. The Data set contains the EEG brain signals for a subject in two states as mentioned above. I combine methodologies from Signal processing and Graph Theory to understand the brain signal and connectivity pattern in the brain. The list of the specific tasks is present in Table 1. Analysis comprises of the brief comparison of two eye states involving each methodology.

Description of Dataset & Tasks

The Data Set was obtained from https://physionet.org/physiobank/database/eegmmidb/ involving only the subject '064' and two eye states namely S064R01.edf & S064R02.edf. As mentioned in the Homework assignment, the dataset was read through pyedflib library. The dataset consists of 64 columns corresponding to different channels for EEG as mentioned in the handout of homework assignment.

For the Spectral Analysis, the mandatory task along with three other optional ones of class C and higher were performed. For the rest of the sections, only the mandatory tasks were implemented. The following table shows the tasks were performed in the analysis.

Task number	Category	Class
1.1	Mandatory	-
1.3	Optional	C
1.4	Optional	D
1.5	Optional	E
2.1	Mandatory	-
3.1	Mandatory	-
4.1	Mandatory	-
5.1	Mandatory	-

Table 1. Table showing serial number of tasks performed

Data Analysis

In this section I will talk about each of the individual tasks in detail and will also mention the relevant results under each subsection.

1.Spectral Analysis

1.1 Select relevant channel and estimate Power Spectral Density (PSD) using one of the methods introduced during the course. Justify your choice of channel selection and parameters used for PSD estimation

In this task, Power Spectral Density estimation is performed using Welch's Method. I used the Python routine https://docs.scipy.org/doc/scipy-0.14.0/reference/generated/scipy.signal.welch.html to accomplish this. The choice for channel selection is relatively trickier but since the location of visual cortex lies in the occipital lobe, I came up with the channels 'O1','O2','Oz' as the most relevant choice and therefore I perform a PSD estimation for all of these three channels for both states Eyes open and Eyes closed. The parameters involved in the PSD estimation were Sampling Frequency Fs which was 160 Hz. The length of each segment *nperseg* was set at 256. *Hanning window* was applied to the signal before taking the DFT and also the parameter to detrend the signal was set as True.

Parameters for PSD, Fs:160 Hz, Window:Hanning , nperseg : 256 , detrend : constant. Figure 1 displays the PSD plot for channel O1 for both states Eyes Open and closed. Other relevant figures are attached at the very end of the document.

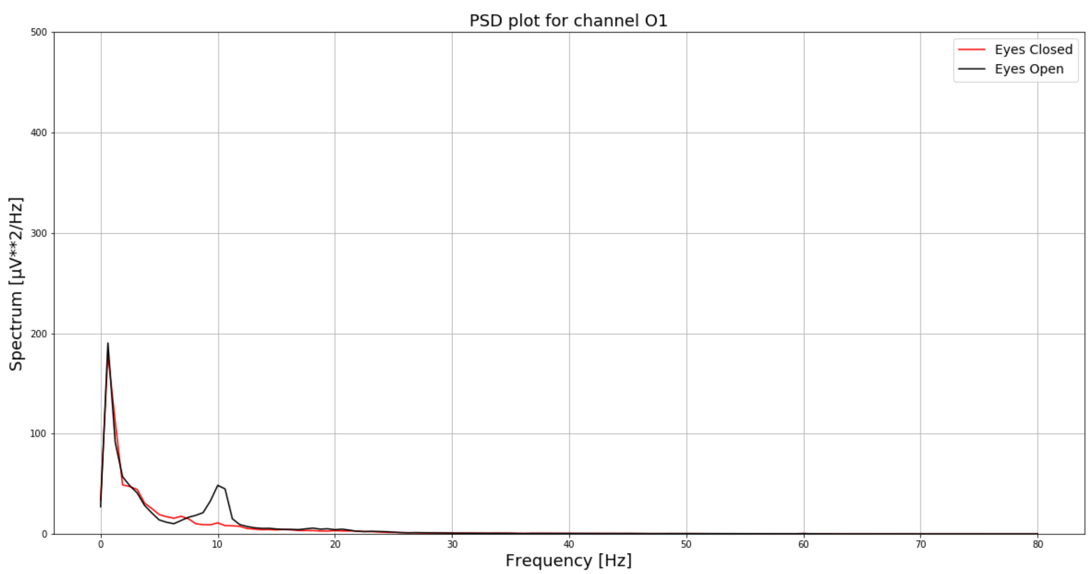


Figure 1. PSD plot for channel O1. Black curve represents Eyes open state and red represents Eyes closed. It can be seen that for the eyes open state the frequency component 10 contributes to the PSD.

1.3 Compare different choices of estimation parameters

To experiment with different choices of hyperparameters, a total of 120 configurations or 120 PSDs were obtained , in a basic gridsearch manner. The parameters which were varied were nfft or the number of points in DFT calculation, choice of windowing function, length of each segment, and overlap between each segment.

- nfft : [512,1000,2000,5000,10000]
- window : ['blackman', 'hamming', 'flattop']
- nperseg : [128, 256, 384, 512]
- overlap: [64, 128, 192, 256, 42, 85, 128, 170]

Figure 2 displays the multiple PSD plots for channel O1 for state Eyes Open. Each of the curve is a PSD with corresponding hyperparameters.

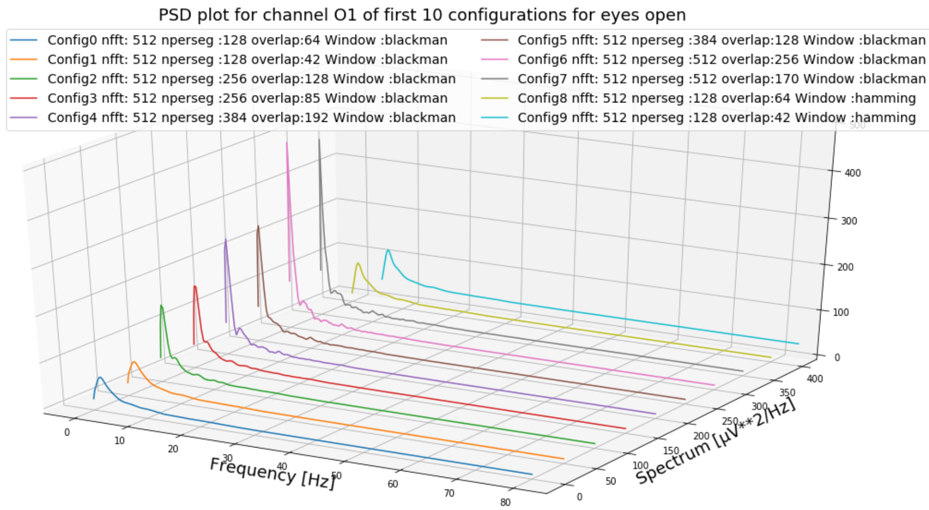


Figure 2. Plot showing PSD with the first 10 configurations for channel O1 and state Eyes Open, each having a different choice of hyperparameter.

1.4 Make a multivariate comparison of PSDs, considering simultaneously two independent variables (channels resting state)

A multivariate comparison of PSDs was done through Principal Component Analysis. I wanted to find out which of the channels can contribute the most for any estimator of this data. PCA expresses the variability of the data through hidden or principal components. It finds out how much the covariates contribute to variance of each of the principal components. To keep things simple, I kept the number of principal components equal to three (as the first components are the most significant.) Figure 3 shows the importance of the channels in PCA analysis.

Variable Importance for Eyes Open

	PCA1	PCA2	PCA3
Po8.	134.039307	70.126839	11.433230
O1..	134.045450	70.107750	11.493594
Oz..	134.044922	70.113401	11.477477
O2..	134.042660	70.111585	11.453044
Iz..	134.026040	70.126709	11.495579

Variable Importance for Eyes Closed

	PCA1	PCA2	PCA3
Po8.	20.641912	10.427421	8.334263
O1..	20.645442	10.424099	8.283701
Oz..	20.646813	10.428414	8.300129
O2..	20.642530	10.408385	8.318402
Iz..	20.626269	10.457613	8.251803

Figure 3. Figure showing the importance of channels or variables in the PCA analysis. A channel with higher value means they contribute more to the Principal Component.

I also make a comparison of PSD for channels O1 and O2, for each of the eyes state. The relevant figure is attached at the bottom of the report to save space.

1.5 Assign a statistical significance value to the differences in PSDs between the two rest conditions.

In statistics, statistical significance essentially means building Confidence Intervals for the estimator of your data. The confidence interval obtained indicates your confidence in the obtained results after repeating the experiment large number of times. The difference between PSD of two channels is thus one kind of an estimator. One efficient technique to construct Confidence Intervals without any assumption of the underlying distribution of your data , is Nonparametric Bootstrapped Confidence Intervals. In this one has to sample multiple number of times from the data obtained during the experiment. These large number of samples are analogous to repetition of the experiment large number of times. For every bootstrapped sample, the favorable point estimate is computed and confidence intervals are constructed using the mathematical definition[1]. I calculated the normal CI and percentile CI both of which had very tight bound on the estimator. One can also calculate pvalues by measuring the area between true mean and right limit of confidence interval. In the next figure 4 , i graphically show the normal confidence interval for channel O1.

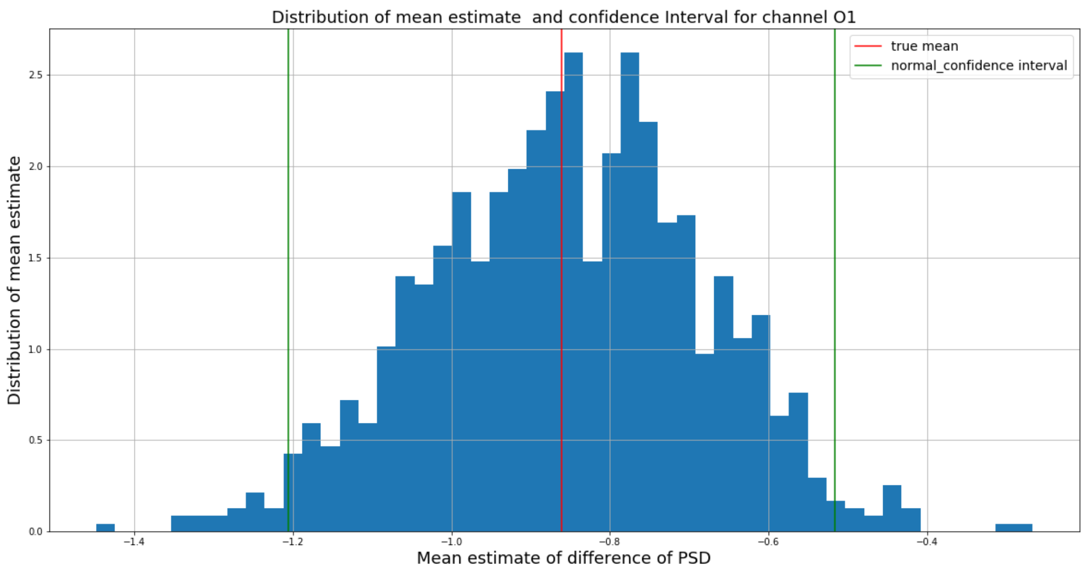


Figure 4. Figure showing the distribution of mean estimate and the obtained confidence interval shown in green and the true mean in red.

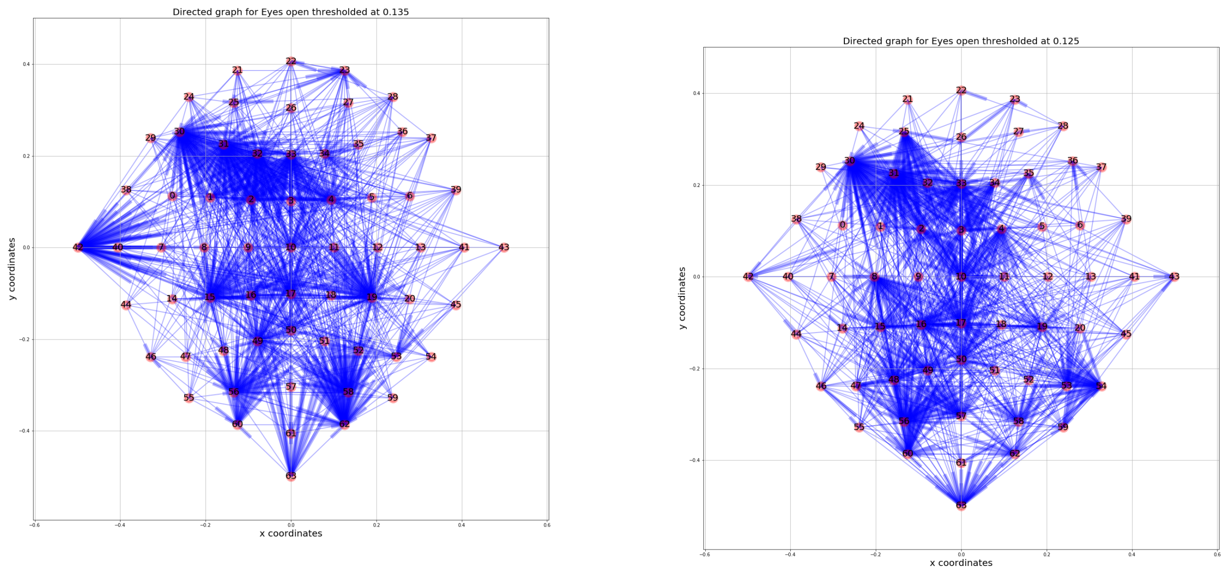
2. Connectivity Graph

2.1 Estimate functional brain connectivity among 64 channels using one of the MVAR estimators: Partial Directed Coherence (PDC), Direct Transfer Function (DTF). Select one relevant frequency value. Apply a threshold so that the resulting binary connectivity matrices have network density equal to 20%

The brain Connectivity estimation was performed using Direct Transfer Function with the help of a Python routine that can be found at http://connectivipy.readthedocs.io/en/latest/tutorial.html. I found out that the optimal order is 5 using AIC model order. The results returned by the Python routine were in the dimension of (160,64,64) as the sampling frequency was 160 Hz.

The threshold which was used to make the network density equal to 20% for eyes open was 0.135 and for eyes closed was 0.125.

Figure 5 shows the connectivity pattern or the network graph among the channels thresholded at 20% for eyes open.



(a) Figure showing the directed graph for the eyes open state. The nodes represent the channels and edges are the values of DTF between those channels.

(b) Figure showing the directed graph for the eyes closed.

Figure 5. Connectivity patter of thresholded graphs for both states eyes open and closed

3. Graph Theory Indices

3.1 Compute binary global (clustering coefficient, path length) and local (degree, in/out-degree) graph indices. List the highest 10 channels for local indices.

Global Characteristics Clustering Coefficient of a graph is the average over all the local clustering coefficients. The local clustering coefficient tells how important is that node in terms of making 'knots' or clusters around itself. Somehow interpreted as a highly important node. In Networkx this can be done by using the function average clustering. For eyes open the measure was found out to be 0.7449 slightly higher than eyes closed which was 0.69.

Average shortest path length is the average over the shortest paths for all possible 2 node combinations. This measure for found out to be almost same for eyes open as 1.65 and eyes closed as 1.66.
Local Characteristics Three measures for every node were calculated namely degree of the node, in degree of the node and out degree of the node. In degree represents the number of edges incident into that node and out degree are the number of edges going out of that node. For all three measures channels 'F5', 'F3' were top two channels for highest degree and highest in degree, while for out degree channel 'Af8' was the top one.

The following figure 6show the top 10 nodes for all three measures.

Eyes open

Top 10 nodes with highest in_degree for eyes open

	channel name	in_degree	index_num
0	F5..	64	30
1	F3..	64	31
2	F1..	56	32
3	Cp4.	54	19
4	Cp3.	53	15
5	Fz..	47	33
6	O2..	46	62
7	T9..	42	42
8	Po4.	40	58
9	P1..	38	49

Top 10 nodes with highest out_degree for eyes open

	channel name	index_num	out_degree
0	Af8.	28	20
1	Fp2.	23	19
2	Ft8.	39	18
3	F6..	36	17
4	F8..	37	17
5	C2..	11	16
6	Fp1.	21	16
7	Af7.	24	16
8	Af4.	27	16
9	F4..	35	16

(a) Top 10 nodes for in degree and out degree for eyes open

Eyes closed

Top 10 nodes with highest in_degree for eyes closed

	channel name	in_degree	index_num
0	F5..	64	30
1	F3..	64	31
2	Fz..	48	33
3	Po3.	45	56
4	Cpz.	38	17
5	Cp1.	35	16
6	Cz..	33	10
7	Pz..	33	50
8	C3..	30	8
9	P3..	30	48

Top 10 nodes with highest out_degree for eyes closed

	channel name	index_num	out_degree
0	P7..	46	22
1	Tp8.	45	18
2	P5..	47	18
3	P2..	51	18
4	P8..	54	18
5	P3..	48	17
6	P1..	49	17
7	P6..	53	17
8	Iz..	63	17
9	C6..	13	16

(b) Top 10 nodes for in degree and out degree for eyes closed

Figure 6. Figure showing the local characteristics for the top 10 nodes for both eye states. Channels 'F5' and 'F3' have most of the incoming edges.

4. Motif Analysis

4.1 Perform motifs analysis to investigate the presence of 3-node configurations in the networks: determine their frequency and statistical significance (motifs, anti-motifs).

Motif Analysis is done to detect the presence of important node configurations in a network. These node configurations maybe 3 node or 4 node. For example, given a directed graph and three set of nodes of subgraph, A,B,C. Because these nodes are connected there are many possible ways these can interact with each other. There might be a path from A to C through B, or B to C through A or both A and B leading to C. The frequency of all such ways of interaction of these three nodes is calculated. This is done for all the possible three node configurations in the network. This process is repeated for many samples of random graphs. Based on this, a statistical significance is assigned to this estimator, in quite the same fashion as that of nonparametric bootstrap. In such a way one can obtain the z scores and P values. The node configurations with very low P values are called Anti Motifs with P values less than 0.05. Such a low P value tells there is very significant difference in the experiment conducted. Statistically speaking , we reject our null hypothesis which was true for P value greater than 0.05.

The output of mfinder software labeled the group 46 and 238 as Antimotifs for both states eyes open and eyes closed. The following figures 7 shows the output obtained along with the statistical measures.

Motifs/Under represented nodes for Eyes Open

	motif_id	freq	stat_significance	+-error	z_score	p_value
0	6.0	464.0	561.7850	30.2958	-3.2277	1.00
1	12.0	635.0	688.6550	23.2709	-2.3057	0.99
2	14.0	32.0	67.2067	7.9771	-4.4135	1.00
3	36.0	8264.0	8462.2450	34.8425	-5.6897	1.00
6	74.0	1343.0	1316.3933	27.1217	0.9810	0.16
7	78.0	25.0	54.2500	6.7170	-4.3546	1.00
8	98.0	4.0	11.5300	4.5609	-1.6510	0.99
9	102.0	62.0	89.6225	8.7139	-3.1699	1.00
10	108.0	1391.0	1398.7900	16.1785	-0.4815	0.71
11	110.0	161.0	160.7900	11.1440	0.0188	0.50

Motifs/Under represented nodes for Eyes closed

	motif_id	freq	stat_significance	+-error	z_score	p_value
0	6.0	550.0	654.0950	24.7139	-4.2120	1.00
1	12.0	1337.0	1342.4200	25.3046	-0.2142	0.57
2	14.0	139.0	212.2600	16.2789	-4.5003	1.00
3	36.0	6104.0	6203.3050	29.8620	-3.3255	1.00
4	38.0	1672.0	1658.2200	26.2104	0.5257	0.24
6	74.0	1345.0	1511.0700	24.9855	-6.6466	1.00
7	78.0	72.0	102.1800	9.1752	-3.2893	1.00
8	98.0	56.0	58.7667	6.6708	-0.4147	0.72
9	102.0	180.0	195.0075	11.5837	-1.2956	0.89
11	110.0	225.0	220.7700	12.3439	0.3427	0.40

(a) Motif ids for both states eyes open and closed.

Anti Motifs/Over represented nodes for Eyes Open

	motif_id	freq	stat_significance	+-error	z_score	p_value
4	38.0	1541.0	1468.8867	32.3666	2.2280	0.02
5	46.0	79.0	48.0950	5.4700	5.6499	0.00
12	238.0	91.0	81.3200	3.4667	2.7923	0.01

Anti Motifs/Over represented nodes for Eyes closed

	motif_id	freq	stat_significance	+-error	z_score	p_value
5	46.0	210.0	176.5475	7.9353	4.2156	0.00
10	108.0	867.0	792.8350	13.7835	5.3807	0.00
12	238.0	54.0	47.4633	3.5483	1.8422	0.03

(b) Anti Motif ids for both states eyes open and closed. These anti motifs indicate these are highly significant.

Figure 7. Motif Analysis using mfinder software

5. Community Detection

5.1 Determine number and composition (i.e. list of nodes) of the communities obtained applying one of the algorithms introduced during the course.

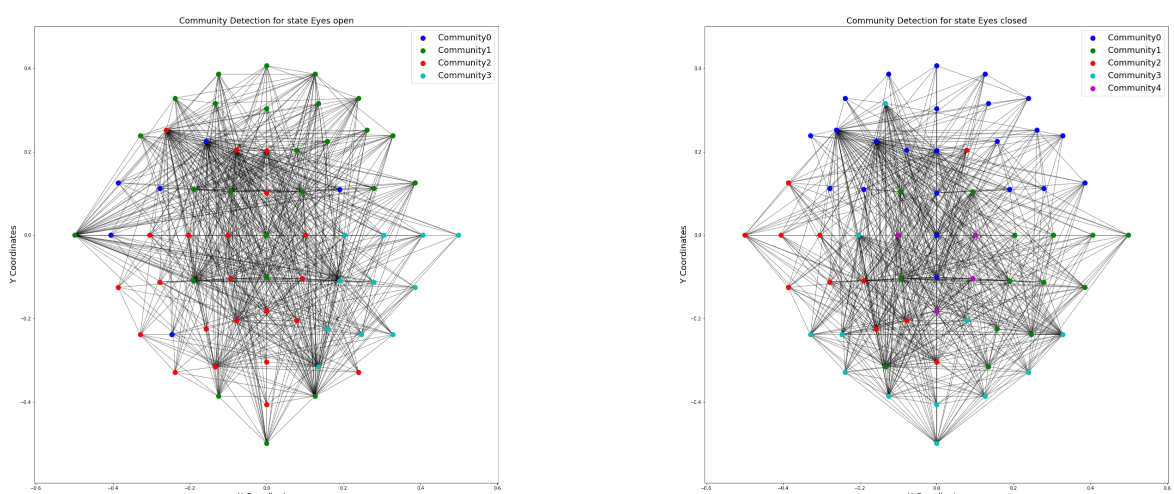
The aim of community detection algorithms is to find out how nodes interact amongst each other in a network. The network is said to have communities within itself if network can be easily grouped into set of nodes where each node is densely connected internally. The clusters so formed are a great deal of importance in the field of social and biological networks. In our analysis, these nodes represent the channels placed on the lobes of the brain. So communities formed after the analysis would essentially tell which areas of the brain interact with which other areas the most. Community analysis is carried out on the thresholded graphs by using the best partition function present in Networkx. The underlying algorithm is Louvain clustering which maximizes the modularity and the partitions are computed using Louvain heuristics.

Figure 8 displays the clusters obtained for both states eyes open and closed.

```
Communities for State eyes open

Community_id 0   Members {0, 5, 38, 40, 47, 31}
Community_id 1   Members {1, 2, 4, 6, 10, 15, 17, 21, 22, 23, 24, 25, 26, 27, 28, 29, 34, 35, 36, 37, 39, 42, 60, 6
2, 63}
Community_id 2   Members {3, 7, 8, 9, 11, 14, 16, 18, 30, 32, 33, 44, 46, 48, 49, 50, 51, 55, 56, 57, 59, 61}
Community_id 3   Members {41, 43, 12, 13, 45, 19, 20, 52, 53, 54, 58}


Communities for State eyes closed

Community_id 0   Members {0, 1, 3, 5, 6, 10, 17, 21, 22, 23, 24, 26, 27, 28, 29, 30, 31, 32, 33, 35, 36, 37, 39}
Community_id 1   Members {2, 4, 41, 43, 12, 13, 45, 16, 19, 20, 52, 53, 56, 58}
Community_id 2   Members {34, 38, 7, 40, 42, 44, 14, 15, 48, 49, 57}
Community_id 3   Members {8, 46, 47, 51, 54, 55, 25, 59, 60, 61, 62, 63}
Community_id 4   Members {9, 18, 11, 50}
```

Figure 8. Figure showing communities obtained by louvain clustering. Notice that the channels belonging to the occipital lobes are in the same cluster.

Visualization of the communities in the graph is presented below in the figure 9. The nodes belonging to the same community have the same color.



(a) Community structure for state Eyes open. Total 4 communities were found out.

(b) Community structure for state Eyes open. Total 5 communities were found out.

Figure 9. Community detection using Louvain Clustering

References

1. Wasserman, L. All of statistics. *Springer* **20**, 129–141 (2004).

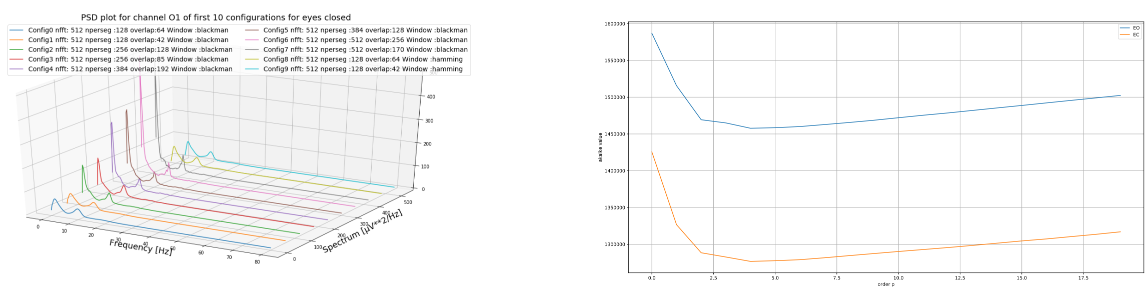
Additional information

Other relevant figures of the analysis are included here.

(a) DTF Values among the channels for state Eyes open

(b) DTF Values among the channels for state Eyes closed

Figure 10. 3D Plots for DTF values of section Connectivity Graph



(a) Figure showing multiple PSD plots for channel O1 of first 1o configurations for section 1.3

(b) Figure showing AIC plot for model selection for section 2.1 connectivity graph

Figure 11. PSD plots and AIC model criteria

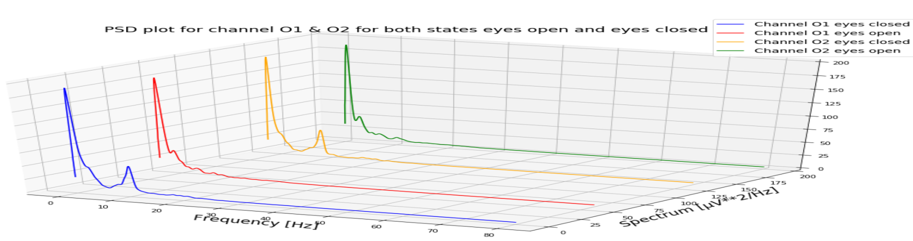


Figure 12. Figure showing comparison of PSD for channel O1 and O2 for both states eyes open and closed. This is for section 1.4