# 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

---

Authors:
Alireza Gargoori Motlagh          alireza.gargoorimotlagh@ee.sharif.edu
Mojtaba Lahijanian                  mojtaba.lahijanian@ee.sharif.edu
Amirali Loghmani                     amirali.loghmani@ee.sharif.edu
S. Moienadin Makkiyan                 moien.makkiyan@sharif.edu

Student, project by:

Borna Khodabandeh                         borna710kh@gmail.com

Spring 2023

# Preface

**Notes on the project:**

- Due date: 1402/04/10

- The project must be done individually. Each individual will present his results in an online session on 1402/04/11.

- Please submit your project report as a `.pdf` file. Include all outputs and final results in the report. Make sure to list the practice text questions and provide a concise explanation of your problem-solving approach in each section.

- Ensure that all codes are provided in a separate `.m/.py/.ipynb` file. If a code cannot be tested accurately upon submission, the reported results will be considered invalid, and no points will be awarded in such cases.

- You have the flexibility to utilize either `MATLAB` or `Python` for your project. However, please be aware that `MATLAB` is recommended since certain aspects of the project rely on `MATLAB` toolboxes.

- Ensure that you save all files, including your report, codes, helper functions, and any additional outputs, if required, in a compressed file format such as `.zip` or `.rar`. This compressed file should then be uploaded to the Coursework CW submission platform.

- Your file names must be in the following format:

  `Project_#StudentID.zip/.rar/.pdf/.m/.py/.ipynb`

- The details of the grading system of this project will be provided in the coming days. Generally, the project is worth a total of 1 point, with an additional 1 point allocated for the bonus section. Part 5.1 carries 0.5 points, and another 0.5 points are assigned to part 5.2.

- In this project, it is essential to uphold the principles of academic integrity and refrain from any form of cheating or copying. Cheating undermines the learning process, diminishes personal growth, and compromises the trust placed in us as students/researchers/professionals. It is crucial to recognize that engaging in dishonest practices not only tarnishes our own reputation but also has serious consequences, both ethically and academically. We want to emphasize that if anyone is found to have cheated, their results will not be accepted in this project, and they will receive a zero mark.

# Contents

# 1 Introduction

## 1.1 Neurodegenerative Diseases

Neurodegenerative diseases, including Alzheimer's Disease (AD) and Mild Cognitive Impairment (MCI), pose significant challenges to individuals, families, and healthcare systems worldwide. These conditions are characterized by progressive deterioration of cognitive functions, leading to severe impairment and loss of independence. Understanding the importance of these diseases and the urgency to find reliable biomarkers for their early detection and accurate diagnosis is critical for effective intervention and disease management and has become a priority in the field of neurodegenerative research.

Alzheimer's disease, the most common form of dementia, affects millions of individuals globally, and its prevalence is expected to rise with the aging population. MCI, often considered a transitional stage between normal aging and AD, is characterized by subtle cognitive decline that does not severely impact daily functioning. Scientists recognize that the combination of a person's genes and environment contributes to their risk of developing a neurodegenerative disease. For example, someone might have a gene that makes them more susceptible to Parkinson's disease, but their environmental exposures can affect whether, when, and how severely they are affected. [4]

## 1.2 Olfactory Dysfunction

The sense of smell is today one of the focuses of interest in aging and neurodegenerative disease research. In several neurodegenerative diseases, such as Parkinson's disease and Alzheimer's disease, the olfactory dysfunction is one of the initial symptoms appearing years before motor symptoms and cognitive decline which manifests as a decreased ability to detect, identify, or differentiate odors and thus, being considered a clinical marker of these diseases' early stages and a marker of disease progression and cognitive decline. [5]

One of the primary reasons olfactory dysfunction is prominent in neurodegenerative diseases is the presence of pathological changes in the olfactory system. In AD, for example, amyloid plaques and neurofibrillary tangles, the hallmark pathological features of the disease, are found not only in brain regions associated with memory and cognition but also in areas involved in olfaction, such as the olfactory bulb and olfactory cortex.

## 1.3 Goal of the Project

Understanding the significance of olfactory dysfunction in neurodegenerative diseases is important as it can serve as a potential biomarker for early detection and help unravel underlying disease mechanisms. The study of olfactory dysfunction in neurodegenerative diseases is an active area of research. Researchers are investigating the potential of olfactory testing as a diagnostic tool and exploring the mechanisms underlying olfactory dysfunction. They are also examining the role of olfactory dysfunction in disease progression and exploring therapeutic interventions targeting the olfactory system.

In this project, we want to identify early biomarkers for related brain disorders through olfactory stimulus.

# 2 Electroencephalography (EEG)

## 2.1 What is EEG?

There are different tools for collecting data from the brain. One of the methods of capturing brain signals is called Electroencephalography (EEG). These signals are changes in voltage level caused by changes in brain signals captured by some electrodes. These voltages are microVolt-level, so they can be sensitive to small noises.

One of the EEG advantages compared to other methods is its high temporal accuracy (i.e. high sampling frequency) while it suffers from low spatial accuracy. Another benefit of EEG devices is their smaller size compared to other devices like fMRI (functional Magnetic Resonance Imaging). While fMRI devices occupy the whole room, you can use EEG via portable devices.

EEG headsets are devices built to save EEG signals. These headsets could contain many electrodes. One internationally recognized electrode placement method is the **10-20 system**. This method was developed to maintain standardized testing methods ensuring that a subject's study outcomes (clinical or research) could be compiled, reproduced, and effectively analyzed and compared using scientific methods. It is called 10-20 because the distance between adjacent electrodes is 10% or 20% of the skull's total front–back or right–left distance.
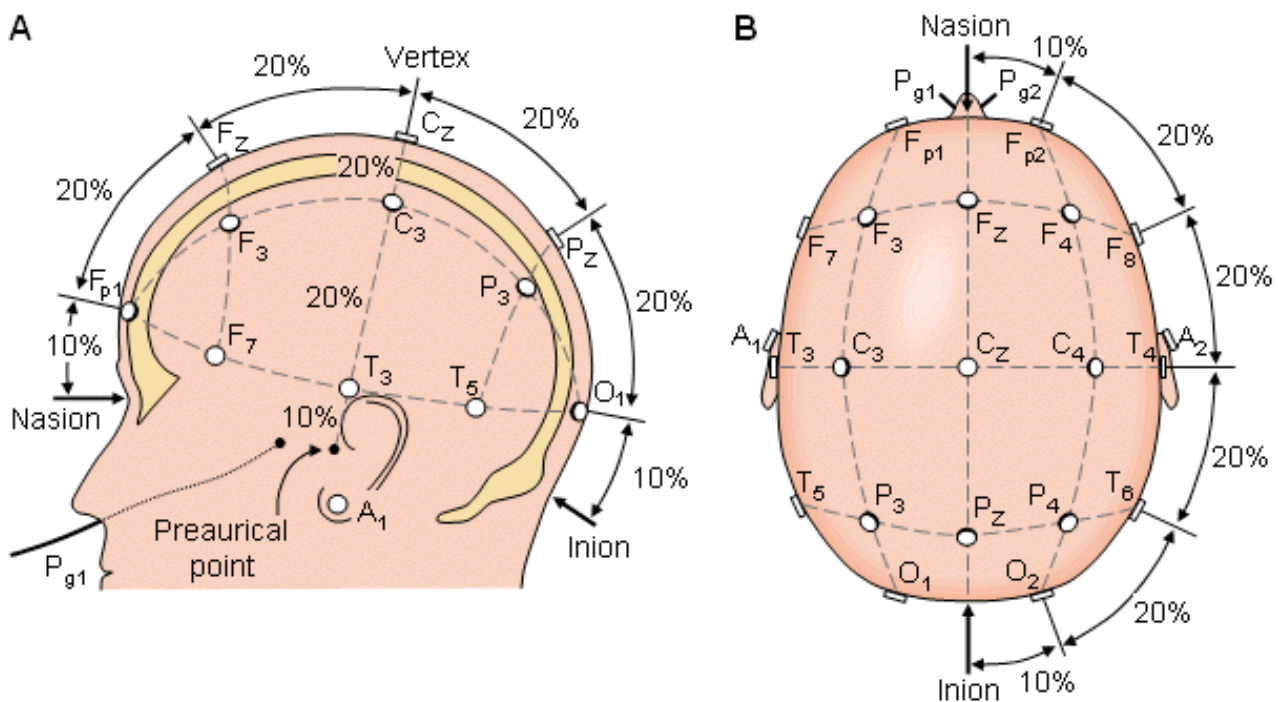


Figure 1: EEG 10-20 Electrode Placement System

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 10-20 EEG electrode placement system is a standardized method for positioning electrodes on the scalp for EEG recordings. Electrodes are named based on their anatomical locations using a combination of letters and numbers. The letters F, T, C, P, and O represent frontal, temporal, central, parietal, and occipital regions, respectively. The numbers indicate the distance between electrode placements and are based on percentages of specific measurements on the skull. For example, Fp1 and Fp2 electrodes are placed at 10% and 20% of the distance between specific landmarks on the skull. This naming method helps researchers and clinicians identify electrode locations and ensures consistent recording and comparison of EEG data across individuals and studies.

## 2.2  Alzheimer's Disease

Alzheimer's Disease (AD) is a progressive and irreversible neurological disorder that affects the brain, primarily causing problems with memory, thinking, and behavior. It is the most common cause of dementia, a general term for a decline in cognitive ability severe enough to interfere with daily life.

The exact cause of Alzheimer's disease is not yet fully understood, but it is believed to involve a combination of genetic, lifestyle, and environmental factors. The staging of the AD is associated with the accumulation of Amyloid- beta ($A\beta$) proteins in the brain. These depositions cause synaptic and neuronal loss, which leads to major cognitive dysfunction in the advanced levels of the disease.

While EEG is not currently used as a primary treatment for Alzheimer's disease, it can be a valuable tool in the diagnosis and monitoring of the disease. EEG can help in the diagnosis of Alzheimer's by detecting abnormal patterns of brain activity that are characteristic of the disease. In individuals with AD, EEG often shows changes such as a reduction in certain brainwave frequencies and an increase in others. These patterns can aid in differentiating Alzheimer's from other types of dementia or cognitive disorders.

## 2.3  Frequency Bands of EEG

In the frequency domain, EEG signals are divided into 5 bands. [3]
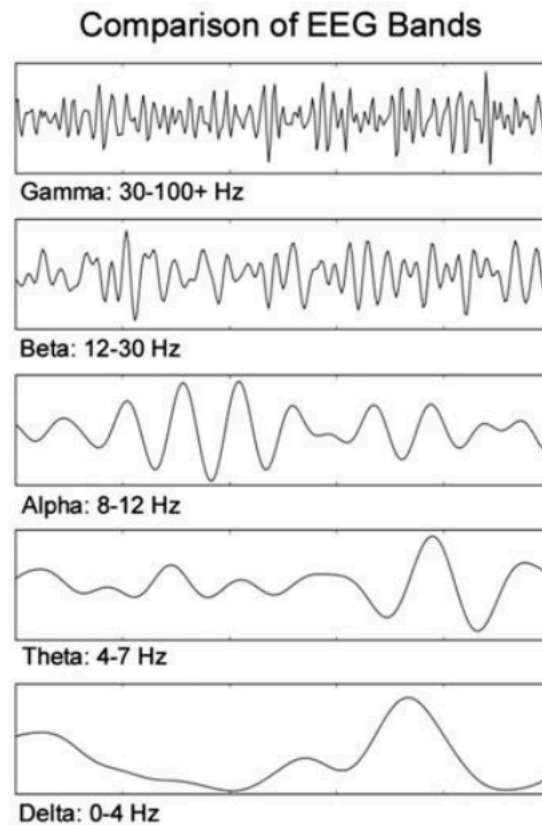Determine the activities each frequency band is associated with.

Figure 2: EEG Frequency Bands

EG signals are divided into five frequency bands associated with specific activities:

1. Delta (0.5-4 Hz): Deep sleep, restorative processes, unconscious mind.

2. Theta (4-8 Hz): Light sleep, relaxation, meditation, memory formation, creativity.

3. Alpha (8-12 Hz): Relaxation, closed eyes, wakeful relaxation, inhibition of distractions.

4. Beta (12-30 Hz): Active mental states, alertness, problem-solving, decision-making.

5. Gamma (30-100 Hz): Cognitive processing, attention, memory retrieval, perception.

These frequency bands represent different neural activities and cognitive processes observed in the brain. Understanding the associations between EEG frequency bands and specific activities provides insights into brain function and cognitive states.

## 2.4 Sampling frequency

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

> In order to accurately capture and analyze the intricate details of brain waves without introducing aliasing effects, it is essential to consider the maximum frequency of these waves. According to Figure 2, the maximum frequency of brain waves is stated as 100Hz. To meet the requirements of the Nyquist rate, which states that the sampling frequency must be at least twice the maximum frequency of the signal, a minimum sampling frequency of 200Hz is necessary.

# 3   EEG Signal Processing

In this section, firstly you would get familiar with the task and the structure of the data.

## 3.1   Task Definition

[9] To identify the effect of olfactory dysfunction among different brain health states, the following task was performed to collect the data. The same sequence of stimuli was presented to all participants. The stimulation sequence was composed of two different odors, one occurring frequently (standard) with a probability of 0.75 and the other presented rarely (deviant) with a probability of 0.25. Each trial consisted of a 2s stimulus presentation followed by 8s of rest (pure water vapor). The odors were delivered to the participants using a laboratory olfactometer. The experiment involved 120 trials in which 90 frequent and 30 rare stimulation cycles were presented in a predetermined, randomized order. Lemon essence was used as the frequent odorant and rose essence was used as the rare odorant. These odors were selected to avoid trigeminal system activation as the olfactory and trigeminal systems are interconnected and may interact with each other during exposure to certain stimuli [10]. The duration of odor presentation was set at 2s to enable regular breathing cycles for the participants.

## 3.2   Data Description

[9] The dataset consists of three files as follows:

- `AD.mat`: Contains data for Alzheimer's disease patients.

- `Normal.mat`: Contains data for healthy elderly participants.

- `MCI.mat`: Contains data for mild cognitive impairment patients. (Described in part 5.1)

The structure of the files is the same. Each file is organized as a structure array, in which each row contains information of one participant and the three columns correspond to the "epoch", "odor" and "noisy" fields as described in Table 1.

| Field | Description |
|---|---|
| epoch | This is a 3D array structured as $4 \times 600 \times$ Num_trials. The first dimension indicates EEG channels respectively from the first column as Fp1, Fz, Cz, and Pz. The second dimension contains EEG samples from 1 s pre stimulus to 2 s post stimulus, which at a 200 Hz sampling rate amounts to 600 samples. The last dimension shows the number of trials. This could be different for each participant as some trials were deleted during preprocessing. |
| odor | This is a 2D binary array shaped as Num_trial $\times$ 1. This array shows the odorant type (lemon/rose) the participant was exposed to in each trial. The value = 1 indicates the rose odor and the value = 0 indicates the lemon odor. |
| noisy | This is a 2D array with the size $1 \times$ Num_noisy. This array indicates noisy trials identified based on comparing the instantaneous and average trial amplitudes. These noisy trails can be ignored in processing and were included for the dataset completeness. |

Table 1: Description of each structure array (.mat file) in the dataset.

## 3.3   Pre-Processing

Using a standard pipeline in EEG signal preprocessing is crucial for ensuring consistency, reproducibility, and objectivity in research. It reduces bias, enhances the reliability of results, and provides established best practices for addressing common challenges. A popular and widely used pipeline for EEG signal preprocessing is Makoto's pipeline (Makoto's preprocessing pipeline - SCCN).

The collected raw data from all participants were preprocessed following the full pipeline of Makoto with the use of EEGLAB and posted as a dataset, as described in the following steps:

1. Apply 1 Hz high pass filter to remove baseline drifts.

2. Apply relevant notch filter to remove the 50 Hz line noise.

3. Reject bad channels as a critical step before average referencing with the use of `clean_rawdata()` EEGLAB plugin.

4. Interpolate the removed channels.

5. Re-reference the data to the average of all channels to obtain a good estimate of referenceindependent potentials.

6. Apply `clean_rawdata()` for cleaning the data by running artifact subspace reconstruction(ASR).

7. Re-reference the data to the average again to compensate for any potential changes in the data caused by the previous step.

8. Run independent component analysis (ICA) to identify EEG sources as well as the sources associated with noise and artifacts.

9. Fit single and bilateral (if available) current dipoles.

10. Further clean the data by source (dipole) selection using `IClabel()` plugin in EEGLAB.

In the `Dataset/Preprocess` folder you can find the raw data for 2 subjects with the corresponding additional information provided. In this section you are required to preprocess these data and save your final preprocessed cleaned data.

However, there is no need to fully implement the Makoto's pipeline and a simplified version of this is as follows; follow the instructions below and provide the required results in each step:
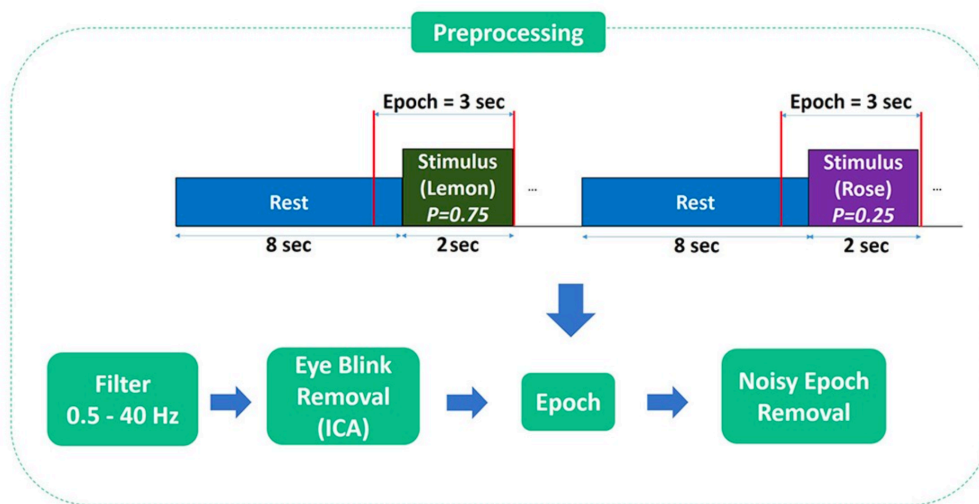


Figure 3: Task and Preprocessing Steps [7]

- **Step 1**: To preproccess using EEGLAB, first re-reference data to the mean of the channels. Then use a bandpass filter to filter 0.5 - 40.5 Hz frequencies. As we have filtered to 40.5 Hz, there is no need to apply a 49.9 - 50.1 Hz notch filter to remove the line noise (However, keep this step in mind as this is a crucial step in EEG signal preprocessing!). Using FFT function or EEGLAB, plot the frequency spectrum of `Fz` channel data. (Just to note, your data will be saved at EEG struct in MATLAB workspace.)

- **Step 2**: In this part you would remove the artifacts of the signal. Artifacts include blinking, eye movement, muscle movement, heart rate and etc. For this, load your data at EEGLAB. Now load `Standard-10-20-Cap19.loc` file from `edit-channel loacations` menu that contains locations of channels. 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`. By clicking on
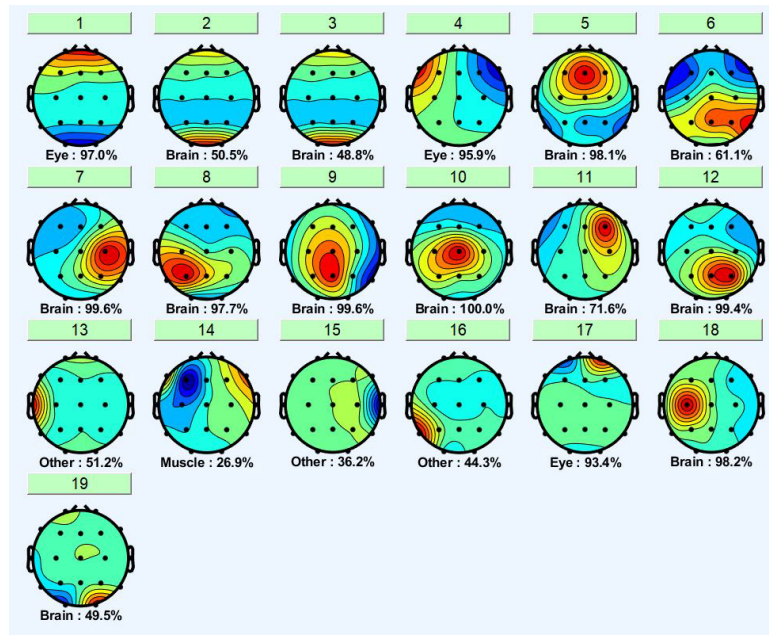
Figure 4: An example of ICA components

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

Now remove all non-brain components. For this purpose, from `Tools-remove components from data` enter the number of components that must be removed.

- **Step 3**: Epoch the data of each subject. Epoch is a 3D matrix of the shape {`Num_Channels` × `Samples` × `Num_Trials`}. In fact, all data must be reshaped as the following figure suggests:
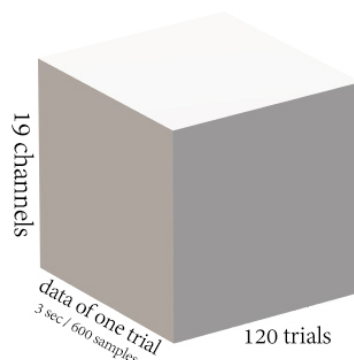


Figure 5: Epoch

For epoching the data, starting point of the experiment is required. This is provided in

the `help` file for each subject. Please note that you must epoch the data by considering this time as the start. Also, the data after 120 trials should be neglected as well.

- **Step 4**: In this step, you need to remove noisy trials. There are two ways to achieve this:

  1. Observe data at EEGLAB and remove any trial that seems noisy. (PREFERRED!)

  2. Using power spectrum of each trial, remove trials that their standard deviation of their power spectrum is bigger than 3.5 .Create a 3D matrix by each trial's power spectrum for each channel using `pspectrum` in MATLAB. You can use the following commands to find noisy trials:

```matlab
vr = sum(nanstd(p,[],2).^2,2);
noisy_trials = find(abs(zscore(vr))>3.5);
```

  In these commands, `p` is a matrix of frequency spectrums of all trials of a channel. `noisy_trials` contains the number of noisy trials of that channel. These commands must run for each channel individually and the resultant noisy trials must be accumulated over all channel. Then remove all `noisy_trials` from your epoch.

- **Step 5**: In the final step, only subsample the data corresponding to the `Fp1, Fz, Cz & Pz` channels. You can find the channels' orders in the `Channels.jpg`.

Do these 5 steps for each subject and save the final data through an `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 process is followed and the final data is stored in the project `preprocessing` directory. snapshots of the process for subject 1 follow:

```
#1: subject1

    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
```
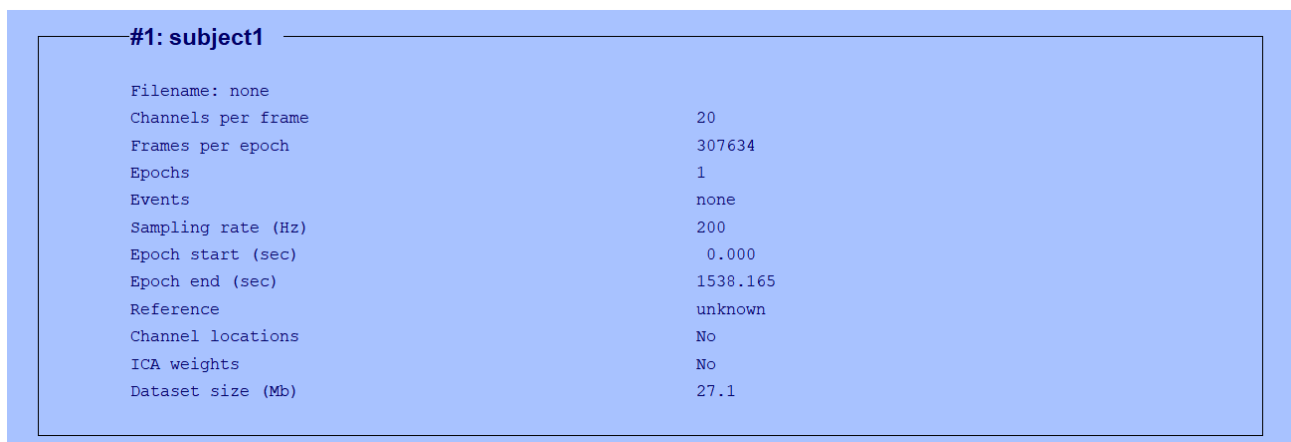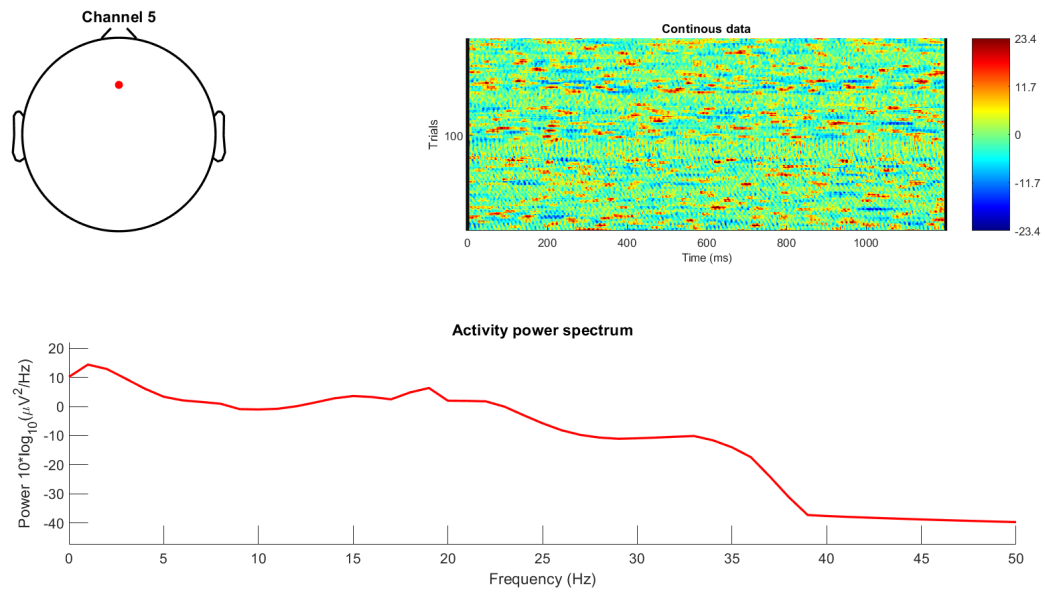
Figure 6: starting point
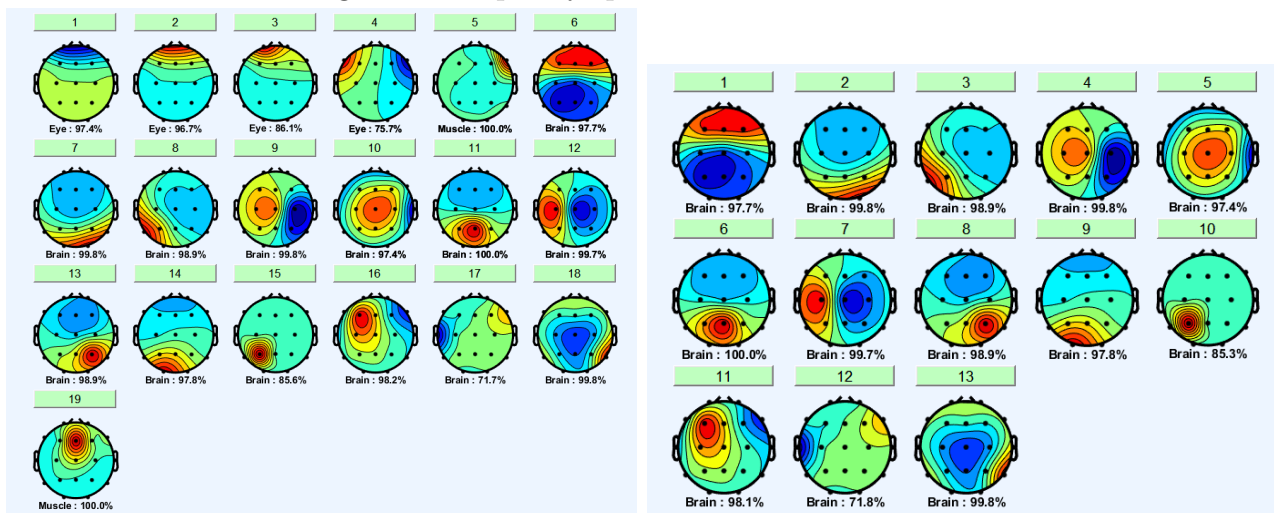
Figure 7: Frequency spectrum for the Fz channel



Figure 8: ICA labeling

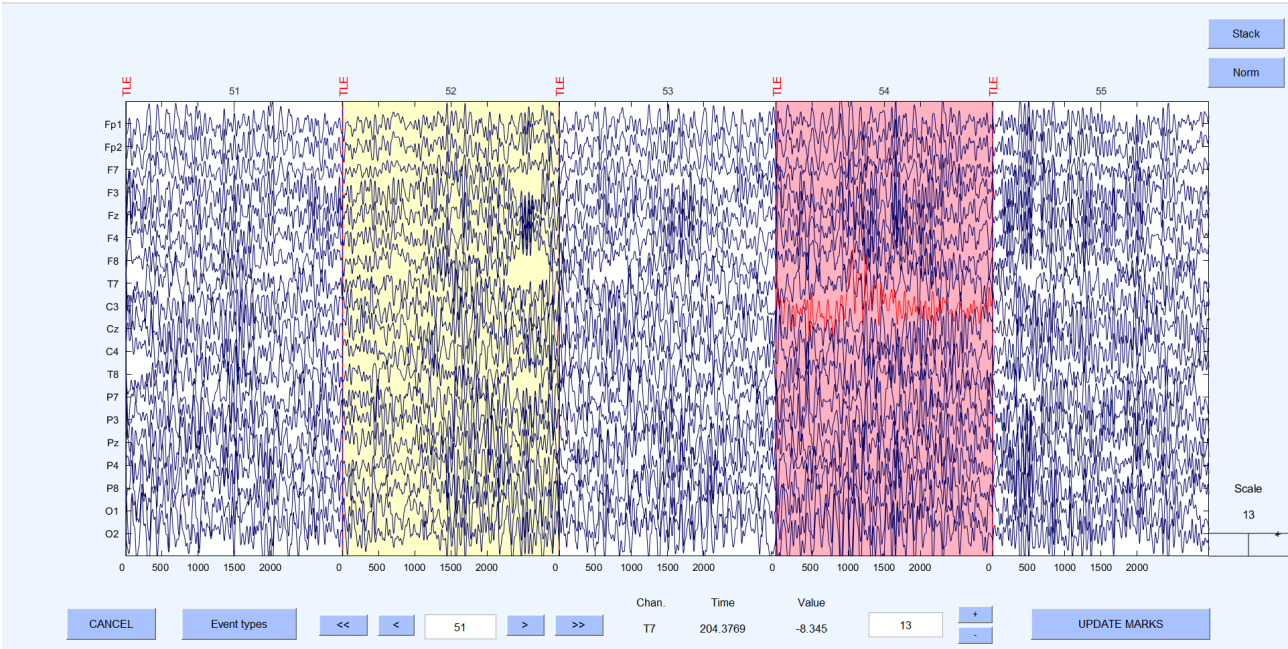we remove noisy epochs using the given toolboxes in EEGLAB

Figure 9: noisy epoch removal

## 3.4 Phase Locking Value (PLV)

Phase Locking Value (PLV) is a metric used to quantify the degree of phase synchronization or phase consistency between two oscillatory signals. It assesses the relationship between the phases of two signals at a specific frequency range. PLV is commonly used in the analysis of neural signals, including electroencephalography (EEG) and magnetoencephalography (MEG), to investigate the synchronization of oscillatory activity between different brain regions or across different frequency bands within a single region.. It provides insights into the functional connectivity and coordination of neural activity.

PLV ranges from 0 to 1, where a value of 1 indicates perfect phase synchronization, while a value close to 0 represents a lack of synchronization. High PLV values suggest that the phases of the two signals are consistently aligned or coupled, indicating strong synchronization. This synchronization can reflect functional interactions between brain regions or coordinated activity within a network. In contrast, low PLV values indicate weaker or desynchronized activity, suggesting less functional coupling between the signals.

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

> Phase synchronization, as indicated by the Phase Locking Value (PLV), carries important functional implications in the study of neural activity and brain function. It provides insights into the coordination and communication between different brain regions or within a single region. By analyzing the phase synchronization between oscillatory signals, researchers can uncover the functional connectivity and interactions that underlie various cognitive processes.
>
> The functional significance of phase synchronization lies in its association with information transfer and integration across brain regions. Synchronized oscillatory activity enables the coordination of neural firing and facilitates the communication and integration of information between different regions.

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

> The PLV is typically calculated using the following formula:
>
> $$PLV = |E(e^{i\Delta\Phi(t)})|, \quad \Delta\Phi(t) = \Phi_1(t) - \Phi_2(t)$$
>
> where $\Phi$ is the phase of the signal, more percsisely the phase of the analytical signal, where the analytical signal is defined by the following formula:
>
> $$z(t) = s(t) + ih(t), \quad h(t) = \mathbb{H}(s(t)) = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{s(\tau)}{t-\tau}d\tau$$
>
> [1, reference]

- Implement a function which finds the PLV between two channels in a specific frequency range. This function is going to be needed in the section 4. (NOTE: You are allowed to define this function with any required input arguments.)

> here is the implementation of this function in matlab.

```matlab
function plv = PLV(signals1, signals2, fs, bandwidth)
    if nargin < 4
        bandwidth = [0.001 fs/2-0.001]; % Default to all-pass
    filtering
    end

    numSignals = size(signals1, 3);
    plv = zeros(1, numSignals);

    for i = 1:numSignals
        signal1 = signals1(:,:,i);
        signal2 = signals2(:,:,i);

        % filtering the signal to the desired bandwidth
        filteredSignal1 = bandpass(signal1, bandwidth, fs);
        filteredSignal2 = bandpass(signal2, bandwidth, fs);

        % calculating phase difference
        phaseDiff = phase_diff(filteredSignal1, filteredSignal2);

        % Calculate the PLV
        plv(i) = abs(mean(exp(1j * phaseDiff)));
    end
end

function phaseDiff = phase_diff(signal1, signal2)
    analytic1 = hilbert(signal1);
    analytic2 = hilbert(signal2);
    phase1 = angle(analytic1);
    phase2 = angle(analytic2);
    phaseDiff = phase1 - phase2;
end
```

# 4    Results

In this section, you need to present the required results to assess the difference of Phase Locking Values (PLV) among two groups, namely AD and Normal in the slow gamma frequency range, which is 35 to 40 Hz.

To fairly compare your results in this part, you do not need to use your preprocessing data from section 3.3 and the preprocessed data of 15 healthy (normal) (age = 69.27±6.65, female = 53.33%) individuals and 13 AD patients (age = 75.31±9.90, female = 61.54%) are availabe through `Dataset/Normal.mat` and `Dataset/AD.mat`.

## 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 .

> To calculate the Phase Locking Value (PLV) between the specified channels for each participant, we employ the following procedure:
> Firstly, we utilize different epochs and determine the average signal for Fz and Cz in each trial. This step helps in mitigating any temporal effects by averaging out activities other than the stimulus. By computing the mean of the signal, we obtain a signal that solely represents the impact of the stimulus, effectively removing extraneous variations.
> Subsequently, we proceed to calculate the PLV between the Fz and Cz channels for each participant, distinguishing between rare and frequent stimuli. The resulting PLV values are then stored in an array, enabling further analysis and interpretation.
> For instance, when considering normal patients, the code implementation follows as follows:

```matlab
for p = 1:numNormalPatients
    % Extract relevant data for the current Normal patient
    person = normal(p);
    rare = person.epoch(:,:,logical(person.odor));
    frequent = person.epoch(:,:,~logical(person.odor));
    Fz_Cz_rare = squeeze(rare(2:3,:,:));
    Fz_Cz_frequent = squeeze(frequent(2:3,:,:));
    % Calculate PLV for rare trials
    Fz_mean = mean(Fz_Cz_rare(1,:,:),3);
    Cz_mean = mean(Fz_Cz_rare(2,:,:),3);
    mean_plv_rare_Normal(p) = PLV(Fz_mean,Cz_mean,sampling_rate,
    bandwidth);
    % Calculate PLV for frequent trials
    Fz_mean = mean(Fz_Cz_frequent(1,:,:),3);
    Cz_mean = mean(Fz_Cz_frequent(2,:,:),3);
    mean_plv_frequent_Normal(p) = PLV(Fz_mean,Cz_mean,sampling_rate,
    bandwidth);
end
```

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

> after finding the plv for each event, we simply draw a box plot of the plv values to see if the values are different.
> after that, we just use the matlab `fitdist` and then `ttest2` functions in MATLAB to test significance and to fit a gaussian distribution to our findings.



Mean PLV of AD and Normal Patients
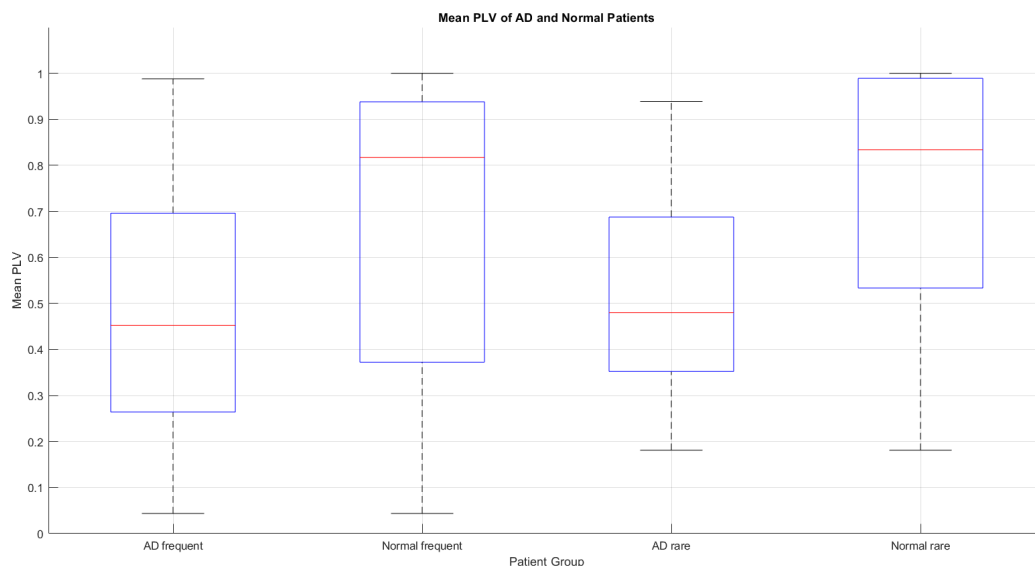
## 4.3    Statistical Significance

Based on the `p-values` you founded 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".

> by the results of the `ttest2`, we could say that:
>
> " **for rare trials, PLV is significantly different among AD and Normal subjects in the slow gamma frequency range ( p = 0.037 )** "
> " **for frequent trials, PLV is most likely different among AD and Normal subjects in the slow gamma frequency range ( p = 0.156 )** "
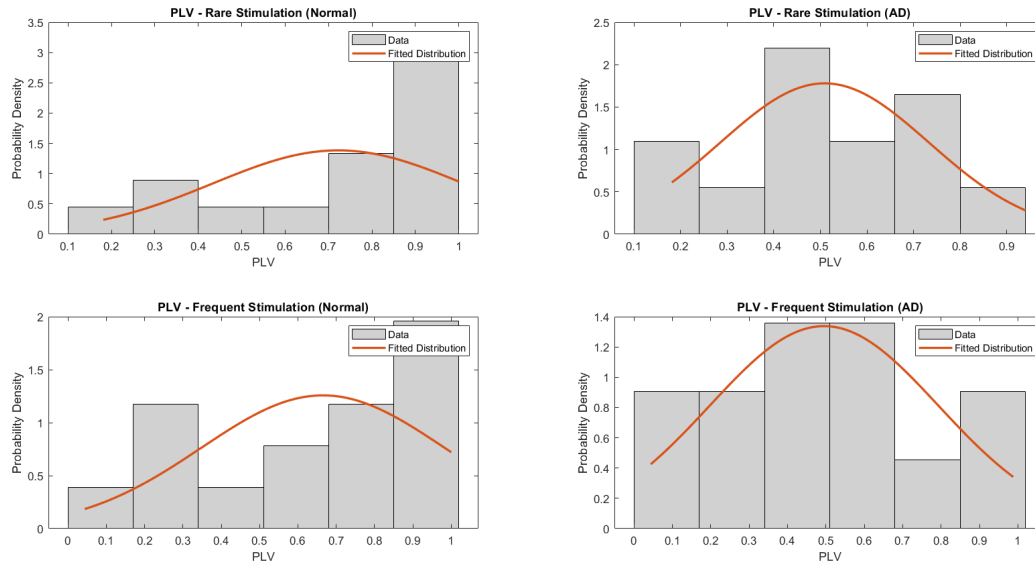
Figure 10: fitting gaussian distributions

## 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.

Using the previously defined phase difference for calculating the PLV between the two signals, we construct a histogram that represents the distribution of this phase difference. This histogram captures the accumulated phase data for all participants within each group, pertaining to each odor condition.

Essentially, this histogram provides a visual representation of the collective phase information for all participants in a given group, highlighting the variations and patterns in the phase differences between the signals. By examining the histogram, we can gain insights into the synchronization or desynchronization of the signals, allowing for a deeper understanding of the neural responses to different odors within each group.

as seen from the data, the AD patients have more sparse phase patterns.
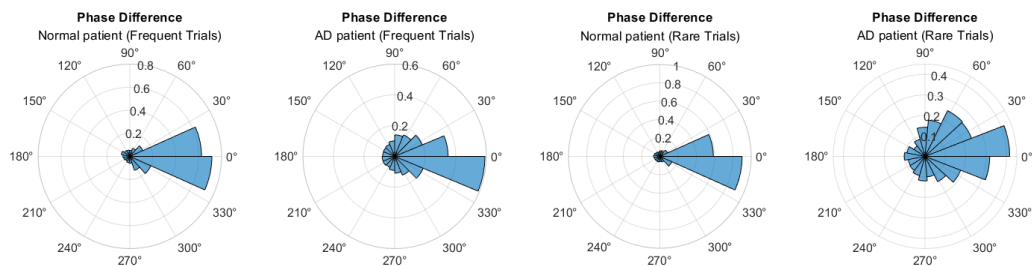
## 4.5   Heatmaps

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. (NOTE: You need to provide `p-values` for your hypothesis if you found any significantly different channel pairs apart from (`Fz,Cz`).)

here, we do the same kind of analysis we did before for calculating the plv between Fz, Cz channels. here we take the mean of the PLV between all participants.
The PLV matrix represents the average Phase Locking Value (PLV) between each pair of channels for a group of normal patients. Mathematically, the PLV between two channels, denoted as channels $i$ and $j$, is calculated as follows:

$$PLV_{ij} = \frac{1}{M} \sum_{m=1}^{M} \left| \frac{1}{N} \sum_{n=1}^{N} e^{i(\phi_i^{m,n} - \phi_j^{m,n})} \right|$$

where:

- $PLV_{ij}$ represents the PLV value between channels $i$ and $j$.

- $M$ is the total number of normal patients.

- $N$ is the number of trials per patient.

- $\phi_i^{m,n}$ and $\phi_j^{m,n}$ are the instantaneous phase values of channels $i$ and $j$ in the $n$th trial of the $m$th patient, respectively.

as seen from the results, the Fz, Cz channels indeed show the most significant difference in the PLV, so we conclude that Fz, Cz are the most different channels. the results follow:
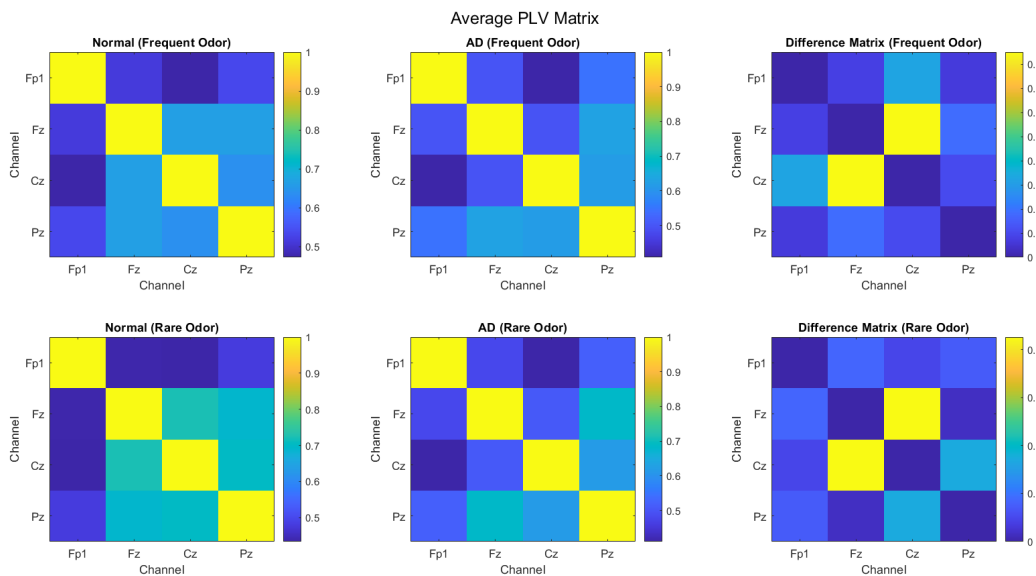
# 5 *Bonus

## 5.1 Mild Cognitive Impairment (MCI)

Mild Cognitive Impairment (MCI) is the stage between the expected decline in memory and thinking that happens with age and the more serious decline of dementia. MCI may include problems with memory, language or judgment. People with MCI may be aware that their memory or mental function has slipped. [2]
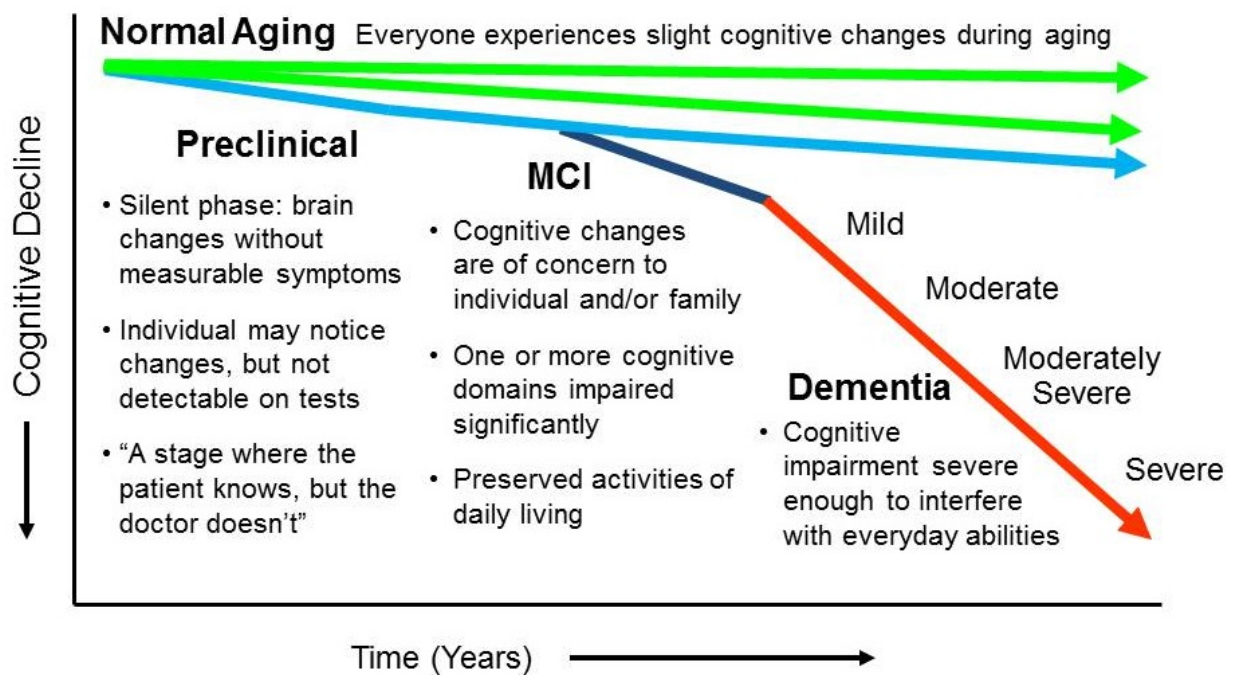


Figure 11: Normal Aging to Demantia Process

### 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.

> MCI is a stage characterized by mild cognitive decline that is greater than expected for age but does not meet the criteria for dementia. While MCI increases the risk of developing AD, it does not always lead to AD, and some individuals with MCI may remain stable or even improve over time. The causes of MCI are diverse and can include age-related changes, vascular issues, neurodegenerative diseases, genetic factors, and lifestyle/environmental factors.

### 5.1.2   MCI Data Processing

In the provided dataset, you can find `MCI.mat` file. This dataset contains preprocessed cleaned EEG recording of the same task described in sections 3.1 and 3.2 for 7 MCI patients.

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`.

---

here we apply the same methods we used in the previous section, the results follow: due to limited number of MCI subjects, the p-values don't show much significance.

" **for rare trials, PLV is most likely different among AD and MCI subjects in the slow gamma frequency range ( p = 0.17 )** "
" **for frequent trials, PLV is significantly different among AD and MCI subjects in the slow gamma frequency range ( p = 0.045 )** "
" **for rare trials, PLV might be different among Normal and MCI subjects in the slow gamma frequency range ( p = 0.74 )** "
" **for frequent trials, PLV might be different among Normal and MCI subjects in the slow gamma frequency range ( p = 0.69 )** "

---
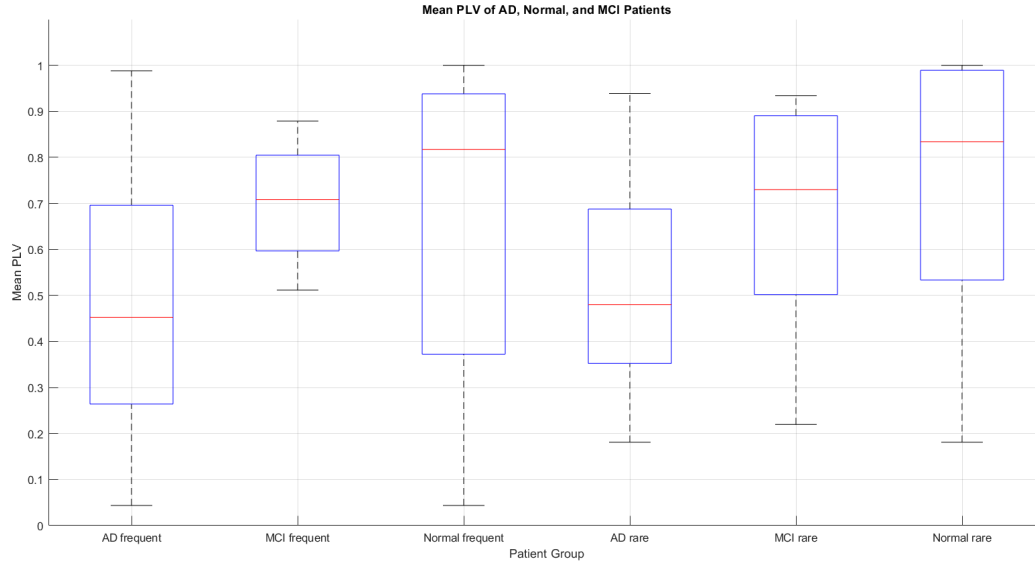


## 5.2   Phase-Amplitude Coupling (PAC)

PLV was just one instance of the Phase-Amplitude Coupling (PAC) metrics. PAC is a form of cross-frequency coupling where the amplitude of a high frequency signal is modulated by the phase of low frequency oscillations. PAC is the most-studied type of cross-frequency coupling and is thought to be responsible for integration across populations of neurons. Low

---

frequency brain activity controls the information exchange between brain regions by modulating the amplitude of the high frequency oscillations. [8]

### 5.2.1 Metrics

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

- Phase-locking value (PLV): This is a measure of the consistency of the phase difference between two signals at a given frequency. It ranges from 0 to 1, with 1 indicating perfect phase locking.

- Mean vector length (MVL): This is a measure of the strength and direction of the coupling between two signals. It ranges from 0 to 1, with 1 indicating perfect coupling.

- Modulation index (MI): This is a measure of the degree to which the amplitude of one signal is modulated by the phase of another signal. It ranges from -1 to 1, with 0 indicating no modulation.

- Generalized linear modeling cross-frequency coupling (GLM-CFC): This is a method that uses a generalized linear model to estimate the strength and direction of coupling between two signals at different frequencies.

[6] here we arbitrary chose to implement the modulation index.

### 5.2.2 Implementation

Implement one of the metrics mentioned earlier as a biomarker for distinguishing between AD and Normal groups. Present the relevant results through plots and provide a discussion regarding the efficacy of the selected metric.

here we arbitrary chose to implement the modulation index, sadly the results were not very significant for any of the channels.

```matlab
function MI = modulation_index(x, fs, bandwidth, num_bins)
    if nargin < 3
        bandwidth = [0.001 fs/2-0.001]; % Default to all-pass filtering
    end if nargin < 4
        num_bins = 18; % Default to 18 bins
    end
    numSignals = size(x, 1);
    MI = zeros(1, numSignals);
    for i = 1:numSignals
        signal = x(i,:,:);
```

```
11          % Filtering  the  signal  to  the  desired  bandwidth
12          filteredSignal  =  bandpass(signal, bandwidth, fs);
13          analyticSignal  =  hilbert(filteredSignal);
14          % Calculate  the  phase  and  amplitude  of  the  analytic  signal
15          phase  =  angle(analyticSignal);
16          amplitude  =  abs(analyticSignal);
17          % Bin  the  phase  values  into  the  specified  number  of  bins
18          bins  =  linspace(-pi, pi, num_bins+1);
19          phase_bins  =  discretize(phase, bins);
20          % Convert  phase_bins  to  numeric  array
21          phase_bins  =  double(phase_bins);
22          % Calculate  the  average  amplitude  in  each  phase  bin
23          avg_amplitude  =  zeros(num_bins, 1);
24          for  j  =  1:num_bins
25              values  =  amplitude(phase_bins  ==  j);
26              if  ~isempty(values)
27                  avg_amplitude(j)  =  mean(values);
28          end end
29          % Normalize  the  average  amplitudes
30          normalized_amplitude  =  avg_amplitude  /  sum(avg_amplitude);
31          % Calculate  the  Shannon  entropy
32          entropy=-sum(normalized_amplitude.*log2(normalized_amplitude));
33          % Calculate  the  Kullback-Leibler  distance
34          KL_distance  =  log2(num_bins)  -  entropy;
35          MI(i)  =  KL_distance  /  log2(num_bins);
36      end
37  end
```
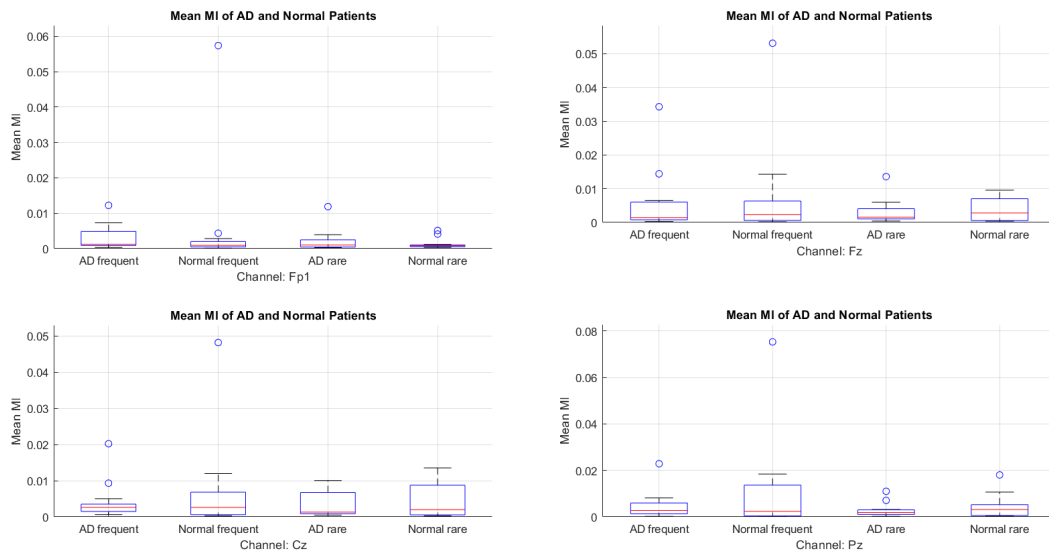


Figure 12: findings for the MI index of each channel

# 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.

**Abstract**

In summary, our study aimed to investigate phase-amplitude coupling using different metrics, including the Phase Locking Value (PLV), Modulation Index (MI), and Generalized Linear Model-Phase Coupling Factorization (GLM-CFC).

When analyzing the PLV, we observed significant differences between groups in the slow gamma frequency range. Specifically, for rare trials, the PLV was significantly different between AD and Normal subjects (p = 0.037), while for frequent trials, the PLV showed significant differences between AD and MCI subjects (p = 0.045). However, the significance of the findings was limited due to the small number of MCI subjects.

Although we implemented the modulation index, the results did not yield significant findings for any of the channels studied. This suggests that phase-amplitude coupling, as measured by the modulation index, may not play a prominent role in the observed differences between groups.

Furthermore, we analyzed the PLV matrix to examine the average PLV between channels for a group of normal patients. The Fz and Cz channels exhibited the most significant differences in PLV, indicating that these channels are the most distinct among the measured pairs.

Due to the limited number of MCI subjects in our study, the p-values obtained did not show significant differences. However, the results suggest potential differences in PLV between AD and MCI subjects in the slow gamma frequency range. Specifically, for rare trials, the PLV is most likely different between AD and MCI subjects (p = 0.17), while for frequent trials, the PLV is significantly different between AD and MCI subjects (p = 0.045).

It is important to acknowledge the limitations of our study, including the small number of MCI subjects and the arbitrary choice of the modulation index. Future research should explore alternative measures and methods to further investigate phase-amplitude coupling and its relevance to cognitive disorders.

# References

[1] Aydore s, pantazis d, leahy rm. a note on the phase locking value and its properties. neuroimage. 2013 jul 1;74:231-44. doi: 10.1016/j.neuroimage.2013.02.008. epub 2013 feb 19. pmid: 23435210; pmcid: Pmc3674231.

[2] Mild cognitive impairment (mci). https://www.mayoclinic.org/diseases-conditions/mild-cognitive-impairment/symptoms-causes/syc-20354578#:~:text=Overview,mental%20function%20has%20%22slipped.%22. Last Reviewed: Jan. 18, 2023.

[3] Neural oscillations – interpreting eeg frequency bands. https://imotions.com/blog/learning/best-practice/neural-oscillations/.

[4] Neurodegenerative diseases. https://www.niehs.nih.gov/research/supported/health/neurodegenerative/index.cfm. Last Reviewed: June 09, 2022.

[5] Marin C, Vilas D, Langdon C, Alobid I, López-Chacón M, Haehner A, Hummel T, and Mullol J. Olfactory dysfunction in neurodegenerative diseases. In *Curr Allergy Asthma Rep*, volume 18, 2018 Jun 15.

[6] Mareike J. Hülsemann, Ewald Naumann, and Björn Rasch. Quantification of phase-amplitude coupling in neuronal oscillations: Comparison of phase-locking value, mean vector length, modulation index, and generalized-linear-modeling-cross-frequency-coupling. *Frontiers in Neuroscience*, 13, 2019.

[7] Sedghizadeh MJ, Hojjati H, Ezzatdoost K, Aghajan H, Vahabi Z, and Tarighatnia H. Olfactory response as a marker for alzheimer's disease: Evidence from perceptual and frontal lobe oscillation coherence deficit. *PLOS One*, 15(12), December 15, 2020.

[8] T.T.K. Munia and S. Aviyente. Time-frequency based phase-amplitude coupling measure for neuronal oscillations. In *Scientific Reports 9, 12441*, 27 August 2019.

[9] Mohammad Javad Sedghizadeh, Hamid Aghajan, and Zahra Vahabi. Brain electrophysiological recording during olfactory stimulation in mild cognitive impairment and alzheimer disease patients: An eeg dataset. *Data in Brief*, 48:109289, 2023.

[10] Cécilia Tremblay and Johannes Frasnelli. Olfactory and Trigeminal Systems Interact in the Periphery. *Chemical Senses*, 43(8):611–616, 07 2018.