E.T.S. of Industrial Engineering,
Informatics and Telecommunications

Phase synchrony modulations in pre-term infants

BIOMEDICAL ENGINEERING

Bachelor's degree final thesis

Íñigo Herrero Vidaurre

Gabriele Arnulfo, PhD & Monica Roascio

Genova (Italy), June 2022

UPNA (Home University)

upna
Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa

BACHELOR FINAL THESIS

UNIVERSITÀ
DEGLI STUDI
DI GENOVA

BIOMEDICAL ENGINEERING

Phase synchrony modulations in preterm infants

Author: Íñigo Herrero Vidaurre

Tutor: Prof. Gabriele Arnulfo (phD), Monica Roascio

Genova (Italy), June 2022

UNIVERSITÀ DEGLI STUDI DI GENOVA (Mobility University)

ACKNOWLEDGMENTS

Thanks to Prof. Gabriele Arnulfo, and future PhD Monica Roascio, for helping me in every step of the process, and for tutoring the thesis in Genova.

Also, thanks to the UPNA organization and professors for their resources, knowledge and continuous availability.

But a special thanks to my family and friends for their constant support through these last four years. I could not have done it without you. Grateful, as well, to those that have made my stay abroad, in Italy, an unforgettable experience.

'In research science, the process is more important than the outcome' - Emilio Muñoz Ruiz.

ABSTRACT

Phase synchrony analysis has been developed since years ago to test and measure the statistical connectivity of two regions. Specifically, in this thesis, a code to get one metric of phase synchrony is developed, the Phase Locking Value or PLV. This metric allows us to quantify the connectivity of two regions of the brain. Which could have an immense impact if correlated with clinical outcomes, with the potential of predicting the outcomes of preterm neonates. Being a preterm neonate, is still in 2022, the major cause of mortality for children under five years old around the world.

Hence, learning about the connections of these patients will help understand their neural development while helping them in the near future. Here, by analyzing the PLV and iPLV for preterm and term patients, we learn how those populations relate to each other.

Key Words: Phase Locking Value, neonate, preterm, EEG.

INDEX

1. INTRODUCTION

1.1. PROJECT FRAMEWORK

The following final thesis for the Degree in Biomedical Engineering at the Public University of Navarre, is done during a mobility exchange at the University of Genoa at DIBRIS (*Dipartimento di Informatica, Bioingegneria, Robotica e Ingegneria dei Sistemi*).

Although the goal of the project is just to analyze and compare the results of PLV values of EEG data, it is framed under a research work to develop a method to predict the outcome of preterm neonates. In order to do so, a python script is created and developed, with the aim of taking out of the raw EEG signals some determinative features to characterize them.

Besides, the EEG signals are acquired and classified in the *Istituto Giannina Gaslini* (Genova), by Dra. Sara Uccella (MD), during the first trimester of 2022.

1.2. MOTIVATION AND JUSTIFICATION

According to the WHO [1], the words 'preterm neonates', is defined for babies with a gestation of no more than 37 weeks. There are sub-categories, such as extremely preterm (less than 28 weeks), very preterm (28 – 32 weeks) and moderate to late preterm (32 – 37 weeks).

There are more than 15 million premature births each year, that sums up to be more than 10% of the total births in the world. The main problem remains in the outcome of those preterm babies, since every year, around 1 million die. Nowadays, being preterm is the main cause of death in children under 5 years old, due to complications during labor, or developing problems.

As will be explained in the following paragraphs, controlling, and preventing preterm births is complicated, thus, the main health resources are being expended in taking care of those neonates, to increase their chance of survival, decreasing the +10% chance of dying.

In that sense, the justification of this project is to understand more correctly the preterm neonate EEG, to help develop a method to predict the positive or negative outcomes of those premature babies, to assist medical professionals in their daily care and decisions.

1.3. OBJECTIVES

- Develop the necessary code to obtain a feature that characterizes the EEG signal.
- Understand the correlation between those features and each type of patient.
- Aid in the development of an algorithm to predict the neural outcome of neonates.

1.4. MATERIALS

As it has been previously exposed, the EEG data is acquired and sent to DIBRIS from the Gaslini Hospital in Genova, Liguria, from preterm - neonates. Specifically, 25 patients' eeg data are sent, each one divided in one hour edf (European Data Format) files.

There are patients that, because of their medical condition, present one or more EEG recordings, at 30-35 weeks (of pregnancy), and 40 weeks. It might be said that normally, each recording has a total duration of 20h, but, for external reasons, some of them are shorter.

Hence, this is the amount of data received and used for the project:

CATEGORY	NUMBER OF FILES
PRETERMS WITH ONE RECORDING BEFORE 35 WEEKS	17
PRETERMS WITH TWO RECORDINGS	4
TERM PATIENTS	4

Table 1.1. Summary of the amount of EEG data received.

Apart from that, for three patients, the university received proper information regarding its sleep classification (hypnograms). This was contained in a matlab file, from which the information was extracted. This file contained a folder called ‘ipnog’, with several numbers corresponding to a sleep phase, each one corresponding to a 20 seconds epoch with one sleep phase. Sleep phases and its numbers:

SLEEP PHASE	NUMBER IN MATLAB FILE
Move	1
Wake	2
Active Sleep (AS)	3
Sleep onset AS	4
Undetermined	5
Quiet Sleep (QS)	6

Table 1.2. Summary of EEG classification

2. STATE OF ART AND ANATOMY

2.1. GENERAL BIOLOGY

2.1.1. BRAIN

The brain, with the marrow, creates what is called the central nervous system, which controls the majority of the functions of the body and the mind.

We can divide the anatomy of the brain in different regions such as forebrain (telencephalon and diencephalon), midbrain and hindbrain (metencephalon and myelencephalon).

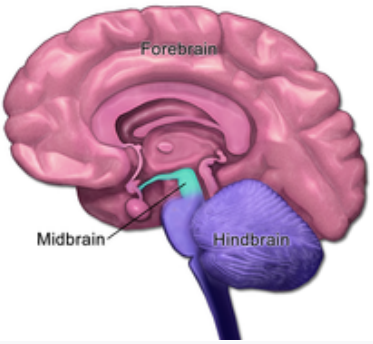

Figure 2.1. Midbrain and hindbrain [2].

If we talk about EEG as we are going to do in this thesis, the motor cortex is the most determinant part. It is the exterior part of the brain, with a height of 2-3 mm, and 2,5 m^2 of area. It is not regular, as it has convolutions (to allocate many neurons), as seen in the following figure.

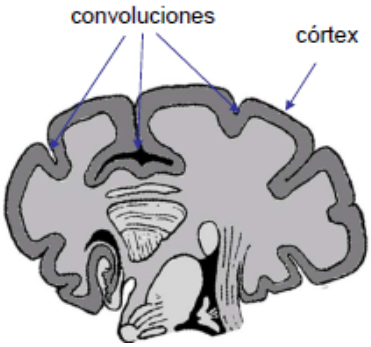

Figure 2.2. Brain convolutions and cortex [3].

It includes more than 10^{10} neurons [3], which are the most important part of the central system, as they produce the majority of the cognitive functions, like memory, sense, learning, voluntary movement, talk, perception…

Also, the cortex is divided in two hemispheres, and each one in four lobes (those are part of the telencephalon) [4]:

- Frontal Lobe: Importance in voluntary movement, expressive language and managing hard tasks.

- Parietal Lobe: This part is specified in receiving and interpreting somatosensory input.

- Temporal Lobe: It is related to processing auditory inputs, and memory.

- Occipital Lobe: As it is located in the front part, it is related with vision perception.

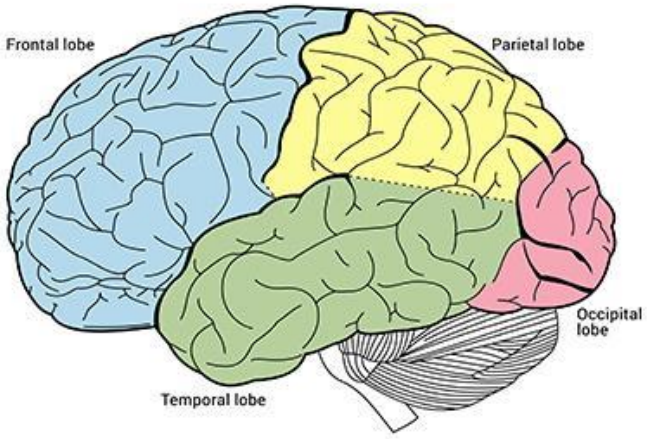

Figure 2.3. Brain Lobes [5].

The cortex is protected by three layers of meninges [6], dura mater, pia mater and arachnoid mater. They create a structure that gives a supportive framework to the brain vasculature and protects the brain from mechanical damage.

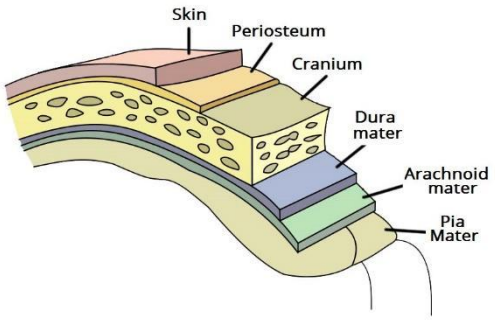

Figure 2.4. Brain Meninges [6].

In the cortex, each part of the body has its own area where all the sensory and motor processing is done. The size of this area depends on the precision required in that part of the body to do all the sensory or motor function. For example, in the cortex of the human brain, the area related to the hands is the biggest, to allow high precision in sensory and motor activities.

2.1.2. NEURON

The neuron is the basic element of the nervous system, and is its activity, what we are capturing with techniques like EEG. The neuron, as a system, receives inputs (from the dendrites), decides (in the soma), and gives an output through its axon.

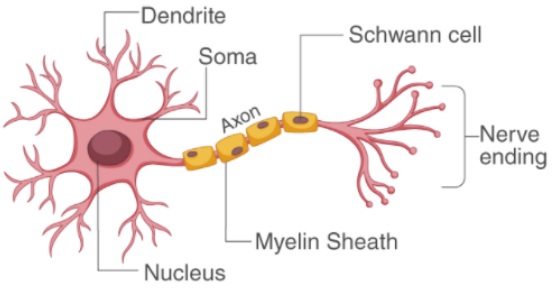

Figure 2.5. Neuron physiology [7].

Although there are many types of neurons, in an EEG, the technique in which this thesis will be based on, we detect the activity of the pyramidal neurons. Its morphology is multipolar, which means that it has many branch prolongations (for inputs) in the dendrites. It has an apical dendrite that extends itself towards the surface of the cortex. Also, the axon is some centimeters long, as it has to reach the spinal cord.

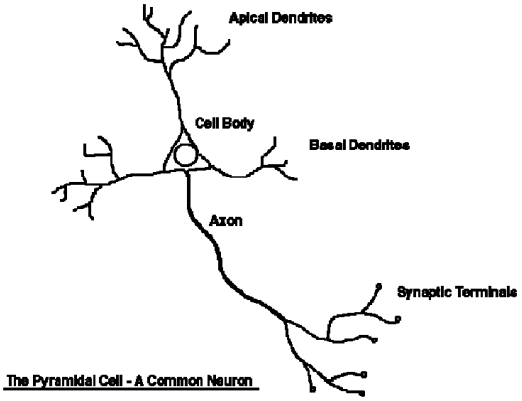

Figure 2.6. Pyramidal Neuron [8].

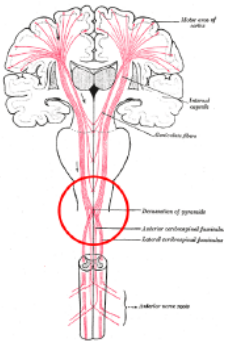

Figure 2.7. Corticospinal tract [3].

The corticospinal tract is the main via for the motor impulses. All the axons are from motor neurons, they go from the motor cortex to the spinal cord, where a synapsis is

produced, with the inferior motor neuron. Around 80% of the axons are decussated, near the medulla oblongata, that is the explanation for why the movements of one side of the body are controlled by the opposite side of the brain.

Electrical behavior of the neuron

The presynaptic neuron sends an action potential through its axon, until it reaches the presynaptic ending. There, it produces an entrance of calcium ions, which finishes in the liberation of neurotransmitters to the space between neurons (synaptic cleft).

The neurotransmitters arrive at the postsynaptic membrane and are bonded to some receptors. This produces an opening in the ion channels of the membrane, changing its potential. If positive ions enter the postsynaptic membrane, the action potential formed is excitatory, however, if the ions are negative, the action potential is inhibitory.

The previous explanation is not real, as normally, many impulses arrive to each neuron at the same time, it is a whole system. Throughout all the dendrites, the soma receives both inhibitory and excitatory action potential. The soma has to decide whether to send a positive or negative action potential. Normally it has a threshold, that the difference between the inhibitory and excitatory inputs should have, to decide.

Sum of all the excitatory and inhibitory inputs -> Decision in the soma using thresholds -> Send an action potential.

The flow of positive and negative ions during synapsis, can be detected in the surface of the head by using electrodes, that is the basic idea of an EEG, the technique used in this thesis, it would be explained later.

The ion flow produces an electric field that can be compared to a dipole when registered at the cortex.

2.1.3. SLEEP IMPORTANCE

Sleep is important in every stage of life, as it helps maintain brain health, through memory consolidation, emotional processing, and maintaining neural network and synaptic plasticity.

It has been discovered to be even more important in the first stages of life, as it helps develop the brain, hence, analyzing sleep (sleep ontogenesis), as it has been previously said, is the best way of knowing and controlling a child's evolution.

However, for some preterm babies, sleep is constantly disrupted in the ICUs, which might translate into future neural problems.

2.2. EEG DESCRIPTION

2.2.1. BRIEF HISTORY OF THE EEG

The first ever EEG recording of a human occurred in 1924 [9], thanks to the work of Hans Berger, a German psychiatrist, although it might be said that 50 years before, Richard Caton was able to perform the first EEG in history, to an animal. Throughout the following years, the technique arrived in different countries such as the USA or UK, where its potential was viewed [10]. Laboratories and societies were created to investigate this new technology, as it was valued as an indicator of brain function. In these early stages, it was used to evaluate illnesses or behaviors, sometimes in an inappropriate way, like many other psychiatric techniques during the XX century.

Until the 1960s, EEG data was examined visually, but thanks to the beginning of the digital era, Fourier, spectral analysis and other tools were available for doctors and researchers, opening a new way of understanding the EEG.

2.2.2. EEG DEVELOPMENT

EEG stands for Electroencephalography, it is a technique that uses electrodes to capture the electrical activity of the brain, which gives relevant information about the activity and functionality of the brain.

Specifically, the EEG signal is formed out of the sum of the cortical pyramidal neurons in the motor cortex [3], that are oriented perpendicularly to the surface of the brain. This means, the electrodes are not able to capture the electrical activity of the whole brain, but the sum of the activity of the cortical neurons.

It is possible to detect this activity because of three main facts:

- The pyramid neurons are parallel to each other, producing a signal that is the sum of all of them, a dipole of big amplitude. Other neurons are oriented randomly, and their signals are canceled.

- The temporal activity of these types of neurons has a different type of synchronization, producing for example an additive sum, when high synchronization, and destructive sum when low.

- These neurons are the first neurons of the brain, which means, they are within a short distance of the electrode, allowing it to capture the activity.

The second and third facts, manifest the idea that thanks to synchronization and the short distance, the potential generated by the electric field is bigger than the noise, which means that we are able to differentiate it.

Following the second fact, it is directly obtained that the amplitude of the EEG signal depends on the quantity of synchronization. The frequency of the signal is more complicated to guess as it depends on some brain parts that we have talked about

before, like the thalamus. In any case, there are some recognizable patrons such as high frequency and low amplitude (low synchronization), which means dreams with reverie, although we will talk about the main patrons and frequencies in the following paragraphs.

There are five main rhythms [11]:

- **Delta (0,5 – 4 Hz):** Deep sleep state.
- **Theta (4 – 8 Hz):** Self-absorbed individual.
- **Alfa (8 – 11 Hz):** Individual relaxed with eyes closed, but awake.
- **Beta (11 – 30 Hz):** Awake and focus, but also in some sleep stages.
- **Low Gamma (30 – 55 Hz)** and **High Gamma (55 – 80 Hz):** Active state with brain processing.

2.2.3. EEG IN POLYSOMNOGRAPHY ANALYSIS

A polysomnography or sleep study is a medical test required to study and discover, generally, sleep disorders, commonly known as the Gold Standard of neurophysiology [12]. However, as it has been previously explained, sleep is a decisive stage in neural development for neonates, hence, this study helps following the growth and development of those children.

EEG analysis is the differentiating aspect between polysomnography and other less relevant sleep studies because it allows the visualization and classification of the sleep phases. However, it is composed, also, of other signals:

- <u>EMG:</u> It is recorded in a simple way, it enables information required to characterize some sleep phases, and even, in adult polysomnography study, helps detect illnesses such as restless sleep disease.

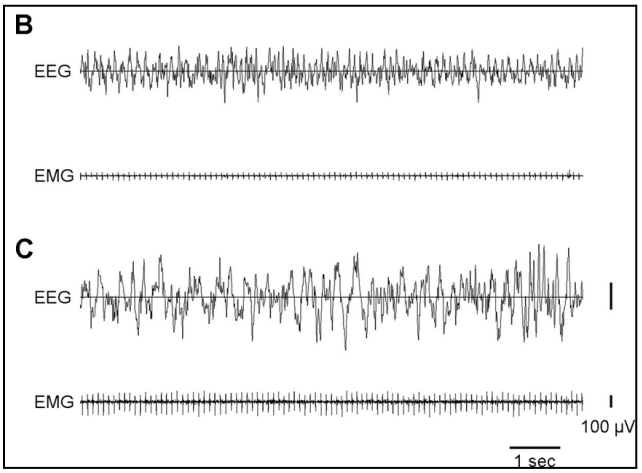

Figure 2.8. Difference in REM (B) and non-REM (C) in muscle activity (EMG) of in an adult polysomnography. [12]

For example, a defined characteristic of adult REM sleep phase or quiet sleep in neonates, is the lack of muscle movement, which will correspond to a flat EMG signal.

Somehow, as it detects muscle activity, it can help visualize muscle artifacts in the EEG signal.

- ECG and blood dynamics tests: In adult pathological sleep disorder polysomnography, is important as many sleep disorders are related to cardiac anomalies.

- Audio and image: Complementary information to aid in the sleep classification.

This allows clinical personnel to do a classification of the different sleep phases. Adult classification is different from neonates. Indeed, for them, which are our patients, we have two main sleep phases, Quiet Sleep and Active Sleep, which could be similar to REM and NREM phases in adults [13].

- Quiet Sleep: In a way, the NREM phase for neonates. Without eye movements, but with chin activations (one of the main differences compared to NREM). Regular breath and heart rates.
- Active Sleep: Resemble to REM sleep phase. In this case, they have eye movements, as well as small limb and face movements. Also, there is no notable chin movement. Besides, breath and heart rates are irregular.

2.2. PREMATURE BIRTHS

As it has been exposed previously, there are more than 15 million preterm births each year globally, which sums up to be more than 10% of the total births in a year. This means, 1 in every 10 newborn babies is born prematurely.

Of those 15 million, 1 million die of direct contribution to the risks that are being born before the 38 week.

Since 2010, the World Health Organization, WHO, has been instructing member countries to develop care methods for preterm neonates, while reporting results and statistics.

According to WHO, more than three quarters of premature births can be saved without many expenses, such as intensive care, steroids injections to help in the development of the body, or the use of antibiotics.

One of the main keys for preterm newborn babies' survival and positive development is the constant surveillance and test, to corroborate how the neurodevelopment is going, and avoid malformations. Testing and reviewing is done according to the risk of the baby, in the following chart there are some of the main risk aspects to consider [14].

1. Less than 32 weeks of gestation and/or less than 1500g of weight.
2. Below 3rd percentile in the gestational week.
3. Hypoxic ischemic encephalopathy.

4. Chronic lung or cardiopulmonary diseases.
5. Main malformations.
6. Abnormal neurological test.

Table 2.1. Main facts for risk in preterm neonates.

For high-risk babies (with the first characteristic of the previous chart), as it has been said, special attention is required. Tests are made regularly, to review the state of the newborn, and to try to predict his or her outcome. In this predictivity field is where this thesis is located. Some of the test that are carry out, not only during IC but also during their first years of life, are:

- Brain echography.
- M – CHAT
- Magnetic Resonance.
- Blood tests.
- Pediatric analysis (tests to confirm child development).
- Auditory potentials.
- Intelligence tests.
- Neurological tests.

Preterm birth occurs for many reasons, but mainly for induced or non-induced labor, for medical or non-medical reasons, such as complications due to illnesses, like diabetes or high pressure, or for not-know situations. Nevertheless, it has been proved that premature births are related to bad gestation behaviors, which is germane to poor countries and areas. That explains why Malawi, Comoros and Congo are the countries with the highest rates of preterm births, 16 to 18%.

2.3. MODERN TECHNIQUES

ERPs or Event Related Potentials [15], the measure of the brain response to a sensory event can be used to characterize brain development in neonates. However, measuring sleep brain connectivity, as in Functional brain Connectivity (FC) [16], as it represents the brain's ability to process and load information, is one of the oldest techniques to analyze neural development.

Functional connectivity is the connectivity between distant regions of the brain, shown by the coincidence of events. Basically, it assumes that if two events occur at the same time, that might be because those areas are connected in some way. This is opposite from the effective connectivity analysis.

One kind of FC technique is the Phase Locking Value (PLV), which is the main metric of this thesis.

Sleep EEG classification is defined for adult humans, but there is a big uncertainty in classifying preterm neonates' sleep, due to its neural development.

For example, it is only from 30 to 31 weeks when rapid-eye movement and time-locked characteristics are recognizable, producing a more common EEG signal. Even during the last weeks of preterm neonates, such as 37-38 weeks, there are still crucial developments, which lead to a more continuous EEG. With that being said, the main help to define the EEG is using cerebral and non – cerebral measures.

3. METHODOLOGY

3.1.GENERAL PROCESS

The code that has been developed, for the datasets with hymnogram information, aims to get a metric out of the signals [17], the Phase Locking Value [18]. After initializing all the necessary libraries, frequencies and paths, the first part of the process is taking out the information of the sleep stages that is contained on the matlab files 'PSGx-mat'. After doing that, thanks to the function *loadmat* of *Scipy*. We can easily get some statistics of the signal, such as how many sleep stages we have and their duration, number of transitions, etc. One important characteristic that can be obtained out of this is the minimum sleep stage duration of the recording, because in view of that information, the minimum epoch length for the analysis is created.

The function 'Results' accomplishes that. After *loadmat*, a python dictionary is achieved, and it is possible to navigate through the different folders until getting the 'ipnog' folder, in which the information regarding the sleep phase is contained.

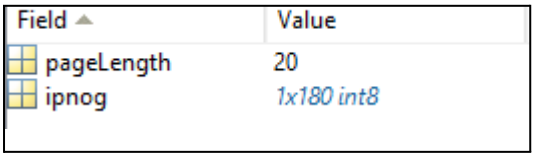

Figure 3.1. 'Ipnog' folder inside matlab file. With 180 sleep phases, each one corresponding to an epoch of 20 seconds (pageLength).

Specifically, the medical doctor that acquired and revised the hypnogram, uses epochs of 20 seconds to classify the signal, which means, there are [number of seconds of the signal] / 20, numbers in that folder, each number corresponding to a sleep phase, as it can be seen in figure 3.1. in the materials part of the document.

As it has been previously discussed, each patient is divided into .edf files of normally one hour, so the same process is done on all the files thanks to a 'for cycle'.

In that cycle, the signal is read from the right path, thanks to mne *read_raf_edf* function, and the bipolar reference is set:

- Fp1 – C3
- C3 – O1
- Fp1 – T3
- T3 – O1
- Fp2 – C4
- C4 – O2
- Fp2 – T4
- T4 – O2

A notch filter could be applied later in order to eliminate the electrical artifacts that might be hidden in the signals, and then filter the signal between 0,5 and 70 Hz (preprocessing [19]). Nevertheless, because of using a wavelet filter between 2 and

20Hz, there is no necessity for that preprocessing, but it can be useful for viewing purposes, for plotting the raw signal, in case it is necessary.

Due to having the hypnogram information, we can analyze only the stages we are interested in, which are 'Wake', 'Quiet Sleep' and 'Active Sleep', discharging the others. If we reach a stage of the signal that is interesting to us for being one of those stages, we get the metric out of it.

We cut that stage out of the signal with the self-made function *select_stage*, to which a Morlet Wavelet (30 points between 2 and 20 Hz) will be applied, to get a 'complex' output as this is what is more interesting for our analysis, to measure the Phase Locking Value, our metric, that will be explained in the following paragraphs (section 3.3).

After having the interest stage separated from the rest of the signal, the stage is epoched, time – windowed, following the minimum stage length of the signal (as it has been shown before, the minimum sleep phase, acquired from the 'Results' function). At this moment, the data has 4 dimensions,

[Epochs, channels, frequencies, samples]

From this, the cPLV metric (Complex Phase Locking Value) is acquired, comparing all the epochs of all the channels of each frequency. To do that, two 'for cycles' are required, one for going through all the frequencies (upper cycle) and another to walk through the epochs (lower cycle).

The metric, cPLV, measures the difference of phase of two signals. For that, two epochs are needed. It should be remarked that the epochs have the same length (minimum sleep phase duration of the signal), since it is a metric that is affected by length. Precisely, the values are acquired for every pair of EEG electrodes (after re - referencing), at each epoch at a time, going through every frequency available.

Afterwards, the mean across epochs is obtained, to have one PLV value for every pair of electrodes, for every sleep phase. Again, the process is repeated, but with the following frequency.

Therefore, out of the function which get the cPLV, you have the following dimensions:

[frequency, channels, channels]

For these datasets, since we compute 30 logarithmically spaced frequencies from 2 to 20 Hz, and there are 8 EEG channels, we have the following shape of the output -> [30, 8 , 8].

For every frequency point, there is a 8x8 matrix. The first line will have 8 numbers, it will be the cPLV mean of the epochs of the first electrode against the other 8, in order. First with itself, then with the second electrode, and so on. The second line is the second electrode against the others. Which means, that the second value of the first raw, and the first of the second raw will be the same, cPLV of first electrode and second

one. Also, when it computes the metric with itself it gets a value of 1 + 0j, which is a value of 1 for PLV (absolute value of cPLV) and 0 for the imaginary part or iPLV (absolute value of the imaginary part of cPLV).

0.000000000000000 0.102679382173930 0.102037974673134 0.106006367065687 0.098410215960259 0.104300944043941 0.111220889539400 0.103193730755805
0.102679382173930 0.000000000000000 0.109277534387233 0.097955892935108 0.101838032389843 0.082990047775287 0.111664425312709 0.107721763105976
0.102037974673134 0.109277534387233 0.000000000000000 0.114491154536764 0.101694520962323 0.097161690711414 0.109414412458040 0.111004925480632
0.106006367065687 0.097955892935108 0.114491154536764 0.000000000000000 0.095812723790802 0.099570304376050 0.114495947854263 0.094898219154268
0.098410215960259 0.101838032389843 0.101694520962323 0.095812723790802 0.000000000000000 0.099820435973116 0.094918206787052 0.110680199663745
0.104300944043941 0.082990047775287 0.097161690711414 0.099570304376050 0.099820435973116 0.000000000000000 0.096331138161044 0.103966151783346
0.111220889539400 0.111664425312709 0.109414412458040 0.114495947854263 0.094918206787052 0.096331138161044 0.000000000000000 0.089841080800045
0.103193730755805 0.107721763105976 0.111004925480632 0.094898219154268 0.110680199663745 0.103966151783346 0.089841080800045 0.000000000000000

Table 3.1. Results for iPLV of frequency 2 Hz, patient number 2, first recording, Active Sleep.

In the previous chart you see what it has just explained, how in the diagonal there are all 0s, since that is the iPLV from one electrode with itself (there is no difference in phase). Also it is easy to see how the numbers repeat through the rows.

For that last reason, there are only 28 unique values on every 8x8 matrix, made out of the different combinations of all the re-referenced electrodes. A mask is needed to get those unique values to get the mean of all electrodes pairs. You can do other masks, to get only a piece of those values, to, for example, plotting only the metric in one area of the brain. In fact, this is what is done to plot the inter and within - hemispheric values, that will be shown later.

After analyzing every stage of the signal, a mean across conditions is computed to have only one 30x8x8 matrix for each of them (AS, QS and Wake).

The previous process changes a bit for analyzing the data for the 5 minutes recordings of the datasets without hypnogram information. Indeed, it is the same methodology but, without taking into account the sleep phases. A 5 minutes raw signal is introduced, and the cPLV is calculated. The 5 minutes signal is picked visually from the complete raw signal, taking 5 minutes without major artifacts.

3.2. SURROGATES

For the purpose of knowing the significance of each coupling, we should acquire information regarding distribution under null – 0. Hence, a surrogate analysis is made. The null distribution of one test is the probability distribution when the null hypothesis is true. The null hypothesis says that no statistical significance exists in a set of given observations. It is used to test hypotheses, in this case to test that the connectivity measured is higher than the intrinsic connectivity that noise represents.

The process explained previously is repeated N times, but with temporal rotation [20]. For each pair of signals, one is rotated from a random point, breaking the inter channel time coherence, which is a way of simulating the coherence that noise produces. The mean of the five repetitions is acquired, to get the surrogates values.

In order to do the temporal rotation, a function is used, in which every raw signal is inverted from every point.

Specifically, N equals 5 for the datasets with hypnogram information, and a 100 for the 5 minutes signals, since we used a most powerful computer.

3.3. THE METRIC: PLV

The main metric that is used in the project for analyzing the patients is PLV, which stands for phase locking value [21][22]. Phase analysis is a useful method for determining brain connectivity, two different brain areas have different oscillation properties unless there is connectivity between them, hence, the oscillation properties are related in some way.

Phase locking value is one of the most used methods, due to its simplicity. Specifically, it gets the instantaneous phase difference of two regions. If two areas are connected, following the hypothesis of this metric, they should evolve together, as they are 'locked' in.

The computation of PLV is simple, mathematically speaking:

$$PLV_{i,j}(t) = \frac{1}{N}\left|\sum_{n=1}^{N} e^{-i(\varphi_i(t,n)-\varphi_j(t,n))}\right| \qquad (1)$$

As it appears in the (1) formula, PLV can be directly obtained from the phase difference of two regions (i and j), in a moment n. N is the number of trials.

Translated into python code, the steps to follow are:

- Filter data to a given frequency band and extract analytic signal (FIR + Hilbert or Morlet wavelet transform).
- Create the epochs necessary for the time evolution analysis.
- Get the phase difference, and then the complex phase difference.
- Get the absolute value (for PLV) or absolute value of the imaginary part (for iPLV) of the previous step and divide it with the number of samples of the shape.

In the first annex you can find the code implementation.

The problem with this implementation is that it is not fast enough for processing big quantities of data. For that reason, and alternative implementation is presented:

- Get the complex Morlet wavelet results of the data.
- Get the epochs of the necessary time length.
- Normalize data by dividing it by its absolute value.
- Get the inner product between the signal and its complex conjugate.
- Afterwards, for the proper iPLV, it is needed to get the imaginary part of the absolute of the previous step.

With this method, we can process bigger quantities of data, operating with the whole matrix at the same time. The key to this optimization is the main mathematical characteristic of the metric, its simetricy.

With the imaginary part of the PLV (iPLV), we are able to discard the information regarding the volume conduction, as it automatically introduces artificial linear correlation between brain areas. Indeed, what you are accomplishing is discarding the 0 - lag, which also means you are discarding the instantaneous coupling.

The PLV metric gets a value of the functional connectivity across the brain. And that is what we measure when the mean across all channel pairs is done. However, that can be changed to analyze how the brain works in some areas separately. In that sense, we measure the inter and within - hemispheric values, by doing the averaging on some channel pairs only.

The distribution of the electrodes is as seen in the following figure.

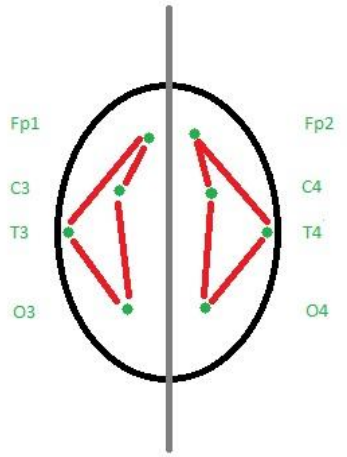

Figure 3.2. Distribution of electrodes (green) and its re - referencing (red), getting pre-Frontal (Fp) - Temporal (T), pre-Frontal - Central (C), Temporal - Occipital (O), and Central - Occipital relationships.

Therefore, if the mean across the 8 re - referenced channels is acquired, the PLV is calculated for the whole scalp. However, to measure the interhemispheric values, the mean should be done across four pairs:

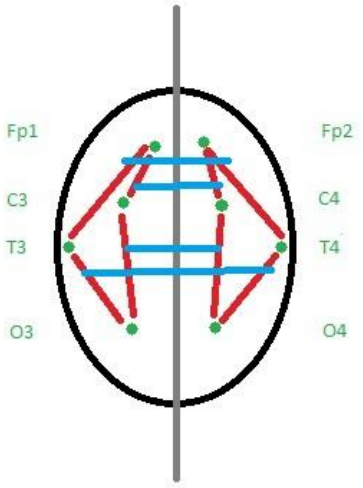

Figure 3.3. Blue lines represent the 4 pairs, each electrode pair with its own homologous of the other hemispheric, to get the interhemispheric value.

Hence, for this metric, the following pairs are computed:

- Fp1-C3 / Fp2-C4
- Fp1-T3 / Fp2-T4
- T3-O3 / T4-O4
- C3-O3 / C4-O4

On the other hand, the within - hemispheric PLV gets the connectivity across each hemisphere. In order to achieve that, the mean is done across the pairs of each side, which means, no inter hemispheric pairs are taken into account. Thus, we have:

- Fp1-C3 / Fp1-T3
- Fp1-C3 / T3-O3
- Fp1-C3 / C3-O3
- C3-O3 / T3-O3
- C3-O3 / Fp1-T3
- Fp1-T3 / T3-O3
- And the homologous for the other hemisphere.

It might be said that the process is repeated with the amount of data that does not have hypnogram information. For this reason, a extraction of 5 minutes (visually inspected to get 5 minutes of continuous data without major artifacts) is achieved. The process is similar, but the PLV metric acquisition is done regarding the sleep phase.

4. RESULTS

4.1. CONDITION ANALYSIS

The main study is taking part with the information of the sleep classification available and given by the hospital. This means that this part of the study is done with the hypnogram information. Hence, as it has been shown, we have the following materials:

- 3 patients with 2 recordings (born - week recording (33 or 35 weeks) and 40 weeks).
- Hypnogram classification.

After analyzing only the sleep phases of interest (Active Sleep, Quiet Sleep and Wake), numerous graphics are plotted to get the results. Following the previous bibliography (Yrjölä, 2021)[22], it was expected to see an evolution on the phase locking value, comparing both recordings.

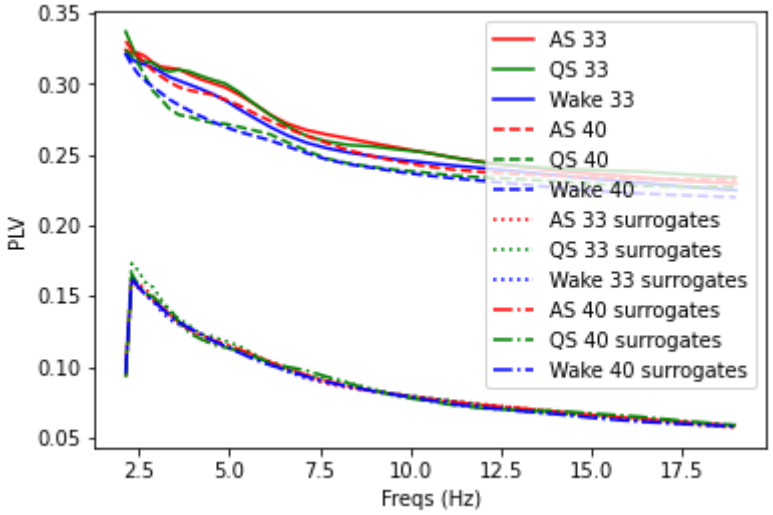

Figure 4.1. Comparison of PLV value across frequency, between sleep conditions. Surrogates values are included too.

The main difference between surrogates values and observation values prove that there is a significance in the coupling, with values above the basic noise that the surrogates represent. In the previous figure, we can see the PLV of the mean of all subjects of all channels for each condition.

It should be remarked that AS 33, is the mean of the Active Sleep for the first recording of the three patients. '40' is for the second recording. Also, QS stands for Quiet Sleep, and Wake, is the phase of Wake during the sleep recordings.

For viewing purposes, in the next figure we can see a zoom on the observation values.

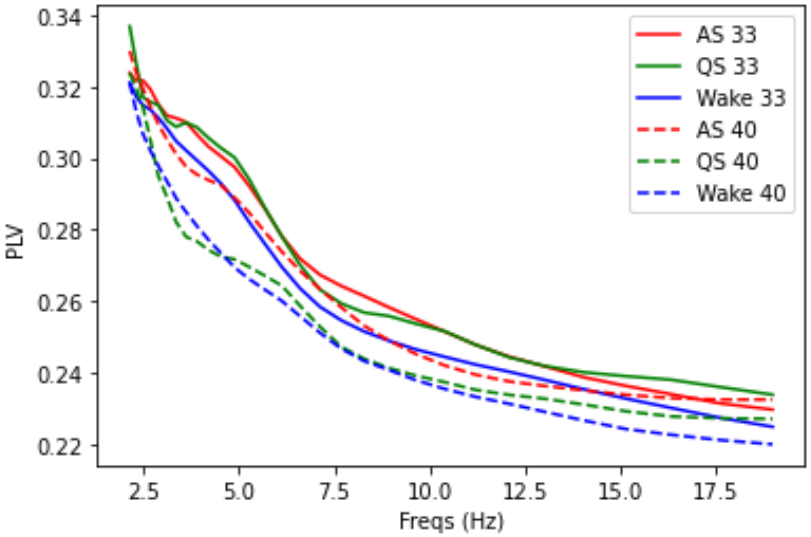

Figure 4.2. Zoom on the observation values for the same data as in Figure X.

In Figure 4.2., the expected hypothesis of having an evolution of PLV across time, it can be seen. There is a visual difference between the first recording mean, and the second one. The first ones, as continuous lines in the figure, have a higher value than the second ones (dashed lines). This happens in the three conditions. As Yrjölä described in the mentioned paper, the biggest difference exists in the QS values, in the delta band, specifically, a difference of almost 0.03 in PLV.

Nonetheless, these results are achieved with the mean of the available samples available, which are only three patients. This fact makes it impossible to generalize these hypotheses to every neonate child. Besides, it creates a problem regarding the confidence interval.

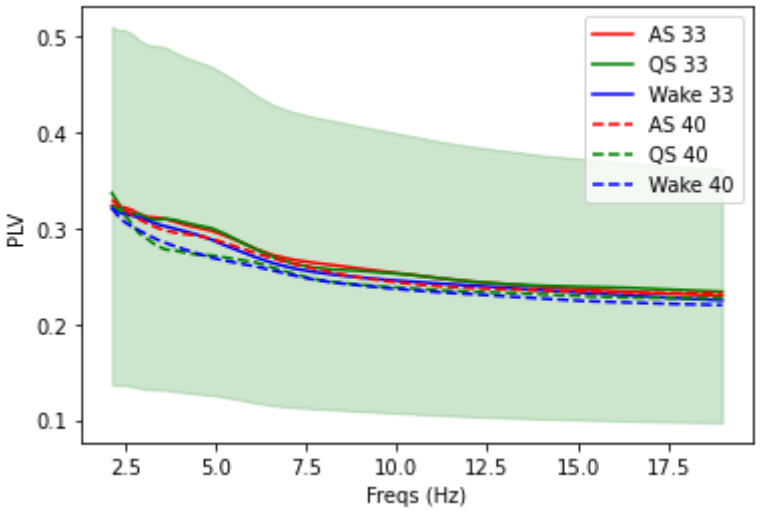

Figure 4.3. Confidence interval of QS for both recordings. The other CI are omitted for viewing purposes.

These big confidence intervals are derived from the idea of using that small number of samples. It should be remarked that the other conditions suffer from the same fact, all of them have big CI.

Apart from analyzing every pair of electrodes, on the other hand, we can analyze the results on inter and within-hemispheric values.

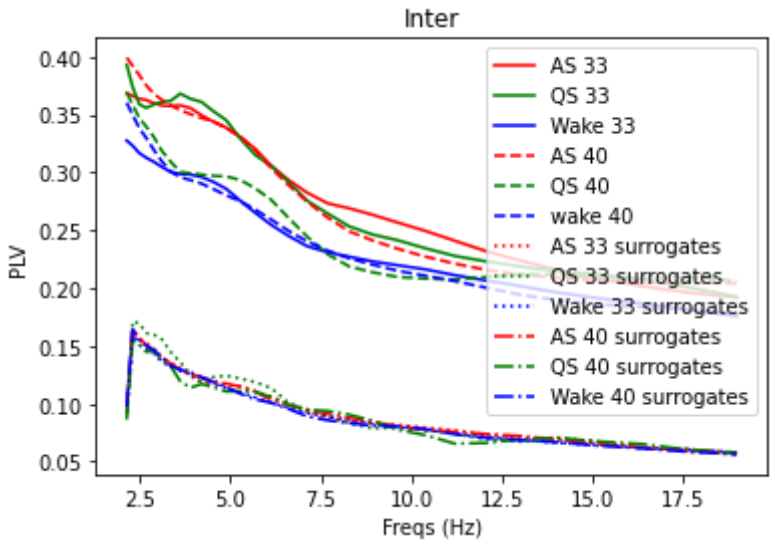

Figure 4.4. Interhemispheric plot with surrogate values.

Again, the surrogate values are below the observation values, which speaks about the robustness of the metric.

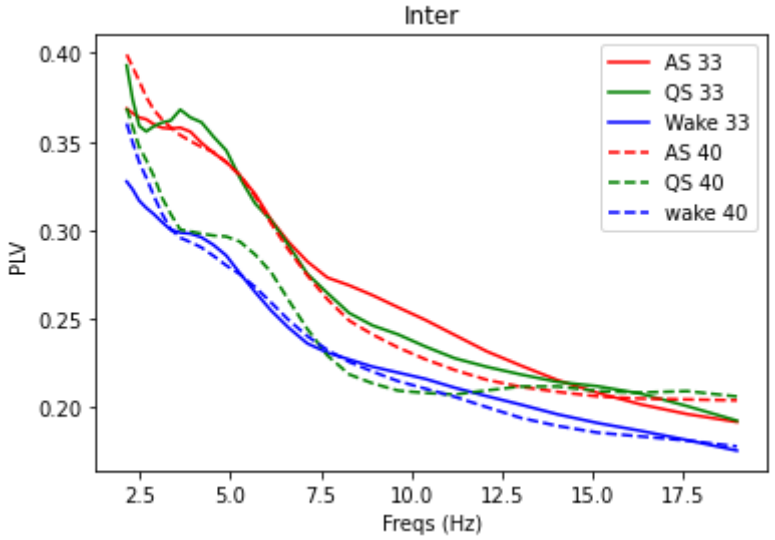

Figure 4.5. Zoom on the observation values of the last figure.

Applying a zoom to the observation values, it can be seen that there is no evolution on active sleep or Wake phase between the first and the second recording. Nevertheless, Quite Sleep is different, with its bigger decrease, again in the delta band, for more than 0.05 in PLV.

Again, in this case, the confidence intervals are big.

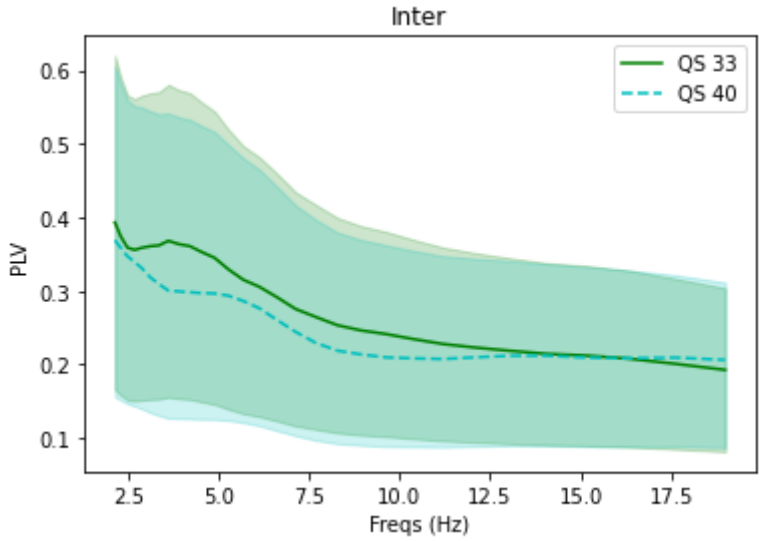

Figure 4.6. Confidence interval for Quiet Sleep conditions, in the inter - hemispheric representation.

From now on, the CI will be omitted for this section, since for all conditions and representations are as just shown.

On the other side, we can compare the results regarding the within - hemispheric analysis.

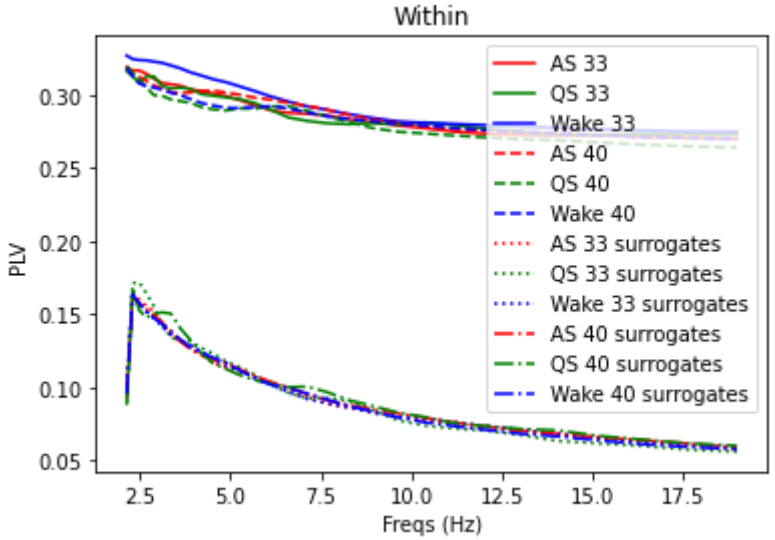

Figure 4.7. Within-hemispheric observation and surrogates values.

In this case, as before, surrogate values are lower than the observation ones.

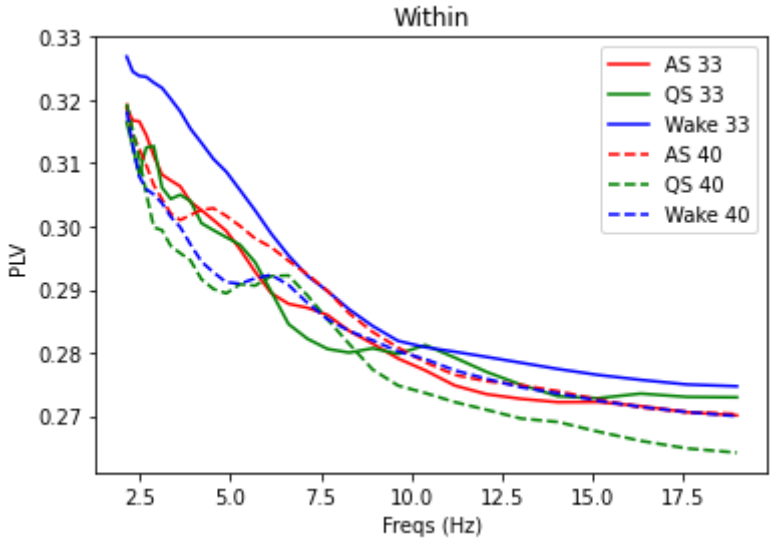

Figure 4.8. Within - hemispheric observation values.

Nonetheless, in this case, the main difference between the two recordings happens in the Wake phase., with a decrease that is bigger in the delta phase, for 0.02 in PLV.

Following the importance of Quit Sleep in the neural maturation of the brian and neural system, it is interesting to highlight these results. In the subsequent figure it is analyzed how QS PLV varies throughout the composition of compiling the whole system, or within or inter - hemispheric as explained.

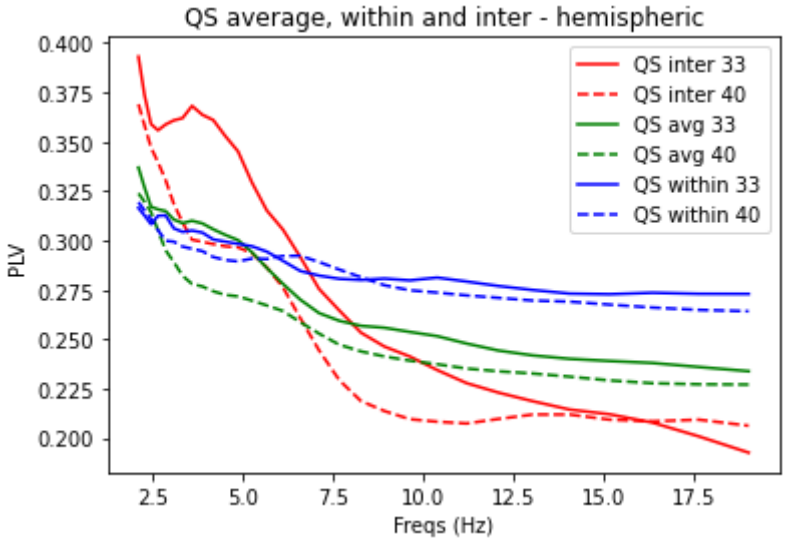

Figure 4.9. Quiet Sleep phase analysis of first (continuous line, 30 - 33 weeks) and second recording (dashed line, 40 weeks), computing PLV for all the pairs, inter - hemispheric and within - hemispheric.

As can be seen in the previous figures, the main difference between the first and second recording is in the average of the whole system, and in the inter - hemispheric distribution. On the other hand, for the within - hemispheric distribution, the

difference is less quantitative. In that sense, we can conclude that in the pre - term birth, the majority of the connectivity between regions, for these patients, seems to happen as inter - hemispheric links, rather than within the same hemispheric coupling.

Also, inter - hemispheric values decrease faster and longer, in contraposition to within values that are more constant, which produces a stabilization of the average PLV value for the whole system in both recordings.

There is not much knowledge on quiet or active sleep in the sense of what is happening in the brain. But, supposedly, as a non - activity sleep phase, there is a:

- Optimization and creation of neural paths for learnings.
- Improvement of brain memory.
- Physical growth and repair.
- Maturation of the brain.

This might be the reason for having that much connection between both hemispheres.

In opposition, you can create the same graphics with the imaginary part of the PLV, however, there are no distindible results (see results on Annex A).

Apart from this, it can be easy to calculate the nPLV or normalized PLV, which consists of:

$$nPLV = \frac{PLV}{mean(surrogates)} \quad (2)$$

For this experiment it is not possible to calculate the normalized iPLV since the SD of the surrogates is omitted.

The results, across all representations (normal, inter and hemispheric) give the same results, a value over 2,48 for all the bands except the 2.5 Hz one. This gives the idea that the signal, in the other frequencies, has a p value of less than 0.05, which makes it statistically significant.

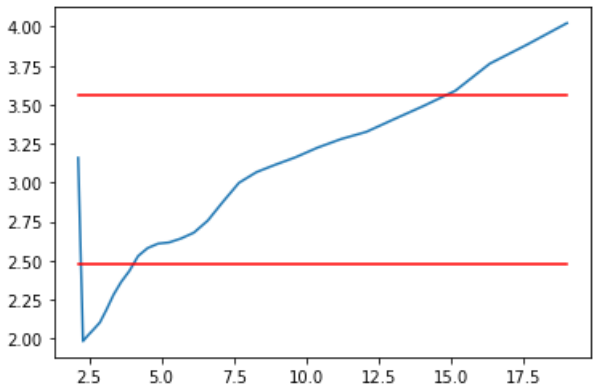

Figure 4.10. nPLV representation across frequencies for AS first recording. It is always above 2,56 (first threshold), which means a p value of less than 0.05. The second threshold is 3.56, which is a p value of 0.001. The other figures can be found in the annex A.

4.2. FIVE MINUTES SIGNALS RESULTS

With the rest of the data, which its hypnogram information is not given to DIBRIS, as explained, a 5 min extraction of clean signal is achieved. Although this does not allow to compare and discuss sleep differences, it creates the possibility of comparing more data.

It is clean data because the 5 minutes signal is extracted manually after visualization.

For that reason, this allows us to compare the two recordings of preterms, or the second recording of preterms with the first of term babies.

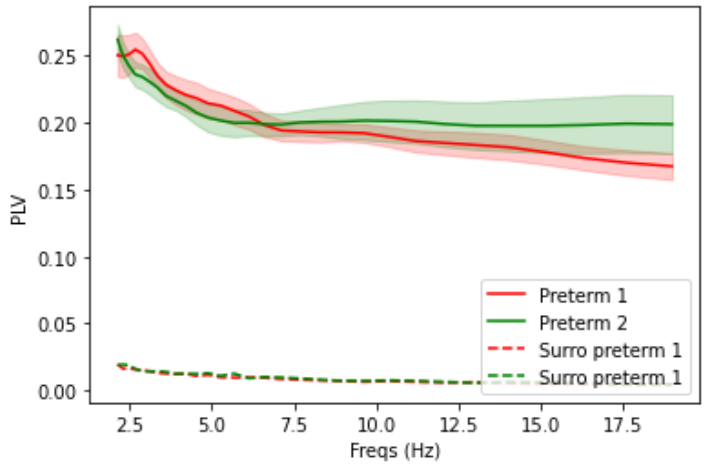

Figure 4.11. Comparison of first recording of preterm (less than 35 weeks), and second recording (40 weeks), with surrogates information and confidence intervals (as std/sqrt(N)).

As seen in the condition results (section 4.1), from the first to the second recording, a decrease in PLV is noted. Another proof of the evolution of this metric through neural development.

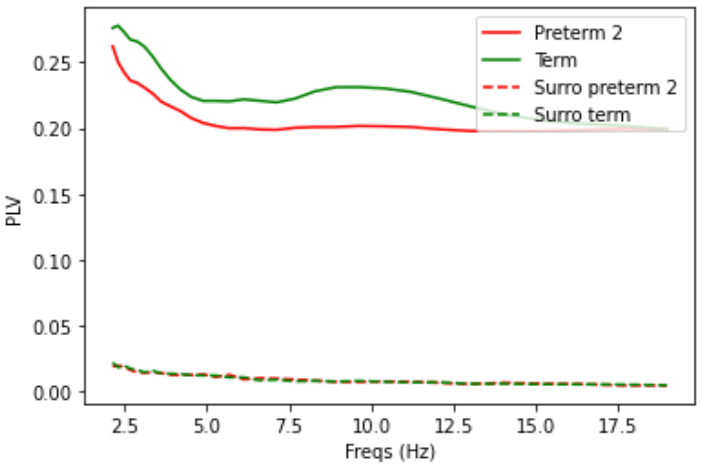

Figure 4.12. Plot of mean of second preterm recordings (40 weeks) and the recording of term babies at born age (40 weeks approximately).

These two recordings, of the previous figure, hypothetically, should be similar, as the preterm second recording is done at fictitious birth time (40 weeks). Indeed, although the amplitude is a bit different, it is only about 14% of the difference in its biggest part.

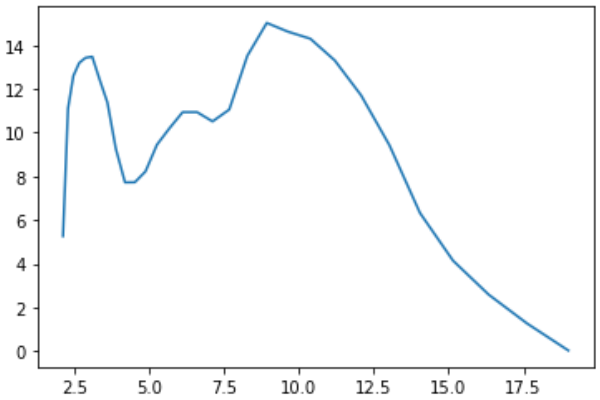

Figure 4.13. Percentage of difference in amplitude between both curves.

Ignoring the slight difference in amplitude, if we do an analysis on slope, the visual difference is represented, with its biggest distinction on the 10 - 12 Hz band.

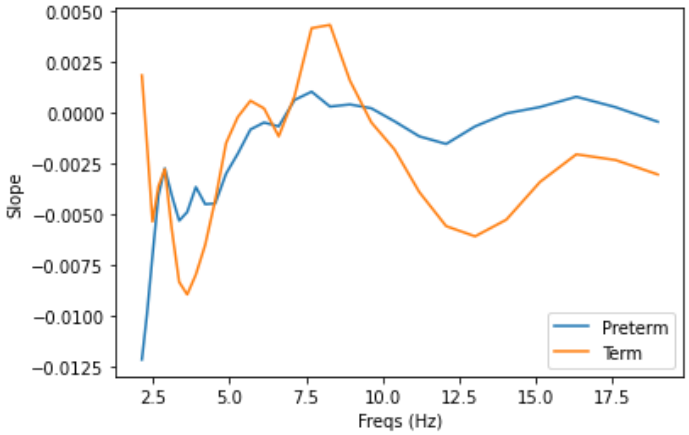

Figure 4.14. Slope of both curves, observation PLV for mean of preterm 2nd recording and 1st recording of term patients, across frequencies.

The main difference is in that band, with its explanation on the within - hemispheric connectivity.

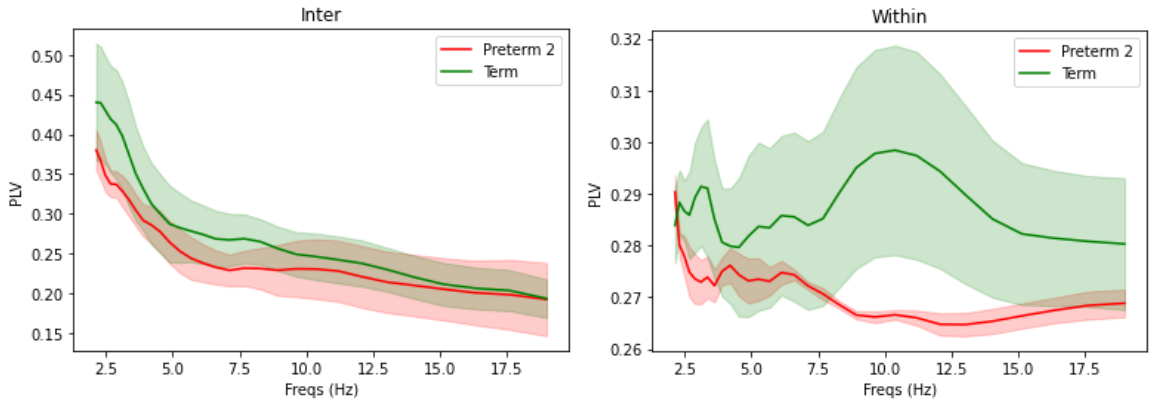

Figure 4.15. Inter and within - hemispheric PLV representation for means of preterm second recording patients, and term ones.

Following the last figure, the main difference can be seen in the within - hemispheric connectivity, since, in the inter - hemispheric one, it happens the same as in figure 4.14, almost the same amplitude and slope throughout the curves.

Hence, due to the small number of samples (4 patients for both populations), as before, it is not possible to assume a patron. Indeed, the confidence values of the within - hemispheric representation (calculated as std / sqrt(N), being N the sample population), indicate how different the 4 samples are. Although, all of them, apparently, have an increase of connectivity in that frequency band compared to the preterms patients represented here.

The iPLV plot can be found in Annex B.

Besides, to test if the results are statistically significant a nPLV is produced for both PLV and iPLV:

$$niPLV = Trf\ Z = \frac{iPLV - mean(surrogates}{std(surrogates)} \quad (3)$$

In this case, for iPLV, the threshold is in 1,58 of the Z transform for a p value of 0.05. The signals do not overcome this value, except in the first frequencies, as seen in the annex, showing the poor statistical significance. On the other hand, PLV is significant since it always overtakes the 0.05 p value line.

5. CONCLUSION AND FINAL DISCUSSION

The code developed and explained in these lines tries to extract a metric that characterizes EEG signals. This metric is the Phase Locking Value or PLV, specifically the complex PLV or cPLV. After which analysis based on its absolut part (PLV) and imaginary part are done.

As previously demonstrated by fellow researchers [22][23], the obtained results show a decrease in the PLV metric comparing first recording of preterm - neonates, at birth time, and a second recording at 40 weeks. Hence, this could be an indicator of the neural development of the pre - term brain.

Besides, the exposed plots and analysis, tries to find a relationship between the second recording of preterm patients, and a normal birth time (40 weeks) of term patients. Discovering the evolution of the metric throughout the development of the preterm neonate could be an aid for clinical personnel to predict how the patient is evolving, neuronally.

Indeed, the future research lines of this project is to use this information, with more samples, and correlate it with clinical outcomes, general ones such as good or bad evolution, with that goal.

Regarding the amount of data used for the thesis, for external reasons of the University of Genova, only 3 preterm patients with both recordings and sleep phase classification were received, which has limited the range of action and research for this work. This means that, with such a small number of samples, every affirmation said here, is based only on our samples, making it impossible to generalize these results on all preterm medical basis.

However, the outcome of this work, and its confirmation of previous work, opens some promising questions and propositions of this line of research, in which there are still many elements to discover, with the only mission of trying to help patients.

6. REFERENCES

[1] Who.int. 2022. *Preterm birth*. [online] Available at: https://www.who.int/en/news-room/fact-sheets/detail/preterm-birth#:~:text=Se%20considera%20prematuro%20un%20beb%C3%A9,prematuros%20(28%20a%2032%20semanas).

[2] En.wikipedia.org. 2022. *Midbrain - Wikipedia*. [online] Available at: <https://en.wikipedia.org/wiki/Midbrain>.

[3] Javier Rodriguez Falces. 2020. Apuntes de Bioseñales (Grado en Ingeniería Biomédica – UPNA).

[4] Health.qld.gov.au. 2022. *Brain Map | Queensland Health*. [online] Available at: <https://www.health.qld.gov.au/abios/asp/brain>.

[5] Qbi.uq.edu.au. 2022. *Lobes of the brain*. [online] Available at: <https://qbi.uq.edu.au/brain/brain-anatomy/lobes-brain>.

[6] 2022. *Teach Me*. [online] Available at: <https://teachmeanatomy.info/neuroanatomy/structures/meninges/>.

[7] Google.com. 2022. *Google Image Result for https://cdn1.byjus.com/wp-content/uploads/2020/02/STRUCTURE-OF-NEURON.png*. [online] Available at: <https://www.google.com/imgres?imgurl=https://cdn1.byjus.com/wp-content/uploads/2020/02/STRUCTURE-OF-NEURON.png&imgrefurl=https://byjus.com/biology/neurons/&tbnid=Sfc15T2QXeEdYM&vet=1&docid=Uh4nogxP0JaoZM&w=750&h=425&source=sh/x/im>.

[8] Chen W., 2008. The Performance of Associative Memory Models

with Biologically Inspired Connectivity. PhD thesis at University of Hertfordshire.

[9] Brainclinics. 2022. *History of the EEG and QEEG - Brainclinics history of eeg qeeg*. [online] Available at: <https://brainclinics.com/history-of-the-eeg-and-qeeg/>.

[10] Schirmann, F., 2014. ¨'The wondrous eyes of a new technology' - a history of the early electroencephalography (EEG) of psychopathy, delinquency, and immorality¨. *Frontiers in Human Neuroscience*, 8.

[11] Abhang, P., Gawali, B. and Mehrotra, S., n.d. *Introduction to EEG- and speech-based emotion recognition*.

[12] Mayoclinic.org. 2022. *Polysomnography (sleep study) - Mayo Clinic*. [online] Available at: <https://www.mayoclinic.org/tests-procedures/polysomnography/about/pac-20394877#:~:text=Polysomnography%2C%20also%20called%20a%20sleep,leg%20movements%20during%20the%20study>.

[13] Sleepfoundation.org. 2022. *What Happens During NREM Sleep? | Sleep Foundation*. [online] Available at: <https://www.sleepfoundation.org/stages-of-sleep/nrem-sleep#:~:text=its%20different%20components.-,What%20Is%20NREM%20Sleep%3F,%2C%20heartbeat%2C%20and%20brain%20waves> .

[14] Pediatriaintegral.es. 2022. *Seguimiento del recién nacido prematuro y del niño de alto riesgo biológico*. [online] Available at: <https://www.pediatriaintegral.es/publicacion-2014-07/seguimiento-del-recien-nacido-prematuro-y-del-nino-de-alto-riesgo-biologico/>.

[15] Friedrich, M., Wilhelm, I., Mölle, M., Born, J. and Friederici, A., 2017. The Sleeping Infant Brain Anticipates Development. *Current Biology*, 27(15), pp.2374-2380.e3.

[16] Grafton, S. and Volz, L., 2019. From ideas to action: The prefrontal–premotor connections that shape motor behavior. The Frontal Lobes, pp.237-255.

[17] Stancin, I., Cifrek, M. and Jovic, A., 2021. A Review of EEG Signal Features and Their Application in Driver Drowsiness Detection Systems. *Sensors*, 21(11), p.3786.

[18] Arnulfo, G., Hirvonen, J., Nobili, L., Palva, S. and Palva, J., 2022. *Phase and amplitude correlations in resting-state activity in human stereotactical EEG recordings*.

[19] NeurotechEDU. 2022. *Preprocessing*. [online] Available at: <http://learn.neurotechedu.com/preprocessing/#references>.

[20] Roascio, M., Canessa, A., Trò, R., Mattioli, P., Famà, F., Giorgetti, L., Girtler, N., Orso, B., Morbelli, S., Nobili, F., Arnaldi, D. and Arnulfo, G., 2021. Phase and amplitude electroencephalography correlations change with disease progression in people with idiopathic rapid eye-movement sleep behavior disorder. *Sleep*, 45(1).

[21] Bruña, R., Maestú, F. and Pereda, E., 2018. Phase locking value revisited: teaching new tricks to an old dog. *Journal of Neural Engineering*, 15(5), p.056011.

[22] Yrjölä, P., Stjerna, S., Palva, J., Vanhatalo, S. and Tokariev, A., 2021. Phase-based cortical synchrony is affected by prematurity.

[23] Uchitel, J., Vanhatalo, S. and Austin, T., 2021. Early development of sleep and brain functional connectivity in term-born and preterm infants. *Pediatric Research*, 91(4), pp.771-786.

7. ANNEX

7.1. ANNEX A: iPLV AND nPLV FIGURES FOR CONDITION ANALYSIS

In this annex the figures for the iPLV will be shown. It has to be remarked that the results do not offer any visual information in contrast to the graphics of the PLV.

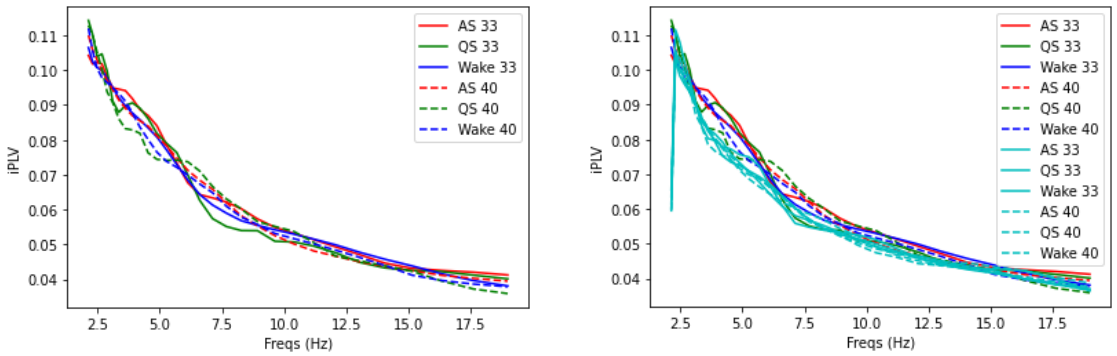

Figure 7.1. iPLV for all conditions, and both recordings. Graph on the right show surrogate values too.

As seen in the previous figure, there is no visual difference or evolution between the first and the second recording of any condition. Also, because of creating only 5 surrogates, and usually more are needed (for following experiments more computational capacity is used, getting up to a 100 surrogates), the surrogate values (in cian) are high.

This fact does not change along inter or within hemispheric analysis.

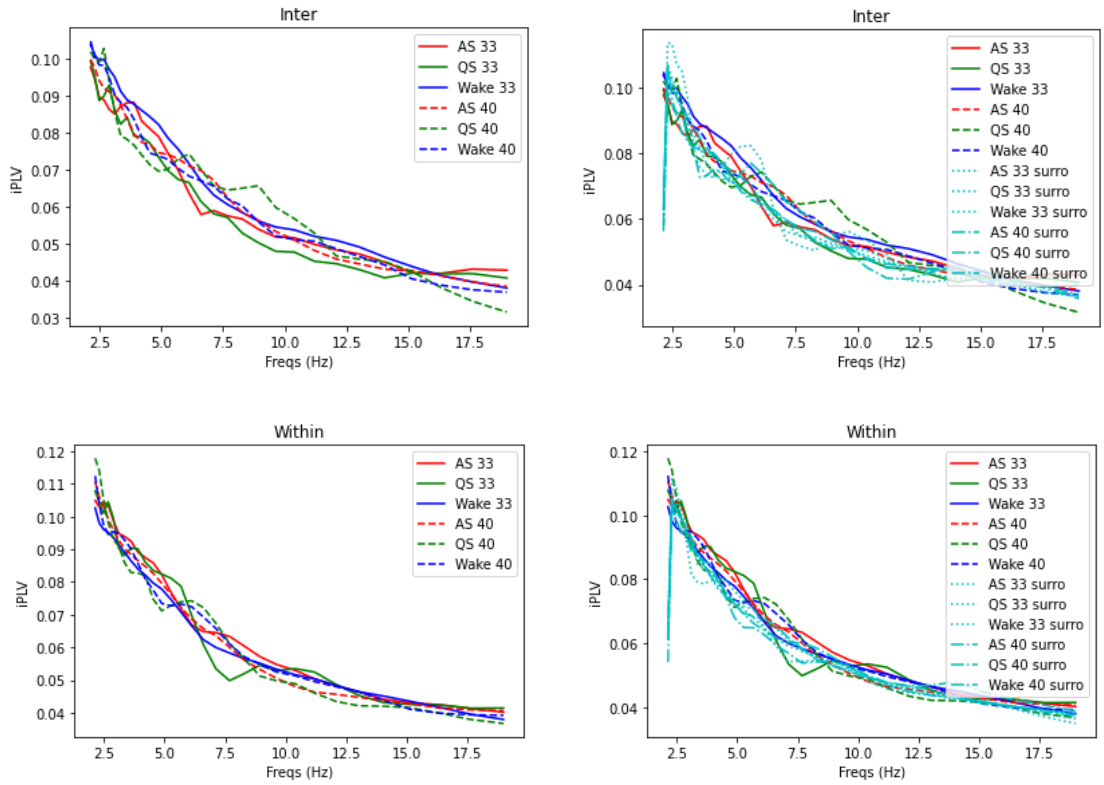

Figure 7.2. iPLV for inter - hemispheric and within - hemispheric analysis.

It has to be said, as before, that the confidence intervals for this analysis are also omitted for viewing purposes, for being big.

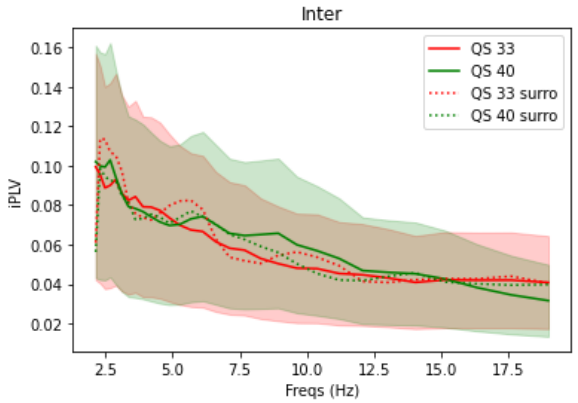

Figure 7.3. CI for inter - hemispheric analysis of both recordings for condition ‘Quiet Sleep’.

For the nPLV, we find the following results:

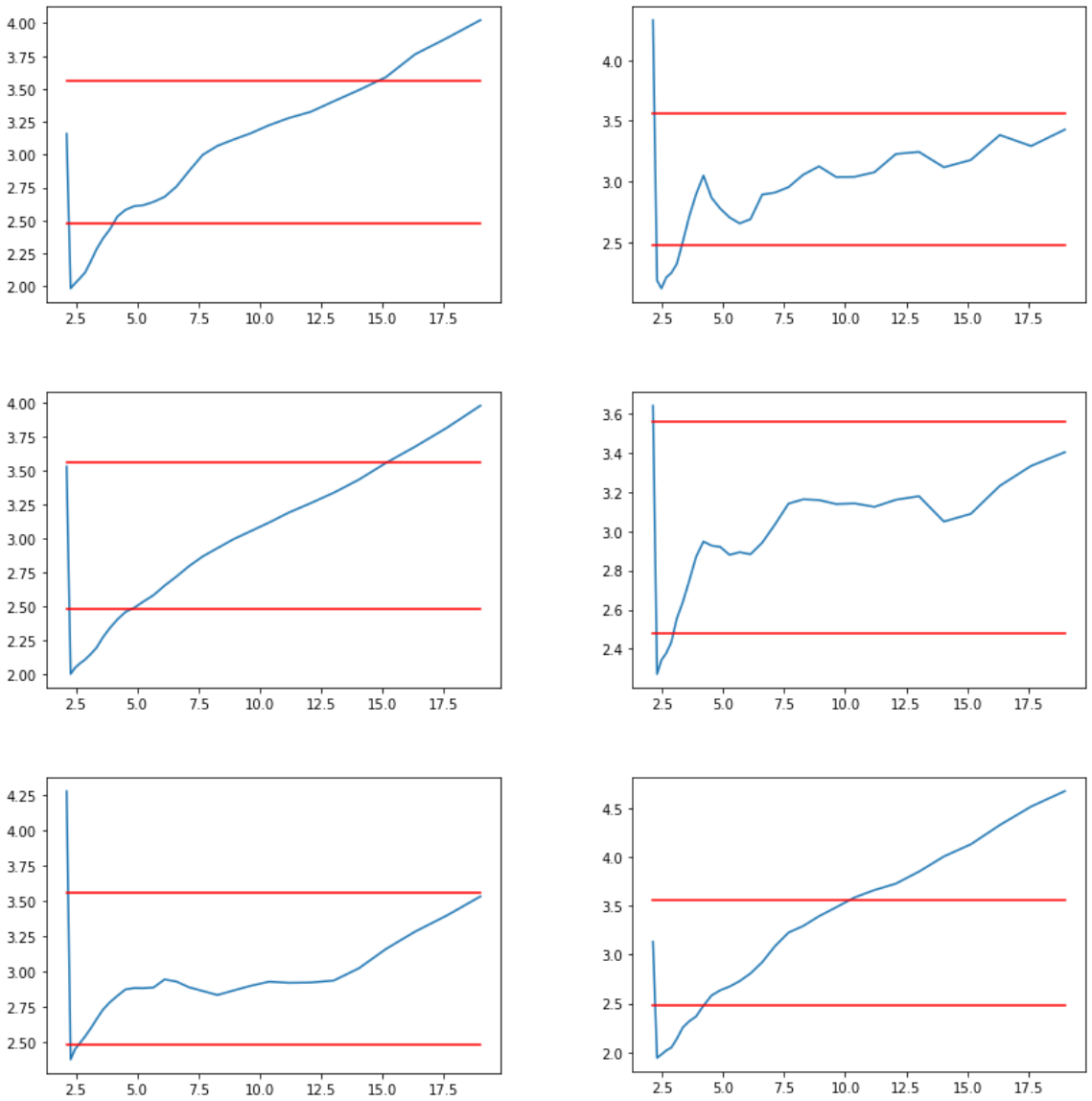

Figure 7.4. nPLV for AS. From left to right and upper side to bottom: AS first recording (30), AS second recording (40), interhemispheric AS 30 and AS 40, and withinheispheric AS 30 and 40.

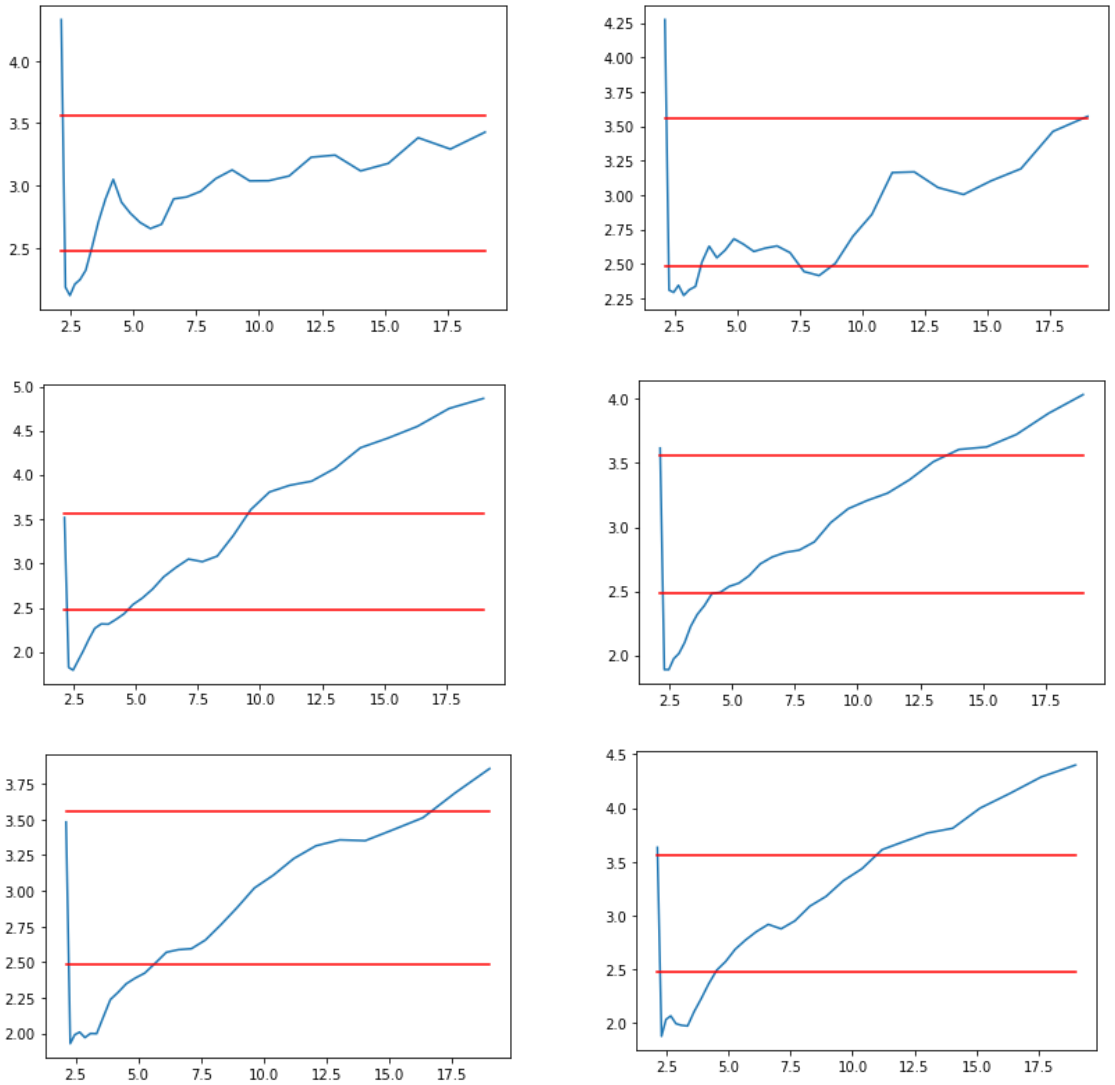

Figure 7.5. nPLV for QS. From left to right and upper side to bottom: AS first recording (30), AS second recording (40), interhemispheric AS 30 and AS 40, and withinheispheric AS 30 and 40.

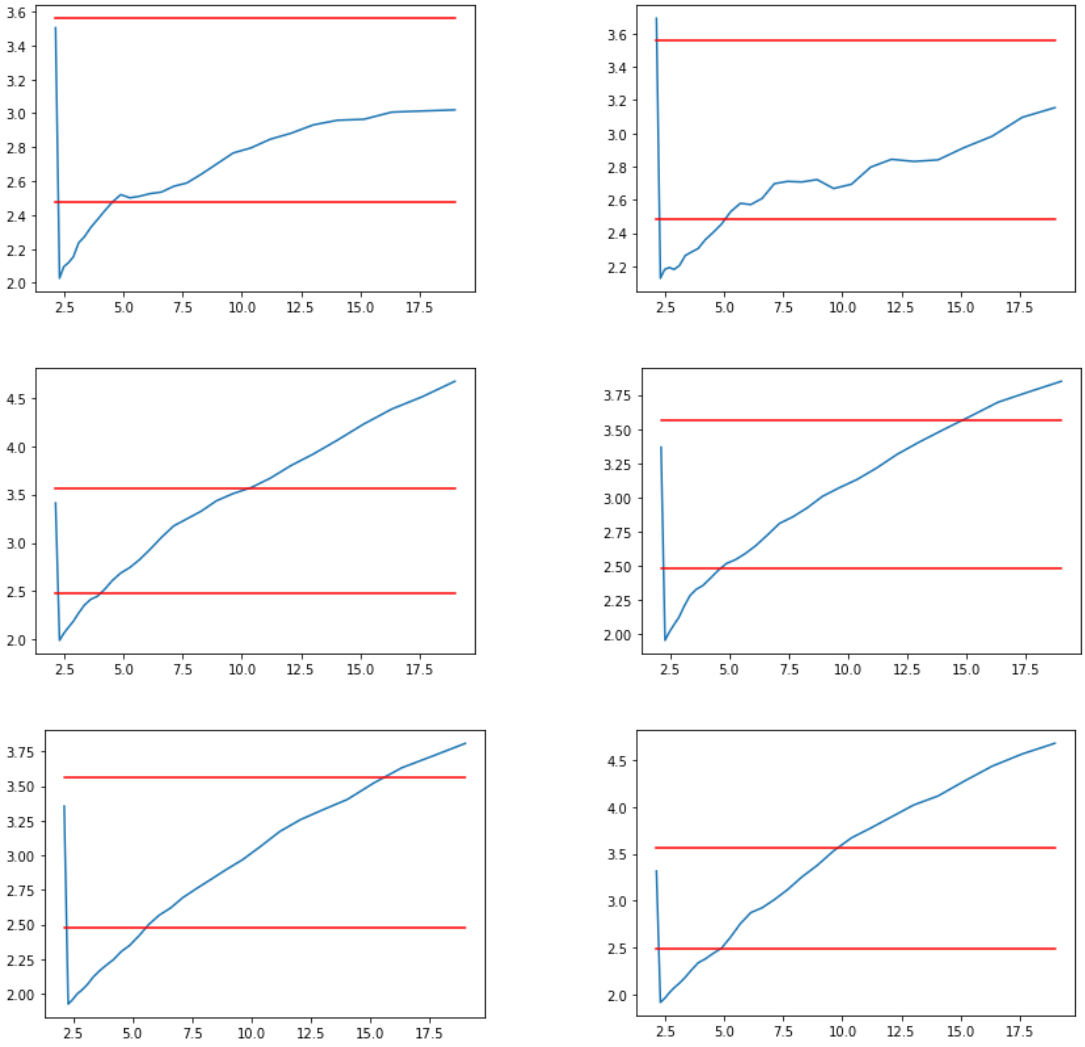

Figure 7.6. nPLV for Wake. From left to right and upper side to bottom: AS first recording (30), AS second recording (40), interhemispheric AS 30 and AS 40, and withinheispheric AS 30 and 40.

As previously commented, all the lines are above 2.48 (p value = 0,05), excepto the band around 2.5 Hz.

7.2. ANNEX B: iPLV, nPLV AND niPLV FIGURES FOR ANALYSIS OF 5' SIGNALS

Next, the figures for iPLV of the 5 minutes signals analysis will be shown. First for the first recording against the second recording of the preterm experiment.

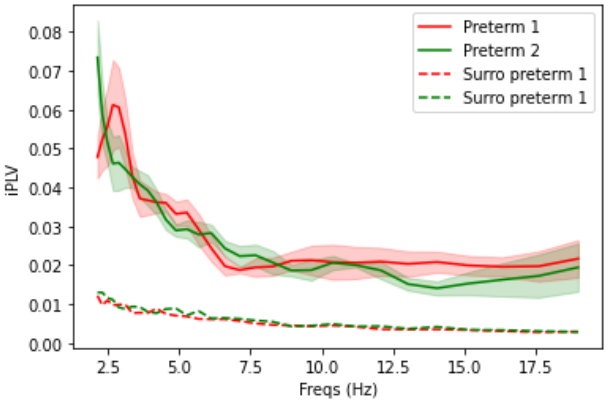

Figure 7.7. iPLV for Preterm first recording against second recording.

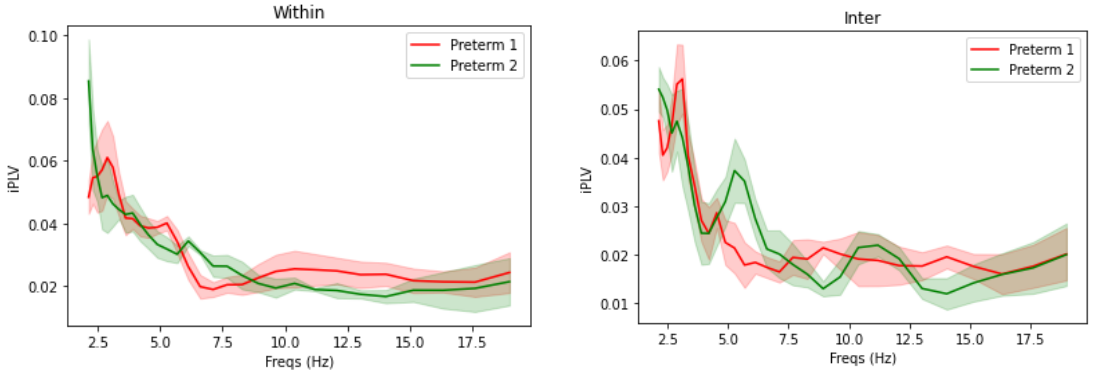

Figure 7.8. iPLV for Inter and within - hemispheric analysis.

Secondly, the figures regarding the second recording of preterm patients against the term ones.

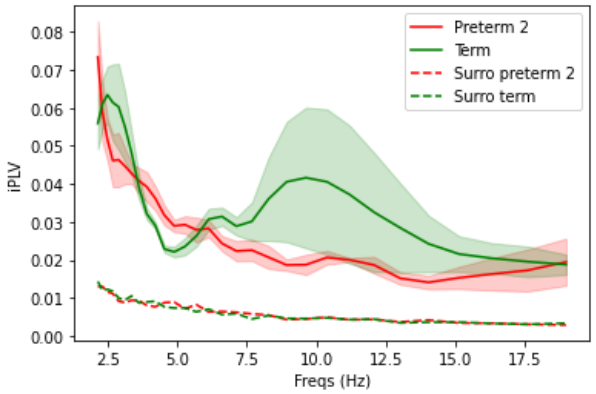

Figure 7.9. iPLV for second recording of preterms vs Term first recording.

There is a major difference around the 10 Hz band that can be explained in the hemispheric analysis.

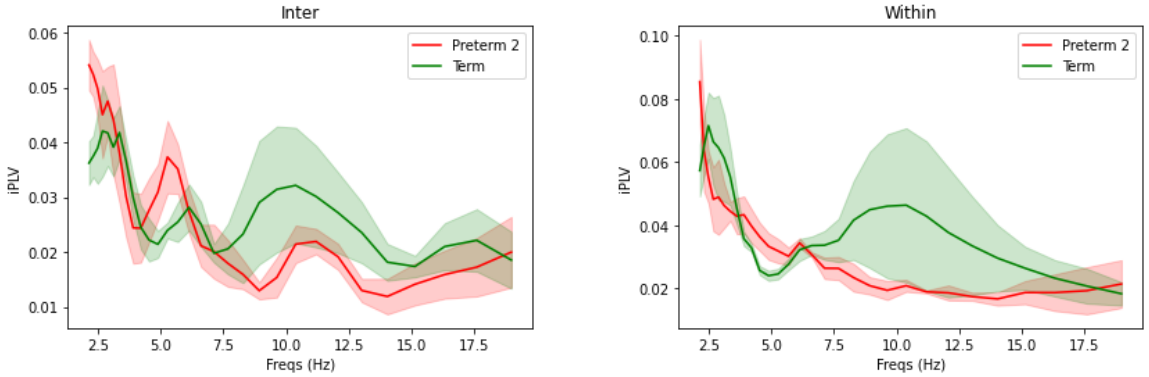

Figure 7.10. iPLV inter - hemispheric and within - hemispheric analysis.

This increase in the value for that band is also found in PLV.

All the nPLV show values over 3.56 (p value = 0.001), which means they have an important statistical significance. For example:

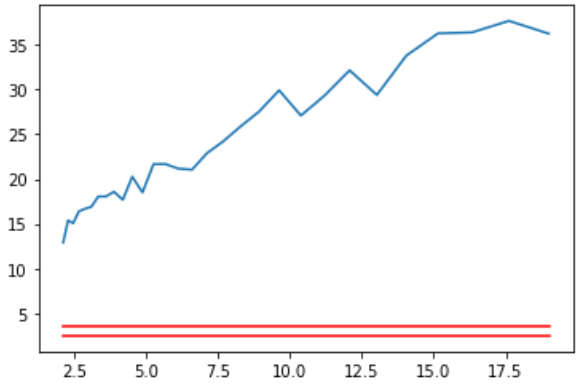

Figure 7.11. nPLV for the second recording of preterm patients across frequencies.

The niPLV figures, on the other hand:

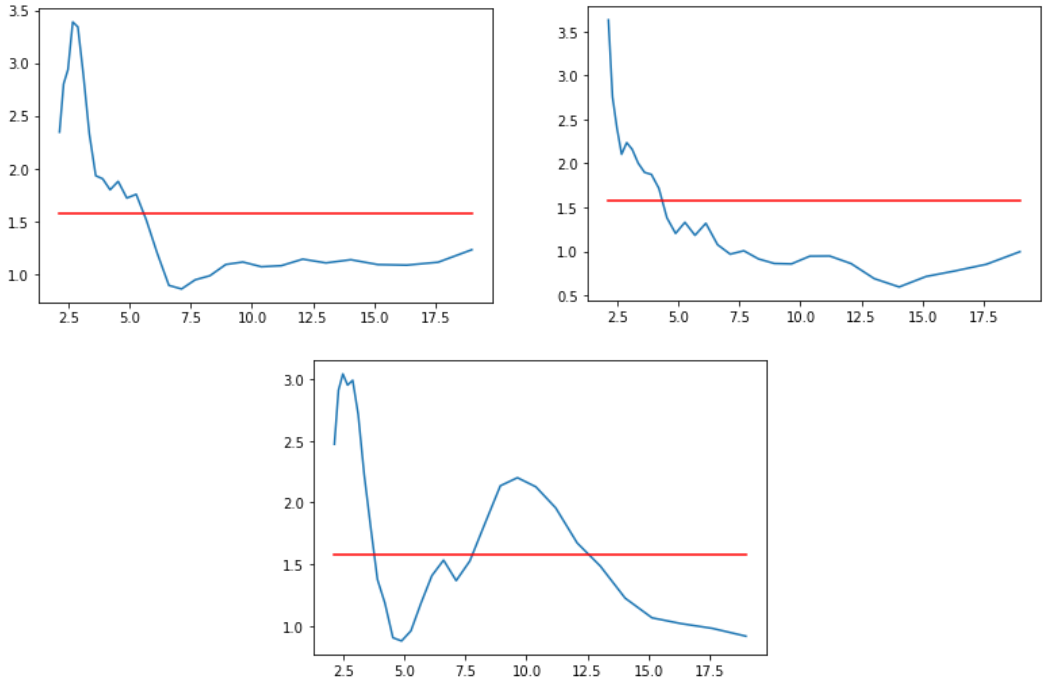

Figure 7.12. From left to right and up to bottom, preterm first recording, second recording and term.

Generally, statistical significance mainly on the first frequencies. Also around 10 Hz for the term patients.

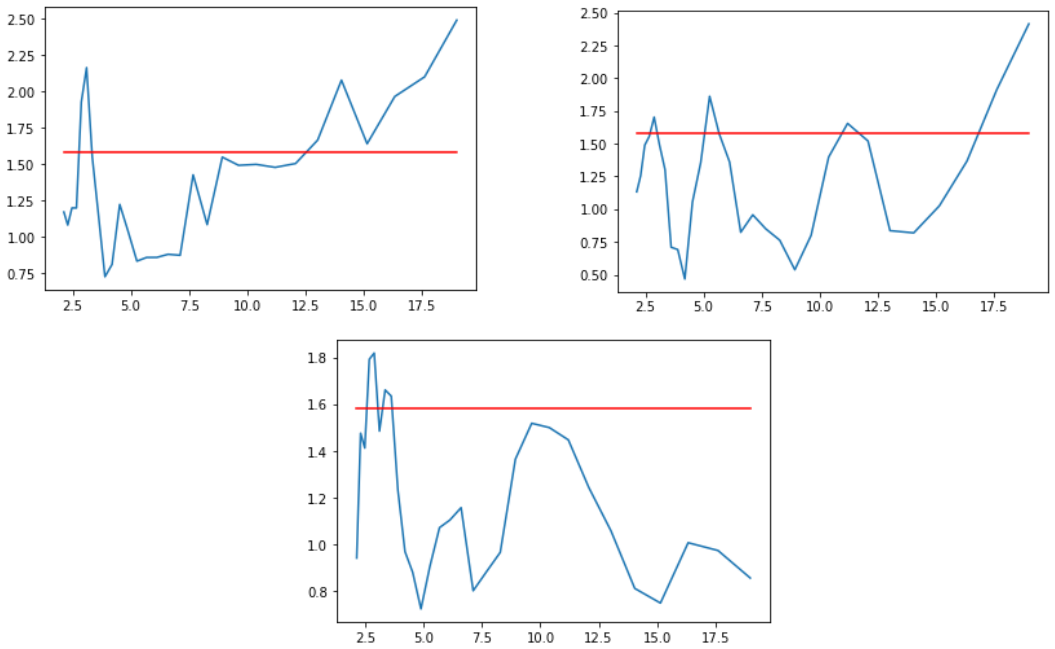

Figure 7.13. Same order as previous figure, but for interhemispheric representation.

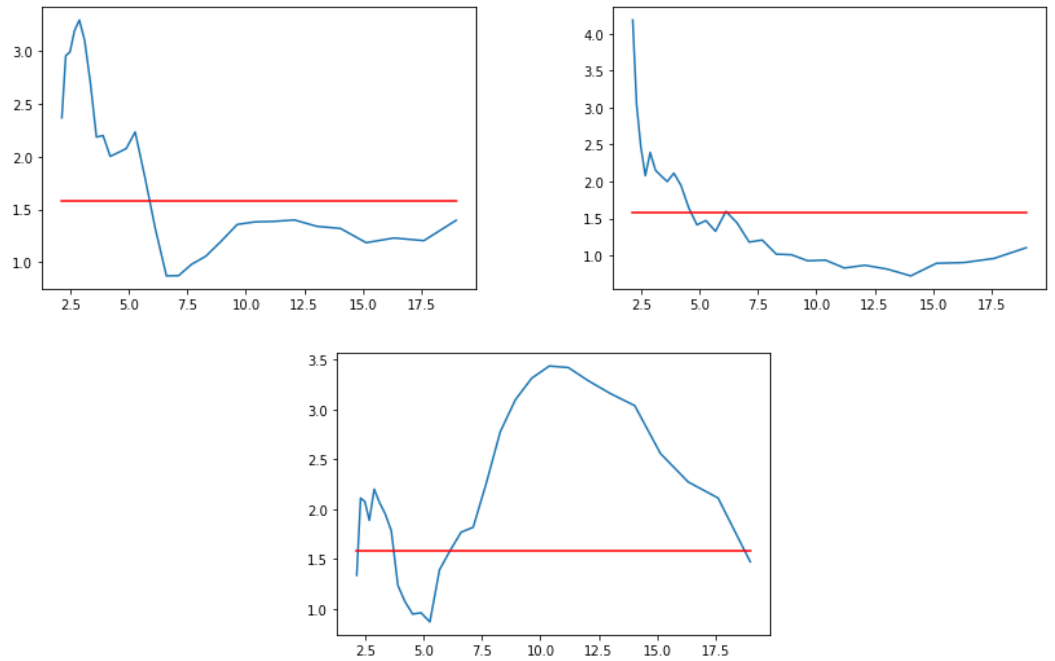

Figure 7.14. Same order but for within - hemispheric representation.

7.3. ANNEX C: CODE

In this section, some important functions and lines of the developed code are shown.

Libraries definition

```
import os

import cv2 as cv

import scipy

import numpy as np

from numpy.fft import fft, fftfreq

from scipy import signal as sig

import matplotlib.pyplot as plt

from mne.time_frequency.tfr import morlet

from mne.viz import plot_filter, plot_ideal_filter

import mne

from mne.preprocessing import compute_proj_ecg

from mne_connectivity import envelope_correlation

from mne.minimum_norm import make_inverse_operator,
apply_inverse_epochs

from mne.preprocessing import compute_proj_ecg, compute_proj_eog

from scipy import signal

from scipy.integrate import simps

import mne_features

import collection

from mne.time_frequency import (tfr_multitaper, tfr_stockwell,
tfr_morlet, tfr_array_morlet)

import math

from typing import Optional
```

Mask to collect unique values of each 30x8x8 matrix

Function which uses a mask to select the unique values in the matrix that correspond to the unique connectivity measure of each pair of electrodes in the 30x8x8 matrix, for each frequency.

```
def unique_values(matrix, freqs):
      posiciones = [(0,1), (0,2), (0,3), (0,4), (0,5),(0,6), (0,7),
      (1,2),    (1,3),    (1,4),    (1,5),    (1,6),    (1,7),    (2,3),
      (2,4),(2,5),(2,6),(2,7),(3,4),(3,5),(3,6),(3,7),(4,5),(4,6),
      (4,7),(5,6),(5,7),(6,7)]
      surro_unique = np.zeros((freqs.shape[0],28))
      for j in range (freqs.shape[0]):
            data = matrix[j,:,:]
            for i in range (28):
                  surro_unique[j, i] = data[posiciones[i]]
      surro_unique2 = np.zeros(freqs.shape[0])
      for i in range (freqs.shape[0]):
            surro_unique2[i] = np.mean(surro_unique[i,:])
      return surro_unique2
```

PLV function 1

First PLV function designed directly from the PLV formula explained in the methodology section.

```
def PLV(y1,y2):
      sig1_hill=sig.hilbert(y1)
      sig2_hill=sig.hilbert(y2)
      phase_y1=np.angle(sig1_hill)
      phase_y2=np.angle(sig2_hill)
      Inst_phase_diff=phase_y1-phase_y2
      complex_phase_diff = np.exp(1j*(Inst_phase_diff))
      plv = np.abs(np.sum(complex_phase_diff))/phase_y1.shape[2]
      avg_phase=np.average(Inst_phase_diff)
      return plv,Inst_phase_diff,avg_phase
```

cPLV function

Second PLV function, and the one that is used for being faster. Given one or two arrays, copilates the cPLV for them, using the second method explained in section 3.3, using the symmetry characteristic of the metric. First normalize, afterwards, getting the inner product of both, divided by the shape, the amount of samples.

```
def cplv2(x: np.ndarray, y: Optional[np.ndarray]=None, is_normed: bool=False) -> float:
    n_ts = x.shape[1]
    if is_normed:
        x_norm = x
        y_norm = x_norm if y is None else y
    else:
        x_norm = x / np.abs(x)
        y_norm = x_norm if y is None else y / np.abs(y)
    avg_diff = np.inner(x_norm, np.conj(y_norm)) / n_ts
    return avg_diff
```

Create surrogates by time rotation

Matrix with time rotation used for surrogate values. It gets a random sample number, from which rotate (roll) the data.

```
def create_surrogate_by_time_rotation(data):
    n_channels, n_freq, n_samples = data.shape
    offsets = np.random.randint(0, n_samples, n_channels)
    surro = data.copy()
    for idx in range(n_channels):
        surro[idx] = np.roll(data[idx], offsets[idx])
    return surro
```

Results function

Function to extract the hypnography information, such as how many sleep phases transitions there are. It checks if following sleep phases numbers of the matlab file are equal, which doesn't add a new sleep phase. However, if the number changes, there is a transition of sleep phase.

```
def results_fun(ipnog):
	count = 0
	results = []
	for i in range (1,len(ipnog)):
		if ipnog[i] == ipnog[i-1]:
			count = count + 1
		else:
			results.append([ipnog[i-1],(count + 1)*20])
			count = 0
	num_trans = np.array(results).shape[0]
	results.append([ipnog[len(ipnog)-1], (count+1)*20]) #add last
	stage
	return results
```

Sleep phase extraction

Function developed to extract the raw signal of just one sleep phase to compute the cPLV. In order to extract it, the transition of phase information acquired from the 'Results' function is needed.

```
def select_stage(sgn, stage, phase_time_matrix): #stage is the number of phase, sorted
    if type(sgn) != np.ndarray:
        sgn = sgn.get_data()
    if stage == 0:
        print()
        raw_stage = sgn[:, 0:(phase_time_matrix[stage][1]*512)]
    else:
        sum = 0
        for i in range (stage):
            sum = sum + phase_time_matrix[i][1]*512
            raw_stage = sgn[:,sum:sum+phase_time_matrix[stage][1]*512]
    return raw_stage
```

‘Get PLV’ function

After applying the morlet wavelet to the sleep phase of interest (previously it has been staged by the ‘select_stage’ function), a normalization is made, and the rotation is applied for the surrogates. Then, a simple split is applied to get the epochs of the minimum sleep phase length. To finish, the function to go through every epoch to get the PLV is called for both the observation values, and the surrogate one (the process is repeated 5 times).

```
def metricOBS_matrix(raw_staged, results, freqs, minimo): #stage: num de stage y channels (EMPIEZAN EN O)
      '''Get the metrics out of the 'evolution_PLVmorlet_matrix function.
            -> Inputs:      raw: raw signal
                            stage: number of sleep stage that we are analyzing. First is 0.
                            results: matrix with sleep stage data
                            freqs: frequencies
           -> Outputs:      Matrix of observed iPLV
           '''
      power = tfr_array_morlet(raw_staged[np.newaxis,...], 512, freqs=freqs, n_cycles=5, output='complex', n_jobs = 25).squeeze()
      power_staged = power[:,:,0:((power.shape[2]//(minimo*512))*minimo*512)]
      power_staged /= np.abs(power_staged)
      #create surro
      surro_data = create_surrogate_by_time_rotation(power_staged)
      surro_data /= np.abs(surro_data)
      power_epoched = np.array_split(power_staged, (raw_staged.shape[1]//(minimo*512)), axis = 2)
      power_epoched = np.array(power_epoched)
      surro_epoched = np.array_split(surro_data, (raw_staged.shape[1]//(minimo*512)), axis = 2)
      surro_epoched = np.array(surro_epoched)
      #OBS VALUE
      matrizPLVobs, matriziPLVobs = evolution_PLVmorlet_matrix(power_epoched, freqs.shape[0], (raw_staged.shape[1]//(minimo*512)))
      #5 TIMES FOR 5 SURROGATES
       matrixsurro = 0
```

```
        matrixsurroPLV = 0

        for i in range (5):

            matrixsurro_interPLV,        matrixsurro_inter        =
            evolution_PLVmorlet_matrix_surro(power_epoched,
            surro_epoched,                          freqs.shape[0],
            (raw_staged.shape[1]//(minimo*512)))

            matrixsurro = matrixsurro + matrixsurro_inter

            matrixsurroPLV = matrixsurroPLV + matrixsurro_interPLV

            matrixsurro = matrixsurro / 5

            matrixsurroPLV = matrixsurroPLV / 5

      return    matrizPLVobs,    matriziPLVobs,    matrixsurroPLV,
matrixsurro
```

‘Evolution function’

Function called from the previous function, to run the cPLV through every epoch (for ‘i’) of every frequency (for ‘j’). Afterwards the mean of the epochs is made, in order to have only one value for every frequency.

```
def evolution_PLVmorlet_matrix(power, num_freqs, num_epochs):
      '''Get iPLV value for each frequency.
      -> Inputs:        power: morlet output
                        num_freqs: number of frequencies that we have
                        num_epochs: number of epochs (windows) that we have
      -> Outputs:       iPLV matrix for echar frequency'''
      matriz_plv = np.zeros((num_freqs, 8, 8))
      matriz_imag = np.zeros((num_freqs, 8, 8))
      for j in range (num_freqs):
            matriz_plv_inter = 0
            matriz_imag_inter = 0
            for i in range (num_epochs):
                  power_2var= power[i,9:17,j,:]
                  avg_dif = cplv2(power_2var)
                  iplv_matrix = np.abs(np.imag(avg_dif))
                  plv_matrix = np.abs(avg_dif)
                  matriz_plv_inter = matriz_plv_inter + plv_matrix
                  matriz_imag_inter = matriz_imag_inter + iplv_matrix
            matriz_plv[j, :, :] = matriz_plv_inter / num_epochs
            matriz_imag[j, :, :] = matriz_imag_inter / num_epochs
      return matriz_plv, matriz_imag
```

‘Evolution function for surrogates analysis’

Same as before, but called in the ‘matrixOBS’ function for getting the cPLV for the surrogates analysis. The difference is calling ‘cPLV’ with two signals, the normal one, and the temporal rotated.

```
def evolution_PLVmorlet_matrix_surro(power, power_surro, num_freqs, num_epochs):
	'''Get iPLV value for each frequency.
	-> Inputs:		power: morlet output
				num_freqs: number of frequencies that we have
				num_epochs: number of epochs (windows) that we have
	-> Outputs:		iPLV matrix for echar frequency'''
	matriz_plv = np.zeros((num_freqs, 8, 8))
	matriz_imag = np.zeros((num_freqs, 8, 8))
	for j in range (num_freqs):
		matriz_plv_inter = 0
		matriz_imag_inter = 0
		for i in range (num_epochs):
			power_2var= power[i,9:17,j,:]
			surro_2var = power_surro[i,9:17,j,:]
			avg_dif = cplv2(power_2var, surro_2var, is_normed=True) #Calling cplv with two powers, one rotated
			iplv_matrix = np.abs(np.imag(avg_dif))
			plv_matrix = np.abs(avg_dif)

			matriz_plv_inter = matriz_plv_inter + plv_matrix
			matriz_imag_inter = matriz_imag_inter + iplv_matrix
		matriz_plv[j, :, :] = matriz_plv_inter / num_epochs
		matriz_imag[j, :, :] = matriz_imag_inter / num_epochs
	return matriz_plv, matriz_imag
```

Summary of the code

After the libraries and the definition of all the functions, the frequency axis should be created, as well as saving all the different paths. Next, after reading the file a bipolar reference is applied.

```
raw = mne.io.read_raw_edf(path_list[i], eog=None, misc=None, stim_channel='auto', exclude=(), infer_types=False, preload=False, verbose=None)

raw.load_data()

raw = mne.set_bipolar_reference(raw, anode = ['EEG Fp1', 'EEG C3', 'EEG Fp1', 'EEG T3', 'EEG Fp2', 'EEG C4', 'EEG Fp2', 'EEG T4'],cathode = ['EEG C3', 'EEG O1', 'EEG T3', 'EEG O1', 'EEG C4', 'EEG O2', 'EEG T4', 'EEG O2']) #Bipolar reference
```

Following that, after reading the transitions (‘Results’ function), ‘metricOBS_matrix’, is called for the condition sleep phases.