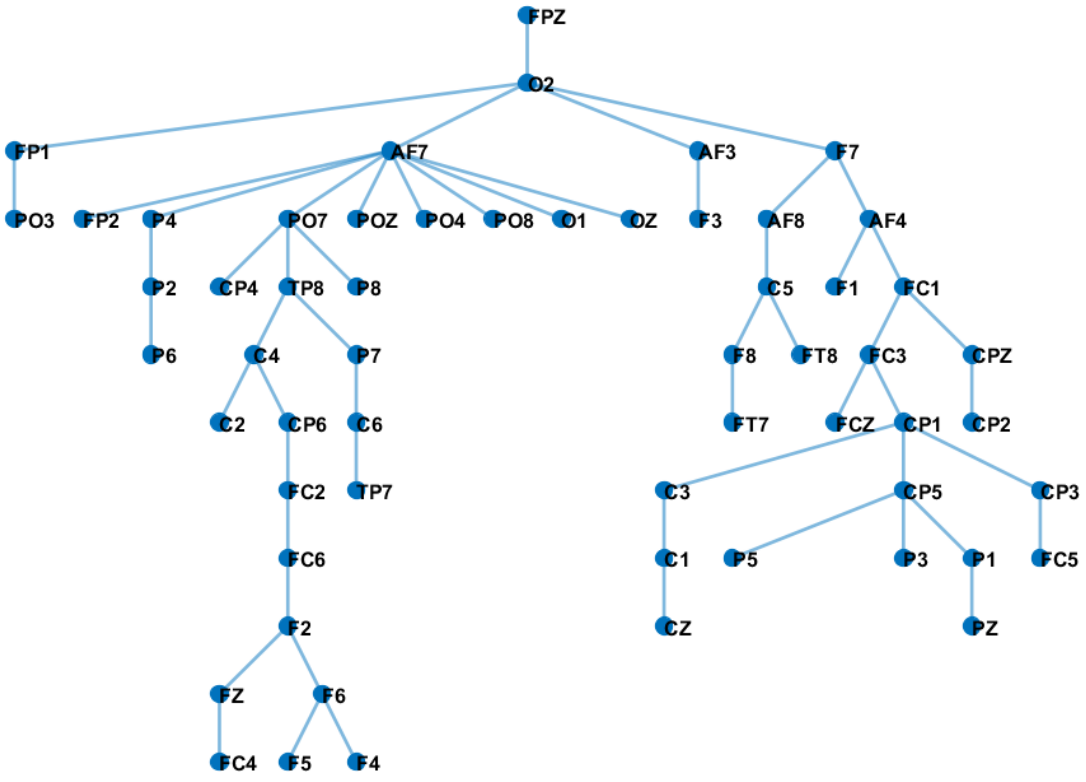
CHALMERS
UNIVERSITY OF TECHNOLOGY



Graph Theory Application to the Study of Functional Brain Networks Underlying Phantom Limb Pain

An exploratory study with electroencephalography.

Master's thesis in Biomedical Engineering

DANIEL FRIDOLFSSON

Department of Electrical Engineering

CHALMERS UNIVERSITY OF TECHNOLOGY

Gothenburg, Sweden 2019

Master Thesis: 28/01/2019 - 04/06/2019

Graph Theory Application to the Study of Functional Brain Networks Underlying Phantom Limb Pain

An exploratory study with electroencephalography.

DANIEL FRIDOLFSSON

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Electrical Engineering

CHALMERS UNIVERCITY OF TECHNOLOGY

Gothenburg. Sweden 2019

Graph Theory Application to the Study of Functional Brain Networks Underlying Phantom Limb Pain.
An exploratory study with electroencephalography.
DANIEL FRIDOLFSSON

Supervisor: Eva Lendaro, PhD student, Department of Electrical Engineering, Chalmers University of Technology.
Examiner: Associate Professor Max Jair Ortiz Catalan, Department of Electrical Engineering, Chalmers University of Technology.

Master Thesis 28/01/2019 - 04/06/2019
Department of Electrical Engineering
Division of Signal Processing and Biomedical Engineering
Biomechatronic and Neurorehabilitation Laboratory
Chalmers University of Technology
SE-412 96 Göteborg
Sweden
Telephone + 46 (0)31-772 1000

Cover: [The figure shows a so-called minimum spanning tree graph that describes the strongest connections between electrodes from EEG data. More information can be found in the Theory section on graph theory.]

Typeset in Microsoft Word
Gothenburg, Sweden 2019

Graph Theory Application to the Study of Functional Brain Networks Underlying Phantom Limb Pain.

An exploratory study with electroencephalography.

DANIEL FRIDOLFSSON
Department of Some Subject or Technology
Chalmers University of Technology

Abstract

Phantom limb pain is a chronic disease that commonly follows the amputation of a limb. Currently the origin of phantom limb pain is poorly understood, and several theories have been suggested. Most of the more recent theories all include a common element, brain plasticity. These plastic changes are believed to come from sensorimotor deprivation with intense behavioural changes[1]. However, how these changes happen and what causes them in the brain is still discussed.

The human brain functions as a network where different parts communicate with one another to achieve complex functions such as cognition and feelings. In this study said network is studied to identify possible changes correlated with phantom limb pain. Electroencephalography data from a resting state condition was used to measure each subject's functional connectivity by computing the phase lag index between each electrode pair. The functional connectivity was then used to create a special kind of graph called a minimum spanning tree. From the minimum spanning tree several different metrics were computed. These metrics reflect different characteristics of the network such as efficiency and shape.

A statistical analysis was then applied on the metrics between control subjects with subjects suffering from phantom limb pain. The statistical significance limit was set to 5% and Welch's t-test or the Wilcoxon rank sum test was used, depending on the distribution of the data, to compare the results.

Several metrics were found to be significant at the 5% significant level. Most of the metrics found were in the Delta frequency band which indicate that the networks differ more at lower frequencies. However, after multiple comparison correction on one metric, the mean phase lag index, in the Beta frequency band were considered statistically significant between the groups.

Keywords: Phantom limb pain, electroencephalography, graph theory.

Acknowledgements

First and foremost, I would like to thank all the participants of this study, without all of you none of this would have been possible. I would also like to thank my supervisor Eva Lendaro whom have helped me immensely and guided me through this project. Another important person in this project is my examiner Max Ortiz-Catalan who helped me deciding upon what type of project I would like to do as my master's thesis and his role as examiner in this process. For this I am extremely grateful. A special thanks to all the lovely people at the Biomechatronics and Neurorehabilitation Laboratory at Chalmers for the amazing work environment you all have provided me with. Lastly, I would like to thank my friends and family who have survived me during stressful times and who never failed to improve my mood when it was most needed, especially you Amanda!

Daniel Fridolfsson

Gothenburg, June 2019

List of Figures

Any Figures not referenced in their captions is made by the author of the thesis.

List of Tables

List of Abbreviations

Table of Contents

1 Introduction

In this sections a short background to the project is presented covering the uncertain origin of phantom limb pain (PLP) followed by the aim and scope of the project. Lastly the outline for the rest of the thesis is presented.

1.1 Background

Technological improvements in the field of non-invasive neuroimaging have opened vast possibilities in the study of the structure and function of the human brain. The ongoing quest to fully understand complex aspects such as cognition is in full progress. However, when it comes to understanding how the brain changes following a traumatic event such as an amputation many questions are still unanswered. One of the most challenging problems in this field is the understanding of the origin and maintenance of PLP, which is a severe form of chronic neuropathic pain arising from the missing limb.

Many attempts have been made at explaining PLP. An early paper in the field written by Flor et. al.[2] theorized that cortical reorganization could account for some non-painful phantom limb phenomena and have an adaptive function. It had been observed in patients suffering from chronic back pain that the amount of cortical alteration reduced the magnitude of pain hence they predicted that the same would be true in the relationship between cortical reorganization and PLP. However, with non-invasive magnetic source imaging (MEG) a strong correlation between cortical reorganization and PLP was found. Their data indicated that PLP is related to plastic changes in the primary somatosensory cortex.

Makin et. al.[3] analysed resting state functional connectivity of the brain of arm amputees. Functional magnetic resonance imaging (fMRI) was used to identify large scale reorganisation besides the primary sensorimotor cortex compared to a two-handed control group. They specifically examined changes in functional connectivity values between the cortical territory of the missing hand in the primary sensorimotor cortex and both the sensorimotor network and the default mode network (DMN). The sensorimotor network is normally strongly associated with the hand cortex while the DMN is typically dissociated. In this study, Makin et. al.[3] show that the functional connectivity values between the missing hand cortex and the sensorimotor network were reduced in amputees and connectivity was weaker in patients who had been amputated for a longer period of time. Together with lower levels of functional coupling between the missing hand cortex and the sensorimotor network there was also a stronger connectivity between the missing hand cortex and the DMN. Their result demonstrated that plasticity following arm amputation is not restricted to local remapping but rather leads to a cascade of cortical reorganisation at a network level scale. It could however not identify a clear statistical relationship between increased coupling with the DMN and PLP which had been seen in other types of chronic pain. However, this could also be a limit of the type of data used, namely fMRI, which due to the physiological constraints of the temporal resolution of the hemodynamic response, might not give a complete representation of how brain networks change due to pain.

An encompassing theory has been brought forwards by Max Ortiz-Catalan[4] who proposes the stochastic entanglement of the pain neurosignature with impaired sensorimotor circuitry as a possible origin of PLP. He argues that following amputation or sensorimotor impairment, the related motor and somatosensory circuitry become susceptible to perturbation and could wire together with other networks. Hence stochastic entanglement could occur, and networks of sensorimotor processing and pain perception begins to activate together. This would also explain why not all amputees suffers from PLP even though they experience almost identical conditions.

All these theories and studies share a common denominator, they all discuss the neural network in the human brain.

This project intends to analyse brain networks for both healthy patients and patients suffering from PLP with a graph theory approach based on electroencephalography (EEG) data to evaluate the topological differences in said networks. If several networks in an PLP patient wire together the connectivity[5] between the networks should differ compared to a healthy control. By computing the phase lag index (PLI) the functional connectivity can be determined. EEG data have a high temporal resolution in the range of milliseconds[6] which could be important because the functional connectivity is highly time dependent and constantly changing[5]. An EEG amplifier is also very mobile and cheap compared to an MR scanner, alternatively used for functional connectivity studies. By using graph theory with each electrode as a node and the connections between them as edges a model can be created to evaluate whether PLP can be related to the connectivity between each node. The final aim is to find metrics that can show a difference between amputees with PLP and healthy controls and find metrics that correlate reliably to some aspect of PLP.

1.2 Aim

Develop a functional system to create brain networks and evaluate its potential to discern healthy patients from patients suffering from PLP. To evaluate the potential to discern the two groups graph theoretical measures are to be computed in both groups. From these metrics identify the ones that relate to the presence or absence of PLP.

1.3 Scope and Limitations

In this study only two groups are compared, subjects with PLP and subjects without PLP. The comparison is limited to graph theory and the metrics introduced in the

Method section. Different pain rating scales, time since onset etc. of PLP are not considered in the analysis. All analyses is done on EEG data recorded in the Biomechatronic and Neurorehabilitation Laboratory at Chalmers university of technology.

1.4 Thesis Outline

The thesis is divided into eight major sections, Introduction, Theory,

Method, Result,

Discussion,

Conclusion,

References and Appendix. The first section, *Introduction*, lays out the reasons why this study was carried out and how this aims to contribute to the scientific field. The introduction is followed by section 2, *Theory*, which covers the basics of the theory necessary to understand the methodology used in the study. Section 3,

Method, presents how data are acquired, processed and how graph analyses were performed. It also gives relevant information about the participants of the study. *Result,* in section 4, summarizes the results of the graph theory analyses and the statistical comparisons. Section 5,

Discussion, intends to address further considerations regarding the development and the execution of the different parts of this master thesis project, for instance clarifying why certain methods were adopted, what are the limitations of this study and how these could be address by extending this work. The project is shortly summarized in section 6,

Conclusion.

2 Theory

This project intertwines two disciplines. The first discipline being electrophysiological monitoring and brain imaging, more specific EEG, and the other the mathematical field of graph theory.

2.1 Electroencephalography

EEG is an electrophysiological recording technique used to measure the flow of neural ionic currents. The currents are recorded using pairs of electrodes either inside the scalp, electrocorticography (ECoG), or on the outside of it, scalp EEG. In this study the noninvasive scalp EEG was recorded and is hence focused upon.

EEG measures the differences over time of the electrical potential at two electrodes. These electrodes may either be directly attached to the scalp surface or fitted into a cap to simplify the attachment. The positions in the cap or directly at the surface do most often abide to the international 10-20 system, see Figure 1 for labels and positions of the electrodes. This system aligns the electrodes based on a percentage of the distance of the scalp from the nasion to the inion and between the mastoids. The "10" and "20" refers to how many percentages of the distance it is between the adjacent electrodes. The 10-20 system has over the years developed higher resolution versions of itself as well as the 10-10 system[7] and the 10-5 system[8] to include more and more electrodes placed around the scalp.

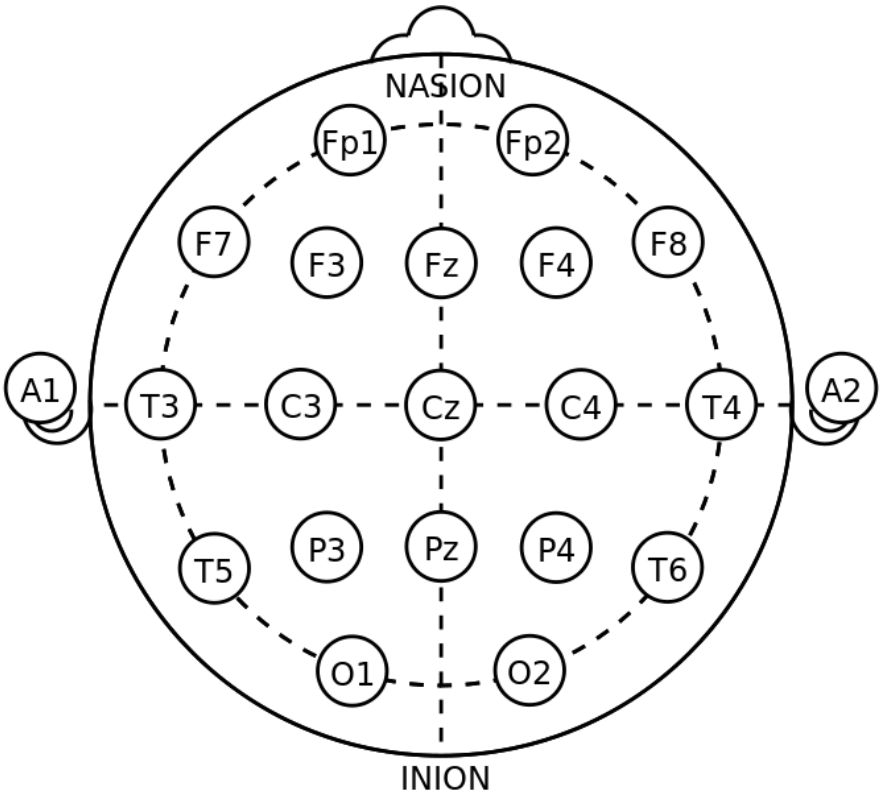


Figure 1: Electrode locations of International 10-20 system for EEG recordings. The image is taken from Wikimedia Commons and is made by Asanagi[9].

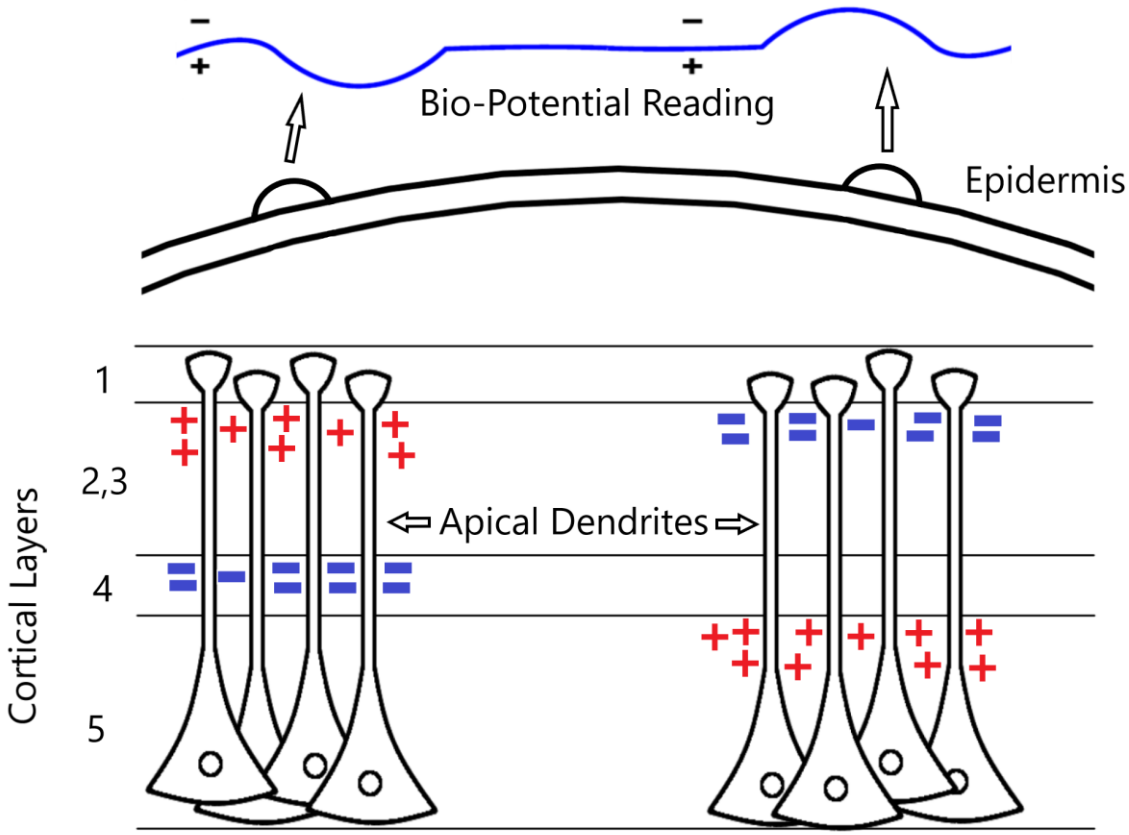


Figure 2: The figure shows how different dipoles contribute to the potential measurement.

The potential differences are generated by neurons, in particular cortical pyramidal cells, that form synapses at their dendrites, see Figure 2: The figure shows how different dipoles contribute to the potential measurement. Due to their unique anatomical structure as a long apical dendrite perpendicular to the cortical surface, cortical pyramidal neurons are excellent dipoles. The direction of the dipole is determined by the superficial or deep location of the synaptic input. Due to the direction of the dipole it can be separated from EMG-signals which have a horizontal orientation. In Figure 2 it can be seen how the charge of the dipole affects the EEG-recording.

There are two different intracellular potentials that potentially can generate scalp EEG signals, action potentials and postsynaptic potentials. An action potential is generated by a sudden change in transmembrane resting potential due to movements of intracellular and extracellular ions. When the action potential propagates to a synapse the postsynaptic potential is generated across a pair of neighboring neuronal membranes. If the postsynaptic potential exceeds a threshold level the action potential of the first neuron is delivered to the next. It is believed that the postsynaptic potential contributes to a higher degree to the generation of measurable extracranial electric fields. This is because contrary to the action potential the synaptic potential can be generated synchronously in a large number of neurons. It is possible because of its relatively longer duration (~30ms compared to ~5ms).[10]

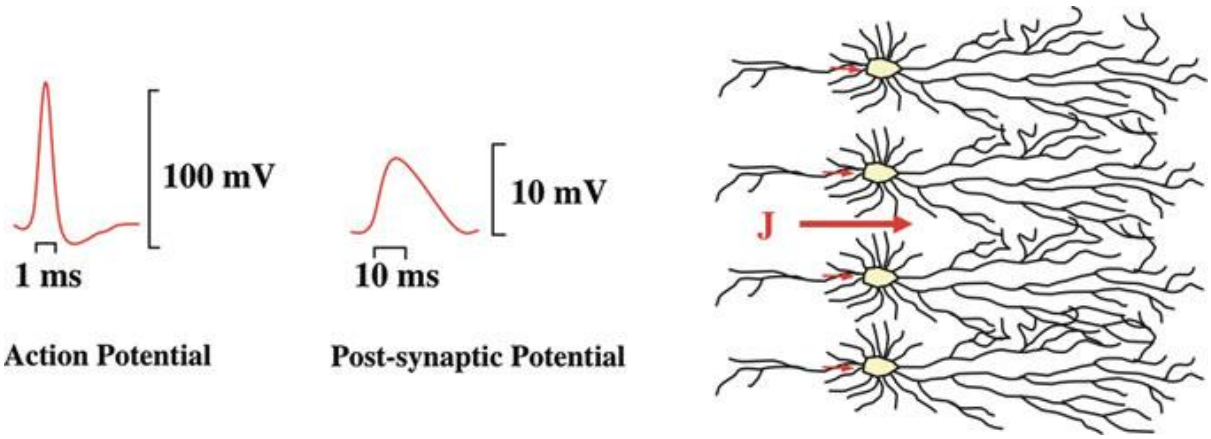


Figure 3: The image displays the difference between the waveforms of action potentials and postsynaptic potentials. Synchronous occurrence of postsynaptic potentials can produce a current flow large enough to be detected from outside of the head. The image was taken from Chang-Hwan's book Computational EEG Analysis[10].

When numerous cortical neurons within a small area are activated synchronously a so called unidirectional neural current flow is formed. Such a current is called *primary* or *impressed currents*. Because of the dielectric properties of the human body the extracellular currents induced by the primary currents can flow anywhere in the human body. These extracellular currents are known as *secondary, volume*, or *return currents*. The flow of these secondary currents results in nonuniform potential distributions on the scalp. These are the potential differences between two scalp positions over time that defines the EEG.[10]

There are a few other techniques that could be used to study the brain and its networks such as MEG and fMRI. Both MEG and EEG have high temporal resolution while fMRI have a higher spatial resolution. Because the dynamics of cognition are very fast[11] EEG is very well suited when analyzing brain networks.

2.2 Preprocessing

Preprocessing is any transformation or reorganization that occurs between collection and analysis of data. Preprocessing steps are meant to facilitate the analysis and remove unwanted distortions like noise. Hence some preprocessing steps merely organize the data without changing them, other steps involve removing bad or artifact-ridden data without changing clean data, and some involve modifying the data with the purpose to clean them

EEG data contain the signal, i.e. the potential differences discussed in the previous section, and noise. Proper preprocessing will attenuate the noise without altering the data. The data and noise are in most cases mixed together and the attenuation or removal of the noise may come at the cost of data. On the other hand, one may not be interested in all the data but only a certain frequency range which means data lost outside this range does not affect the result. In other words which preprocess routines that are applied in a project is heavily affected by what is being studied.[11]

2.2.1 Filtering

A filter removes undesired components or features of a signal. This oftentimes involves removing some spectral characteristics of a signal, such as the powerline noise around 50-60 Hz[11]. In this study both infinite impulse response (IIR) and finite impulse response (FIR) filters were considered.

2.2.1.1 Infinite Impulse Response Filter

IIR filters are characterized by having an impulse response which never becomes zero but continues indefinitely. In practice, the impulse response approaches zero in most cases and can be disregarded past a certain point. The transfer function of an IIR filter is often defined in its Z-transformed state:

$$H(z) = \frac{\sum_{i=0}^{P} b_i z^{-i}}{\sum_{j=0}^{Q} a_j z^{-j}}.$$

However, in most IIR filter designs the coefficient a_0 is 1. Hence the IIR filter transfer function is often expressed as:

$$H(z) = \frac{\sum_{i=0}^{P} b_i z^{-i}}{a_0 + \sum_{j=1}^{Q} a_j z^{-j}}.$$

P is the feedforward order, b_i is the feedforward filter coefficients, Q is the feedback filter order and a_i are the feedback filter coefficients.

An IIR filter is very efficient compared to a FIR filter in order to meet the specifications in terms of passband, stopband, ripple and roll-off. These specifications can be achieved with a lower filter order. In other words, IIR filters are less computationally expensive than FIR filters. However, even though they can be implemented with a lower rank the implementation itself is more complicated. It is simpler to implement specific characteristics in a FIR filter especially when not one of the more common cases are of interest. IIR filters can also diverge due to their definition. Lastly an IIR filter have a non-linear phase and could cause phase distortions. [12]

2.2.1.2 Finite Impulse Response Filter

A FIR filter is, as the name suggests, a filter whose impulse response have a finite duration. FIR filters are designed by finding the coefficients and filter order that matches the needs of the application. The impulse response of a FIR filter can be defined as:

$$h[n] = \sum_{i=0}^{N} b_i \cdot \delta[n-i] = \begin{cases} b_n & 0 \leq n \leq N \\ 0 & \text{otherwise} \end{cases}$$

N is the filter order, an Nth-order filter has $(N+1)$ terms in the sum. b are the filter coefficients, i.e. the value of the impulse response at the specified instant. δ is the Kronecker delta function.

The advantages of FIR filters are that they require no feedback which reduces the amount of errors due to summed iterations. They are also naturally stable. Most importantly FIR filters can easily be designed to have linear phase. This property is desirable when working with phase-sensitive applications.[12]

2.2.1.3 *Zero-phase Filtering*

As mentioned earlier FIR filters can be designed to have a linear phase i.e. introduce a linear phase shift most often perceived as a time lag, see Figure 4. In this project phase-lag will be used as a measure of similarity, see section Graph Theory. Hence an alteration in the phase is undesirable.

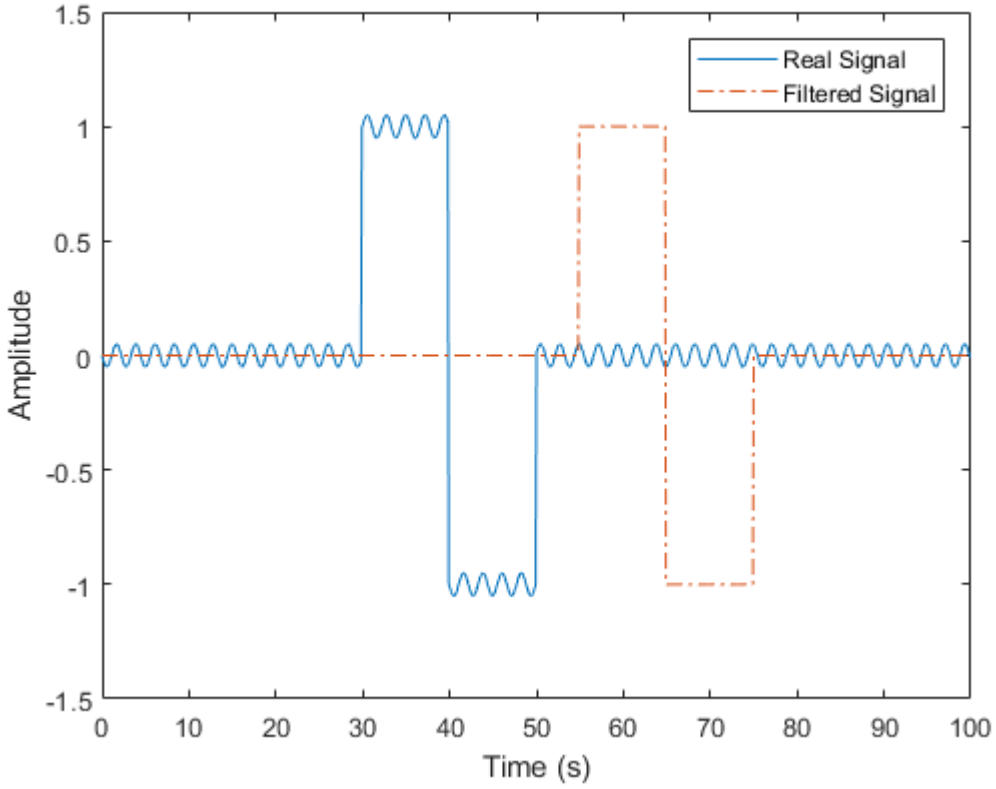


Figure 4: The figure shows how a signal gets delayed in time after filtering. The time lag in this case is very long.

Zero-phase filtering is filtering done without affecting the phase of the signal. This is achieved when the frequency response of the filter is a real and even function. In some cases, the filters themselves are designed in a manner to achieve a real and even frequency response but this is almost impossible to do with IIR filters. Instead what can be done, both with IIR and FIR filters, is to apply the filter from both ends of the signal. If $x[n]$ is the input signal and $h[n]$ is the filter's impulse response, then the result of the first application is in frequency domain:

$$X(e^{j\omega})H(e^{j\omega}),$$

where X and H are the Fourier transforms of x and h. Time reversal corresponds to replacing ω with $-\omega$ resulting in:

$$X(e^{-jw})H(e^{-jw}).$$

For real-valued filter coefficients $H(e^{-jw}) = H^*(e^{jw})$, which means when the filter is applied again to this signal the result is:

$$X(e^{-jw})H(e^{-jw})H(e^{jw}) = X(e^{-jw})|H(e^{jw})|^2.$$

Which after time-reversal means that the filtering process is the same as filtering with a frequency response $\left|H\left(e^{jw}\right)\right|^2$ which is a real and even function.

2.2.2 Artifact Removal

There are several ways of removing artifacts. If the shape of an artifact is well known and enough data is acquired, then segments containing said artifact can manually be avoided. An example of this can be seen in Figure 5 where three blinks are marked in the data. Problems could occur if the artifact is present if not in all but in most of the epochs created to analyze the data. In EEG data these artifacts could be blinking, eye movements or swallowing. Although eye movements and swallowing are not frequent enough to infect most of the data, blinking is. Instead of removing all data epochs where blinking is found several other methods could be used to remove them. In this study independent component analysis (ICA) is used to filter our such artifacts.

2.2.2.1 Independent Component Analysis

ICA decomposes the signals into n independent components where n is the dimension of the data. In EEG data, n would represent the number of available channels. The data are separated not by direct spatial filtering for activities generated in a predefined cortical location. But instead by using the information content of the data itself to separate parts of the EEG data from each active cortical and artifact area. Based on simple but statistically and physiologically plausible assumptions[13] that over time these activities should be nearly independent. This assumption also includes that most of the far-field potentials detected at the scalp are generated within the skull.

This approach has the major advantage that EEG sources will be grouped together into a single independent component (IC) that include all projections to the EEG channels while unrelated EEG sources will be rejected from said IC and stored into other ICs. Under favorable circumstances ICA transforms the recorded scalp data into a set of source recordings thereby discovering *what* distinct signals are present in the data instead of *where* the data is generated.[14]

The data submitted to ICA are an n by t matrix (X) where n is the number of channels and t the time points of the data. Unlike similar spatial filtering procedures ICA requires no channel location information. Based on the criterion that resulting source time courses (U) are maximally independent ICA finds a component unmixing matrix (W) that, when multiplied by the original data (X), result in the matrix (U) of IC time courses:

$$U = WX.$$

X and U are $n \times t$ matricies while W is an $n \times n$ matrix. By inverting the unmixing matrix the mixing matrix (W^{-1}) can be computed. Each column of the mixing matrix represents the projection weight at each electrode, i.e. the IC scalp map. These projections are then mapped on a 2-D model to allow visualization of the scalp projection of each source. The source locations are presumed to be stationary, which means that the brain source locations and projection maps (W^{-1}) are assumed to be spatially fixed. Hence their activations (U) reveal their activity time courses throughout the input data. Hence what is being done is rather:

$$X = W^{-1}U.$$

Where W is trained to create as independent components as possible while constrained by the assumptions of ICA.

As mentioned earlier ICA makes a key assumption: In the far-field signals produced by the cortical and non-cortical EEG sources are temporally distinct, and over sufficient input data, near

temporally independent of each other. Thus, the quality of the decomposition is highly dependent on two factors.[13]

First, the number of time points of the n-channel data used in the decomposition must be enough to learn the n^2 weights in the ICA unmixing matrix. A common rule of thumb is that the amount of data needed is related to the number of electrodes squared, n^2, times a factor k. k is usually set to be 25 or higher for a dataset with a high number of electrodes. In this project a sampling rate of 2400 Hz was used and a maximum of 62 channels, 62 when eye-channels were included otherwise 58, which means at least 40 second of filtered data is needed. However, ICA decompositions where $k > 25$ tends to be more regular and produce more dipolar component maps. Hence the more data the better as long as the condition of the recording don't change, i.e. going from a resting-state eyes-open condition to resting-state eyes-closed condition in the same data set.

Second, as is the case with all signal processing: "*garbage in, garbage out*". If the EEG data is still contaminated by high noise or the recording itself is faulty in some way the ICA decomposition will be of the same quality. ICA may be able to separate the typical noises such as powerlines. However, as it is very simple to remove that type of noise before the decomposition with filters it is recommended.

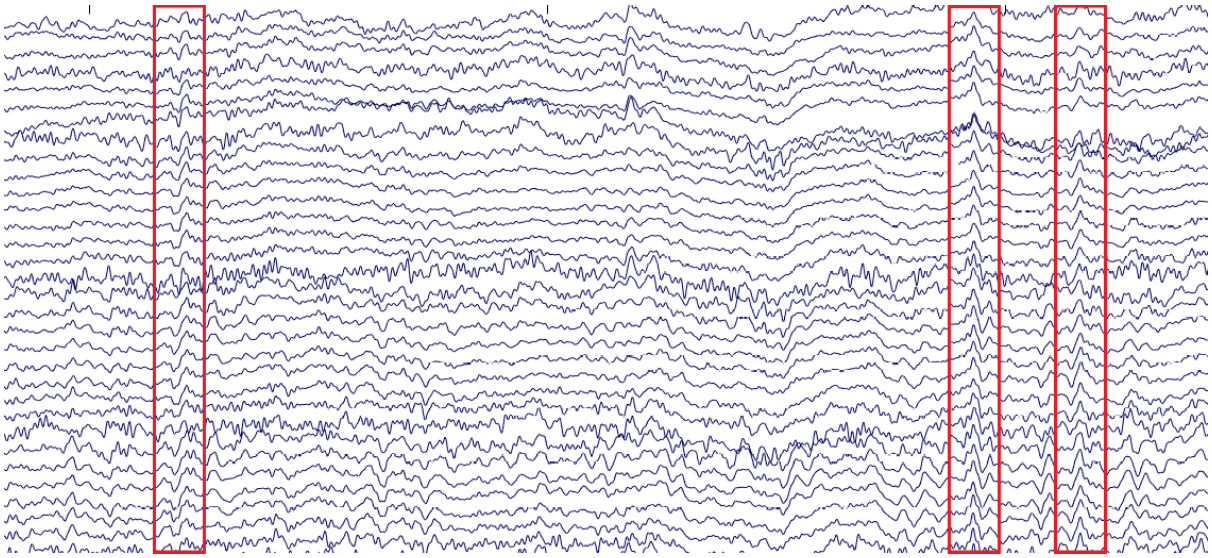


Figure 5: The Figure displays EEG-data where three blinks are marked with red rectangles. A clear peak can be seen in all channels which is a characteristic trait of a blink artifact.

2.2.3 Epoching

EEG data is recorded continuously and is usually represented as a two-dimensional matrix where the rows are electrodes and columns are time. Epoching means dividing the data set in segments that are often locked around task-related changes in the EEG. It is not necessary in resting-state datasets but can be used to sort the data into non-overlapping segments to easily discard or keep epochs that are heavily affected by noise or artifacts. As can be seen in the Figure 6 epoch after event 2 and after event 3 is selected while the epoch after event 1 was considered un-useable. One thing to consider when choosing epochs is to make sure their length is appropriate. For ERPs the epochs can be as long as the time period of interest plus a baseline period. However, if the

analysis involve filtering edge artifact may become a problem due to transients created from filtering.

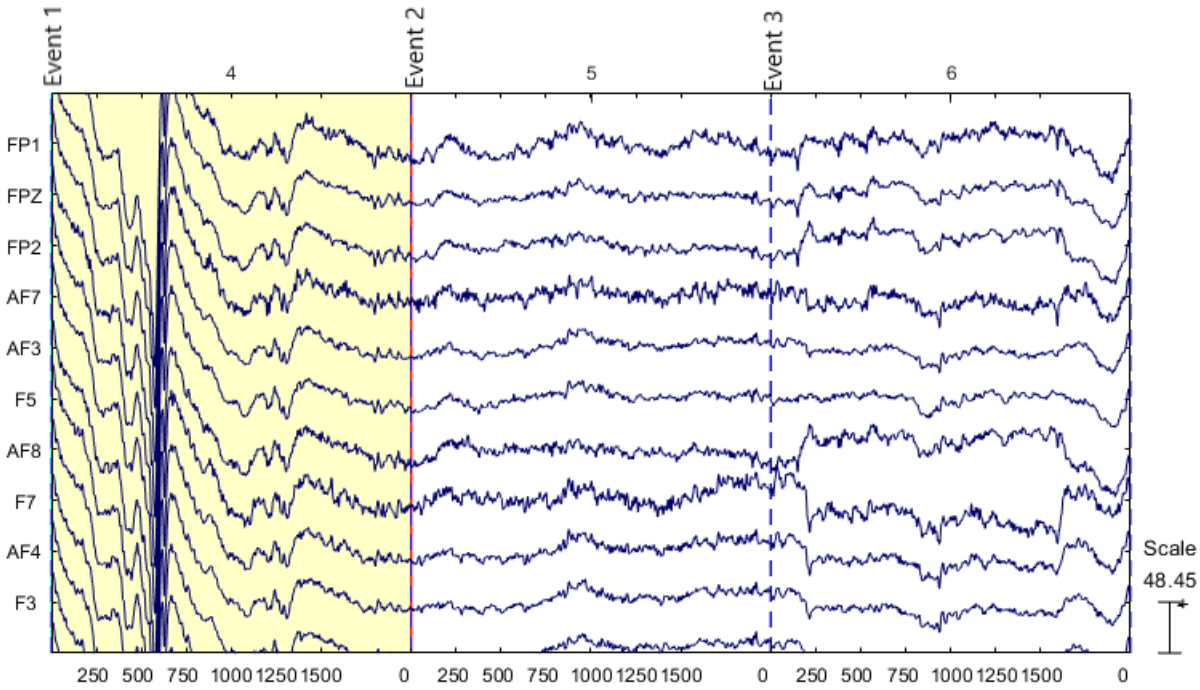


Figure 6: The figure shows two EEG signals being epoched into three 2-second-long epochs where two of them are marked while the first one is unmarked. Due to the large artifact, which is likely from some type of movement, the first epoch is considered un-useable and is hence not marked.

2.3 Connectivity

Brain networks can be derived from different types of observational techniques Specifically they can be distinguished in either *structural connectivity*, *functional connectivity* or *effective connectivity networks*.[15]

Structural connectivity describes the anatomical connections among a set of neural elements. When used to analyze the human brain, structural connections most often correspond to white matter projections linking cortical and subcortical regions. This kind of connectivity is relatively stable on time scales of seconds to minutes but may change during longer time scales of hours to days due to plastic experiences. In neuroimaging studies structural connectivity consists of a set of undirected links due to the impossibility to evaluate the directionality of projectionsfssd.[15]

Functional connectivity corresponds to temporal correlations of activity and may occur between pairs anatomically unconnected regions. These times series data may be derived from various techniques such as EEG, MEG or fMRI. The data can be analyzed in numerous ways and are all based upon a correlation measure. The presence of a statistical relationship between two neural elements is often interpreted as a sign of functional coupling however, the presence of such a correlation does not imply a causal relationship[15]. Functional connectivity is a highly dynamic and time-dependent measure. It often changes in matters of tens or hundreds of milliseconds[15] as the connections are continually modulated by stimuli and task context.[10,16,17]

Effective connectivity attempts to measure directed causality in a network and identify which mechanisms causes which reactions. Thus, effective connectivity networks are time- and task-dependent in nature. Effective connectivity may be view as the union of structural and functional

connectivity. However, most studies are still carried out on either structural or functional connectivity data.[10,15,16,17]

In Figure 7 the different connectivity types are described. The top images display how the elements interact with each other while the bottom part illustrate the connectivity matrix from the same network. In the effective connectivity the connection matrix is not symmetric which can be seen that the connection between two elements not necessarily is the same in a causal network.

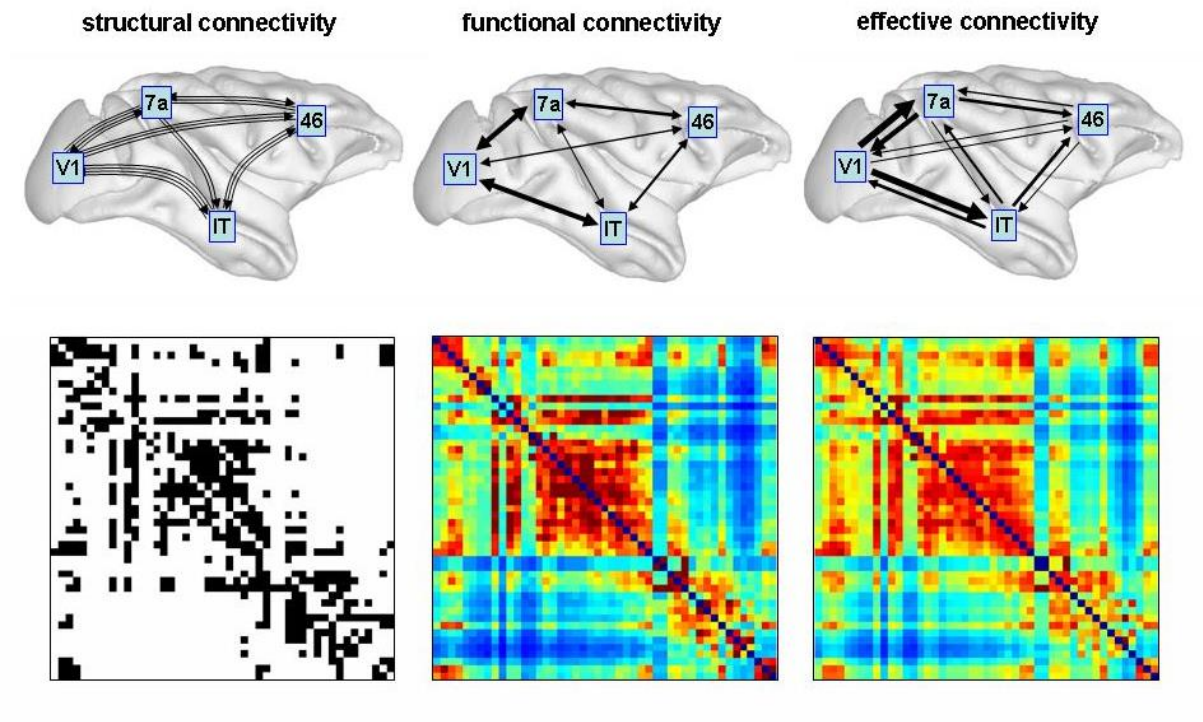


Figure 7: The image displays how the different connectivity measures interact among elements in the network. It also displays the connectivity matrix created from said networks. The image is taken from Sporns article about brain connectivity on Scholarpedia[17].

2.4 Graph Theory

This theory section is based on several books about graph theory and graph theory in spectral analysis. If there is an interest for a deeper understanding they are all referenced in the bibliography of the thesis.[18,19,20,21,22]

A graph may be used to illustrate any information that can be modelled both as objects and relationships between said objects. Graph theory is the study of those graphs consisting in the definition, computation and analysis of different metrics that reflect and summarize inherent properties such as nodal importance or topological properties.

A *graph*, $G(N, E)$ consists of a set of *nodes*, $N(G)$, and a set of *edges*, $E(G)$, connecting the nodes to each other. The two nodes associated with an *edge*, e, are usually called *end-nodes* of e. The edge between two nodes, u and v, are denoted by $e(u, v)$. The number of nodes in a graph is often denoted n and the number of edges as m. Generally, graphs are drawn by representing each node

with a point or a small circle and each edge by a line segment between its two end-nodes, see Figure 8.

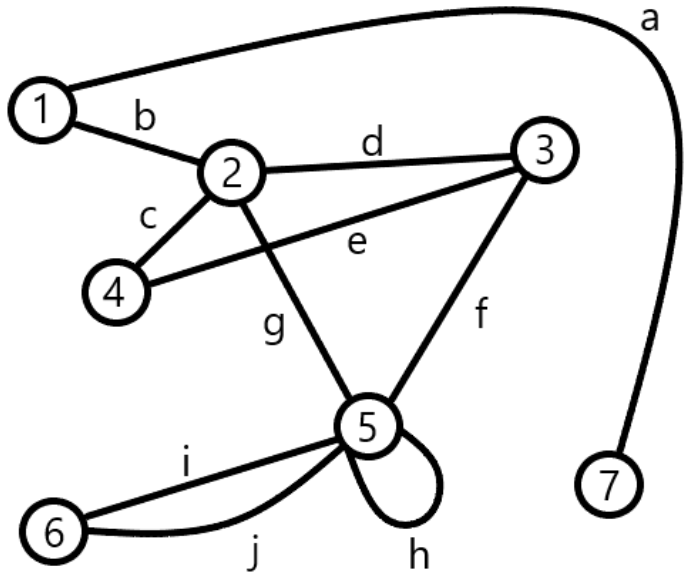


Figure 8: The figure shows a typical graph with 7 nodes and 9 edges where edge h is a loop. Due to the loop at edge h and the multi edge between node 5 and 6 this is not a simple graph but a multigraph instead.

A *loop* is an edge whose end-nodes are the same node, see edge h in Figure 8. Other types of looped structures are *multiple edges*. These are edges with the same pair of end-nodes, edge i and edge j are an example of multiple edges. A graph that contains neither loops nor multiple edges is called a *simple graph*. If there are any loops or multiple edges the graph is instead called a *multigraph*. A graph can also be *directed* or *undirected* which reflects whether the edges have a direction, i.e. in a directed graph the connection is only valid in the direction assigned to it. In an undirected graph on the other hand the connection goes both ways. Whether a graph is directed or undirected is most often decided by the type of information that is being modeled. A graph is called *weighted graph* if a weight is assigned to each node or edge. If the weight is assigned to the edges it could reflect the distance or connectivity between nodes or if the weight is assigned to the nodes it can reflect nodal importance instead. If no weight is assigned to the edges the graph is often referred to as *binary*. The last definitions involve the nodes connected to each other. If $e = (u, v)$ then the two nodes u and v are said to be *adjacent* in the graph G or to be *neighbours* in graph G. On the other hand, edge e is said to be *incident* to node u and v. With clean resting-state EEG data with N-channels the graph theoretical analysis is usually done in one of the following ways. [23]

The first approach calculates the connectivity matrix and considers a threshold T. The connectivity matrix, i.e. the graph, gets binarized by setting all weights below T to zero and the weights equal or over T to one, see Figure 9.

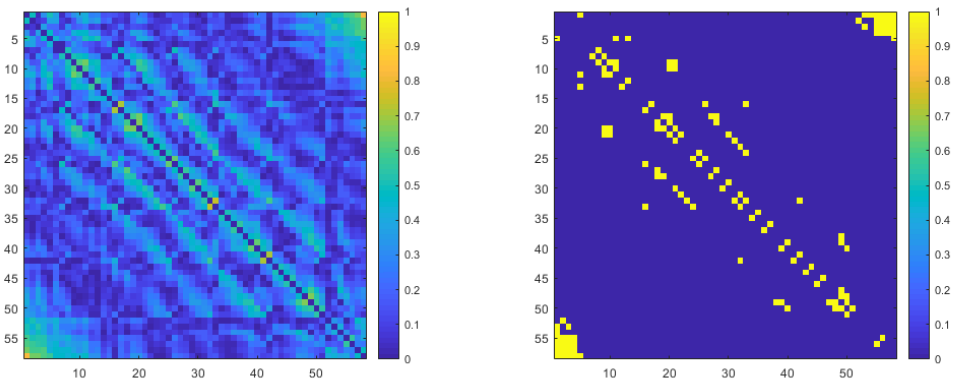


Figure 9: The left image displays the phase lag index between each electrode in a weighted network where the weight corresponds to the color in the color bar. The image to the right is the same network but binarized at 0.5.

If the weight matrix was symmetric the corresponding graph will be undirected whereas if the weight matrix was asymmetrical the corresponding graph will be directed. This entirely depends on the type of information the weight matrix is based upon. The number of nodes in the graph will always be N but the number of edges, m, depends on what threshold is used.[23]

The second approach to graph analysis is to not binarize the weight matrix and instead use all edges with their respective weights. In this case the maximum number of edges is $(N-1)N/2$. This results in a weighted graph and requires specific metrics to characterize it which may or may not be applicable to binary graphs.[23]

Alternatively, one could also combine both methods and keep the weights assigned to edges over a specified threshold. This is more common when larger networks are analyzed, and a weighted network is preferable where it becomes impractical to keep all edges.[23]

A problem with these design methods is the threshold T. When comparing two networks with the same N and T it is very unlikely that they'll have the same m. The value m does however influence several common graph theoretical metrics like the clustering coefficient and path length. This means that the chosen threshold functions as a bias in further calculations because it changes m. Another problem is the choice of T, which is essentially arbitrary. One could address this by considering a range of values of T but then these ranges become arbitrary and the value of each T is still biased. A range of T values also causes problems with the statistical analysis due to separate tests will have to be done for each value of T, increasing the likelihood of type 1 error.[23]

Two solutions have been proposed, fixing m and comparison with random control networks to create normalized metrics. To fixate m, T can be chosen separately for each graph to be compared such that all graphs have the same m. This will solve the bias arising from different values of m but introduce a new problem, proper choice of the value of m. Another option would be to consider random graphs with the same N, T and m. These graphs can be obtained by randomly shuffling the original graph and create a normalization factor by computing the graph theory metrics of interest on the randomized graphs as well and use the mean of those as a normalization value, see Figure 10. Normalization does however not solve all the bias problems[24] and leaves an open choice of proper choice of m. A bias is still present due to normalization effects on different types of network's properties, meaning that depending on the type of network the normalization process increases or decreases certain metrics more.[23,24]

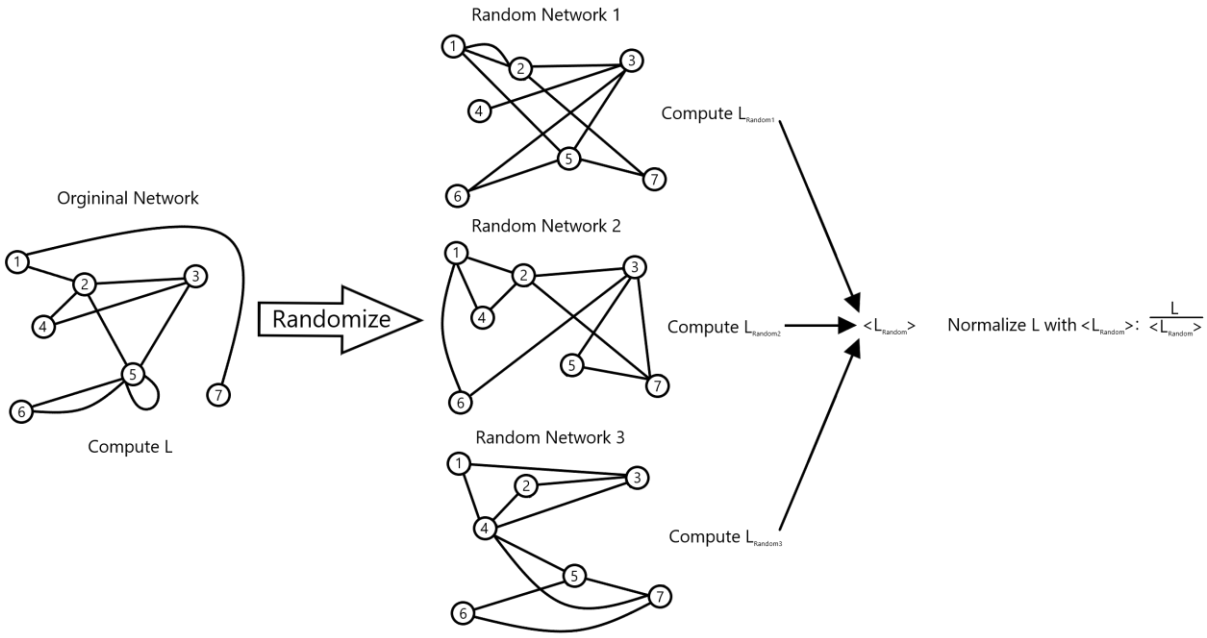


Figure 10: The image illustrates how to generate a normalized value of the characteristic path length (L), a common graph theory metric. Note that the number of edges is kept constant in the randomization process. This is to not introduce a bias in the normalization factor.

In the case of weighted graphs, problems occur based on the sum of the weights, W. The total weight will affect both the weighted clustering coefficient and the weighted path length for both the original graph and randomly shuffled graphs if that 'solution' is applied. A higher W results in both higher clustering coefficient and higher path length which means a bias will be present if two graphs have different W. The normalization process discussed in the previous paragraph reduces the bias but does not solve it completely.[23]

Stam et. al[23] proposed creating a so called minimum spanning tree (MST) as a means to obtain an unbiased representation of a weighted network.

2.4.1 Minimum Spanning Tree

A MST is an acyclic subgraph, i.e. a simple graph, that connects all nodes and minimizes the weights between the nodes in the network[25]. MST is designed to avoid biases caused by the differences in connectivity in networks with the same number of nodes[23]. The MST always contains $m = N - 1$ edges, where N is the number of nodes. An MST is created by applying Kruskal's algorithm which iteratively selects the edges in the connectivity matrix with the lowest weights and adds the edge to the tree only if no loops are created. The result is hence a graph with no cycles or loops in which all nodes are connected[23], see Figure 11.

An MST constructed from a weighted graph with unique weights is unique. This uniqueness discards the need to choose and arbitrary threshold or value of m to reconstruct the graph. MST have the advantage of focusing on the most important subgraph and avoid biases due to differences in W since after the tree has been constructed the edges are binary. A tree is also a much simpler structure and simplifies the analysis of the graph. However, the downside with this approach is that some properties of a graph is not reflected by the MST, particularly those that are based on cycles.

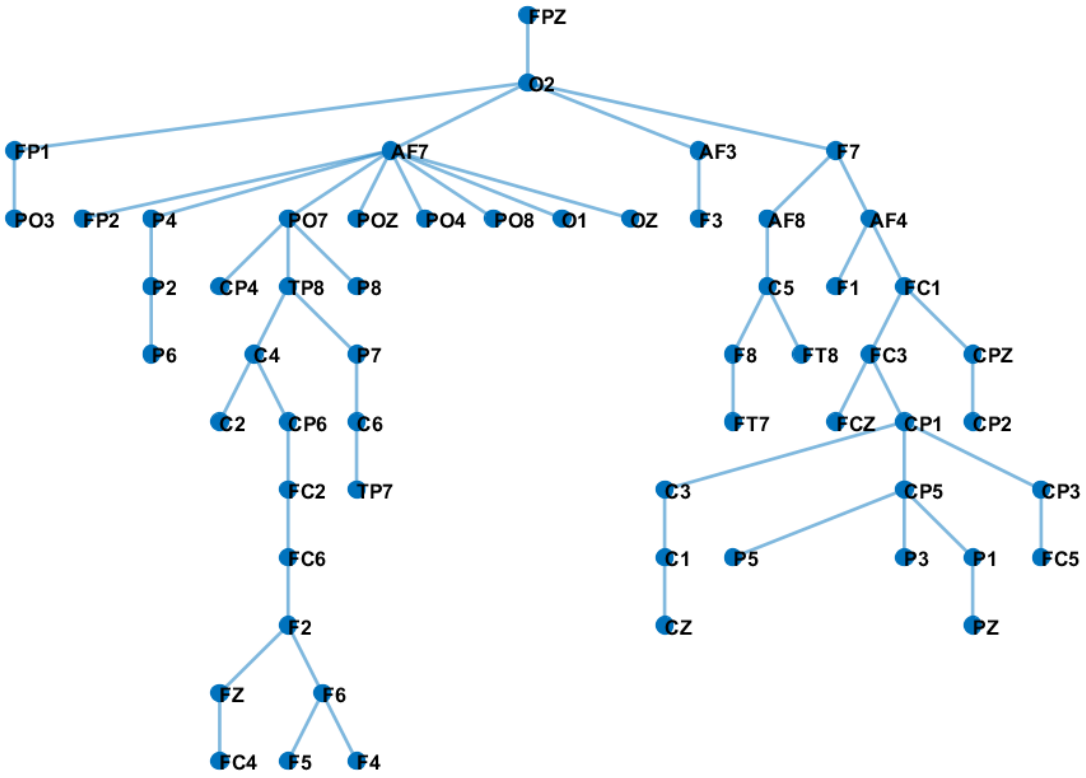


Figure 11: The figure shows a minimum spanning tree made from the phase lag index from data in the beta frequency band. Each electrode is represented by one node in the network.

2.5 Statistical Analysis

There are several methods to evaluate whether a result is statistically significant or not. The type of test that should be used depends on the type of data analyzed and its distribution.[26]

The most common way to use statistical analysis is to use hypothesis testing. This is a method of making statistical decisions using experimental data. This data is evaluated regarding two different hypotheses, the null hypothesis (H_0) or the alternative hypothesis (H_1). An example of null hypothesis is for example that the samples under analysis come from to the same distribution: this can be evinced by showing that the mean of the two sample are equal at a certain level of confidence. The opposite case, the alternative hypothesis, affirms that the two samples come from two different distributions instead. The statistical analysis then evaluates the probability that the means are different. Usually a threshold, α, is decided upon which the probability needs to be less than. This probability is usually 5%, 1% or 0.1%.[26]

To begin with, the number of groups to compare between is relevant. Depending on whether there are one, two, three or more groups, different statistical tests can be employed. In this study the number of groups are two, as there are two conditions (pain and no pain). [26]

Next the distribution of the data is relevant. Depending on the normality of the distribution either a parametric or nonparametric test is suitable. Many tests such as ANOVA or t-tests assume that the data follows a normal distribution. Tests that assume the data is normal distributed is called parametric tests. This might seem like a restrictive constraint, however surprisingly many types of biological data follow the normal distribution when a large enough sample is available. In some situations, usually when the samples are small, when the data is not normal distributed so-called

nonparametric tests are needed. Nonparametric tests can be applied regardless of the data distribution. Most parametric tests have a nonparametric variant. [26]

Nonparametric tests are thus more robust than its parametric counterpart. However, nonparametric tests usually have less power. This means that the probability that the test will reject the null hypothesis when the alternative hypothesis is true is less likely i.e., they tend to make more type II errors, false negatives. [26]

To decide whether a parametric or nonparametric test is more suitable, a normality test can be used. Common tests are *D'Agostino-Pearson normality test* or *Kolmogorov-Smirnov test* but there are several others that are used as well. Normality tests are also based on statistical analysis which means that they lose power when the samples are not large enough. This is because small samples do not contain enough information to decide whether the distribution is normal or not. [26]

Data can also be dependent or independent between samples. Data is considered dependent when we expect it to vary less between certain groups of the data, for example mothers and daughters in a genetic study. [26]

Further corrections must be done when several metrics are analyzed at the same time due to the cumulative chance of getting false positives. If the tests are done at the 5%-level, there is still a 5% chance of incorrectly rejecting the null hypothesis. When instead 100 tests are done, and all corresponding null hypothesis are true it is expected to be 5 incorrect rejections. If the tests are statistically independent from each other the probability of at least incorrectly reject one null hypothesis is $1 - 0.95^{100} = 0.994$ i.e. 99.6% chance. Hence a multiple comparison is performed when several metrics are analyzed. [26]

A common multiple comparison correction is the so-called *Bonferroni correction*. The Bonferroni correction does not require any specific distribution of the data nor any specific dependences. It is however, one of the more conservative multiple comparison methods which means it have a lot of power when it neglects the null hypothesis but that is at the cost of more false negatives. The Bonferroni correction is done by dividing the threshold by the number of variables being analyzed, $\alpha_{Bon} = \frac{\alpha}{m}$ where m is the number of metrics analyzed. [26]

3 Method

This project intends to analyze brain networks for both healthy subjects and patients suffering from PLP with a graph theory approach based on EEG data to determine topological differences. EEG data have a high temporal resolution in the range of milliseconds. Functional connectivity is highly time dependent and is constantly changing[15] which makes EEG more suitable than fMRI when the characteristics of the network is to be evaluated.

3.1 Participants

Data from ten amputees with PLP were collected. The age of the amputees with PLP varied from 17 to 66 with a mean age of 47 years old, two of them were females and eight of them were males. The amputee's PLP pain ranged from zero to four on the weighted pain distribution[27] (WPD) pain scale. Data was also gathered from five healthy subjects, two of them were females and three of them were males. These subjects age ranged from 22 to 42 with a mean age of 28 years old. No history of neurological or psychiatric illness were present in the group of amputees or healthy controls. The paradigm of the experiment was explained to all participants and a written informed consent was collected prior to participating in the experiment. This study was approved by the Västra Götalandsregionens ethical committee and conformed to the ethical aspects of the Declaration of Helsinki.

3.2 Procedure

The EEG sessions started with preparation and placement of electrodes. The participants were asked to preform several different tasks during the recordings however only the resting state data were used in this study. The resting state data were collected during two different conditions, eyes-open and eyes-closed. During the eyes-open recording the participant were instructed to watch a white cross with a grey background on a computer screen to prevent the eyes from fatiguing. The screen was placed approximately 80 cm away from the participant. The recordings were at least five minutes long to ensure enough artifact free data were collected. The eyes-closed experiment was executed in a similar fashion but without the screen. During the recording the laboratory was emptied to avoid distractions. The data from resting state eyes-closed was used in the graph theory analysis.

3.3 EEG Recordings

The EEG data was recorded using Guger Technologies active electrode system, g.HIamp-RESEARCH, with 64 channels. Later in the study the system got upgraded to the 128-channel equivalent. Even though data was collected from all 128 channels only the same electrodes were used in the network analysis. The relevant electrodes was placed according to the extended 10-10 system ($FP_{Z,1,2}$, $AF_{Z,3,4,7,8}$, $F_{Z,1,2,3,4,5,6,7,8}$, $FC_{Z,1,2,3,4,5,6}$, $FT_{7,8}$, $C_{Z,1,2,3,4,5,6}$, $CP_{Z,1,2,3,4,5,6}$, $TP_{7,8}$, $P_{Z,1,2,3,4,5,6,7,8}$, $PO_{Z,3,4,7,8}$, $O_{Z,1,2}$, $A_{1,2}$, see Figure 12 for a graphical and an image of the experimental electrode setup) and applied using an elastic electrode cap, g.GAMMAcap. The recordings were done with AF_Z as the ground and no electrode was used as a reference during the recordings. However, in the 64-channel recordings A_2 was recorded and in the 128-channel recordings A_1 and A_2 was recorded to enable offline re-referencing. Four electrodes were also placed around the eyes in the 128-channel recordings to more easily detect eye-movements for artifact removal. No filters were use during the data acquisition and a sampling frequency of 2400 Hz were used. The analog to

digital converter precision was 24 bits. The electrodes used were active Ag/AgCl electrodes, g.SCARABEO.

The positions of the electrodes were measured with the Polaris Krios. A handheld digitizing scanner designed to localize electrodes or sensors for EEG, MEG, NIRS, PSG and ECG. The electrodes positions were saved and labeled in Polaris Krios software and exported.

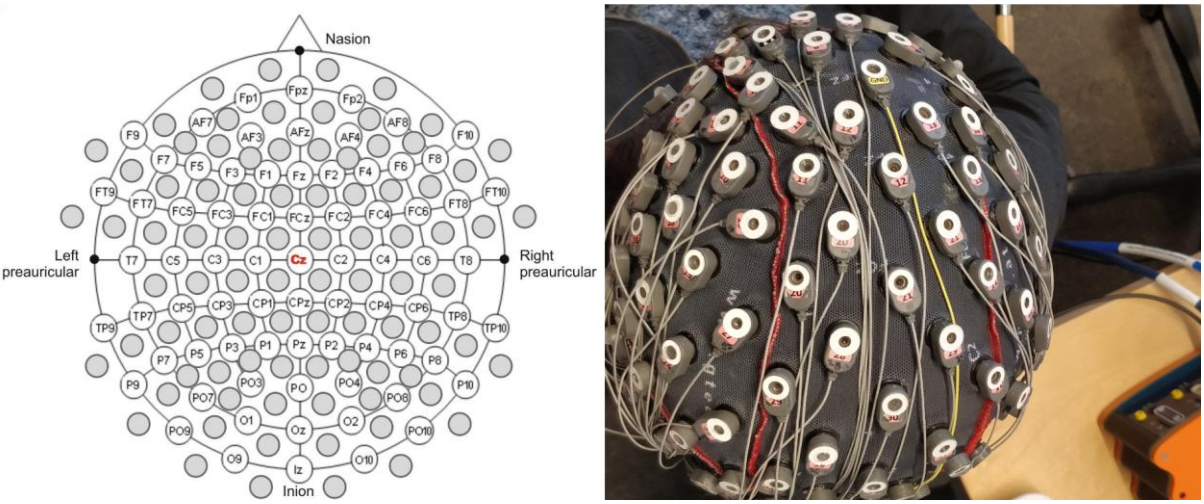


Figure 12: The figure on the left show the possible electrode positions in the g.GAMMAcap. The image on the right is the g.GAMMAcap placed on a participant where 128 electrodes are connected. Note that even though recordings where made with all 128 electrodes only 58 were used in the analysis.

3.4 EEG Preprocessing

The EEG-recordings were imported in EEGLAB v.14.1.2b[28], a MATLAB based open toolbox, together with the electrode positions. The data was band pass filtered between 0.5-90 Hz using a zero-phase FIR filter to remove linear trends and high frequency noise. The data was also notch filtered between 45-55 Hz to remove the powerline noise. The main purpose of the filtering was to improve results from ICA. The first 15 seconds of each data set was then removed to discard data heavily affected by transients from the filtering process and artifacts created from the initiation of the EEG-recording. ICA was performed to remove blinks, eye-movements and other stereotyped artifacts from the data[14]. The "runica" algorithm implemented in EEGLAB was used for improved detection of sub-Gaussian distributions[29]. The independent components related to artifacts was identified with the automatic algorithm ADJUST[30]. ADJUST uses artifact-specific spatial and temporal features that have been optimized to capture blinks, eye movements and generic discontinuities. The channels representing these artifacts was set to zero before retransformation back from independent components. ICA was applied to the full data set to ensure that the algorithm learnt all weights needed for full decomposition. ICA decomposition tends to be more regular and produce more dipolar maps with a higher amount of data points, but it is recommended to use at least $25 \cdot N^2$ data points to learn N weights[13]. In the cases where a channel is completely contaminated this channel is interpolated from the remaining channels.

Once again, the data was filtered but now into the relevant frequency bands, delta (0.5-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), beta (13-30 Hz), low gamma (30-45 Hz) and high gamma (55-90). The first 15 seconds of the recordings had to be removed in order to avoid transients from the filtering process. All filters used during the preprocessing was FIR filters of order 15840.

From the cleaned data 10 artifact free epochs were manually extracted which were 2 s long. Functional connectivity is highly time dependent and constantly changing as mentioned before hence a short time interval is preferred however if a shorter time window than 2 s was to be used frequencies lower than 0.5 Hz would not be possible to register. Therefore, to get results in the common frequency bands a shorter time window was not used. Lastly the data was re-referenced to the average of all scalp channels.

Filters of this order are very slow and creates significant transients lasting several seconds but needed when step cut-off edges are needed as in the 0.5 Hz limit. Since these transients were already present from the first cut-off edge the same order was used for the remaining filters as well. Another argument to use high-order filters is to improve the results from the PLI analysis. Since it requires band-pass filtering over the relevant frequencies steep cut-off edges could be preferable.

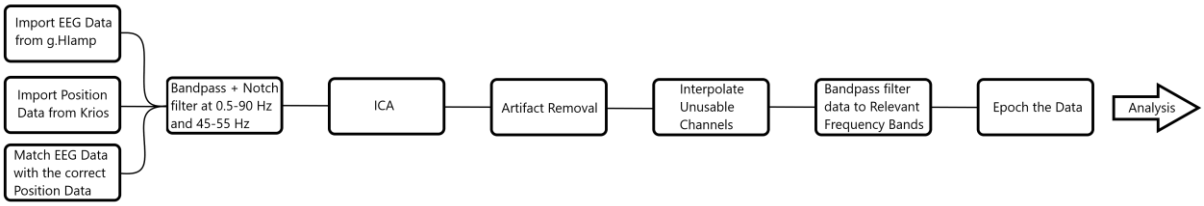


Figure 13: The Figure displays the preprocessing pipeline which describes how data from the g.HIamp becomes processed to enable analysis.

3.5 Functional Connectivity

In order to measure the connectivity between the cortical regions the functional connectivity between all pairs of electrodes were evaluated. In this project PLI was used to determine the strength of the functional connectivity between nodes, electrodes, in the network. PLI is a measure of phase synchronization designed to mitigate spurious phase synchrony resulting from common sources for example volume conduction. PLI is defined to quantify the asymmetry of the distribution of phase differences between two signals. If the phase synchrony is due common sources, then the phase differences are expected to be symmetrically distributed around zero. The calculation of PLI involves first bandpass filtering around the band of interest and then calculating the instantaneous phase angle: $\varphi(t) = arctan\frac{\tilde{x}(t)}{x}$, where $\tilde{x}(t)$ is the Hilbert transform of $x(t)$. The PLI between two signals is calculated as:

$$PLI_{xy} = \frac{1}{N}\left|\sum_{t=1}^{N} sign\left(\varphi_x(t) - \varphi_y(t)\right)\right|$$

The PLI ranged from 0 to 1 where 0 represents no synchronization or a coupling centered around 0 (mod π) and 1 represents perfect synchronization.[10]

3.6 Minimum Spanning Trees

An MST was calculated for each PLI matrix. In this project the weights are defined as $1 - PSI$ to maximize the connection strength instead of minimizing it. Even though the weights in G_w are used to create the MST, the tree itself is binary meaning an edge either do exist or doesn't. There

are numerous MST metrics that can be used to describe the topological properties of the tree[23]. In this study the metrics in Table 1. Summary over graph metrics. were examined. These metrics can be divided into two groups, nodal properties and topological properties. Nodal properties refer to metrics that describes a single node in the network while topological properties describe the structure of the network.

3.6.1 Nodal properties

The degree (D) of a node is the number of edges connected to it. The degree is a simple measurement of connectivity of a specific node as it directly reflects its number of connections.

Eccentricity (E) is the local largest distance i.e. the largest distance between a specific node and any other node. Eccentricity expresses how central a node is to the MST. A central node has a low eccentricity value.

The betweenness centrality (BC) of a given node u is the number of shortest paths between any pair of nodes, i and j, going through that node divided by the total number of paths between i and j. The BC value measures how important a node is to the flow of information between all other nodes in the network and ranges from 0 to 1. The nodes with the highest BC have the highest load in the network. Degree, eccentricity and betweenness centrality measures relative nodal importance and can be used to discern critical nodes in the network. A node with a high centrality is characterized as a network hub.

3.6.2 Topological properties

The diameter of the MST is defined as the largest distance between any two nodes in graph. The upper limit of the diameter becomes $d = m - L + 2$ where m is the number of edges in the network and L is the leaf fraction. The diameter is the graphs maximum eccentricity and hence reflects the efficiency of the network.

Leaf fraction (L) refers to the number of nodes in the MST with degree = 1. Leaf fraction quantifies whether the tree have a more chain-like structure, low leaf fraction, or more star-like structure, high leaf fraction.

Tree hierarchy (T_h) indicates how well balanced a network is with respect to efficient communication and risk of overload of hub nodes.

$$T_h = \frac{L}{2mBC_{max}}$$

T_h is proposed to be related to optimal network performance[31].

The degree correlation R indicate whether the degree of a node is related to the degree of its neighboring nodes to which it is connected. R is computed by calculating the Pearson correlation between the degrees of neighboring nodes. Networks with a negative degree correlation is often called assertive while networks with negative degree correlation is called disassortative.

Kappa (κ) is the width of R and reflect the spread of information in the network. High kappa indicates high-degree nodes which in turn leads to more synchronization in the tree. However, it also makes the network more vulnerable if such a node is damaged.

Finally, the mean of all weights in G_w is computed to extract the mean connectivity in the tree.

All metrics were calculated in MATLAB with either custom made scripts or available toolboxes. The toolboxes used were EEGLAB v.14.1.2b[28], the EEGLAB plugin ADJUST[30,32]and a function to test the normality of a data distribution by Öner et. al.[33].

Table 1. Summary over graph metrics.

Metrics		Definition
D	Degree	Number of neighbors for a given node in the MST.
L	Leaf Fraction	Fraction of nodes with degree = 1 (leaves) in the MST.
d	Diameter	Largest distance between any two nodes of the tree.
E	Eccentricity	Largest distance between a specific node and any other node in the tree.
BC	Betweenness Centrality	Normalized value representing the number of shortest paths going through a certain node.
T_h	Tree Hierarchy	Indicates the balance of the network with respect to efficient communication and risk of overload of hub nodes.
R	Degree Correlation	Topological measure which describes whether a node is prone to connect with other nodes of the same degree
κ	Kappa	Reflect the spread of information in the network. High k indicate high-degree nodes which leads to more synchronization in the tree
	Mean Connectivity	The mean of the connectivity weights used to create the MST.

3.7 Statistical Comparison

The null hypothesis used in the statistical analysis was that the mean of the metrics data distributions was equal while the alternative hypothesis was the opposite, that they were unequal. The comparison was made between participants with PLP and healthy controls without PLP. The evaluation was made using the *Welch's t-test* or the *Wilcoxon rank sum test*. This was because the data acquired were independent, the variance of the two sample groups was unequal, and the two sample sizes were unequal. If the data were normally distributed Welch's t-test was adopted and if not, Wilcoxon's rank sum test was applied instead. To test the normality of the data, the *D'Agostino & Pearson test* was used.[34]

This way a table over the mean of each metrics and its corresponding *p* and test value, *t* or *rank sum*, was computed.[35,36] Multiple comparison corrections were done both as the Bonferroni correction and FDR-correction. The multiple comparison correction is not presented in section 4, *Result,* but is discussed in section 5, *Discussion.*

4 Result

Before the metrics were compared the data-distribution was examined. The result from the normal distribution analysis is presented in Table 2. A one indicates that the data passed the D'Agostino & Pearson test while a zero indicates that the data is not normal distributed. Several other tests were also used to ensure that not only the skewness and kurtosis indicate normal distribution. The result from all 10 tests examined can be found in Appendix A.

Table 2: The table show which of the metrics in the different frequency bands passed the D'Agostino & Pearson test for normality. A one in the table indicates that the data is normal distributed while a zero indicate that it is not.

Frequency band / Metric	Delta	Theta	Alpha	Beta	Low-Gamma	High-Gamma
$std\ of\ D$	0	0	1	0	0	0
L	0	0	1	1	0	0
d	1	1	0	1	1	1
E	1	1	0	1	1	1
BC	1	1	0	1	1	0
T_h	1	1	1	1	0	0
R	1	1	0	0	1	1
κ	0	0	0	0	0	0
Mean PLI	0	1	0	0	0	0

The results of the statistical comparison between the groups, control without pain and amputee with PLP, is presented in Table 3 to Table 8. The section is divided into the different frequency bands analyzed and all results are presented with 3 significant digits.

4.1 Delta 0.5-4 Hz

Below, the result of the statistical comparison is presented for the delta frequency band. The metrics standard deviation of the degree, the leaf fraction, the diameter, the eccentricity, the betweenness centrality, the tree hierarchy and kappa were considered statistically different between the groups at a threshold of $p = 0.05$. Any metric marked as bold is considered statistically significant. Note that this threshold is not Bonferroni corrected.

Table 3: The table shows the mean and variance of the metrics of the network. It also shows how the result of the statistical analysis between the groups. Any metric marked as bold is considered statistically significant at the 0.05 threshold. Note that this threshold is not Bonferroni corrected.

	Control		Amputee (PLP)		Group Comparison		
	M	STD	M	STD	t	Rank Sum	p
$std\ of\ D$	**1.42**	**0.27**	**1.56**	**0.330**	-	**2900**	**0.00669**
L	**28.8**	**3.44**	**30.2**	**3.24**	-	**3020**	**0.0283**
d	**16.7**	**2.87**	**15.5**	**2.74**	**2.36**	-	**0.0201**
E	**13.1**	**2.13**	**12.2**	**2.06**	**2.60**	-	**0.0107**
BC	**0.0521**	**0.00828**	**0.0477**	**0.00787**	**3.09**	-	**0.00264**
T_h	**0.794**	**0.0879**	**0.836**	**0.0905**	**-2.70**	-	**0.00817**
R	-0.296	0.0937	-0.293	0.0866	-0.210	-	0.834
κ	**7.92**	**2.35**	**9.01**	**3.08**	-	**3070**	**0.0467**
Mean PLI	0.370	0.0542	0.364	0.0650	-	3550	0.908

4.2 Theta 4-8 Hz

Below, the result of the statistical comparison is presented for the theta frequency band. No metrics were considered statistically different between the groups at a threshold of $p = 0.05$. Any metric marked as bold is considered statistically significant. Note that this threshold is not Bonferroni corrected.

Table 4: The table shows the mean and variance of the metrics of the network. It also shows how the result of the statistical analysis between the groups. Any metric marked as bold is considered statistically significant at the 0.05 threshold. Note that this threshold is not Bonferroni corrected.

	Control		***Amputee (PLP)***		***Group Comparison***		
	M	**STD**	**M**	**STD**	***t***	***Rank Sum***	***p***
$std\ of\ D$	1.14	0.180	1.12	0.172	-	3600	0.762
L	23.8	3.03	24.0	2.98	-	3540	0.956
d	21.5	3.44	20.9	3.41	0.937	-	0.351
E	16.7	2.53	16.3	2.53	0.705	-	0.482
BC	0.0652	0.00846	0.0635	0.00862	0.505	-	0.614
T_h	0.703	0.106	0.701	0.0981	0.109	-	0.913
R	-0.258	0.109	-0.255	0.0980	-0.138	-	0.890
κ	6.28	1.51	6.19	1.48	-	3630	0.639
$Mean\ PLI$	0.401	0.0280	0.401	0.0319	-0.0462	-	0.963

4.3 Alpha 8-12 Hz

Below, the result of the statistical comparison is presented for the alpha frequency band. The metrics standard deviation of the degree and the leaf fraction were considered statistically different between the groups at a threshold of $p = 0.05$. Any metric marked as bold is considered statistically significant. Note that this threshold is not Bonferroni corrected.

Table 5: The table shows the mean and variance of the metrics of the network. It also shows how the result of the statistical analysis between the groups. Any metric marked as bold is considered statistically significant at the 0.05 threshold. Note that this threshold is not Bonferroni corrected.

	Control		***Amputee (PLP)***		***Group Comparison***		
	M	**STD**	**M**	**STD**	***t***	***Rank Sum***	***p***
$\boldsymbol{std\ of\ D}$	**1.25**	**0.180**	**1.32**	**0.211**	**-2.06**	**-**	**0.0418**
$\boldsymbol{L}$	**25.96**	**3.18**	**27.2**	**3.34**	**-2.19**	**-**	**0.0308**
d	20.2	3.58	19.4	3.63	-	3830	0.189
E	15.9	2.77	15.2	2.63	-	3830	0.189
BC	0.0629	0.0101	0.600	0.00931	-	3900	0.107
T_h	0.773	0.105	0.810	0.115	-1.89	-	0.0608
R	-0.286	0.0732	-0.297	0.0942	-	3680	0.507
κ	7.40	1.87	7.74	2.05	-	3300	0.316
$Mean\ PLI$	0.320	0.0811	0.297	0.0519	-	3880	0.123

4.4 Beta 12-30 Hz

Below, the result of the statistical comparison is presented for the beta frequency band. The mean PLI was considered statistically different between the groups at a threshold of $p = 0.05$. Any metric marked as bold is considered statistically significant. Note that this threshold is not Bonferroni corrected

Table 6: The table shows the mean and variance of the metrics of the network. It also shows how the result of the statistical analysis between the groups. Any metric marked as bold is considered statistically significant at the 0.05 threshold. Note that this threshold is not Bonferroni corrected.

	Control		***Amputee (PLP)***		***Group Comparison***		
	M	**STD**	**M**	**STD**	***t***	***Rank Sum***	***p***
$std\ of\ D$	1.06	0.163	1.07	0.150	-	3410	0.620
L	21.58	2.76	21.6	2.93	0.00447	-	0.996
d	23.18	2.95	23.7	4.05	-0.852	-	0.396
E	17.9	2.12	18.3	2.93	-0.972	-	0.333
BC	0.0693	0.00746	0.0702	0.00950	-0.618	-	0.538
T_h	0.621	0.102	0.619	0.101	0.114	-	0.910
R	-0.225	0.0907	-0.209	0.020	-	3330	0.412
κ	6.10	1.54	6.3	1.38	-	3320	0.352
Mean PLI	**0.441**	**0.0360**	**0.424**	**0.0358**	**-**	**4300**	**0.000831**

4.5 Low-Gamma 30-45 Hz

Below, the result of the statistical comparison is presented for the delta frequency band. No metrics were considered statistically different between the groups at a threshold of $p = 0.05$. Any metric marked as bold is considered statistically significant. Note that this threshold is not Bonferroni corrected

Table 7: The table shows the mean and variance of the metrics of the network. It also shows how the result of the statistical analysis between the groups. Any metric marked as bold is considered statistically significant at the 0.05 threshold. Note that this threshold is not Bonferroni corrected.

	Control		***Amputee (PLP)***		***Group Comparison***		
	M	**STD**	**M**	**STD**	***t***	***Rank Sum***	***p***
$std\ of\ D$	1.06	0.172	1.06	0.202	-	3680	0.491
L	22.5	2.79	22.2	3.29	-	3740	0.349
d	22.0	2.68	21.0	3.28	1.86	-	0.0650
E	16.8	1.97	16.3	2.30	1.50	-	0.136
BC	0.0651	0.00774	0.0644	0.00764	0.458	-	0.648
T_h	0.651	0.0940	0.648	0.109	-	3620	0.694
R	-0.234	0.0983	-0.215	0.113	-1.11	-	0.271
κ	5.82	1.56	5.86	1.75	-	3530	0.968
$Mean\ PLI$	0.535	0.0237	0.524	0.0292	-	3620	0.694

4.6 High-Gamma 55-90 Hz

Below, the result of the statistical comparison is presented for the delta frequency band. No metrics were considered statistically different between the groups at a threshold of $p = 0.05$. Any metric marked as bold is considered statistically significant. Note that this threshold is not Bonferroni corrected

Table 8: The table shows the mean and variance of the metrics of the network. It also shows how the result of the statistical analysis between the groups. Any metric marked as bold is considered statistically significant at the 0.05 threshold. Note that this threshold is not Bonferroni corrected.

	Control		***Amputee (PLP)***		***Group Comparison***		
	M	**STD**	**M**	**STD**	***t***	***Rank Sum***	***p***
$std\ of\ D$	1.18	0.226	1.24	0.358	-	3560	0.891
L	24.16	3.83	24.0	5.30	-	3730	0.363
d	20.5	3.12	20.0	3.63	0.784	-	0.436
E	15.7	2.32	15.4	2.71	0.659	-	0.511
BC	0.0607	0.00767	0.604	0.0100	-	3710	0.427
T_h	0.680	0.107	0.685	0.168	-	3740	0.352
R	-0.270	0.155	-0.242	0.160	-1.02	-	0.309
κ	6.64	1.96	7.19	2.88	-	3440	0.709
$Mean\ PLI$	0.571	0.0312	0.570	0.0479	-	3500	0.929

5 Discussion

The aim of the project was to develop a functional system to create brain networks and evaluate which metrics could possibly be used as a bio marker for PLP. In this section the methods and results will be discussed with that as a basis.

5.1 Method

This section discusses what was taken into consideration when the methods for this project was established. It also discusses the possible changes that could be done in the future if a similar project or an extension of this study is to be made.

5.1.1 Data Collection and Processing

In the

Method section it was explained why EEG was preferred over other techniques such as fMRI and MEG. It can also be argued that EEG is cheaper and a more accessible method which could become an important factor if the research is continued. In the following subsection it is also discussed how handedness corrections could be made to lessen the differences between right- and lefthanded subjects.

The filters in this study have an extremely high order. If the method is to be applied for real-time recordings such high rank filters are most likely not needed. The first band-pass filter and notch-filter could be applied for the data at the same time with similar rank as used now (that's being done in current software without problems) while small epochs get filtered into respective epochs by filters with a rank of around ten, for further analysis.

The epoch selection could also be automated. There are a few functions in EEGLAB that enables automatic artifact detection. However, EEG is very subjective at times and currently there is no algorithms that do a better job than what can be done by simple observing the data, especially if the observer is experianced[11]. The amount of data in this study was not excessive to the point of requiring automation. Even though the author had little experience, with 7-minute-long recordings where a total of 20 seconds of data was extracted, artifact free epochs were rather easy to select.

In the end eyes-closed resting state recordings were used to compute the graph theoretical metrics. There are no theoretical proofs that eyes-closed data would yield better result than eyes-open. However, as mentioned in the Theory section "garbage in, garbage out". The choice was made by simply observing which of the two types contained the least number of artifacts and was the least contaminated by unwanted noise.

5.1.2 Participants

In this project subjects without PLP are compared with subjects with PLP. Among the control subjects there was one participant who was an amputee but did not suffer from PLP. In a comparison study it is usually desired to only study one change between the groups. Hence at the beginning of the project three groups were considered healthy controls, amputees without PLP and amputees with PLP. A comparison between these three groups would have made a more powerful statement about the network changes in specifically PLP as one could claim that the observed changes originates from the PLP itself and not for example the amputation. It is also important to note that such a study might not be feasible due to the lack of possible participants. Today it is hard to recruit amputees with PLP during the duration of a 6-month project. It is even harder to recruit amputees without PLP. This could be because of a lack of interest when the amputee is healthy but more likely because the sheer lack of amputees without PLP.

Another factor that could be eliminated is handedness. Handedness is known to modulate neural responses in the brain[37]. Hence to really change as few attributes as possible between and in the groups all subjects should have their right hand as their dominant one and the amputation should have been done on said hand. Once again due to the lack of available subjects this might not be feasible. In total there are 4827 lower limb amputees in Sweden[38]. What could be done instead is mirror the results to a predetermined handedness, i.e. all left side amputee's EEG-data gets mirrored to the right side after the EEG recording.

5.1.3 Network study

To begin with, the choice to make a global network study, i.e. the entire brain is analyzed, was based on successful studies of other diseases such as dyslexia[29,39] and epilepsy[40,41] among others[42,43,44]. Several studies have also been made on functional connectivity in PLP where they claim to have found a network-level reorganization of the human brain[3,45]. These studies have however not fully explored the entirety of the human brain network but have explored certain regions of interest and their connections, such as the somatosensory cortex and the default mode network. Hence an exploratory study could provide new findings.

5.1.4 Functional Connectivity

Functional connectivity can be measured by several different metrics. The choice of metrics can affect the outcome of a study if it is not well adjusted to the data or experiment. In this project PLI was used. What differs the PLI from other phase-based measure of connectivity is that it discards zero-phase lag. In an EEG study there is bound to be false connectivity between electrodes due to volume conduction or measuring of the same signal. This type of correlation is discarded with the PLI. However, this is not the only kind of signals which could have a zero-phase lag. There could be zero-phase lag between neural signals as well which would not contribute to the analysis when PLI is used. Another type of connectivity metrics broadly used are those based on amplitude rather than on phase. Whether or not to use phase-based or amplitude-based measures of connectivity depends on the application. Phase-based methods are more appropriate to tests hypothesis where phase and moment-by-moment changes in synchronization are considered, which neural communication is at least in the higher frequency bands[46]. An example of this is provided by a study by Kuntzelmam et. al.[47] where they show that the coherence measure, which is based on amplitude, of functional connectivity was better suited for the slower frequency bands, delta and theta, while PLI, which is phase-based, produced more reliability in the faster frequency bands, alpha and beta. The same study also concluded that neither metric produced reliable results in the gamma frequency band.

In a study by Harmeier et. al.[48] the reproducibility of the PLI and the weighted phase lag index (wPLI) were tested on high-resolution EEG-recordings and both were considered to have good long-term test-retest-reliability. This is a very important feature if these methods are to one day be used to identify potential biomarkers for PLP, and another argument why to use the PLI or the wPLI.

What could be considered is using wPLI instead of the original PLI. wPLI is an extension to PLI that accounts for the magnitude of the phase difference as well. It has been reported to be less sensitive to noise[48] but is not as well established as a method yet.

5.1.5 Graph Theory

There is a lot that could be said about the application of graph theory. As mentioned before there are several studies where graph theory and network studies have been a success and biomarkers among other things have been identified for different diseases. However, this doesn't imply that there will always be network differences to be found that can be credited the researched condition, especially when not all possible metrics are evaluated.

One could compute all known metrics in graph theory and search for a statistically significant difference between them. The first problem that occurs is that it is very time consuming. All metrics need to be found and then implemented correctly. Secondly, the more metrics that are

being examined the higher the probability that a special case occurs which would result in false results. In this project metrics was chosen to cover specific characteristics of a network such as centrality, load, efficiency and topology. Since previous studies[3,45] had shown significant differences between healthy controls and amputees with PLP in the mean functional connectivity of the networks it was also considered a metric.

5.1.6 Statistical Comparison

Statistical comparisons lose more of its power the more metrics are examined. This is because it increases the likelihood of encountering a special case among the collected data. Hence, it is important to correct the results of the analyses for multiple comparisons. In every frequency band, nine different metrics are examined and in total there are six frequency bands. If a desired threshold is at $p = 0.05$ before multiple comparison the Bonferroni correction would lie at $p = \frac{0.05}{9 \cdot 6} \approx 9.3 \cdot 10^{-4}$. This threshold is adapted to neglect the possible false positive, type I error, finds due to examining several metrics. However, it decreases the amount of type I errors at the cost of more type II errors, false negatives. In an exploratory study such as this one this might be an over correction.

In exploratory studies it is common, since a lot of metrics are computed, to instead compute the false discovery rate (FDR)[49]. The FDR does instead compute the amount of "false discoveries" we can expect based on the computed p-values and threshold. A q-value threshold is then determined which represents the acceptable proportion of features that turn out to be false leads. In this study the q-threshold was determined to be 0.05 which means that 95% of the discoveries are probably true. The q-value was computed for all metrics and the results of the statistically significant metrics are presented in Table 9:

Table 9:The table show the previous computed p-value, their corrected FDR p-value and the corresponding q-value.

Metric	p	***FDR-corrected p***	q
$std\ of\ D$ *(delta)*	0.00669	0.0956	0.0876
L *(delta)*	0.0283	0.173	0.165
d *(delta)*	0.0201	0.144	0.144
E *(delta)*	0.0107	0.0918	0.0918
BC *(delta)*	0.00264	0.0566	0.0566
T_h *(delta)*	0.00817	0.0876	0.0876
κ *(delta)*	0.0467	0.200	0.200
$std\ of\ D$ *(alpha)*	0.0418	0.199	0.199
L *(alpha)*	0.0308	0.165	0.165
Mean PLI (beta)	0.000831	0.0356	0.0356

The goal of this study was never to determine exactly what metric would be most suitable as a biomarker in PLP. Instead, the goal was to explore the possibility that graph theory can be used to detect network differences and where to continue the exploration. The FDR was computed using the FDR-function in MATLAB which employs the FDR-procedure introduced by Storey[49].

5.2 Results

This section discusses the results from the graph theory application on the functional connectivity network. The discussion is divided into the different frequency bands, the same way they were previously presented.

5.2.1 Delta 0.5-4 Hz

The delta frequency band was the frequency span which included most, seven out of ten, of the statistically significant metrics without taking either the Bonferroni correction or FDR correction into account.

The Bonferroni correction would demand a p-value of below $9.3 \cdot 10^{-4}$ to classify something as statistically significant. With this corrected p-value none of the metrics would be considered statistically significant. The same problem occurs after the FDR-correction as well. Not because they are all considered to be false positive findings but rather that their FDR-corrected p- value falls just short of the 0.05 threshold decided upon in the methods.

The most promising metric in this frequency band in Bonferroni correction and FDR correction aspect were the betweenness centrality which could be an indication of being the most interesting metric to further investigate.

5.2.2 Theta 4-8Hz

The theta frequency band showed no statistically significant metric. None of the metrics showed even a slight trend towards being different between the groups.

This is not a total loss, no results are also results. This indicates that the theta band might not be suitable to be used to identify PLP. This would mean for future studies that the theta frequency band could be left out and lesser amounts of data can be analyzed leaving a more forgiving multiple comparison correction. It is important though to not interpret this as there is no

difference in network differences in the theta frequency band between amputees with PLP and controls. Just as type I errors are more likely to occur when more data is examined so is type II errors. There are new papers that relate chronic pain with a change of neural oscillations in the theta frequency band[50], this study however failed to find the same changes.

5.2.3 Alpha 8-12 Hz and Beta 12-30 Hz

Noxious stimuli have previously been correlated with alpha and beta oscillations in sensorimotor areas[51]. Hence, it is not surprising to find statistically significant metrics in these frequency bands.

However, the standard deviation of the degree and the leaf fraction do not remain statistically significant after either the Bonferroni correction or the FDR-correction. The mean PLI on the other hand is the only metric whose p-value is below the Bonferroni corrected threshold in this study. The FDR analysis comes to the same conclusion that the mean PLI most likely is a metric in the alpha frequency band that is significantly different between the two groups and is unlikely due to multiple comparisons.

The reason why just the mean PLI is different between the groups could be because of the importance pain. When you hurt yourself, the natural reaction is too quickly do something to relieve the pain. In the case of PLP it might be the same, the brain focuses on the pain and other functions and task becomes less important and becomes dulled which result in a lower connectivity.

5.2.4 Low-Gamma 30-45 Hz and High-Gamma 55-90 Hz

The gamma frequency band showed no statistically significant metric. The gamma frequency band have in several studies shown a correlation with chronic pain[50,51,52] and was thus a frequency band with high expectations. Further research showed however that graph theoretical measure become less reliable for such high frequencies, independent of whether power or phase is used to evaluate the functional connectivity[47]. Other papers also show that EEG-data in such high frequencies may be heavily contaminated by EMG noise[53]. Thus, the lack of results in the gamma frequencies might be justified.

5.3 Future Work

As mentioned several times, this project was an exploratory project where graph theory application to EEG data was tested as a promising way to discern controls from amputees without PLP. All in all, it has been shown that several metrics manages to capture the differences between the groups but far from enough proof is available to claim either of them as a biomarker for PLP.

The results of this project show a clear trend of more statistically different results in the lower frequency bands, especially in the delta frequencies. These were however not powerful enough to still show these results after a multiple comparison correction had been applied. However due to the exploratory nature of this project the delta frequency band shows the most promise for future study. Because there was not only a single metric that indicated a network difference between the groups but instead, most of them did. This could be an indication of a severe network differences between the groups in this frequency band.

To further investigate the delta frequency band several adjustments could be done to adapt the methods to this frequency band. As mentioned before the choice of metric to evaluate the functional connectivity depends on the expected frequency of the data. In this study PLI was chosen above other metrics such as coherence due it being more reliable on higher frequencies.

If the delta frequency band is the focus of the study coherence would most likely be the better option.

The MST have shown to be able to capture network differences between the two groups of subjects. Hence future studies should expect so see similar results. One could try to evaluate the same or equal metrics with more classical graph theory to determine which one is superior. However, the bias problems mentioned in the Theory section would still be a factor that have to be dealt with.

The participants of this study are quite diverse, and little is done to close that gap. In a future study a method to properly mirror the data in cases where the amputation is done on different sides of the body could be implemented. One could also try to recruit healthy volunteers to match the amputees in gender, handedness and/or age since these things could affect graph theoretical measures[25,31,35]. Or better yet, create the control group from amputees without PLP, this might however not be feasible.

One aspect that is not taken into consideration in this project is the intensity of PLP in the subjects. This is another parameter which could create a spectrum of differences in the graph theoretical metrics. Possible pain scales could be the Weighted Pain Distribution introduced by Max Ortiz-Catalan[27] or the more common pain rating index formed by the summed contribution of 15 qualities of pain as in the short-form McGill Pain Questionnaire[54].

6 Conclusion

Th results from this study confirms that it is possible to identify network differences between subjects with PLP and subjects without PLP. Most of these differences were found in the delta frequency band but a few were also found in the alpha and beta frequency band. Even though differences were found that were considered statistically significant at the 0.05 threshold it is important to notice that all but one metrics, the mean PLI in the beta frequency band, were not considered statistically significant after either Bonferroni correction or FDR-correction. Even though more statistically significant results are desired as an exploratory study this should still be considered a success with respect to the aims of the project.

7 References

1. Flor H, Andoh J. Origin of phantom limb pain: A dynamic network perspective. *Neuroforum*. 2017;23(3):111-116. doi:10.1515/nf-2017-A018

2. Flor H, Elbert T, Knecht S, et al. Phantom-limb pain as a perceptual correlate of cortical reorganization following arm amputation. *Nature*. 1995;375(6531):482-484. doi:10.1038/375482a0

3. Makin TR, Filippini N, Duff EP, Henderson Slater D, Tracey I, Johansen-Berg H. Network-level reorganisation of functional connectivity following arm amputation. *Neuroimage*. 2015;114:217-225. doi:10.1016/j.neuroimage.2015.02.067

4. Ortiz-Catalan M. The stochastic entanglement and phantom motor execution hypotheses: A theoretical framework for the origin and treatment of Phantom limb pain. *Front Neurol*. 2018;9(SEP):1-16. doi:10.3389/fneur.2018.00748

5. Sporns O. Terzahi_1956_Conference_Submarine Landslides.pdf. 2013:247-262. doi:10.1137/S003614450342480

6. Ploner M, May ES. Electroencephalography and magnetoencephalography in pain research - Current state and future perspectives. *Pain*. 2018;159(2):206-211. doi:10.1097/j.pain.0000000000001087

7. Nuwer MR. 10-10 electrode system for EEG recording. *Clin Neurophysiol*. 2018;129(5):1103. doi:10.1016/j.clinph.2018.01.065

8. Oostenveld R, Praamstra P. The five percentage electrode system for high-resolution EEG and ERP measurements. *Clin Neurophysiol*. 2001;112:713-719. papers3://publication/uuid/F6B21DA5-3642-4EBE-9253-C772263C900D.

9. Asanagi. 21 electrodes of International 10-20 system for EEG.

10. Chang-Hwan I. *Computational EEG Analysis, Methods and Applications.*; 2018.

11. Cohen MX. *Analyzing Neural Time Series Data*. Vol 53.; 2013. doi:10.1007/s13398-014-0173-7.2

12. Sörnmo L, Laguna P. *Bioelectrical Signal Processing in Cardiac and Neurological Applications*. Elsevier; 2005. https://app.knovel.com/hotlink/toc/id:kpBSPCNA02/bioelectrical-signal/bioelectrical-signal.

13. Onton J, Makeig S. Information-based modeling of event-related brain dynamics. *Prog Brain Res*. 2006;159:99-120. doi:10.1016/S0079-6123(06)59007-7

14. Makeig S, Jung T, Bell AJ, et al. Blind Separation of Auditory Event-Related Brain Responses into Independent Components. 1997;94(20):10979-10984.

15. Sporns O. Structures and fucntions of complex brain networks. 2013:247-262. doi:10.1137/S003614450342480

16. Rubinov M, Sporns O. Complex network measures of brain connectivity: Uses and interpretations. *Neuroimage*. 2010;52(3):1059-1069. doi:10.1016/j.neuroimage.2009.10.003

17. Sporns O. Brain connectivity. *Scholarpedia*. 2007;2(10):4695. doi:10.4249/scholarpedia.4695

18. Nica B. *A Brief Introduction to Spectral Graph Theory*.; 2016. doi:10.4171/188

19. Trudeau RJ. Introduction to Graph THeory. 1999.

20. Rahman MS. *Basic Graph Theory*.; 2017. doi:10.1007/978-3-319-49475-3

21. Hartsfield, Nora. Ringel G. Pearls in Graph Theory. 2003.

22. Karin R. Saoub. *A Tour Through Graph Theory*.; 2018.

23. Stam CJ, Tewarie P, Van Dellen E, van Straaten ECW, Hillebrand A, Van Mieghem P. The trees and the forest: Characterization of complex brain networks with minimum spanning trees. *Int J Psychophysiol*. 2014;92(3):129-138. doi:10.1016/j.ijpsycho.2014.04.001

24. van Wijk BCM, Stam CJ, Daffertshofer A. Comparing brain networks of different size and connectivity density using graph theory. *PLoS One*. 2010;5(10). doi:10.1371/journal.pone.0013701

25. Petti M, Toppi J, Babiloni F, Cincotti F, Mattia D, Astolfi L. EEG Resting-State Brain Topological Reorganization as a Function of Age. *Comput Intell Neurosci*. 2016;2016. doi:10.1155/2016/6243694

26. Rutherford A. An Introduction to General Linear Models: Regression, Analysis of Variance, and Analysis of Covariance. *Anova and Ancova*. 2013;(2011):1-15. doi:10.1002/9781118491683.ch1

27. Ortiz-Catalan M, Guðmundsdóttir RA, Kristoffersen MB, et al. Phantom motor execution facilitated by machine learning and augmented reality as treatment for Phantom Limb Pain. *Lancet*. 2016;388(10062):2885-2894. doi:10.1016/S0140-6736(16)31598-7

28. Delorme A, Makeig S. EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods*. 2004;134(1):9-21. doi:10.1016/j.jneumeth.2003.10.009

29. Fraga González G, Smit DJA, van der Molen MJW, et al. EEG Resting State Functional Connectivity in Adult Dyslexics Using Phase Lag Index and Graph Analysis. *Front Hum Neurosci*. 2018;12(August):1-12. doi:10.3389/fnhum.2018.00341

30. Mognon A, Jovicich J, Bruzzone L, Buiatti M. ADJUST: An automatic EEG artifact detector based on the joint use of spatial and temporal features. *Psychophysiology*. 2011;48(2):229-240. doi:10.1111/j.1469-8986.2010.01061.x

31. Boomsma DI, De Geus EJC, Stam CJ, Smit DJA, Delemarre-van de Waal HA, Boersma M. Growing Trees in Child Brains: Graph Theoretical Analysis of Electroencephalography-Derived Minimum Spanning Tree in 5- and 7-Year-Old Children Reflects Brain Maturation. *Brain Connect*. 2012;3(1):50-60. doi:10.1089/brain.2012.0106

32. Mognon A, Buiatti M. ADJUST: An Automatic EEG artifact Detector based on the Joint Use of Spatial and Temporal features: A Tutorial. 2014.

33. Öner M, Deveci Kocakoç İ. JMASM 49: A Compilation of Some Popular Goodness of Fit Tests for Normal Distribution: Their Algorithms and MATLAB Codes (MATLAB). *J Mod Appl Stat Methods*. 2017;16(2):547-575. doi:10.22237/jmasm/1509496200

34. Fay MP, Proschan MA. Wilcoxon-Mann-whitney or t-test? *Stat Surv*. 2010;4:1-37. doi:10.1214/09-SS051.Wilcoxon-Mann-Whitney

35. M. M, V. B. Comparing groups for statistical differences: How to choose the right statistical test? *Biochem Medica*. 2010;20(1):15-32. doi:10.11613/BM.2010.004

36. Keselman HJ, Algina J. Student's t Test. *Encycl Res Des*. 2010:1462-1469.

37. Grey S, Tanner D, van Hell JG. How right is left? Handedness modulates neural responses during morphosyntactic processing. *Brain Res*. 2017;1669:27-43. doi:10.1016/j.brainres.2017.05.024

38. SwedeAmp. 2017 Årsrapport, Amputations and Protesregister. 2017:1-50.

39. Fraga González G, Van der Molen MJW, Žarić G, et al. Graph analysis of EEG resting state functional networks in dyslexic readers. *Clin Neurophysiol*. 2016;127(9):3165-3175. doi:10.1016/j.clinph.2016.06.023

40. Sargolzaei S, Cabrerizo M, Goryawala M, Eddin AS, Adjouadi M. Scalp EEG brain functional connectivity networks in pediatric epilepsy. *Comput Biol Med*. 2015;56:158-166. doi:10.1016/j.compbiomed.2014.10.018

41. Quraan MA, McCormick C, Cohn M, Valiante TA, McAndrews MP. Altered Resting State Brain Dynamics in Temporal Lobe Epilepsy Can Be Observed in Spectral Power, Functional Connectivity and Graph Theory Metrics. *PLoS One*. 2013;8(7). doi:10.1371/journal.pone.0068609

42. Guye M, Bettus G, Bartolomei F, Cozzone PJ. Graph theoretical analysis of structural and functional connectivity MRI in normal and pathological brain networks. *Magn Reson Mater Physics, Biol Med*. 2010;23(5-6):409-421. doi:10.1007/s10334-010-0205-z

43. Franciotti R, Falasca NW, Arnaldi D, et al. Cortical Network Topology in Prodromal and Mild Dementia Due to Alzheimer's Disease: Graph Theory Applied to Resting State EEG. *Brain Topogr*. 2019;32(1):127-141. doi:10.1007/s10548-018-0674-3

44. Xing M, Tadayonnejad R, MacNamara A, et al. Resting-state theta band connectivity and graph analysis in generalized social anxiety disorder. *NeuroImage Clin*. 2017;13:24-32. doi:10.1016/j.nicl.2016.11.009

45. Makin TR, Scholz J, Filippini N, Henderson Slater D, Tracey I, Johansen-Berg H. Phantom pain is associated with preserved structure and function in the former hand area. *Nat Commun*. 2013;4:1570-1578. doi:10.1038/ncomms2571

46. Bastos AM, Schoffelen J-M. A Tutorial Review of Functional Connectivity Analysis Methods and Their Interpretational Pitfalls. *Front Syst Neurosci*. 2016;9(January):1-23. doi:10.3389/fnsys.2015.00175

47. Kuntzelman K, Miskovic V. Reliability of graph metrics derived from resting-state human EEG. *Psychophysiology*. 2017;54(1):51-61. doi:10.1111/psyp.12600

48. Hardmeier M, Hatz F, Bousleiman H, Schindler C, Stam CJ, Fuhr P. Reproducibility of functional connectivity and graph measures based on the phase lag index (PLI) and weighted phase lag index (wPLI) derived from high resolution EEG. *PLoS One*. 2014;9(10). doi:10.1371/journal.pone.0108648

49. Storey JD. A direct approach to false discovery rates. 2002:479-498. http://arxiv.org/abs/1804.08222.

50. Dinh ST, Nickel MM, Tiemann L, May ES, Heitmann H. Brain dysfunction in chronic pain patients assessed by resting-state electroencephalography. 2019:1-46.

51. May ES, Nickel MM, Ta Dinh S, et al. Prefrontal gamma oscillations reflect ongoing pain intensity in chronic back pain patients. *Hum Brain Mapp*. 2019;40(1):293-305. doi:10.1002/hbm.24373

52. Zhou R, Wang J, Qi W, et al. Elevated Resting State Gamma Oscillatory Activities in

Electroencephalogram of Patients With Post-herpetic Neuralgia. *Front Neurosci*. 2018;12(October):1-10. doi:10.3389/fnins.2018.00750

53. Whitham EM, Pope KJ, Fitzgibbon SP, et al. Scalp electrical recording during paralysis: Quantitative evidence that EEG frequencies above 20 Hz are contaminated by EMG. *Clin Neurophysiol*. 2007;118(8):1877-1888. doi:10.1016/j.clinph.2007.04.027

54. Melzack R. The short-form McGill Pain Questionnaire. 1987;30:191-197.

Appendix

A. Tables over Normal Distribution Tests

The following tables indicate which of the 10 normal distribution tests each metric passed in the different frequency bands. It is to be noted that the test defined in this reported to be the deciding factor was the D'Agostino & Pearson test. Each table takes up about one page of space and for easy of reading is hence dedicated a full page each.

Metric / Normality Test	Control Data									Phantom Limb Pain Data								
	D	**L**	**d**	**E**	**BC**	**T_h**	**R**	**κ**	***PLI***	**D**	**L**	**d**	**E**	**BC**	**T_h**	**R**	**κ**	***PLI***
KS Limiting Form	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1
KS Stephens Modification	0	0	1	1	1	1	1	0	0	0	0	0	1	1	1	1	0	0
KS Marsaglia Method	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1
KS Lilliefors Modification	0	0	1	1	1	1	1	0	0	0	0	0	1	1	1	1	0	0
Anderson-Darling Test	0	1	1	1	1	1	1	0	0	0	0	0	1	1	1	1	0	0
Cramer-Von Mises Test	0	1	1	1	1	1	1	0	1	0	0	0	1	1	1	1	0	0
Sharpio-Wilk Test	1	1	1	1	1	1	1	0	0	0	0	1	1	1	1	1	0	0
Sharpio-Francia Test	1	1	1	1	1	1	1	0	0	0	0	1	1	1	1	1	0	0
Jarque-Bera Test	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	0	0
D`Agostino & Pearson Test	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**0**	**0**	**0**	**1**	**1**	**1**	**1**	**1**	**0**	**0**

Table 10: The table show which normal distribution tests each metric passed. This table is over the metrics computed from the delta frequency band, 0.5-4 Hz. A 1 in the table indicate that the test is passed at the 0.05 threshold and a 0 indicate that the test was failed.

Normality Test \ Metric	Control Data									Phantom Limb Pain Data								
	D	**L**	**d**	**E**	**BC**	**T_h**	**R**	**κ**	***PLI***	**D**	**L**	**d**	**E**	**BC**	**T_h**	**R**	**κ**	***PLI***
KS Limiting Form	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1
KS Stephens Modification	1	1	1	0	1	1	0	0	1	0	0	1	1	1	0	1	0	1
KS Marsaglia Method	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1
KS Lilliefors Modification	1	1	1	0	1	1	0	0	1	0	0	1	1	1	0	1	0	1
Anderson-Darling Test	0	1	1	1	1	1	0	0	1	0	0	1	1	1	0	1	0	1
Cramer-Von Mises Test	1	1	1	1	1	1	0	0	1	0	0	1	1	1	0	1	0	1
Sharpio-Wilk Test	0	1	1	1	1	1	1	0	1	0	0	1	1	1	0	1	0	1
Sharpio-Francia Test	1	1	1	1	1	1	1	0	1	0	0	1	1	1	0	1	0	1
Jarque-Bera Test	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	0	1
D`Agostino & Pearson Test	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**0**	**0**	**1**	**1**	**1**	**1**	**1**	**0**	**1**

Table 11: The table show which normal distribution tests each metric passed. This table is over the metrics computed from the theta frequency band, 4-8 Hz. A 1 in the table indicate that the test is passed at the 0.05 threshold and a 0 indicate that the test was failed.

Table 12: The table show which normal distribution tests each metric passed. This table is over the metrics computed from the alpha frequency band, 8-12 Hz. A 1 in the table indicate that the test is passed at the 0.05 threshold and a 0 indicate that the test was failed.

	Control Data									Phantom Limb Pain Data								
Metric / Normality Test	**D**	**L**	**d**	**E**	**BC**	**T_h**	**R**	**κ**	***PLI***	**D**	**L**	**d**	**E**	**BC**	**T_h**	**R**	**κ**	***PLI***
KS Limiting Form	1	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	0	0
KS Stephens Modification	1	1	1	1	1	1	1	0	0	0	0	1	1	1	1	0	0	0
KS Marsaglia Method	1	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	0	0
KS Lilliefors Modification	1	1	1	1	1	1	1	0	0	0	0	1	1	1	1	0	0	0
Anderson-Darling Test	1	1	1	1	1	1	0	0	0	0	0	1	1	1	0	0	0	0
Cramer-Von Mises Test	1	1	1	1	1	1	1	0	0	0	0	1	1	1	1	0	0	0
Sharpio-Wilk Test	1	1	1	1	1	1	0	0	0	0	0	1	1	1	1	1	0	0
Sharpio-Francia Test	1	1	1	1	1	1	0	0	0	0	0	1	1	1	1	1	0	0
Jarque-Bera Test	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0
D`Agostino & Pearson Test	**1**	**1**	**1**	**1**	**1**	**1**	**0**	**0**	**1**	**0**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**0**

Metric / Normality Test	Control Data									Phantom Limb Pain Data								
	D	**L**	**d**	**E**	**BC**	**T_h**	**R**	**κ**	***PLI***	**D**	**L**	**d**	**E**	**BC**	**T_h**	**R**	**κ**	***PLI***
KS Limiting Form	1	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	0	0
KS Stephens Modification	1	1	1	1	1	1	1	0	0	0	0	1	1	1	1	0	0	0
KS Marsaglia Method	1	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	0	0
KS Lilliefors Modification	1	1	1	1	1	1	1	0	0	0	0	1	1	1	1	0	0	0
Anderson-Darling Test	1	1	1	1	1	1	0	0	0	0	0	1	1	1	0	0	0	0
Cramer-Von Mises Test	1	1	1	1	1	1	1	0	0	0	0	1	1	1	1	0	0	0
Sharpio-Wilk Test	1	1	1	1	1	1	0	0	0	0	0	1	1	1	1	1	0	0
Sharpio-Francia Test	1	1	1	1	1	1	0	0	0	0	0	1	1	1	1	1	0	0
Jarque-Bera Test	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0
D`Agostino & Pearson Test	**1**	**1**	**1**	**1**	**1**	**1**	**0**	**0**	**1**	**0**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**0**

Table 13: The table show which normal distribution tests each metric passed. This table is over the metrics computed from the beta frequency band, 12-30 Hz. A 1 in the table indicate that the test is passed at the 0.05 threshold and a 0 indicate that the test was failed.

Table 14: The table show which normal distribution tests each metric passed. This table is over the metrics computed from the low-gamma frequency band, 30-45 Hz. A 1 in the table indicate that the test is passed at the 0.05 threshold and a 0 indicate that the test was failed.

	Control Data									Phantom Limb Pain Data								
Metric / **Normality Test**	**D**	**L**	**d**	**E**	**BC**	**T_h**	**R**	**κ**	***PLI***	**D**	**L**	**d**	**E**	**BC**	**T_h**	**R**	**κ**	***PLI***
KS Limiting Form	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	0	1
KS Stephens Modification	0	0	1	1	1	1	1	0	0	0	0	1	1	1	1	1	0	1
KS Marsaglia Method	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	0	1
KS Lilliefors Modification	0	0	1	1	1	1	1	0	0	0	0	1	1	1	1	1	0	1
Anderson-Darling Test	0	0	1	1	1	1	1	0	0	0	0	1	1	1	1	1	0	1
Cramer-Von Mises Test	0	1	1	1	1	1	1	0	0	0	0	1	1	1	1	1	0	1
Sharpio-Wilk Test	0	1	1	1	1	1	1	0	0	0	0	1	1	1	0	1	0	1
Sharpio-Francia Test	0	1	1	1	1	1	1	0	1	0	0	1	1	1	0	1	0	1
Jarque-Bera Test	1	1	1	1	1	1	1	0	1	0	0	1	1	1	0	1	0	1
D`Agostino & Pearson Test	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**0**	**0**	**0**	**0**	**1**	**1**	**1**	**0**	**1**	**0**	**1**

Normality Test \ Metric	Control Data									Phantom Limb Pain Data								
	D	**L**	**d**	**E**	**BC**	**T_h**	**R**	**κ**	***PLI***	**D**	**L**	**d**	**E**	**BC**	**T_h**	**R**	**κ**	***PLI***
KS Limiting Form	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	0	0
KS Stephens Modification	0	1	1	1	1	1	0	0	0	0	0	0	1	1	0	0	0	0
KS Marsaglia Method	1	1	1	1	1	1	1	0	0	0	1	1	1	1	1	1	0	0
KS Lilliefors Modification	0	1	1	1	1	1	0	0	0	0	0	0	1	1	0	0	0	0
Anderson-Darling Test	0	1	1	1	1	1	0	0	0	0	0	0	0	0	0	1	0	0
Cramer-Von Mises Test	0	1	1	1	1	1	0	0	0	0	0	0	0	0	0	1	0	0
Sharpio-Wilk Test	0	1	1	1	1	1	0	0	0	0	0	0	0	0	0	1	0	0
Sharpio-Francia Test	0	1	1	1	1	1	0	0	0	0	0	0	0	0	0	1	0	0
Jarque-Bera Test	0	1	1	1	1	1	1	0	1	0	0	1	1	0	0	1	0	0
D`Agostino & Pearson Test	**0**	**1**	**1**	**1**	**1**	**1**	**1**	**0**	**0**	**0**	**0**	**1**	**1**	**0**	**0**	**1**	**0**	**0**

Table 15: The table show which normal distribution tests each metric passed. This table is over the metrics computed from the low-gamma frequency band, 45-90 Hz. A 1 in the table indicate that the test is passed at the 0.05 threshold and a 0 indicate that the test was failed.