



Signals and Systems Project  
Instructor: Prof. Hamid Aghajan  
Sharif University of Technology

Analysis of Phase Locking Value during Olfactory Stimulation  
as a Biomarker for Alzheimer's Disease in EEG Signals

Author : Pantea Amoie 400101656

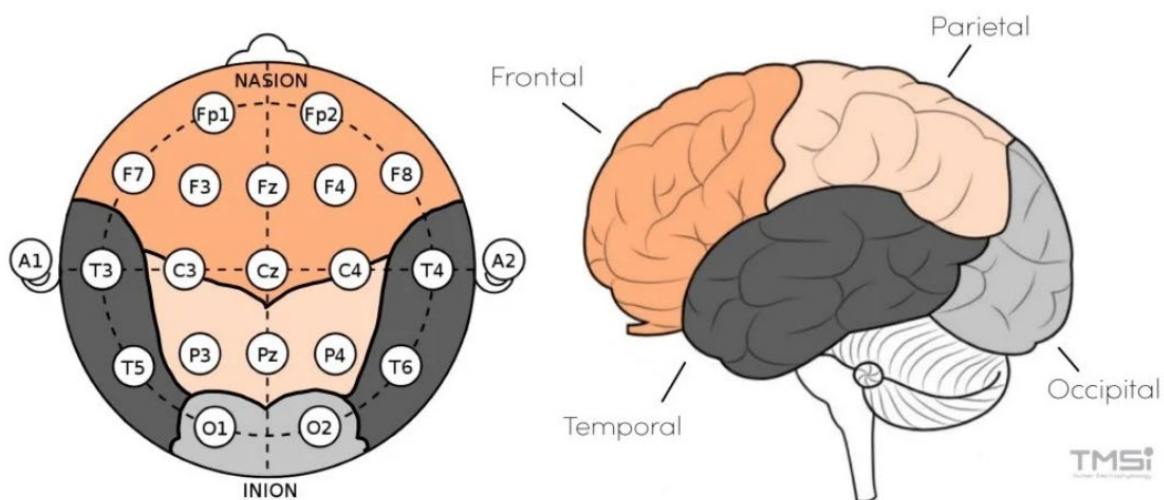
## 2 Electroencephalography

### 2.1 What is EEG?

Based on the picture above, What does each electrode's name stand for? Explain the naming method used in the 10-20 EEG system.

The letter of the electrode stands for the general brain region that the electrode covers. From front to back, the electrode letter labeling is as follows: Fp (pre-frontal or frontal pole), F (frontal), C (central line of the brain), T (temporal), P (parietal), and O (occipital). Electrodes lying between these lines combine multiple letters, ordered from front to back. In addition, the letters M and A are sometimes used to refer to the mastoids or earlobes respectively. Typically, these locations are included to serve as a (offline) reference for signal analysis.

The number of the electrode gives information about the distance from the electrode to the midline of the brain. At the midline, the electrodes are labeled with a 'z' to represent zero. The electrode numbers increase as you move away from the midline. Odd numbers represent electrodes on the left hemisphere and even numbers represent electrodes on the right hemisphere.<sup>[1]</sup>

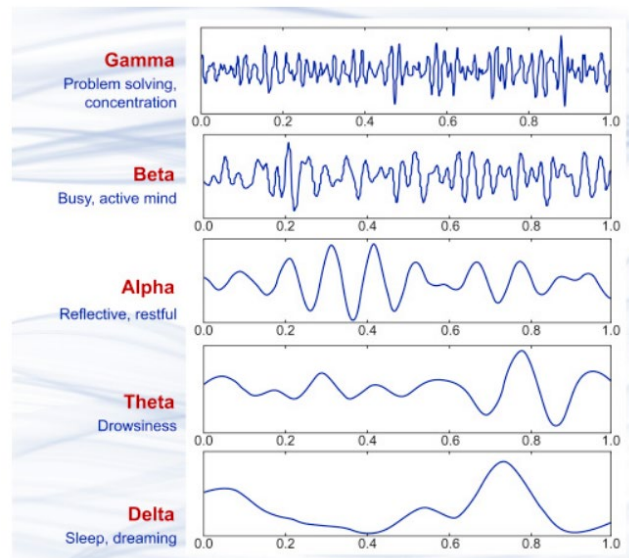


## 2.3 Frequency Bands of EEG

Determine the activities each frequency band is associated with.

- **Delta:** has a frequency of 3 Hz or below. It tends to be the highest in amplitude and the slowest waves. It is normal as the dominant rhythm in infants up to one year and in stages 3 and 4 of sleep. It may occur focally with subcortical lesions and in general distribution with diffuse lesions, metabolic encephalopathy hydrocephalus or deep midline lesions. It is usually most prominent frontally in adults (e.g. FIRDA - Frontal Intermittent Rhythmic Delta) and posteriorly in children e.g. OIRDA - Occipital Intermittent Rhythmic Delta).
- **Theta:** has a frequency of 3.5 to 7.5 Hz and is classified as "slow" activity. It is perfectly normal in children up to 13 years and in sleep but abnormal in awake adults. It can be seen as a manifestation of focal subcortical lesions; it can also be seen in generalized distribution in diffuse disorders such as metabolic encephalopathy or some instances of hydrocephalus.
- **Alpha:** has a frequency between 7.5 and 13 Hz. Is usually best seen in the posterior regions of the head on each side, being higher in amplitude on the dominant side. It appears when closing the eyes and relaxing, and disappears when opening the eyes or alerting by any mechanism (thinking, calculating). It is the major rhythm seen in normal relaxed adults. It is present during most of life especially after the thirteenth year.

- **Beta:** beta activity is "fast" activity. It has a frequency of 14 and greater Hz. It is usually seen on both sides in symmetrical distribution and is most evident frontally. It is accentuated by sedative-hypnotic drugs especially the benzodiazepines and the barbiturates. It may be absent or reduced in areas of cortical damage. It is generally regarded as a normal rhythm. It is the dominant rhythm in patients who are alert or anxious or have their eyes open.[2]



Frequency band	Frequency	Brain states
Gamma ( $\gamma$ )	$>35$ Hz	Concentration
Beta ( $\beta$ )	12–35 Hz	Anxiety dominant, active, external attention, relaxed
Alpha ( $\alpha$ )	8–12 Hz	Very relaxed, passive attention
Theta ( $\theta$ )	4–8 Hz	Deeply relaxed, inward focused
Delta ( $\delta$ )	0.5–4 Hz	Sleep

## 2.4 Sampling frequency

Based on frequency bands and Nyquist criterion, which sampling frequencies are preferred for EEG signals?

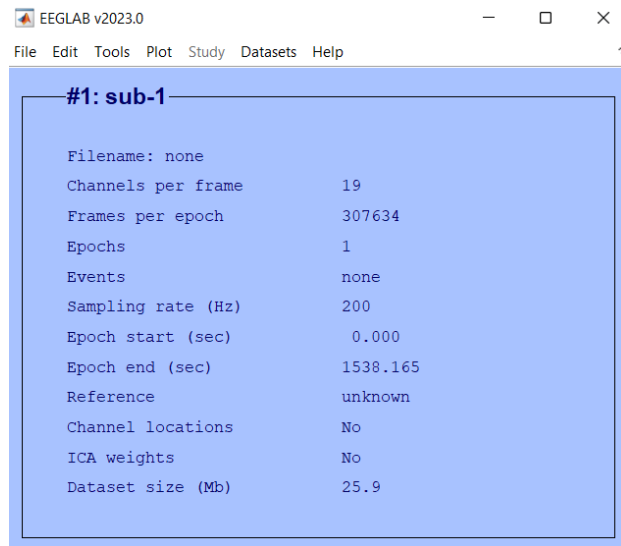
Due to Nyquist criterion:  $f_{dig} > 2f_{max}$  where  $f_{dig}$  is the digitization rate or sampling rate and  $f_{max}$  is the highest frequency present in the time series.

Since the effective bandwidth of EEG signals is  $\sim 100$  Hz, a minimum frequency of 200 Hz is typically enough to sample the EEG for most applications. (It is sometimes better to use a bit higher sample frequencies because of the noises on the signal.)

## 3 EEG Signal Processing

### 3.3 Pre-Processing

Importing data to eeglab :



EEGLAB v2023.0

<b>#1: sub-1</b>	
Filename:	none
Channels per frame	19
Frames per epoch	307634
Epochs	1
Events	none
Sampling rate (Hz)	200
Epoch start (sec)	0.000
Epoch end (sec)	1538.165
Reference	unknown
Channel locations	No
ICA weights	No
Dataset size (Mb)	25.9



EEGLAB v2023.0

<b>#1: sub2</b>	
Filename:	none
Channels per frame	19
Frames per epoch	308547
Epochs	1
Events	none
Sampling rate (Hz)	200
Epoch start (sec)	0.000
Epoch end (sec)	1542.730
Reference	unknown
Channel locations	No
ICA weights	No
Dataset size (Mb)	25.9

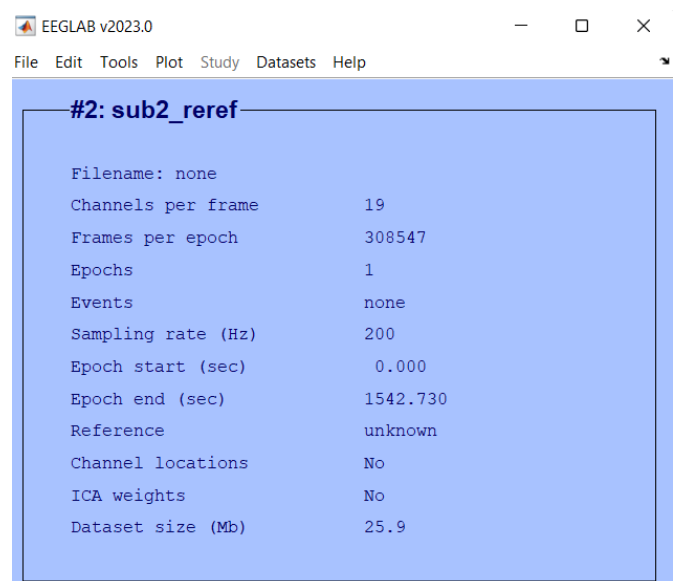
Step 1:

To preprocess using EEGLAB, first re-reference data to the mean of the channels.



EEGLAB v2023.0

<b>#2: sub-1_reref</b>	
Filename:	none
Channels per frame	20
Frames per epoch	307634
Epochs	1
Events	none
Sampling rate (Hz)	200
Epoch start (sec)	0.000
Epoch end (sec)	1538.165
Reference	unknown
Channel locations	No
ICA weights	No
Dataset size (Mb)	27.1



EEGLAB v2023.0

<b>#2: sub2_reref</b>	
Filename:	none
Channels per frame	19
Frames per epoch	308547
Epochs	1
Events	none
Sampling rate (Hz)	200
Epoch start (sec)	0.000
Epoch end (sec)	1542.730
Reference	unknown
Channel locations	No
ICA weights	No
Dataset size (Mb)	25.9

Then use a bandpass filter to filter 0.5 - 40.5 Hz frequencies.

Filter the data -- pop\_eegfiltnew()

Lower edge of the frequency pass band (Hz)

Higher edge of the frequency pass band (Hz)

FIR Filter order (Mandatory even. Default is automatic\*)

\*See help text for a description of the default filter order heuristic.  
Manual definition is recommended.

☐ Notch filter the data instead of pass band

☐ Use minimum-phase converted causal filter (non-linear!; beta)

☒ Plot frequency response

Channel type(s)

OR channel labels or indices

EEGLAB v2023.0

File Edit Tools Plot Study Datasets Help

**#3: sub-1\_reref\_bandfir**

Filename:	none
Channels per frame	19
Frames per epoch	307634
Epochs	1
Events	none
Sampling rate (Hz)	200
Epoch start (sec)	0.000
Epoch end (sec)	1538.165
Reference	unknown
Channel locations	No
ICA weights	No
Dataset size (Mb)	25.9

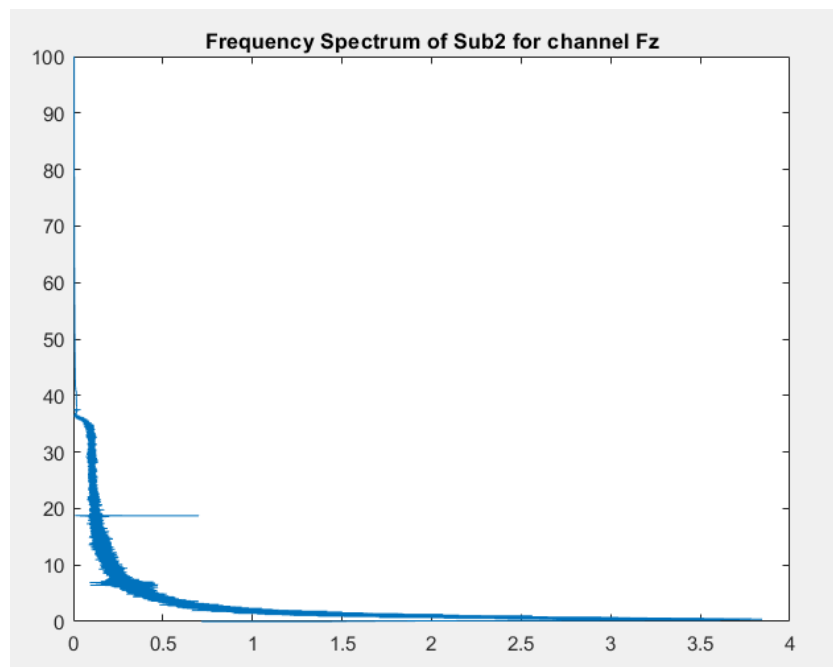
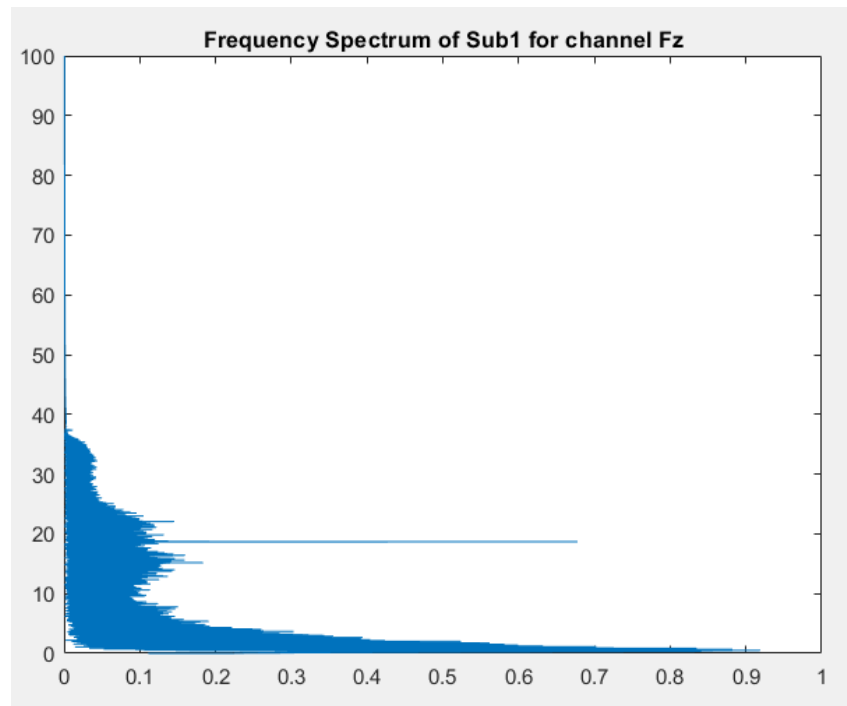
EEGLAB v2023.0

File Edit Tools Plot Study Datasets Help

**#3: sub2\_reref\_bandfir**

Filename:	none
Channels per frame	19
Frames per epoch	308547
Epochs	1
Events	none
Sampling rate (Hz)	200
Epoch start (sec)	0.000
Epoch end (sec)	1542.730
Reference	unknown
Channel locations	No
ICA weights	No
Dataset size (Mb)	25.9

Frequency spectrum plot of channel Fz:



## Step 2:

Now load Standard-10-20-Cap19.loc file from edit-channel locations menu that contains locations of channels.

**Edit channel info -- pop\_chanedit()**

**Channel information ("field\_name"):**

Channel label ("label")	Fp1	Opt. head center
Polar angle ("theta")	-18	Rotate axis
Polar radius ("radius")	0.51111	Transform axes
Cartesian X ("X")	80.7906	
Cartesian Y ("Y")	26.2504	XYZ -> polar & sph.
Cartesian Z ("Z")	-2.9662	Sph. -> polar & xyz
Spherical horiz. angle ("sph_theta")	18	Polar -> sph. & xyz
Spherical azimuth angle ("sph_phi")	-1.9998	
Spherical radius ("sph_radius")	85	Set head radius
Channel type		Set channel types
Reference		Set reference
Index in backup 'urchanlocs' structure	1	
Channel in data array (set=yes)	<input checked="" type="checkbox"/>	

**Channel number (of 19)**

Delete chan    Insert chan    <<    <    1    >    >>    Append chan

Plot 2-D    Plot radius (0.2-1, []=auto)    Nose along +X    Plot 3-D (xyz)

Read locations    Read locs help    Look up locs    Save (as .ced)    Save (other types)

☐ Overwrite Original Channels

Help    Cancel    Ok

**EEGLAB v2023.0**

File Edit Tools Plot Study Datasets Help

**#3: sub-1\_reref\_bandfir**

Filename:	none
Channels per frame	19
Frames per epoch	307634
Epochs	1
Events	none
Sampling rate (Hz)	200
Epoch start (sec)	0.000
Epoch end (sec)	1538.165
Reference	unknown
Channel locations	Yes
ICA weights	No
Dataset size (Mb)	25.9

**EEGLAB v2023.0**

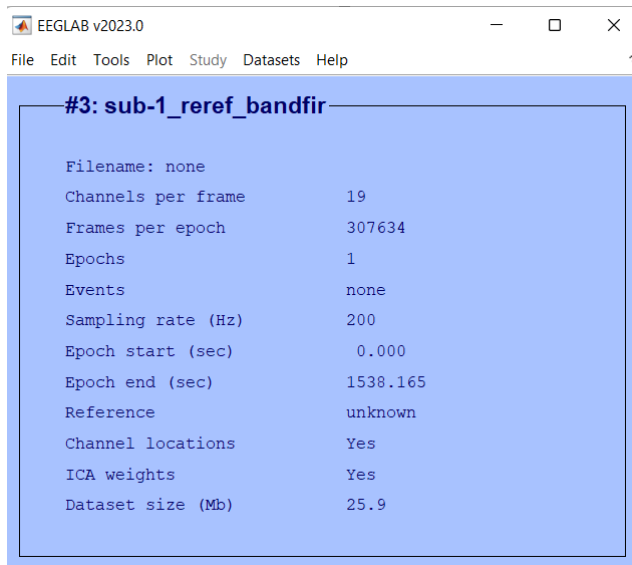
File Edit Tools Plot Study Datasets Help

**#3: sub2\_reref\_bandfir**

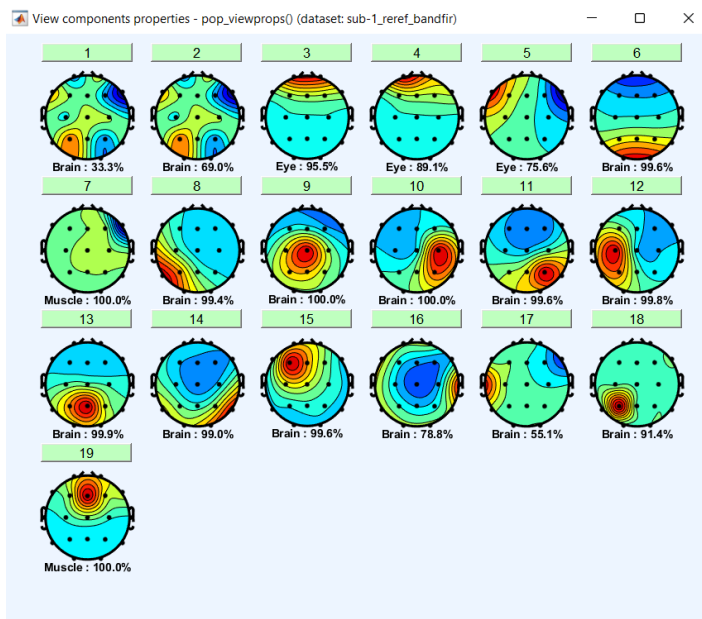
Filename:	...23.0\sub2_reref_bandfir.set
Channels per frame	19
Frames per epoch	308547
Epochs	1
Events	none
Sampling rate (Hz)	200
Epoch start (sec)	0.000
Epoch end (sec)	1542.730
Reference	unknown
Channel locations	Yes
ICA weights	No
Dataset size (Mb)	26



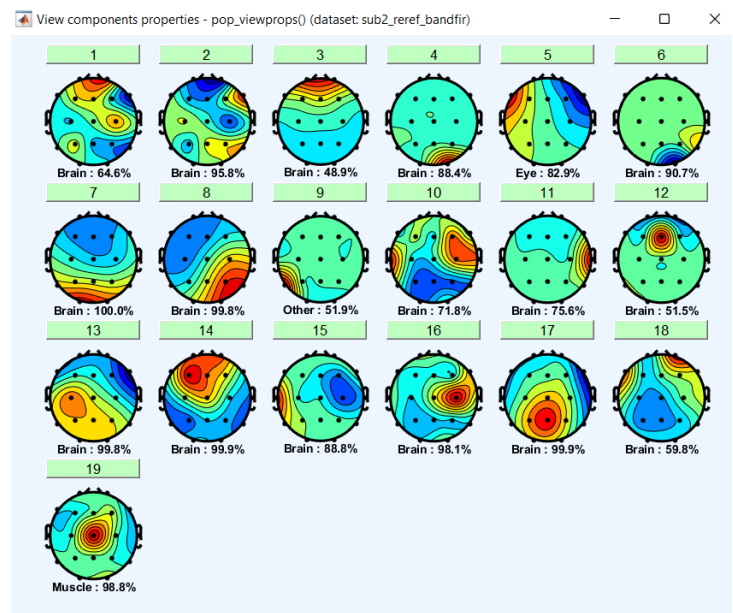
Then run ICA (Independent Component Analysis) algorithm from tools-decompose data by ICA menu. Please note that this part would probably takes more time.



Then you will have the a figure like Figure 4 by running tools → classify components using ICLabel-label components.

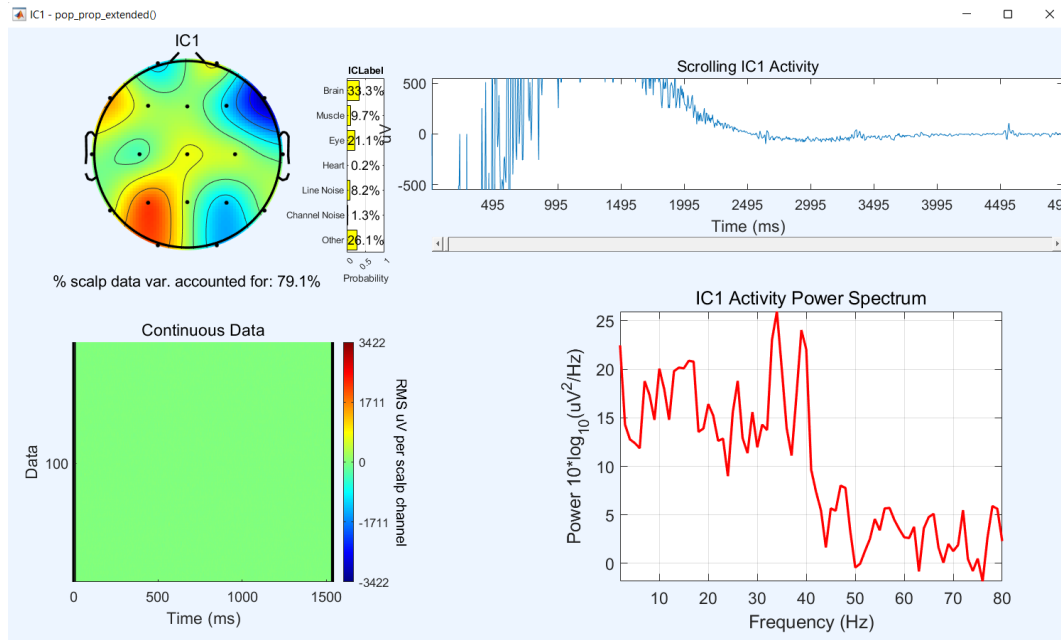


sub1

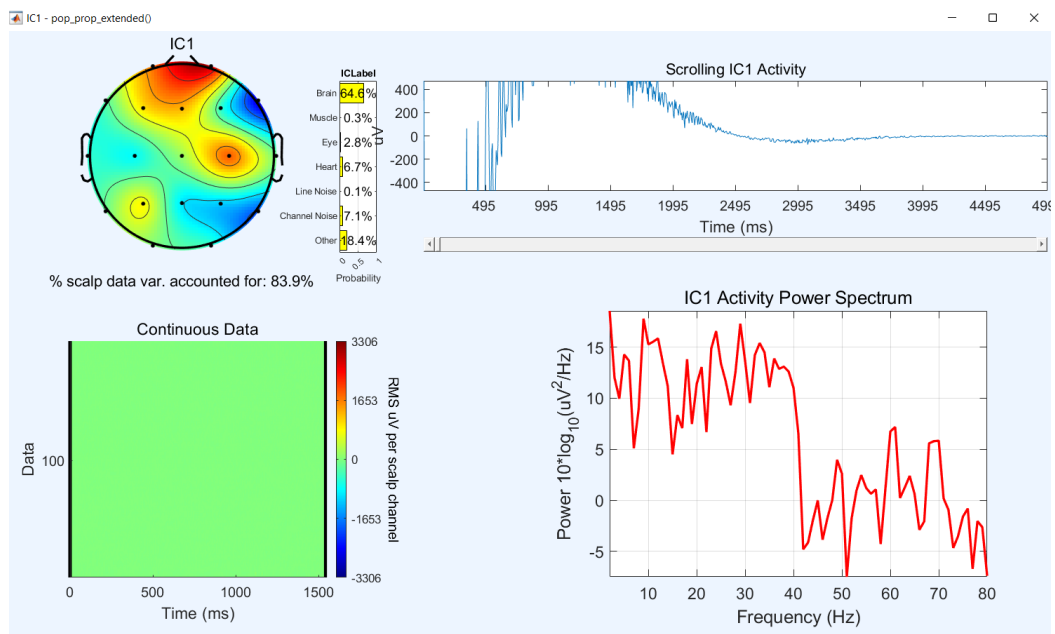


sub2

By clicking on each component, you can see some details about it as well. Present a figure from one of the brain components with its details.

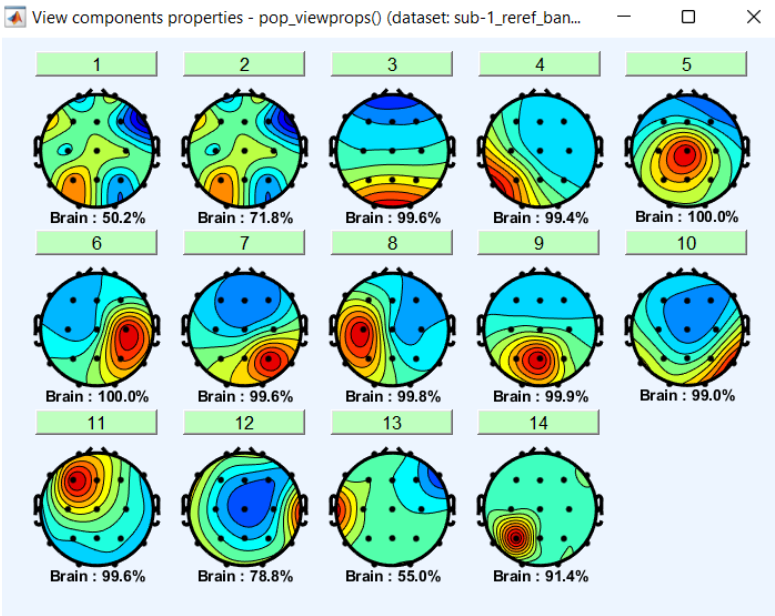
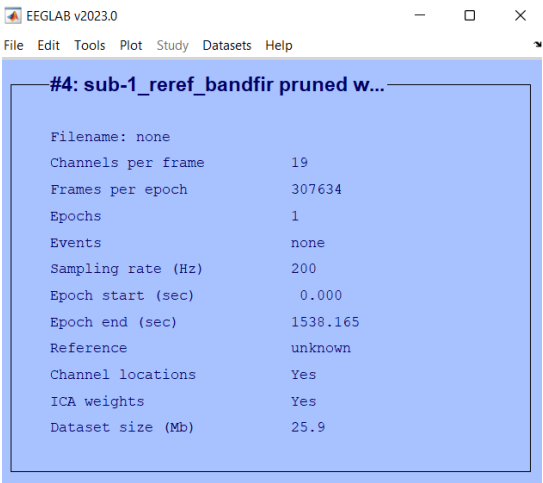


sub1

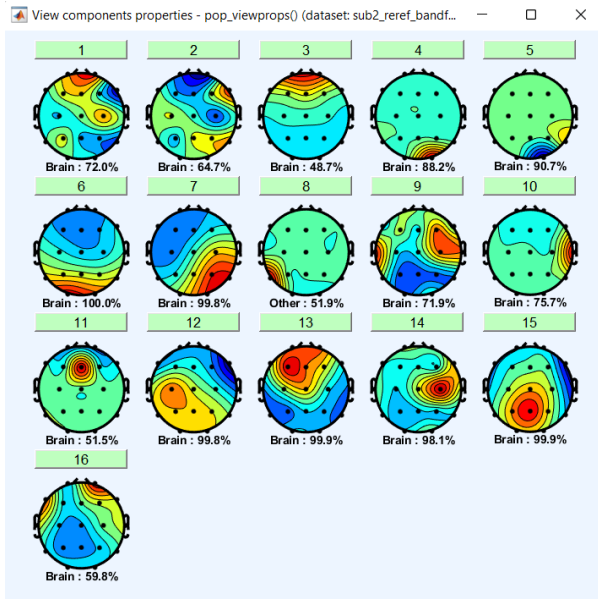


sub2

Now remove all non-brain components.



sub1

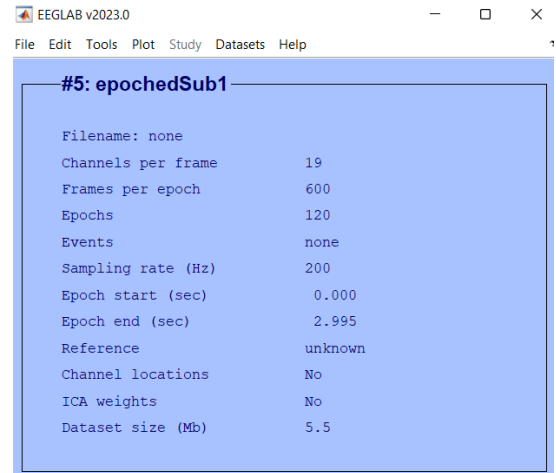


sub 2

### Step 3:

Epoch the data of each subject.

```
%% step 3
sub1=ALLEEG(4).data;
% epoch is a 19*600*120 matrix
epochedSub1=zeros(19 , 600 , 120);
% looping on the channels
for i=1:1:19
    % looping on the trials
    for j=1:1:120
        % skip the first 14 secs , epoch starts from 7th sec and goes on
        % for 3 secs
        startIndex=14*200+j*7*200;
        % 3secs * 200 samples per sec = 600 samples
        endIndex=startIndex+600-1;
        % extract epoched data (from the starting index to the ending
        % index of original data)
        epochedSub1(i,:,j)=sub1(i,startIndex:endIndex);
    end
end
save("epochedSub1","epochedSub1");
```

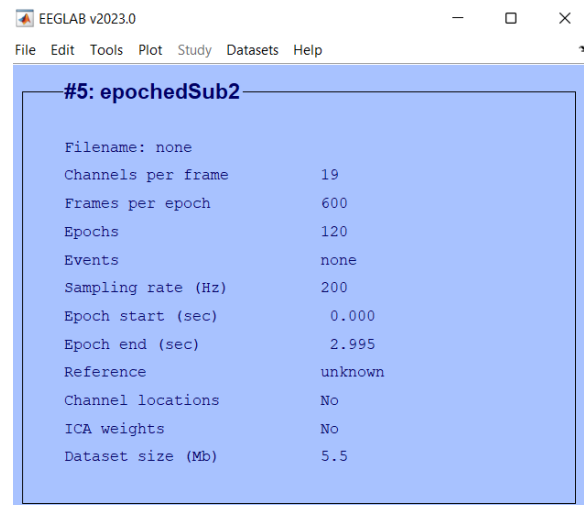


EEGLAB v2023.0

File Edit Tools Plot Study Datasets Help

#5: epochedSub1	
Filename:	none
Channels per frame	19
Frames per epoch	600
Epochs	120
Events	none
Sampling rate (Hz)	200
Epoch start (sec)	0.000
Epoch end (sec)	2.995
Reference	unknown
Channel locations	No
ICA weights	No
Dataset size (Mb)	5.5

```
sub2=ALLEEG(4).data;
% epoch is a 19*600*120 matrix
epochedSub2=zeros(19 , 600 , 120);
% looping on the channels
for i=1:1:19
    % looping on the trials
    for j=1:1:120
        % skip the first 16.4 secs , epoch starts from 7th sec and goes on
        % for 3 secs
        startIndex=16.4*200+j*7*200;
        % 3secs * 200 samples per sec = 600 samples
        endIndex=startIndex+600-1;
        % extract epoched data (from the starting index to the ending
        % index of original data)
        epochedSub2(i,:,j)=sub2(i,startIndex:endIndex);
    end
end
save("epochedSub2","epochedSub2");
```



EEGLAB v2023.0

File Edit Tools Plot Study Datasets Help

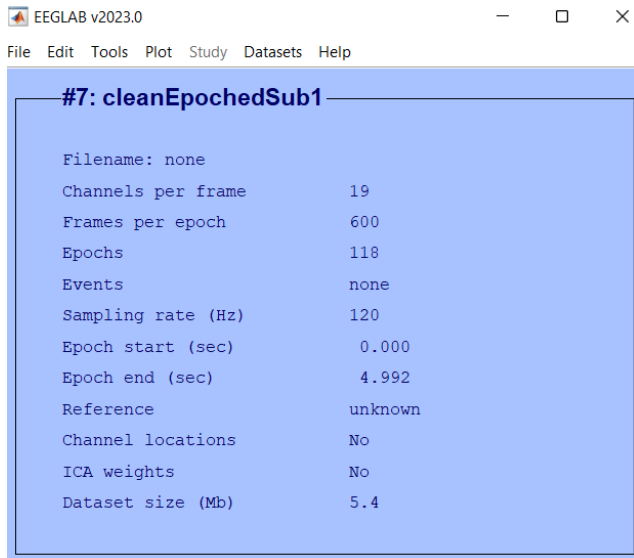
#5: epochedSub2	
Filename:	none
Channels per frame	19
Frames per epoch	600
Epochs	120
Events	none
Sampling rate (Hz)	200
Epoch start (sec)	0.000
Epoch end (sec)	2.995
Reference	unknown
Channel locations	No
ICA weights	No
Dataset size (Mb)	5.5

## Step 4:

In this step, you need to remove noisy trials. Using the second method :

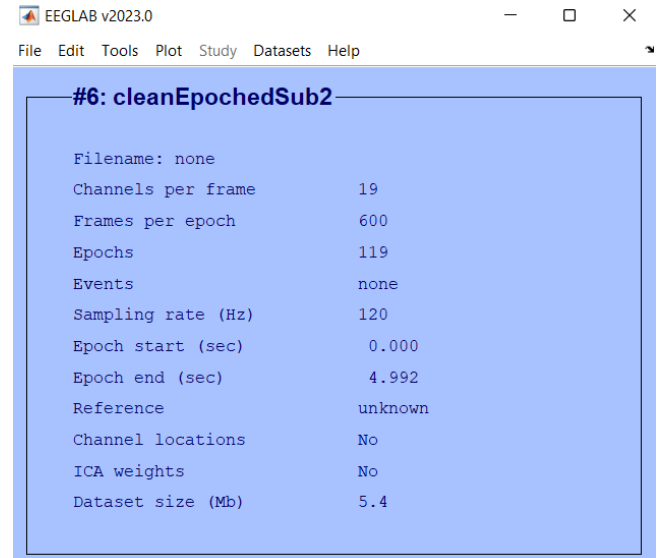
```
%% step 4
noisyTrials = [];
% looping on the channels
for i = 1:1:19
    % p = a matrix of frequency spectrums
    p = nan(120 , 600);
    % looping on the trials
    for j = 1:1:120
        % data of the jth trial
        trialData = squeeze(epochedSub1(:, :, j));
        % frequency spectrum of the trial
        p(j, :) = abs(fft(trialData(i, :))).^2;
    end
    % finding noisy trials (due to the given condition)
    vr = sum(nanstd(p, [], 2).^2, 2);
    noisy_trials = find(abs(zscore(vr)) > 3.5);
    % accumulate over all channels
    accum_noisy_trials = union(noisyTrials, noisy_trials);
end

% remove noisy trials
cleanEpochedSub1 = epochedSub1(:, :, setdiff(1:120, accum_noisy_trials));
save('cleanEpochedSub1', 'cleanEpochedSub1');
```



EEGLAB v2023.0

Filename:	none
Channels per frame	19
Frames per epoch	600
Epochs	118
Events	none
Sampling rate (Hz)	120
Epoch start (sec)	0.000
Epoch end (sec)	4.992
Reference	unknown
Channel locations	No
ICA weights	No
Dataset size (Mb)	5.4



EEGLAB v2023.0

Filename:	none
Channels per frame	19
Frames per epoch	600
Epochs	119
Events	none
Sampling rate (Hz)	120
Epoch start (sec)	0.000
Epoch end (sec)	4.992
Reference	unknown
Channel locations	No
ICA weights	No
Dataset size (Mb)	5.4

## Step 5:

In the final step, only subsample the data corresponding to the Fp1, Fz, Cz & Pz channels.

```
%% step 5
```

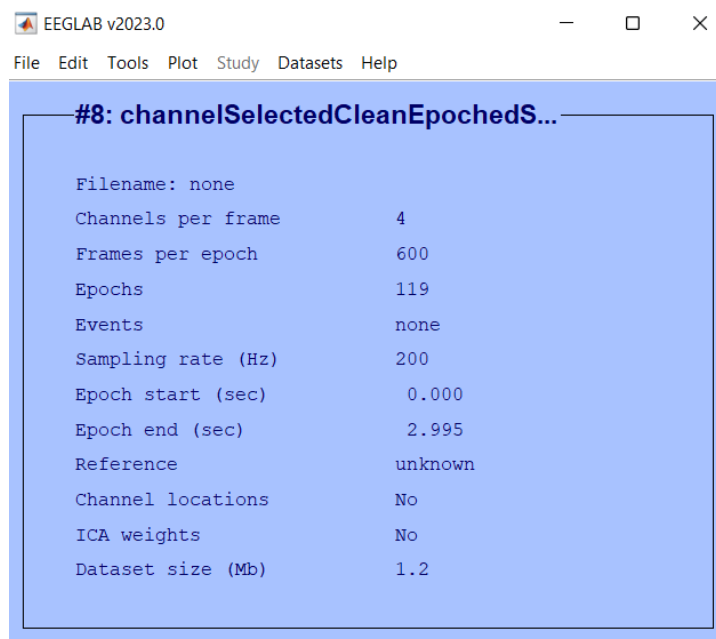
```
% keep the data corresponding to the channels: Fp1,Fz,Cz,Pz
```

```
channelSelectedEpochData = cleanEpochSub1([1 , 5 , 10 , 15] , : , :);
```

```
save("channelSelectedEpochData","channelSelectedEpochData");
```



sub 1



sub 2

Do these 5 steps for each subject and save the final data through a struct with the same format as described in Table 1. Also, consider the order of odor being the same as the ones used for normal participants.

The resulted Structs :

sub1Struct	
1x1 struct with 3 fields	
Field ^	Value
selectedCleanEpoched	4x600x118 double
odor	120x1 double
noisy	[114,116]

sub2Struct	
1x1 struct with 3 fields	
Field ^	Value
selectedCleanEpoched	4x600x119 double
odor	120x1 double
noisy	2

### 3.4 Phase Locking Value (PLV)

What does phase synchronization indicate from a functional point of view? Discuss its importance with valid references.

\*Borrowed from the field of physics, the concept of phase synchronization is widely used in neurosciences. Given two EEG signals in time-domain, we can transform them to be represented in the complex plane to reflect their amplitude and phase.

If we compute the cross-spectrum (by multiplying the complex representation of one signal with the conjugate of the other), we obtain a vector whose length is the product of the two amplitudes and the angle between the vector and real axis represents the phase difference between the signals. A statistical definition can then be given to phase synchronization, which is the stability or consistency of phase differences across EEG time segments, or trials or epochs. Thus measures such as PLV reflect phase synchronization in true sense as it only looks at phase difference.

\*If one observes a phase difference of 0, this could mean that a neuronal source is seen by two sensors at the same instant and the phase observed at these sensors is the same (resulting in the phase difference of 0) . This could be due to volume conduction, although other alternatives cannot be ruled out purely on this basis . Due to such volume conduction, the individual phases observed at these two sensors can also have a difference of 180 degrees or  $\pi$  radians if one of the sensors sees the other end of the source dipole (i.e., records a negative voltage) ! [3]

\*Phase synchronization — here referring to the synchronization of oscillatory phases between different brain regions — supports both working memory and long-term memory and acts by facilitating neural communication and by promoting neural plasticity. There is evidence that processes underlying working and long-term memory might interact in the medial temporal lobe.[4]

Formulate the definition of PLV and briefly discuss the mathematical tools needed to calculate it.

Consider a pair of real signals  $s_1(t)$  and  $s_2(t)$ , that have been band-pass filtered to a frequency range of interest. Analytic signals  $z_i(t) = A_i(t)e^{j\varphi_i(t)}$  for  $i = \{1, 2\}$  and  $j=\sqrt{-1}$  are obtained from  $s_i(t)$  using the Hilbert transform:



$$z_i(t) = s_i(t) + j\text{HT}(s_i(t))$$

where  $\text{HT}(s_i(t))$  is the Hilbert transform of  $s_i(t)$  defined as

$$\text{HT}(s_i(t)) = \frac{1}{\pi} \text{P.V.} \int_{-\infty}^{\infty} \frac{s_i(\tau)}{t-\tau} d\tau$$

and P.V. denotes Cauchy principal value. Once the analytic signals are defined, the relative phase can be computed as :

$$\Delta\varphi(t) = \arg\left(\frac{z_1(t)z_2^*(t)}{|z_1(t)||z_2(t)|}\right).$$

The instantaneous PLV is then defined as

$$\text{PLV}(t) \triangleq |E[e^{j\Delta\varphi(t)}]|$$

where  $E[.]$  denotes the expected value. The PLV takes values on  $[0, 1]$  with 0 reflecting the case where there is no phase synchrony and 1 where the relative phase between the two signals is identical in all trials. PLV can therefore be viewed as a measure of trial to trial variability in the relative phases of two signals. In this work we use the Hilbert transform but the continuous Morlet wavelet transform can also be used to compute complex signals, producing separate band-pass signals for each scaling of the wavelet .

In practice PLV is typically estimated by averaging over trials and/or time .

For notational convenience, we will drop the explicit dependence on  $t$  in the following. A nonparametric estimate of PLV can be computed by approximating the equation above by averaging over trials:

$$PLV_{\text{sample}} \triangleq \left| \frac{1}{N} \sum_{n=1}^N e^{j\Delta\varphi_n(t)} \right|$$

where  $n$  indexes the trial number and  $N$  is the total number of trials. The estimator generalizes in an obvious way to incorporate averaging over multiple time samples.[\[5\]](#)

Implement a function which finds the PLV between two channels in a specific frequency range.

```
function plv = PLVCal(data1 , data2 , fs , frequencyRange)
% calculate plv for each frequent odor trial(looping on the trials)
% apply bandpass filter to extract the desired frequency range
lowFreq = frequencyRange(1)/(fs/2);
highFreq = frequencyRange(2)/(fs/2);
[b , a] = butter(1 , [lowFreq , highFreq] , 'bandpass');
filteredData1 = filtfilt(b , a , data1);
filteredData2 = filtfilt(b , a , data2);
% apply Hilbert transform to extract the phases
analyticSignal1 = hilbert(filteredData1);
analyticSignal2 = hilbert(filteredData2);
% calculate the phases
phase1 = angle(analyticSignal1);
phase2 = angle(analyticSignal2);
% the phase differences
phaseDiff = abs(phase1-phase2);
% calculate PLV due to the formula
plv = abs(mean(exp(1i*phaseDiff)));
end
```

## 4 Results

### 4.1 Values

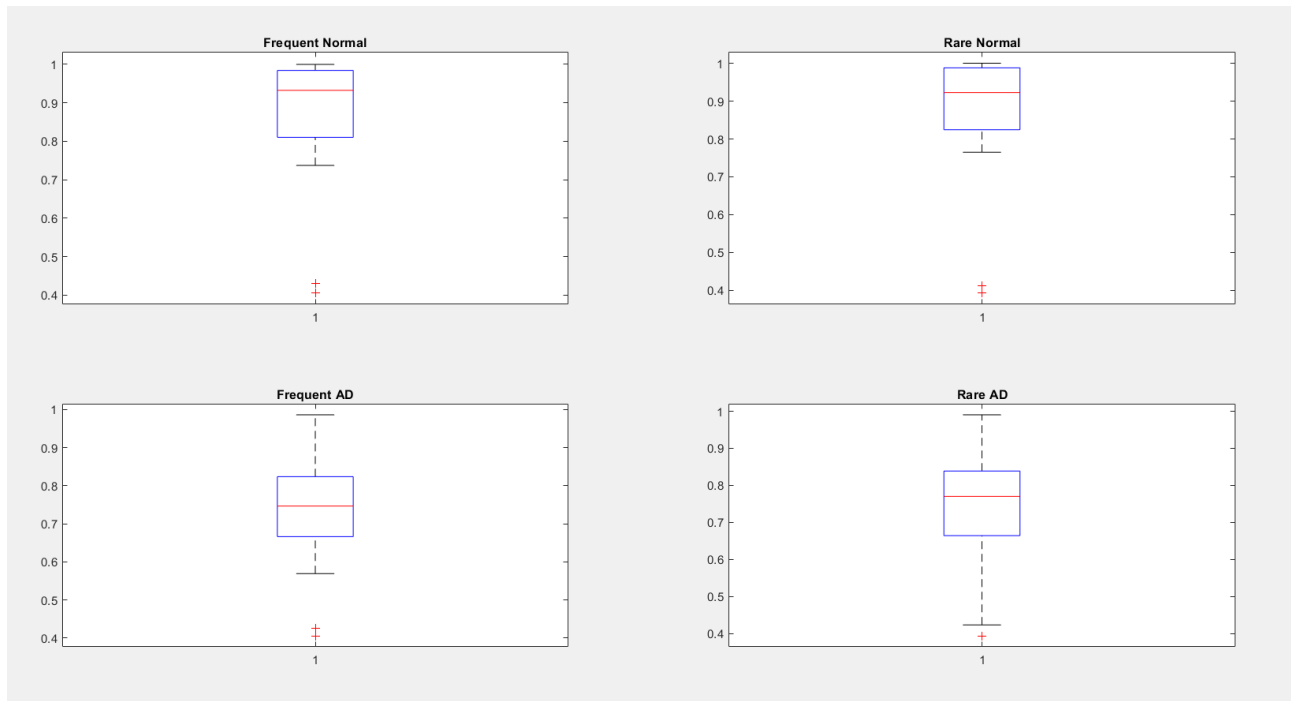
Find the PLV for all participants of both groups on both frequent and rare odors between the Fz and Cz channels using the function you implemented in section 3.4 .

The resulted PLVs:

plvFreqMeanAD			plvRareMeanAD			plvFreqMeanNormal			plvRareMeanNormal		
13x1 double			13x1 double			15x1 double			15x1 double		
	1	2		1	2		1	2		1	2
1	0.9788		1	0.9513		1	0.8113		1	0.8215	
2	0.6986		2	0.6951		2	0.9991		2	1.0000	
3	0.5693		3	0.5879		3	0.9174		3	0.9104	
4	0.7164		4	0.6894		4	1.0000		4	1.0000	
5	0.8231		5	0.8099		5	0.9994		5	0.9993	
6	0.7465		6	0.7787		6	0.9004		6	0.8893	
7	0.9863		7	0.9905		7	0.9363		7	0.9461	
8	0.8171		8	0.8277		8	0.8100		8	0.8334	
9	0.8259		9	0.8713		9	0.9397		9	0.9534	
10	0.7758		10	0.7701		10	0.9994		10	0.9997	
11	0.7055		11	0.6978		11	0.9324		11	0.9250	
12	0.4059		12	0.3937		12	0.7373		12	0.7651	
13	0.4265		13	0.4230		13	0.9367		13	0.9228	
						14	0.4059		14	0.3937	
						15	0.4307		15	0.4130	

## 4.2 Distributions

Draw the box plots of PLVs you found in the previous part among two groups and two odors.



Also, fit a gaussian distribution on these PLVs and present you results.

```
NormalFreqDist =
```

```
NormalDistribution
```

```
Normal distribution
```

```
mu = 0.8504 [0.744177, 0.956623]  
sigma = 0.191814 [0.140432, 0.302509]
```

```
NormalRareDist =
```

```
NormalDistribution
```

```
Normal distribution
```

```
mu = 0.85151 [0.743432, 0.959589]  
sigma = 0.195164 [0.142885, 0.307793]
```

```
ADFreqDist =
```

```
NormalDistribution
```

```
Normal distribution
```

```
mu = 0.728897 [0.621343, 0.83645]  
sigma = 0.177983 [0.127629, 0.293802]
```

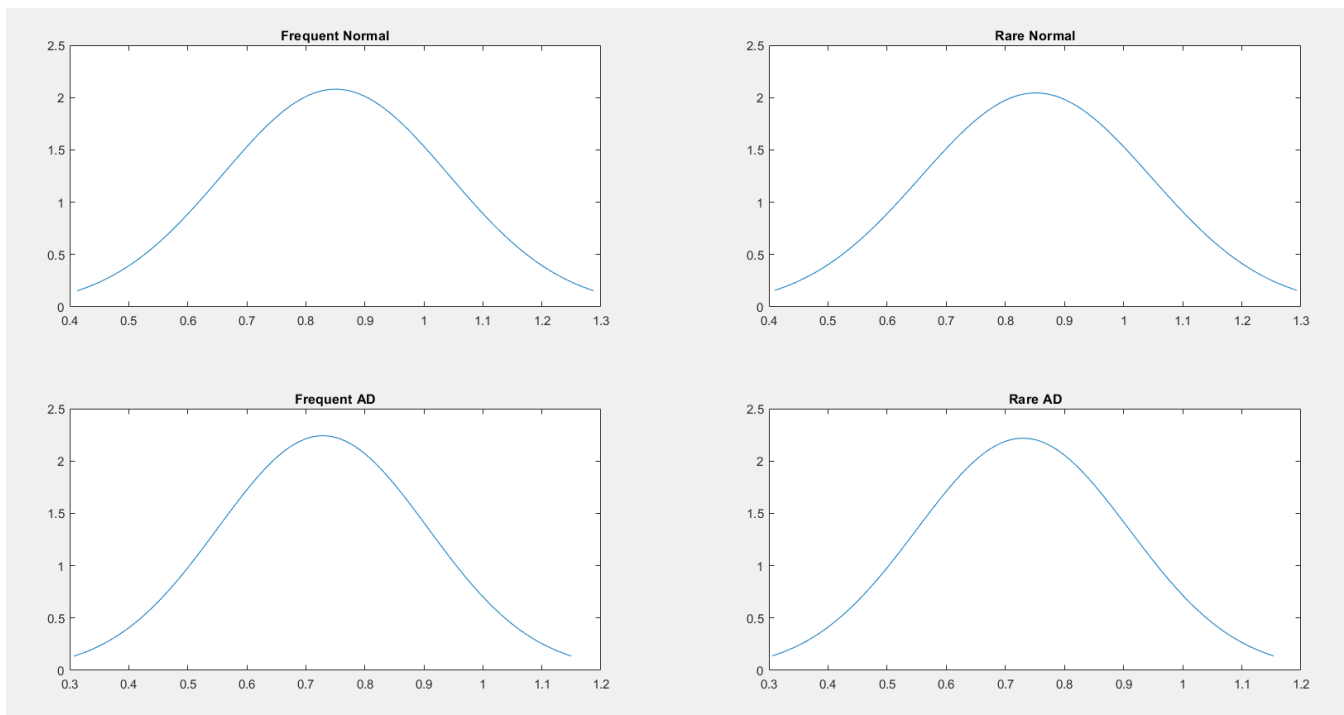
```
ADRareDist =
```

```
NormalDistribution
```

```
Normal distribution
```

```
mu = 0.729726 [0.621036, 0.838417]  
sigma = 0.179864 [0.128978, 0.296907]
```

Plots of the resulted gaussian PDFs:



You need to specify the corresponding p-values to evaluate the statistical significance of your findings.

```
ans =  
    "P_values for frequent odor = 0.095847"  
  
ans =  
    "Hypothesis test result for frequent odor = 0"  
  
ans =  
    "P_values for rare odor = 0.099706"  
  
ans =  
    "Hypothesis test result for rare odor = 0"
```

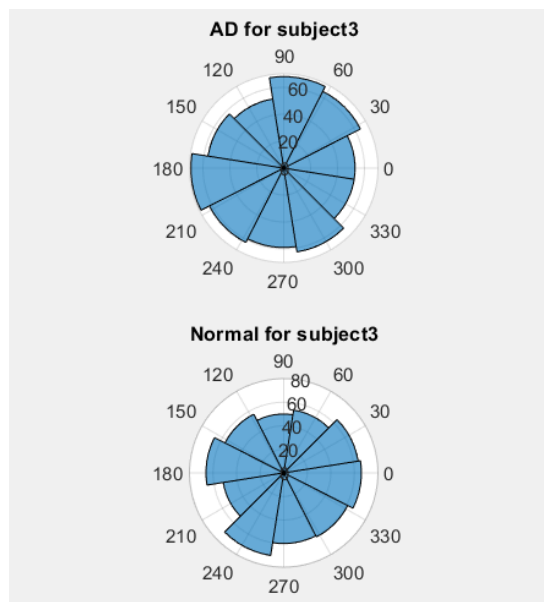
### 4.3 Statistical Significance

Based on the p-values you found in the previous part, discuss whether we could state that the “PLV is significantly different among AD and Normal subjects in the slow gamma frequency range”.

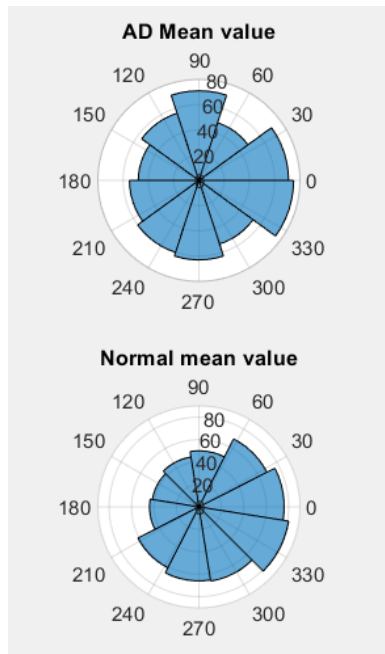
As we can see from the previous part , the p\_values between AD and Normal subjects are very small numbers on both frequent and rare odors and are below the threshold , therefore  $H_0$  is accepted .As the null hypothesis stands for the situation in which AD and Normal PLVs have the same mean or in other words are not significantly different , we can conclude PLVs are not significantly different among AD and Normal subjects on both odors .

### 4.4 Phase Difference

Draw a polar histogram of the phase difference between Fz and Cz channels during frequent odor trials for a random subject in each group and compare the results.



Also, plot the mean value of this quantity among all the subjects of each group and discuss the results.















Variance is higher among AD patients as expected.

## 4.5 Heatmaps

Now you need to plot a heatmap which has the PLVs between each pair of the channels.



Find whether PLV between other channel pairs are significantly different among two groups in the slow gamma frequency range and test your results.

 freqH_Cz_Pz	0	 rareH_Cz_Pz	0
 freqH_Fp1_Cz	0	 rareH_Fp1_Cz	0
 freqH_Fp1_Fz	0	 rareH_Fp1_Fz	0
 freqH_Fp1_Pz	0	 rareH_Fp1_Pz	0
 freqH_Fz_Cz	0	 rareH_Fz_Cz	0
 freqH_Fz_Pz	0	 rareH_Fz_Pz	0

As we could see from the colors in the heatmap , there is no significant difference between the two groups . We can also deduce this result from the p and H values , since the H values have all become zero(which means  $H_0$  is accepted) , we can again conclude that PLV between channel pairs are not significantly different among the two groups.

## 5 \*Bonus

### 5.1 Mild Cognitive Impairment (MCI)

#### 5.1.1 Additional Information

Describe the relationship between MCI and AD. Explain whether MCI would always result in AD and briefly investigate the causes of MCI.

Mild cognitive impairment (MCI) is a precursor to Alzheimer's disease (AD), but not all MCI patients develop AD. Biomarkers for early detection of individuals at high risk for MCI-to-AD conversion are urgently required.

Mild cognitive impairment (MCI) is an intermediate stage between normal aging and dementia, and its presence is associated with a higher risk of progression to clinically probable Alzheimer's disease (AD). The annual conversion rate from MCI to AD has been reported as 10 to 15%. After 6 years of follow-up, approximately 80% of MCI patients will have converted to AD (MCI converters [MCI-C]), although some MCI patients remain stable or convert back to normal (MCI non-converters [MCI-NC]). To date, there are no curative treatments for patients who already have AD, and available treatments are only able to postpone the progression of the disease. Therefore, biomarkers for early detection of MCI-C and prognosis prediction models are both desperately required. These will allow early treatment of patients with MCI before they convert to AD, which could reduce the number of patients with AD. [6]

There is no single cause of MCI. The risk of developing MCI increases as someone gets older. Conditions such as diabetes, depression, and stroke may increase a person's risk for MCI.[7]

### 5.1.2 MCI Data Processing

Based on the significantly different coupled channels you found for differentiation between AD and Normal groups, find the Phase-Locking-Value (PLV) for the MCI subjects and provide the required results by comparing all the the 3 states (Normal, MCI, AD). Your findings must include the significance testing by providing the corresponding p-values

The resulted PLVs:

plvFreqMeanMCI			plvRareMeanMCI		
7x1 double			7x1 double		
	1	2		1	2
1	0.8204		1	0.4186	
2	0.9307		2	0.4186	
3	0.8364		3	0.4186	
4	0.6673		4	0.4186	
5	0.5430		5	0.4186	
6	0.8010		6	0.4186	
7	0.7709		7	0.4186	

The resulted P-values:

freqH1	0	rareH1	0
freqH2	0	rareH2	1
freqH3	0	rareH3	1

The results show that : there is significant difference between channels Cz and Fz among MCI and Normal data , and also among MCI and AD data when exposed to rare odor.



## 5.2 Phase-Amplitude Coupling (PAC)

### 5.2.1 Metrics

Conduct a search about other PAC measures and briefly provide an explanation about two of them.

There are other PAC measures such as MVL(Mean Vector Length) , MI(Modulation Index) , KL(Kullback-Leibler) and etc.

MVL : For the phase-amplitude coupling measure MVL, introduced by Canolty et al. (2006), phase is extracted from the low frequency filtered analytic signal and amplitude is extracted from the high frequency filtered analytic signal. MVL utilizes phase angle and magnitude of each complex number (i.e., each data point) of the corresponding analytic signal in a quite direct way to estimate the degree of coupling. Each complex value of the analytic time series is a vector in the polar plane. Phase-amplitude coupling is present, when the magnitude M of a fraction of all vectors is especially high at a specific phase or at a narrow range of phases. Averaging all vectors creates a mean vector with a specific phase and length.

The length of this vector represents the amount of phase-amplitude coupling. The direction represents the mean phase where amplitude is strongest. When no coupling is present, all vectors cancel each other out and the mean vector will be short. Then its direction does not represent any meaningful phase. The MVL is calculated by the following formula:

$$MVL = \left| \frac{\sum_{t=1}^n a_t e^{i\theta_t}}{n} \right|$$

where n is the total number of data points, t is a data point,  $a_t$  is the amplitude at data point t and  $\theta_t$  is the phase angle at data point t. This value cannot become negative because it represents the length of the mean vector. The length of a vector cannot be negative.

MI: Tort et al. (2008) suggest a very different way of computing phase-amplitude coupling, which anyways is based on the same parameters of the analytic signal, amplitude magnitude and phase angle. For calculating the MI according to Tort et al. (2008), all possible phases from -180 to 180° are first binned into a freely chosen amount of bins. Tort et al. (2008) established to use 18 bins of 20° each, which many authors follow. The amount of bins can influence the results, as will be explained below. The average amplitude of the amplitude-providing frequency in each phase bin of the phase-providing frequency is computed and normalized by the following formula:

$$p(j) = \frac{\bar{a}}{\sum_{k=1}^N \bar{a}_k}$$

where  $\bar{a}$  is the average amplitude of one bin,  $k$  is the running index for the bins, and  $N$  is the total amount of bins;  $p$  is a vector of  $N$  values. With the help of these calculations, one obtains the data for the phase-amplitude plot, which depicts the actual phase-amplitude coupling graphically.

Subsequently Shannon entropy is computed; a measure that represents the inherent amount of information of a variable. If Shannon entropy is not maximal, there is redundancy and predictability in the variable. Shannon entropy is maximal, if the amplitude in each phase bin is equal .

Shannon entropy is computed by the following formula:

$$H(p) = - \sum_{j=1}^N p(j) \log p(j)$$

where  $p$  is the vector of normalized averaged amplitudes per phase bin and  $N$  is the total amount of bins. It does not matter which logarithm base is used if permutation testing is applied later on (Cohen, 2014). Like in Tort et al. (2008) the natural logarithm is used here. Shannon entropy is dependent on the amount of bins used and this is why the MI is likewise dependent on the number of bins. The higher the amount of bins, the larger Shannon entropy can become. Complying with the original study and most other studies, 18 bins have been employed here.

Phase-amplitude coupling is defined by a distribution that significantly deviates from the uniform distribution. Kullback–Leibler distance, a measure for the disparity of two distributions is calculated by the following formula:

$$KL(U, X) = \log N - H(p)$$

where  $U$  is the uniform distribution,  $X$  is the distribution of the data,  $N$  is the total amount

$$MI = \frac{KL(U, X)}{\log N}$$

of bins, and  $H(p)$  is the Shannon entropy according to Eq. 4. The uniform distribution is represented by  $\log(N)$ . The final raw MI is calculated by the following formula:

where  $KL(U, X)$  is the Kullback–Leibler distance according to Eq. 5 and  $N$  is the total amount of bins.[8]

## 6 Conclusion

In this section, you are required to thoroughly examine and analyze the results you have obtained throughout this project. You must provide a comprehensive discussion of your findings, highlighting their significance and relevance to the research question. You should also present any limitations or weaknesses in your study and suggest possible areas for future research.

Overall, this section is critical to demonstrating the quality and validity of your research and should be approached with careful attention to detail and clarity of expression.

In the beginning , we tried to preprocess a raw data which was collected in lab . We removed its artifacts , and applied a band-pass filter on it to use it in the needed frequency range .Then we epoched the data to the desired shape , removed its noisy trials and kept 4 of its channels.

At the time we had a clean and epoched data so we could do the processing . We first looked at a metric called PLV . PLV quantifies the degree of phase synchronization between two signals . We used it to examine the relationship between the phases of the signal of two different channels in each group of the patients .By the box plot we could conclude that

normal people's PLV between two channels were higher in average . Which is the expected result , because normal people have higher brain connectivity .

Then we used another processing method by finding the p-values .Using p-values we could check if PLV was significantly different among AD and Normal subjects in the slow gamma frequency range . We also used heatmaps to examine PLV between each two channels .

By examining the results , we concluded that PLVs are not significantly different among AD and Normal subjects . We also did the same thing for MCI subjects . The results showed that there is significant difference between channels Cz and Fz among MCI and Normal data , and also among MCI and AD data when exposed to rare odor.

We did gain some invalid results on this project as well as valid , but we tried to focus on the right answers.

## References :

[1]:

[https://info.tmsi.com/blog/the-10-20-system-for-eeg#:~:text=The%20letter%20of%20the%20electrode,%2C%20and%20O%20\(occipital\).](https://info.tmsi.com/blog/the-10-20-system-for-eeg#:~:text=The%20letter%20of%20the%20electrode,%2C%20and%20O%20(occipital).)

[2]:

[https://www.medicine.mcgill.ca/physio/vlab/biomed\\_signals/eeg\\_n.htm#:~:text=Signal%20frequency%3A%20the%20main%20frequencies,3%20and%204%20of%20sleep.](https://www.medicine.mcgill.ca/physio/vlab/biomed_signals/eeg_n.htm#:~:text=Signal%20frequency%3A%20the%20main%20frequencies,3%20and%204%20of%20sleep.)

[3]:

<https://sapienlabs.org/lab-talk/eeg-connectivity-using-phase-lag-index/>

[4]:

<https://nature.com/articles/nrn2979>

[5]:

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3674231/>

[6]:

<https://alzres.biomedcentral.com/articles/10.1186/s13195-020-00716-0>

[7]:

[https://www.alz.org/alzheimers-dementia/what-is-dementia/related\\_conditions/mild-cognitive-impairment](https://www.alz.org/alzheimers-dementia/what-is-dementia/related_conditions/mild-cognitive-impairment)

[8]:

<https://www.frontiersin.org/articles/10.3389/fnins.2019.00573/full>