# InBrAiN:
# An Interactive Tool for
# Brain Analysis and Visualization

By
Theofanis Oikonomou

SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
AT
UNIVERSITY OF CRETE
HERAKLION, GREECE
JANUARY 2007

UNIVERSITY OF CRETE
DEPARTMENT OF
COMPUTER SCIENCE

The undersigned hereby certify that they have read and recommend to the Faculty of Graduate Studies for acceptance a thesis entitled ***"InBrAiN: An Interactive Tool for Brain Analysis and Visualization"*** by ***Theofanis Oikonomou*** in partial fulfillment of the requirements for the degree of ***Master of Science***.

Dated: January 2007

Supervisor:

Ioannis G. Tollis, Professor

Readers:

Constantine Stephanidis, Professor

Martin Reczko, Principal Researcher at BMI-FORTH

Approved by the Chairman of the Graduate Studies Committee:

Panos Trahanias, Professor

UNIVERSITY OF CRETE

Date: **January 2007**

Author: **Theofanis Oikonomou**

Title: **InBrAiN: An Interactive Tool for Brain Analysis and Visualization**

Department: **Computer Science**

Degree: **M.Sc.** Convocation: **January** Year: **2007**

---

Signature of Author

*Στους γονείς μου Νικόλαο και Μαρία*

*και στον αδερφό μου Λάμπρο*

# Contents

# List of Figures

# List of Tables

# Abstract

Over the years researchers have tried to frame the issue of brain connectivity, a task which has proven to be difficult. The data and methods used to evaluate such connectivity have varied enormously. Furthermore, the actual computational algorithms and concepts used to define the values of brain connectivity differ between studies and researchers. This is very confusing since there is no easy way to combine the results from all these methods and algorithms.

In this thesis we present an alternative way to study brain connectivity. Actually, we describe how the fields of graph theory and graph drawing can assist us produce a tool that can bridge all these studies. By using our framework we provide the means to efficiently analyze, model and visualize brain networks. The networks were produced by clinical experiments (Electroencephalogram - EEG) on populations with schizophrenia and apilepsy.

The developed tool was useful to the doctors as they were able to verify already known properties of brain networks and discover new ideas concerning the reasons for some disorders.

# Περίληψη

Τα τελευταία χρόνια οι ερευνητές προσπάθησαν να οριοθετήσουν το θέμα της διασυνδεσιμότητας του εγκεφάλου, μια προσπάθεια η οποία αποδείχτηκε δύσκολη. Τα δεδομένα και οι μέθοδοι που χρησιμοποιήθηκαν για τον υπολογισμό αυτής της διασυνδεσιμότητας ποικίλουν αρκετά. Επιπρόσθετα, οι υπολογιστικοί αλγόριθμοι που χρησιμοποιήθηκαν για τον ορισμό των τιμών της διασυνδεσιμότητας του εγκεφάλου διαφέρουν μεταξύ μελετών και ερευνητών. Το γεγονός αυτό προκαλεί αρκετή σύγχηση αφού δεν υπάρχει κάποιος εύκολος τρόπος για να συνδιαστούν όλοι οι μέθοδοι και οι αλγόριθμοι.

Σε αυτή την εργασία παρουσιάζεται ένας διαφορετικός τρόπος μελέτης της διασυνδεσιμότητας του εγκεφάλου. Συγκεκριμένα, περιγράφουμε πώς οι τομείς της θεωρίας και της οπτικοποίησης γράφων μπορούν να μας βοηθήσουν να παράγουμε ένα εργαλείο που θα γεφυρώσει όλες αυτές τις μελέτες. Χρησιμοποιώντας το εργαλείο παρέχουμε τα μέσα για την αποτελεσματική ανάλυση, μοντελοποίηση και οπτικοποίηση των δικτύων του εγκεφάλου. Τα δίκτυα παρήχθησαν μετά από κλινικά πειράματα (Εγκεφαλογράφημα - EEG) σε πληθυσμούς με σχιζοφρένεια και επιλειψία.

Το εργαλείο που αναπτύξαμε βοήθησε τους γιατρούς να επιβεβαιώσουν ήδη γνωστές ιδιότητες των δικτύων του εγκεφάλου και να ανακαλήψουν καινούργιες ιδέες για τις αιτίες καποιων δυσλειτουργιών.

# Acknowlegements

First of all, I would like to thank my supervisor Professor Ioannis G. Tollis for his support, scientific guidence and interest in producing original work. Our afternoon meetings put me into creative thinking and helped me overcome difficulties that appeared during this work.

I would also like to thank Professor Sifis Micheloyannis for participating in many of our meetings with useful suggestions and for providing us with the medical background that was necessary. Next, I would like to thank principal researcher Martin Reczko for his useful comments on the developed tool and the final version of this thesis and Professor Constantine Stephanidis for his comments. Finally, I want to thank Assistant Professor Natasa Przulj for her contribution in conducting the experiments and her interest in my work.

Many members of the algorithmic analysis group in the Computer Science Department of the University of Crete (UOC) and the Biomedical Informatics Laboratory of the Institute of Computer Science (ICS) in the Foundation of Research and Technology-Hellas (FORTH) have also helped me with many useful suggestions that improved the quality of my work.

I have also the need to thank the Computer Science Department of the UOC and the Biomedical Informatics Laboratory of the ICS-FORTH for finacially supporting me during the last two and a half years in Crete and for all the infrastructures I have used.

Above all, I owe a lot to my friends (new and old) for always being there for me in difficult and nice moments. Thank you all. Last, but not least, I thank my family for their carry and support in all aspects of my life. This work is dedicated to them.

Heraklion, Crete — Theofanis Oikonomou
January 2007

# Chapter 1

# Introduction

Over the years researchers have tried to frame the issue of brain connectivity, a task which has proven to be difficult. However, the data and methods used to evaluate such connectivity have varied enormously. Techniques to assess brain connectivity have different spatial and temporal resolutions and have focused on different levels of description [25]. The situation is complicated still further since the actual computational algorithms and concepts used to define the values of brain connectivity differ between studies and researchers. Nevertheless, functional brain connectivity has become one of the most influential concepts in modern cognitive neuroscience.

Friston et al [27] defined functional connectivity as the temporal correlation between spatially remote neurophysiological events expressed as deviation from statistical independence across these events in distributed neuronal groups and areas. This is a central and challenging conception of brain connectivity for theories about neural interactions, when analyzing functional neuroimaging data and when developing computer simulations.

The further understanding of brain dynamics will obviously rely upon the development of suitable techniques to analyze neural data for evidence of significant functional correlations. Which measure of functional connectivity is the most useful for studying brain functioning? Treating the brain as a dynamic system from which we are able to observe certain physiological parameters over time, a temporal resolution in the order of milliseconds is of special interest [43]. Here, the Electroencephalogram (EEG) provides a satisfactory scale for accessing temporal evolution of the brain activity associated with cognitive processes in health and disease [45, 59, 38].

According to the definition of functional connectivity the issue that is central to this concept is correlated neurophysiological events, which can be directly derived from EEG data. The question thus arises: which of the known EEG measures of functional connectivity derives the information about operations (discrete events) and estimates the inherent temporal/dynamical correlations among them? Traditionally, coherence has been

the main method to assess the degree of functional connectivity between brain areas [76]. However, mathematically, the coherence value indicates only the linear statistical link between time-series curves in a frequency band [11]. So, the use of methods that capture the linear and non-linear statistical links would be of great importance. In Section 2.6 there is a further descipton of these methods.

The next step should be to find a model that would best describe brain EEG correlation networks, or brain networks for short. Many real-life complex networks have a small-world topology [10] characterized by dense local clustering or cliquishness of connections between neighboring nodes yet a short path length between any (distant) pair of nodes due to the existence of relatively few long-range connections, for details on small-world networks see Section 3.4.2. This is an attractive model for the organization of brain functional networks because a small-world topology can support both segregated/specialized and distributed/integrated information processing.

In humans, there have been several recent reports of small-world [7] or even scale-free [18] (for details on scale-free networks see Section 3.4.4) brain functional networks, but there is much still to learn about the presumably small-world, topology of these networks. For instance what is the real topology of brain networks and what new metrics could help us identify important or emergent properties of these networks.

This thesis is organized as follows: Chapter 2 presents background useful to familiarize the reader with medical concepts. Chapter 3 presents some basic information about graph theory and describes various models used in the thesis. Chapter 4 discusses some results about the way the tool can be used to analyse brain networks and presents results about the model that could best describe these networks. Chapter 5 provides an overview of the tool and the capabilities it provides. Finally, Chapter 6 discusses open problems and future work.

# Chapter 2

# Medical and Signal Processing Background

This chapter will give a brief introductory review of the origins, the physiology and techniques of performing quantitative analysis of the EEG. It presents certain neurophysiology topics necessary for better understanding and describing the electroencephalogram as applied to the study of epilepsy, schizophrenia and electrical brain oscillations in general.

## 2.1 The Origins and evolution of the EEG

Technical developments in the field of electrical measurement and recording in the last quarter of the $19^{th}$ century, made possible one of the greatest triumphs of modern neuroscience; the Electroencephalogram. The EEG was discovered in the $1920's$ and applied in humans by the German psychiatrist Hans Berger [79]. He managed to measure small potential differences at the scalp by using two large sheets of tinfoil, which served as electrodes. The first was placed on the forehead and the other on the back of the head. Currents inside the brain generate these potential differences.

The field was touted as likely to give the psychiatric professionals an insight into the brain's function. Since the times of Berger, it was known that the characteristics of EEG activity change in many different situations, particularly with the level of vigilance (alertness, rest, sleep and dreaming). In the $1930's$, the electrical patterns were found to have neurological correlates for some disorders, such as epilepsy and tumors [16]. The $1940's$ and $50's$ found the instrumentation advancing, with commercially available equipment, fostering establishment of laboratories throughout the world.

EEG can be roughly described as the measurement of the mean electrical activity of the brain in different sites of the head. It is the sum of extracellular current flows of a large group of neurons, as described in the next section. EEG recording became possible by placing high conductivity

electrodes ($impedance < 5000\Omega$) in a fixed head mesh.

Electric potentials may be acquired either between pairs of active electrodes, using the so called bipolar montage or in the case of monopolar montage, with respect to a passive electrode known as the reference. The placement of the electrodes can be either non-invasive superficial on the head (scalp EEG) or introduced into the brain after surgical intervention (intracranial EEG). To perform scalp EEG there have been established standards of electrode placement in 1949 when Rasmussen convened an international committee and designed the international 10/20 electrode placement system. Such system consists of 20 electrodes uniformly distributed along the head generally referenced to two earlobe (A1, A2) electrodes [57]. Normal scalp EEG are routinely taken with the subject at relaxed state, that is with eyes closed or open, or while the subject performs predefined cognitive tasks. The major difficulty one encounters while performing scalp EEG recordings are the artifacts. A variety of different in morphology artifacts arise mainly due to head movements, eye blinking, muscle activity, electrocardiogram (ECG) recording and so on. On the other hand, intracranial recordings using implanted electrodes achieve a significantly higher resolution since possible deviations or alterations of the source EEG signal in its way from the originating neurons towards the scalp are overcome and artifacts are diminished. Mainly two types of intracranial electrodes are used. The subdural electrodes typically arranged in grids or strips placed on the brain and the deep electrodes arranged on a needle entering deep brain structures. The latter involve a major surgery and their use is restricted to pathological cases like the epileptic patients who undergo surgical resection.

## 2.2 The digital EEG processing era

In the $1960's$ and the $1970's$ the digitizing of the EEG was first attempted and the computer was a valuable tool for data analysis. At that moment the original aspiration of Berger and the whole psychiatric community began to see the correlation of the EEG with psychiatric conditions, as well as the brain's detailed response to medication intended to treat these disorders. Even though such medication effects where reported in the previous years, they were often misleadingly interpreted as artifacts; but at this time EEG begins to be systematically studied.

In order to take advantage of the superior computational power and flexibility of computers, a fundamentally crucial step was missing; the analog-digital converter (DAC). This electronic device takes the input of the continuous variable wave and transforms it to discrete values with respect to the amplitude of the input wave. This process is of repeated and is described as the sampling operation. Sampling is performed in real time and each EEG channel is processed in parallel. The sampling rate can vary according

to specific measurement needs. Once the discrete signal is stored, one can perform any mathematical operation like filtering, frequency and amplitude analysis, as well as color mapping. The latter approach is known as quantitative EEG and it differs from any previous analysis attempts to assess the overall appearance of the waves in a qualitative manner.

## 2.3 EEG signal patterns

Even from his first publication on EEG, Berger mentioned the presence of certain signal patterns (rhythms) that he called alpha and beta oscillations. The patterns seen in the EEG, initiating from changes in the frequency amplitude with respect to time, have been divided into the standard bands of delta, theta, alpha and beta, which are somewhat empirical frequency limits. These rhythms are classified according to their location, frequency, amplitude, morphology, periodicity, and behavioral and functional correlates. Since the beginnings of electroencephalography these patterns have been related with different brain arousal states, functions or pathologies [57]. Although these patterns may found to vary across different individuals, the frequency was shown to remain similar. This remarkable note was the spark that led to the definition of frequency bands. The frequency bands are centered around the alpha range and it can be said that they are logarithmically scaled. However, these bands are to this day still not set as a standard, resulting in noting the detailed frequency limits together with the band names, in order to avoid any confusion.

**Delta ($\delta$) band, (0 - 4 Hz):** Commonly, the delta band starts as low as the bandpass filter will allow, with the upper limit set at 3.5 or $4Hz$. It was described in 1936 by W. Gray Walter, while a patient was undergoing neurosurgery for a malignant tumor. Electrodes placed over the related area recorded very slow, high voltage potentials. Since that time, focal delta activity was a reliable indicator of localized disease of the brain. In general, delta waves are not present in the normal adult waking, resting EEG. However, it may occur in elderly subjects in relatively limited amounts, particularly in the temporal regions. Also, it is considered a normal component in infants and young children. It is also characteristic of deep sleep stages, while as noted before specific delta morphologies and localizations are correlated with different pathologies.

**Theta ($\vartheta$) band, (4 - 8 Hz):** It Starts at 4 and goes to $7Hz$ or $8Hz$ and is enhanced during sleep. It is of paramount importance during infancy and diffused theta is usual during childhood; whereas in the vigilant adult high theta activity found in only one location or predominant over one hemisphere is likely considered abnormal and is related to

an underlying brain structural disease. The lesion in this case usually is less malignant or extensive than in the case of delta band focal activity. However symmetrically distributed theta is sometimes considered a normal non-pathological variant, especially when it appears as a rhythmic activity in mid-temporal regions.

**Alpha ($\alpha$) band, (8 - 13 Hz):** Starts at $7-8Hz$ and goes to $12-13Hz$. Neurologists commonly refer to Alpha as the background activity, since it constitutes the principal background feature of the normal adult brain. Its frequency limits are not determined in detail; rather it is defined as the rhythmic posterior activity that attenuates with sensory stimulation. It appears spontaneously during wakefulness, under relaxation and mental idle conditions. It is mostly located in occipital locations while the subject remains at rest with eyes closed, but often distributes to the adjacent parietal and posterior temporal areas. Usually, alpha is more or less symmetrical but often is of higher amplitude over the non-dominant hemisphere. If the subject is tensed alpha may not even be recorded (may reach 2:1 ratio). Alpha asymmetry is always pathological, meaning possible remote infraction in older subjects or brain damage such as congenital hemiatrophy in younger subjects.

**Beta ($\beta$) band, (13 - 30 Hz):** Being the desynchronized faster activity above alpha, it is occasionally divided into beta subtypes, Beta-1 ($\beta 1$) and Beta-2 ($\beta 2$). It is mostly located in central and frontal lobes and has less amplitude compared to alpha waves but is enhanced upon expectancy states or tension. If completely absent it may represent an abnormality. Inter- hemispheric asymmetry and in particular the side of reduced amplitude usually points to the pathological hemisphere, but should always be considered in concert with other background frequency asymmetries.

**Gamma ($\gamma$) band (30 - 95 Hz):** It was initially not regarded to play an important role in interpreting EEG, but became of interest after the cellular level experiments of Gray and Singer [33] and Gray et al. [32] showing the link of stimulus features to a common perceptual information known as the binding theory. Gamma band is also often divided into Gamma-1 ($\gamma 1$) and Gamma-2 - ($\gamma 2$) subtypes.

## 2.4 Measuring the EEG

### 2.4.1 Electrodes

The electrodes play a significant role in the proper and accurate data acquisition. The metal cap sitting at the end of the electrode wire can be made

from a variety of materials. Gold disks are expensive, work well but must be disposed when their surface is scratched. Electrode scratches introduce artifacts. Plated chloride on silver is a very stable electrode but surface scratches again degrade their performance, whereas the more modern silver-silver chloride pellets of amalgamates small pieces avoids this problem. Attention must be paid on the proper skin contact and preparation. Electrode-skin contact below $10k\Omega$ is essential. Skin preparation may be done with numerous commercial surface cleaners. In everyday practice, cap systems are used. These provide small holes to apply conductivity gel through them. Moist gel application helps skin hydration and avoids skin damage. Finally one must ensure the stability of the electrodes, because every movement changes the conductivity and produces unreliable results.

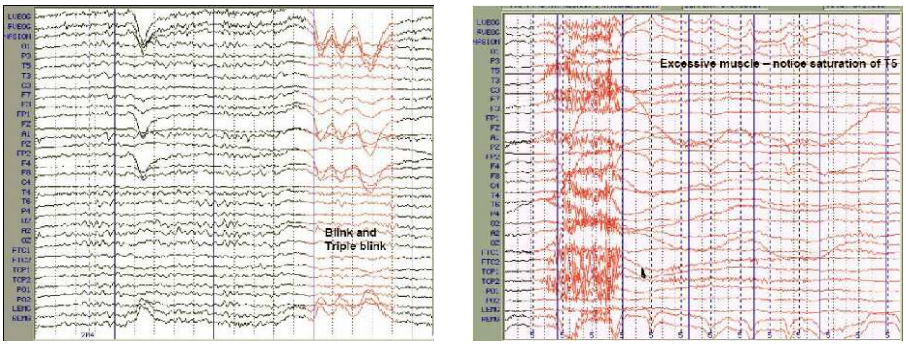
### 2.4.2 Amplifiers and Filters

The potentials measured on the skin surface are in the order of a microvolt. Hence amplification is needed in order for the signals to be scaled and drawn in the display. A certain electrode is used as a reference, like the $CZ$ electrode or maybe a montage of electrode mountings that mostly include the linked ears or even more complex electrode formations that use all available or adjacent electrodes. Amplifier inputs are shielded with impedance values capable of diminishing direct or static potentials contacted from the patient. However, high impedance settings increase the sensitivity to field effects.

Additionally, the amplifiers need to filter out those frequencies that are higher or lower than the EEG spectrum (0 to over $1000Hz$). High and low pass filters combined (band pass filters) take care of this requirement. Additionally, a notch filter is applied, in order to remove the $50Hz$ electrical interference frequency from power lines.

### 2.4.3 Analog to Digital (ADC) conversion

An analog to digital converter is the device needed to convert the continuous EEG signal (analog) to discrete (digital), which can be further processed in computers. The digital signal consists of a number of equally spaced samples. The resolution of the ADC specifies the accuracy of the conversion. A typical ADC resolution for the EEG is of 12 bits. However, a continuous signal can be approximated from the samples under the conditions specified by the sampling theorem [70]. This theorem assumes that a signal can be reconstructed sufficiently, if the signal to be sampled is band limited and the sampling frequency is greater than twice the highest frequency in the signal to be sampled (Nyquist principle). If the sampling theorem is not fulfilled the reconstructed signal is distorted. This effect is referred as aliasing [61].




(a) Eye blinking artifacts (b) Muscle artifacts

Figure 2.1: Examples of EEG artifacts

### 2.4.4 Artifacts

Although EEG is designed to record cerebral activity, it also records electrical activities arising from sites other than the brain. An EEG artifact can be defined as the activity that even though registered in the EEG, does not reflect cortical activity. EEG artifacts can be classified in three broad categories; those arising from physiological processes, the exogenous ones and those coming from the Analog to Digital (A/D) conversion. Physiological artifacts include eye blink and movement (Figure 2.1(a)), as well as muscle (Figure 2.1(b)), pulse, cardioballistic and electrodermal artifacts. The exogenous artifacts are introduced from movement, electrodes, electrical interference, field effect and instrument related artifacts. Finally, the artifacts coming from the A/D conversion include quantum noise, non-linearity of the converter and thermal noise. In addition one should add the spectral leakage or Gibbs's artifact [31] that is introduced when a Fast Fourier Transform (FFT) is attempted.

Several remedies are applied to ensure that the data will be arifact-free. In everyday practice though it is utopic to assert that the data will not contain artifacts. In our study, visual inspection is an extra step included that ensures the rejection of the affected data. This manual process of checking and carefully selecting epochs assures the credibility and accuracy of any results based upon the EEG data.

## 2.5 The abnormal EEG

An EEG is characterized as abnormal not because it does not contain any normal activity patterns, but because it does include abnormal EEG patterns. Such EEG abnormal patterns are usually considered to be either epileptiform activity, slow waves, amplitude abnormalities or specific devia-

tions from normal motifs. Note that the basic abnormal EEG patterns even though they correspond fairly well with a few anatomical and pathophysiological kinds of cerebral lesions, they do not correlate directly with specific neurological diseases. This is true for the reason that similar EEG patterns may be produced by an extensive variety of neurophysiological diseases. Furthermore, many diseases may produce more than one abnormal pattern. Hopefully, there are a few specific patterns that considerably narrow the possible diagnosis. However, etiological diagnosis should not be drawn by the EEG alone, but should be used in combination with complementally clinical examinations such as psychiatric testing. The EEG only helps to establish the presence, severity and cerebral distribution of any present neurological disorders. In the next sections only two syndromes correlated with atypical EEG are briefly described since only these phenomena were addressed and analyzed in this thesis. For extensive information and findings related to numerous EEG abnormalities the reader may refer to clinically oriented references and atlases [22, 67].

### 2.5.1 Epilepsy Syndrome

Epilepsy is a brain disorder that affects about 1% of the population and is characterized by seizures. A seizure is an episode of sudden relatively brief disturbances (lasting one or a few minutes) of mental, motor, sensory or autonomic activity caused by an abnormal paroxysmal cerebral activity. Chronic recurrent seizures characterize an epileptic patient. The cause of epilepsy may be known (i.e., symptomatic epilepsy) or unknown (i.e., idiopathic or cryptogenic epilepsy). EEG analysis has been historically the most useful tool for evaluating the nature of this disorder. The analysis is mostly focused on interictal findings since ictal recordings (during the seizure) are rarely recorded. In the clinical environment seizures may be provoked using different methods, such as photostimulation, hyperventilation and so on, but they do not necessarily share the same properties with the spontaneous ones. However, proper counseling and selection of therapeutic plans is facilitated by making the correct epilepsy syndrome diagnosis among different epilepsy classifications.

Many classifications have been proposed comprising a debatable topic. The most widely used is the one proposed by the Commission on Classification and Terminology of the International League against Epilepsy, which is based on the study of videotape recordings of simultaneously recorded EEG together with the epileptic seizures. The classification is divided into the following major categories.

**Partial seizures:** are those in which the first clinical and electrographic changes indicate the initial involvement of a group of neurons limited to part of one cerebral hemisphere.

**Simple partial seizure:** is noted when consciousness is not impaired during the attack i.e., amnesia for some or all of the events happened during the attack.

**Complex partial seizure:** is noted when consciousness is impaired during the attack.

**Generalized seizures:** are those in which the first clinical and electrographic changes indicate initial involvement of both cerebral hemispheres. Loss of consciousness, bilateral motor activity or both is usually produced. Generalized seizures may be further divided into: absence, myoclonic, clonic, tonic, tonic-clonic and atonic seizures [60].

**Unclassified epileptic seizures:** include all kinds of seizures that cannot be classified because of inadequate or incomplete data, or due to lack of a consensus of opinion among investigators.

### 2.5.2 Schizophrenia Syndrome

Schizophrenia is a mental disorder characterized by impairments in the perception or expression of reality based on physiciatric diagnosis. The term schizophrenia coined from Eugene Bleuler in 1908 comes from the Greek words "σχίζω" and "φρήν" that translate to split or divide and mind, respectively [9]. However, schizophrenia does not necessarily correlate with dissociative identity disorder (multiple personality illusion) and there is no predisposition toward aggressive behavior.

An untreated schizophrenic demonstrates disorganized thinking and experience dilutions or auditory hallucinations. Positive and negative symptoms are possible. As positive symptoms are characterized those diagnosed additional to the normal experience and behavior such as dilutions, auditory hallucinations, thought disorder and other psychosis manifestations. As negative symptoms are characterized those diagnosed to decline in normal experience or behavior, such as flat or constricted affect and emotion and lack of motivation.

Diagnosis is mostly based on the self-reported experiences of the patient, in combination with signs observed be a psychiatrist or a clinical physiologist. Differences in brain structure have been found between schizophrenics and their healthy counterparts, but these tend to be reliable only on the group level and due to the significant variability between each individual, may not reliably represent a particular subject. However, the lack of objective laboratory test makes the possibility of EEG a prospective diagnostic test at least for stabilized patients.

Schizophrenics often face social or occupational isolation and typically live 10 to 12 years less than the healthy subjects due to high suicide rate and increased physical health problems.

## 2.6 Computing Correlations

As stated in Chapter 1 there is a need for measures that capture the linear and the non-linear links between time-series curves in a frequency band in order to assess the degree of functional connectivity between brain areas. Due to the lack of a globally accepted measure to quantify synchronous oscillatory activity based on different underlying assumptions we used several linear and non-linear ones. Namely, the most widely used coherence, wavelet coherence, a robust phase coupling measure known as PLV, a reliable way of assessing generalized synchronization also in state-space also known as RSS-GS and an unbiased alternative called Synchronization Likelihood. For more details about these methods see Appendix A.

# Chapter 3

# Graph Theory Background

When speaking about graph theory in mathematics and computer science, one has to define the basic object which is the graph. Informally speaking, a graph is a set of objects called points or vertices connected by links called lines or edges. In a graph proper, which is by default undirected, a line from point $A$ to point $B$ is considered to be the same as a line from point $B$ to point $A$. In a digraph, short for directed graph, the two directions are counted as being distinct arcs or directed edges. Typically, a graph is depicted in diagrammatic form as a set of dots (for the points, vertices, or nodes), joined by curves (for the lines or edges). Brain and other networks are conveniently described as a graph.

## 3.1 Definitions

A graph $G$ is an ordered pair $G = (V, E)$ such that:

- $V = v_1, v_2, ...v_n$ is a set of nodes.
- $E$ is a set of unordered pairs of distinct nodes, called edges, where $e_{ij}$ denotes an edge between nodes $v_i$ and $v_j$.
- The nodes belonging to an edge are called the endpoints of the edge.

We use $n$ and $m$ to denote the number of nodes and edges, respectively. Hereafter, we assume $v_i \in V$, where $1 \leq i \leq n$.

We define the neighborhood for a node $v_i$ as its immediately connected neighbors, namely $N_i = \{v_j\} : e_{ij} \in E$. The degree $k_i$ of a node is the number of nodes in its neighborhood $|Ni|$. The average degree of a graph is the average degree over all nodes, thus

$$K = \frac{\sum_{v_i \in V} k_i}{n} \tag{3.1}$$

The clustering coefficient $C_i$ for a node $v_i$ is the proportion of links between the nodes within its neighborhood divided by the number of links that could possibly exist between them. For an undirected graph, if a node $v_i$ has $k_i$ neighbors, at most $\frac{k_i(k_i-1)}{2}$ edges could exist among the nodes within the neighborhood, thus

$$C_i = \frac{2|\{e_{jk}\}|}{k_i(k_i-1)} : v_j, v_k \in N_i \tag{3.2}$$

This measure is 1 if every neighbor connected to $v_i$ is also connected to every other node within the neighborhood, and 0 if no node that is connected to $v_i$ connects to any other node that is connected to $v_i$. The clustering coefficient for a graph is given by Watts and Strogatz [84] as the average of the clustering coefficients over all nodes,

$$C = \frac{\sum_{i=1}^{n} C_i}{n} , \tag{3.3}$$

and is a measure of the tendency of graph nodes to form local clusters.

We define a path from $v_i$ to $v_j$ as an alternating sequence of nodes and edges, beginning with $v_i$ and ending with $v_j$, such that each edge connects its preceding with its succeeding node. The shortest path (distance or geodesic distance) $d_{ij}$ between two nodes $v_i$ and $v_j$ is the minimum number of edges we need to traverse in order to go from node $v_i$ to node $v_j$. By definition, $d_{ii} = 0$ for every $v_i$ , and $d_{ij} = d_{ji}$ for $v_i, v_j$. There are several other measurements defined in terms of distance such as the eccentricity of a node, the radius and the diameter of a graph.

The eccentricity $\varepsilon(v)$ of a node $v$ in a connected graph $G$ is the maximum graph distance between $v$ and any other node $u$ of $G$

$$\varepsilon(v) = max\{d_{vu} : v, u \in V\} \tag{3.4}$$

For a disconnected graph, all nodes are defined to have infinite eccentricity.

The radius and the diameter of a graph is the minimum and the maximum eccentricity of any node in the graph respectively. The diameter represents the greatest distance between any two nodes.

The average shortest path length

$$L = \frac{\sum_{i,j \in V, i \neq j} d_{ij}}{n(n-1)} \tag{3.5}$$

is the average shortest path (distance) connecting any two nodes of the graph and is a measure of interconnectedness of the graph. Formally, the absence of a path between $v_i$ and $v_j$ implies $d_{ij} = \infty$, but for our experiments we choose $d_{ij} = 1000$. Let $\sigma_{ij} = \sigma_{ji}$ denote the number of shortest paths from $v_i$ to $v_j$ , where $\sigma_{ii} = 1$ by convention. Let $\sigma_{ij}(v)$ denote the number of shortest paths from $v_i$ to $v_j$ that contain $v \in V$.

By embeding a graph into a plane we can visualize it's nodes and edges. This can be done by assigning $(x, y)$ coordinates to every node. By doing so we would place some nodes closer or further to other nodes in terms of cartesian distance. The geometric coefficient $\gamma_i$ for a node $v_i$ is the proportion of "geometric" edges between the nodes within its neighborhood, thus

$$\gamma_i = \frac{|geometric\ edges|}{k_i}\ , \tag{3.6}$$

where "geometric" edges are the edges that connect a node with nodes that are close to it in terms of cartesian distance.

This measure is 1 if all edges connecting $v_i$ with its neighbors are "geometric", and 0 if no edge connecting $v_i$ with its neighbors is "geometric" or if $k_i = 0$. The geometric coefficient for a graph is given as the average of the geometric coefficients over all nodes,

$$\gamma = \frac{\sum_{i=1}^{n} \gamma_i}{n}\ , \tag{3.7}$$

and is a measure of the tendency of graph nodes to connect with nodes that are geographically reachable.

## 3.2 Node Centrality Measures

Within graph theory and network analysis, there are various measures of the centrality of a node that determine the relative importance of a node within the graph. For example, the importance of a person in a social network, or a room within a building or a road in an urban network. In other words, the centrality of a node in a network is a measure of the structural importance of the node. Such measures are the degree, closeness, betweenness and bridging centrality which are analysed below. These measures attempt to quantify the prominence of an individual node embedded in a network. A central node, presumably, has a stronger influence on other network nodes as it has high degree, close to all other nodes, so it is easily accessible, and lies between other nodes, so it is part of several shortest paths. For interpretability, i.e. to control for the size of the network, the above indices are usually normalized to lie between zero and one. Though their definitions extend naturally to directed or disconnected graphs, normalization then becomes a problem with some of the above measures.

**Degree Centrality:** This is the simplest measure. A node is central in a graph, if it is active enough in the sense that is has a lot of links to other nodes. The degree centrality of $v \in V$ is defined as follows:

$$C_D(v) = \frac{k_v}{n-1} \tag{3.8}$$

**Closeness Centrality:** In 1966, Sabidussi [69] suggested that the most central nodes according to closeness centrality can quickly interact to all others because they are close to them. This measure is preferable to degree centrality, because it does not take into account only direct connections among nodes but also indirect connections. If the graph is not strongly connected, we take only reachable nodes into account, but we weight the result with the number of reachable nodes. The closeness centrality of $v \in V$ is defined as follows:

$$C_C(v) = \frac{n-1}{\sum_{u \in V} d_{vu}} \tag{3.9}$$

The smallest possible distance of the selected node from all other nodes is obtained, if the node has all other nodes for neighbors. In this case the closeness centrality is 1 indicating that a node can reach others on relatively short paths.

**Betweenness Centrality:** In the case of brain or communication networks the most important property of a node is not the distance from other nodes but the the amount of shortest paths it lies on. Such nodes have control over the flow of information in the network. The idea behind betweenness centrality is that a node is central, if it lies on several shortest paths among other pairs of nodes. In 1977, Freeman [23] defined the betweenness centrality of $v \in V$ as follows:

$$C_B(v) = \frac{\sum_{s \neq v \neq t \in V} \frac{\sigma_{st}(v)}{\sigma_{st}}}{(n-1)(n-2)/2} \tag{3.10}$$

Suppose that communication in a network always passes through shortest available paths. Then betweenness centrality of $v \in V$ is the sum of probabilities across all possible pairs of nodes, that the shortest path between $s \in V$ and $t \in V$ will pass through node $v$.

**Bridging Centrality:** A bridging node is a node lying between modules, i.e. a node connecting densely connected components in a graph. The bridging centrality of $v \in V$ is the product of the betweenness centrality $C_B(v)$ and the bridging coefficient $BC$, which measures the global and local features of node $v$ respectively:

$$C_R(v) = C_B(v) \times BC(v) \tag{3.11}$$

The bridging coefficient of a node determines the extent of how well the node is located between high degree nodes. The bridging coefficient of $v \in V$ is defined as:

$$BC(v) = \frac{\frac{1}{k_v}}{\sum_{i \in N_v} \frac{1}{k_i}} \tag{3.12}$$

The bridging coefficient assesses the local bridging characteristics in the neighborhood. The bridging coefficient understands a network as a simple

electrical circuit. Intuitively, there should be more congestion on the smaller degree nodes if an unit electrical current arrives on a node since the smaller degree nodes have lesser number of outlets than the bigger degree nodes have. So, if we consider the reciprocal of the degree of a node as the "resistance" of the node, the bridging coefficient can be seen as the ratio of the resistance of a node to the sum of the resistance of the neighbors. Critical bridging nodes, typically representing rate limiting points in the network and because they connect its densely connected regions, have high "resistance". Thus, higher $C_R(v)$ signifies that more information flows through node $v$, i.e., higher betweenness centrality $C_B(v)$ and more resistance on node $v$, i.e., higher bridging coefficient $BC(v)$, by bridging densely connected regions.

Different centrality measures can give quite different results. Therefore we must be very careful in choosing the appropriate centrality measure for a particular network since some nodes may have low degree, but high betweenness centrality.

## 3.3 Graph Centrality Measures

Nodes can also be aggregated to obtain a group-level centrality indices. Centralization refers to the extent to which the network is concentrated on a group of nodes or just one node. Empirically, a centralized network is one that has few nodes, or just one, with considerably higher centrality scores than others in the network, e.g. a large variability of individual centrality scores. This inhomogeneity of a centrality index is used to define [24] the centralization of a graph with respect to that index as follows:

$$C_A = \frac{\sum_{v \in V} (C_A^* - C_A(v))}{max \sum_{v \in V} (C_A^* - C_A(v))} \tag{3.13}$$

where $C_A^*$ is the highest value of selected node centrality measure over all $C_A(v)$ in the set of nodes of a graph. Graph centralisation index is a number between 0 and 1, thus the index is 0 if all nodes have equal centrality value and 1 when few nodes, or just one node, completely dominates all other nodes. In detail we get:

$$Graph\ Degree\ Centralization: \quad C_D = \frac{\sum_{v \in V} (C_D^* - C_D(v))}{n-2} \tag{3.14}$$

$$Graph\ Closeness\ Centralization: \ C_C = \frac{\sum_{v \in V} (C_C^* - C_C(v))}{(n-1)(n-2)/(2n-3)} \tag{3.15}$$

$$Graph\ Betweenness\ Centralization: \ C_B = \frac{\sum_{v \in V} (C_B^* - C_B(v))}{n-1} \tag{3.16}$$

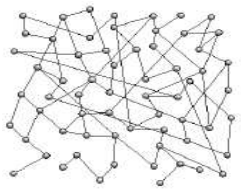


Figure 3.1: An example of a random Erdös and Rényi graph

A theoretical foundation for centrality measures not based on shortest paths is given in [26]. See [81] for further details.

## 3.4 Network Models

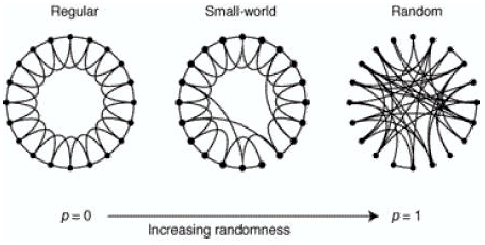
Measuring some basic properties of a complex network is the first step toward understanding its structure. The next step, is to develop a mathematical model with a topology of similar statistical properties, thereby obtaining a platform on which mathematical analysis is possible. We limited our current study to five of the most widely used network models, i.e. Erdös-Rényi (ER), Small-World (SW), Generalized Random (GR), Scale-Free (SF) and Geometric Random (GEO). The reason for this is that these models are very well established and are most extensively used for modeling various real-world phenomena, so we wanted to evaluate how well each of them models brain networks.

Note that other network models exist and new ones will certainly be designed in the future to model the real-world phenomena better. It is not possible to predict how many and which new models will appear in the future; this will largely depend on the new data that need to be modeled.

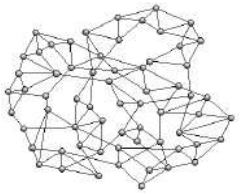
### 3.4.1 Erdös-Rényi Model

The random graph developed by Rapoport [64, 65, 66] and independently by Erdös and Rényi [19, 20, 21] can be considered the most basic model of complex networks. In their paper published in 1959 [19], Erdös and Rényi introduced a model to generate random graphs consisting of $N$ vertices and $M$ edges. Starting with $N$ disconnected vertices, the network is constructed by the addition of $L$ edges at random, avoiding multiple and self connections. Another similar model defines $N$ vertices and a probability $p$ of connecting each pair of vertices. The latter model is widely known as Erdös-Rényi (ER) model. Figure 3.1 shows an example of this type of network. For the ER model, in the large network size limit ($N \to \infty$), the average number of connections of each vertex $\langle k \rangle$, given by

$$\langle k \rangle = p(N-1) \tag{3.17}$$




(a) Small-World construction proccess



(b) A small-world graph with $N = 64$ nodes, $\kappa = 2$ and $p = 0.1$

Figure 3.2: The Small-World model of Watts and Strogatz

diverges if $p$ is fixed. Instead, $p$ is chosen as a function of $N$ to keep $\langle k \rangle$ fixed: $p = \langle k \rangle / (N - 1)$. For this model, the degree distribution $P(k)$ is a Poisson distribution.

### 3.4.2 Small-World Model

Many real world networks exhibit what is called the small world property, i.e. most vertices can be reached from the others through a small number of edges. This characteristic is found, for example, in social networks, where everyone in the world can be reached through a short chain of social acquaintances [82, 83]. This concept originated from the famous experiment made by Milgram in 1967 [46], who found that two US citizens chosen at random were connected by an average of six acquaintances.

Another property of many networks is the presence of a large number of loops of size three, i.e. if vertex $i$ is connected to vertices $j$ and $k$, there is a high probability of vertices $j$ and $k$ being connected (the clustering coefficient, see Equation 3.3, is high); for example, in a friendship network, if $B$ and $C$ are friends of $A$, there is a high probability that $B$ and $C$ are also friends. ER networks have the small world property but a small average clustering coefficient; on the other hand, regular networks with the second property are easy to construct, but they have large average distances. The most popular model of random networks with small world characteristics and an abundance of short loops was developed by Watts and Strogatz [84] and is called the Watts-Strogatz small-world model. They showed that small-world networks are common in a variety of realms ranging from the C. elegans neuronal system to power grids. This model is situated between an ordered finite lattice and a random graph presenting the small world property and high clustering coefficient.

To construct a small-word network, one starts with a regular lattice of

$N$ vertices (Figure 3.2(a)) in which each vertex is connected to $\kappa$ nearest neighbors in each direction, totalizing $2\kappa$ connections, where $N \gg \kappa \gg log(N) \gg 1$. Next, each edge is randomly rewired with probability $p$. When $p = 0$ we have an ordered lattice with high number of loops but large distances and when $p \to 1$, the network becomes a random graph with short distances but few loops. Watts and Strogatz have shown that, in an intermediate regime, both short distances and a large number of loops are present. Figure 3.2(b) shows an example of a Watts-Strogatz network. Alternative procedures to generate small-world networks based on addition of edges instead of rewiring have been proposed [50, 55]. The degree distribution for small-world networks is similar to that of random networks, with a peak at $\langle k \rangle = 2\kappa$.

### 3.4.3 Generalized Random Model

A common way to study real networks is to compare their characteristics with the values expected for similar random networks. As the degrees of the vertices are important features of the network, it is interesting to make the comparison with networks with the same degree distribution. Models to generate networks with a given degree distribution, while being random in other aspects, have been proposed.

Bender and Canfield [8] first proposed a model to generate random graphs with a pre-defined degree distribution called configuration model. Later, Molloy and Reed [48, 49] proposed a different method that produces multigraphs (i.e. loops and multiple edges between the same pair of vertices are allowed).

The common method used to generate this kind of random graph involves selecting a degree sequence specified by a set $k_i$ of degrees of the vertices drawn from the desired distribution $P(k)$. Afterwards, each vertex $i$ is associated to a number $k_i$ of "stubs" or "spokes" (ends of edges emerging from a vertex) according to the desired degree sequence. Next, pairs of such stubs are selected uniformly and joined together to form an edge. When all stubs have been used up, a random graph that is a member of the ensemble of graphs with that degree sequence is obtained [52, 56, 54].

Another possibility, the rewiring method, is to start with a network (possibly a real network under study) that already has the desired degree distribution, and then iteratively choose two edges and interchange the corresponding attached vertices [47]. This rewiring procedure is used in some results presented in Section 16.2.

Due to its importance and amenability to analytical treatment, many works deal with this model, including the papers of Newman [53], Aiello et al. [2], Chung and Lu [12] and Cohen and Havlin [13].

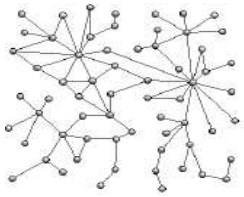


Figure 3.3: An example of a scale-free graph

### 3.4.4 Scale Free Model

After Watts and Strogatz's model, Barabási and Albert [5] showed that the degree distribution of many real systems is characterized by an uneven distribution. Instead of the vertices of these networks having a random pattern of connections with a characteristic degree, as with the ER and SW models (see Figure 3.1), some vertices are highly connected while others have few connections, with the absence of a characteristic degree. More specifically, the degree distribution has been found to follow a power law for large values of $k$,

$$P(k) \sim k^{-\gamma} \tag{3.18}$$

These networks are called scale-free networks. A characteristic of this kind of network is the existence of hubs, i.e. vertices that are linked to a significant fraction of the total number of edges of the network.

The Barabási-Albert network model is based on two basic rules: growth and preferential attachment. The network is generated starting with a set of $m_0$ vertices; afterwards, at each step of the construction the network grows with the addition of new vertices. For each new vertex, m new edges are inserted between the new vertex and some previous vertex. The vertices which receive the new edges are chosen following a linear preferential attachment rule, i.e. the probability of the new vertex $i$ to connect with an existing vertex $j$ is proportional to the degree of $j$,

$$P(i \to j) = \frac{kj}{\sum_u k_u} \tag{3.19}$$

Thus, the most connected vertices have greater probability to receive new vertices. This is known as "the rich get richer" paradigm. (Figure 3.3 shows an example of Barabási-Albert network).

### 3.4.5 Geometric Random Model

Complex networks are generally considered as lying in an abstract space, where the position of vertices has no particular meaning. In the case of several kinds of networks, such as protein-protein interaction networks or networks of movie actors, this consideration is reasonable. However, there are

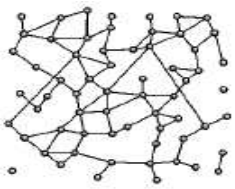


Figure 3.4: An example of a geometric random graph with $N = 64$ nodes

many networks where the position of vertices is particularly important as it influences the network evolution. This is the case for geographical or spatial networks, for example highway networks or the Internet, where the position of cities and routers can be localized in a map and the edges between correspond to real physical entities, such as roads and optical fibers [30]. Other important examples of geographical networks are power grids [3, 42], airport networks [6, 35, 36], subway [44] and neural networks [73]. In such networks, the existence of a direct connection between vertices can depend on a lot of constraints such as the distance between them, available resources to construct the network, territorial limitation and so on. The models considered to represent these networks should consider these constraints.

A simple way to generate geographical networks is to distribute $N$ vertices at random in a two-dimensional space $\Omega$ and link them with a given probability which decays with the distance, for instance

$$P(i \to j) \sim e^{-\lambda s_{ij}} \tag{3.20}$$

where $s_{ij}$ is the geographical distance of the vertices and $\lambda$ fixes the length scale of the edges. This model generates a Poisson degree distribution as observed for random graphs and can be used to model road networks (see Figure 3.4). Alternatively, the network development might start with few nodes while new nodes and connections are added at each subsequent time step (spatial growth). Such a model is able to generate a wide range of network topologies including small-world and linear scale-free networks [39].

# Chapter 4

# Analysing and Modeling Brain Networks

In this chapter we describe how we construct the graphs from the input data, show how we incorporate many of the metrics discussed in Chapter 3 into our framework and how this can be used to derive useful conclusions. Furthermore, we present some of our findings and describe how it can be interactively used in collaboration with the visualization techniques described in Chapter 5.

## 4.1 From Correlations to Graphs

In our study two diseases were encountered, schizophrenia and epilepsy. In the case of schizophrenia three different groups were tested. The first group consists of 20 control university educated subjects (HE), the second one consists of 20 control subjects without higher education (HU) and the third one consists of 20 stabilized patients with schizophrenia (P). For all subjects two different situations are considered: the control (Rest), where subjects had the eyes fixed on a "star" on the computer screen and the cognitive activation during working memory (WM) while performing a two-back[1] test using capital Greek letters. The testing hypothesis suggests that a WM task requires considerable mental effort and the disconnection on neuronal assemblies in patients could be visible. In order to acquire the raw data the EEG signals in all three groups were recorded from $N = 30$ cap electrodes, according to the 10/20 international system [37], referred to linked earlobe electrodes.

In the epilepsy case only one group was tested. It consists of 10 subjects (Subjects) all of which were diagnosed with epilepsy and were scheduled for surgical intervention. All individuals were tested during two different

[1] In the 2-back condition, the target letter was any letter that was identical to the one presented two trials before it.

situations: the control (OB), where subjects were calm and the Seizure where the subjects were having an epileptic seizure. In order to acquire the raw data the EEG signals in the group were recorded from $N = 21$ cap electrodes. These electrodes were implanted under the subjects' skull before the operation took place in order to have a better knowledge regarding the source of the epileptic seizure.

As stated in Chapter 1 there is no universally accepted method for assessing the spatial pattern of functional connectivity. So, we used several linear and non-linear methods to compute it, as discussed in Section 2.6. All these methods yield a statistical measure ranging from 0 to 1, which is an indication of how much a specific electrode is correlated with each of the other electrodes. Thus, we come up with an $N \times N$ correlation matrix (CM) with elements ranging from 0 to 1 formulated per task and subject. In order to obtain a graph from a CM we need to convert it into an $N \times N$ binary adjacency matrix, A. To achieve that we define a variable called threshold T, such that $T \in [0, 1]$. The value $A(i, j)$ is either 1 or 0, indicating the presence or absence of an edge between nodes $i$ and $j$, respectively. Namely, $A(i, j) = 1$ if $CM(i, j) \geq T$, otherwise $A(i, j) = 0$. Thus we define a graph for each value of T. For the purposes of our work we calculated 1000 such graphs, one for every thousandth of T.

## 4.2 Experimental Results

As was described in the previous section 1000 graphs were generated for every subject and its correlation matrix. For each subject there are two states, that is Rest and WM (or OB and Seizure in the epilepsy case), and seven frequency bands, in terms of functional uniformity, namely $delta[0.5-4Hz]$, $theta[4-8Hz]$, $alpha1[8-10Hz]$, $alpha2[10-13Hz]$, $beta[13-30Hz]$, $gamma1[30-45Hz]$ and $gamma2[45-90Hz]$. This is a large number of networks to be analysed.

For every such graph we conducted extended experiments measuring most of the metrics described in Chapter 3 and the results were used in our tool. This way we are able to observe the way these metrics alter as the threshold increases. As we plot the average value, over all subjects belonging to one group, of each metric we used a statistical test (student's t-test) to locate the threshold regions where the differences between the values of different groups were statistically important. We used a dark vertical marker to indicate this difference in the tool.

Our study focuses on $gamma1$ band wavelet coherence analysis. Significant coherent time-regions are transformed to the aforementioned binary matrix A served as input. The values of the average degree $K$, the clustering coefficient $C$ and the average shortest path length $L$ during WM were computed. These three graph measures $K$, $C$ and $L$ actually represent an

| HEALTHY EDUCATED | | | |
|---|---|---|---|
| **T** | **K** | **C** | **L** |
| 0.651 | 26.94 ± 0.27 | 0.97 ± 0.00 | 4.40 ± 3.33 |
| 0.661 | 26.10 ± 0.28 | 0.95 ± 0.01 | 27.60 ± 8.83 |
| 0.671 | 25.25 ± 0.36 | 0.94 ± 0.01 | 37.50 ± 10.10 |
| 0.681 | 24.31 ± 0.48 | 0.92 ± 0.01 | 50.84 ± 9.40 |
| 0.691 | 23.22 ± 0.61 | 0.90 ± 0.01 | 70.72 ± 7.45 |
| 0.701 | 22.25 ± 0.73 | 0.87 ± 0.01 | 80.06 ± 11.79 |
| 0.711 | 21.18 ± 0.85 | 0.85 ± 0.02 | 86.53 ± 12.26 |
| 0.721 | 20.06 ± 0.94 | 0.83 ± 0.02 | 96.09 ± 12.48 |
| 0.731 | 18.73 ± 1.02 | 0.81 ± 0.02 | 117.48 ± 17.13 |
| 0.741 | 17.36 ± 1.11 | 0.78 ± 0.02 | 132.10 ± 21.66 |
| 0.751 | 16.14 ± 1.17 | 0.76 ± 0.02 | 147.98 ± 21.05 |
| 0.761 | 14.89 ± 1.23 | 0.73 ± 0.03 | 150.81 ± 22.31 |
| 0.771 | 13.67 ± 1.26 | 0.70 ± 0.03 | 174.50 ± 25.91 |
| 0.781 | 12.60 ± 1.27 | 0.69 ± 0.03 | 196.24 ± 31.37 |
| 0.791 | 11.56 ± 1.22 | 0.67 ± 0.03 | 210.75 ± 32.57 |
| 0.801 | 10.54 ± 1.19 | 0.63 ± 0.03 | 230.08 ± 35.12 |
| 0.811 | 9.48 ± 1.16 | 0.60 ± 0.03 | 243.48 ± 37.14 |
| 0.821 | 8.45 ± 1.10 | 0.59 ± 0.03 | 271.10 ± 38.40 |
| 0.831 | 7.46 ± 1.00 | 0.55 ± 0.03 | 294.93 ± 40.91 |
| 0.841 | 6.53 ± 0.90 | 0.53 ± 0.04 | 344.56 ± 43.36 |
| 0.851 | 5.63 ± 0.82 | 0.48 ± 0.04 | 395.30 ± 48.06 |
| 0.861 | 4.80 ± 0.74 | 0.46 ± 0.04 | 480.07 ± 53.76 |
| 0.871 | 4.08 ± 0.66 | 0.42 ± 0.04 | 542.51 ± 50.83 |
| 0.881 | 3.35 ± 0.59 | 0.38 ± 0.05 | 632.61 ± 47.67 |
| 0.891 | 2.75 ± 0.51 | 0.35 ± 0.05 | 695.45 ± 44.08 |
| 0.901 | 2.18 ± 0.44 | 0.31 ± 0.05 | 782.01 ± 41.65 |

Table 4.1: Healthy Educated values of K, C and L

overall signature of the graph topology. The values of these metrics for the healthy educated and patient groups for selected thresholds are shown in Table 4.1 and Table 4.2. For each case we show the average metric value over all subjects of the group and the standard error of the mean, which is equal to the standard deviation of the group devided by the square root of the group's size. Furthemore, the $p$-value, where $0 \leq p \leq 1$, derived from the t-test are presented in Table 4.3. In our experiments two sets are thought to be statistically different when $p \leq 0.05$.

| PATIENT | | | |
|---|---|---|---|
| **T** | **K** | **C** | **L** |
| 0.651 | 24.86 ± 0.72 | 0.92 ± 0.01 | 7.80 ± 4.58 |
| 0.661 | 23.19 ± 0.96 | 0.90 ± 0.02 | 17.62 ± 8.01 |
| 0.671 | 21.58 ± 1.11 | 0.88 ± 0.02 | 37.09 ± 12.78 |
| 0.681 | 20.11 ± 1.17 | 0.84 ± 0.02 | 74.55 ± 22.26 |
| 0.691 | 18.66 ± 1.20 | 0.81 ± 0.02 | 108.80 ± 25.68 |
| 0.701 | 17.37 ± 1.20 | 0.79 ± 0.03 | 139.00 ± 28.56 |
| 0.711 | 16.14 ± 1.20 | 0.77 ± 0.03 | 177.01 ± 32.75 |
| 0.721 | 14.96 ± 1.20 | 0.75 ± 0.03 | 194.84 ± 33.44 |
| 0.731 | 13.81 ± 1.17 | 0.72 ± 0.03 | 205.92 ± 35.13 |
| 0.741 | 12.68 ± 1.12 | 0.69 ± 0.03 | 223.07 ± 36.66 |
| 0.751 | 11.67 ± 1.11 | 0.67 ± 0.03 | 244.57 ± 42.90 |
| 0.761 | 10.77 ± 1.10 | 0.66 ± 0.03 | 269.06 ± 44.69 |
| 0.771 | 9.91 ± 1.08 | 0.63 ± 0.03 | 293.20 ± 48.19 |
| 0.781 | 9.09 ± 1.04 | 0.61 ± 0.03 | 327.30 ± 45.81 |
| 0.791 | 8.25 ± 1.00 | 0.58 ± 0.03 | 349.62 ± 45.45 |
| 0.801 | 7.45 ± 0.96 | 0.56 ± 0.03 | 374.56 ± 46.77 |
| 0.811 | 6.66 ± 0.91 | 0.53 ± 0.03 | 418.53 ± 45.09 |
| 0.821 | 5.88 ± 0.84 | 0.50 ± 0.04 | 448.18 ± 48.70 |
| 0.831 | 5.19 ± 0.77 | 0.47 ± 0.04 | 495.11 ± 49.52 |
| 0.841 | 4.48 ± 0.65 | 0.43 ± 0.04 | 530.93 ± 49.30 |
| 0.851 | 3.95 ± 0.60 | 0.39 ± 0.04 | 555.88 ± 51.02 |
| 0.861 | 3.32 ± 0.51 | 0.35 ± 0.04 | 629.29 ± 49.31 |
| 0.871 | 2.80 ± 0.45 | 0.33 ± 0.03 | 701.86 ± 42.76 |
| 0.881 | 2.37 ± 0.40 | 0.29 ± 0.04 | 763.19 ± 43.15 |
| 0.891 | 1.95 ± 0.34 | 0.26 ± 0.04 | 797.72 ± 44.21 |
| 0.901 | 1.65 ± 0.29 | 0.23 ± 0.04 | 824.45 ± 43.56 |

Table 4.2: Patient values of K, C and L

| | | p-value | | |
|---|---|---|---|---|
| $\mathbf{T}_{Healthy}$ | $\mathbf{T}_{Patient}$ | **K** | **C** | **L** |
| 0.651 | 0.651 | 0.01 | 0.01 | 0.55 |
| 0.661 | 0.661 | 0.01 | 0.02 | 0.41 |
| 0.671 | 0.671 | 0.00 | 0.02 | 0.98 |
| 0.681 | 0.681 | 0.00 | 0.00 | 0.34 |
| 0.691 | 0.691 | 0.00 | 0.00 | 0.17 |
| 0.701 | 0.701 | 0.00 | 0.00 | 0.07 |
| 0.711 | 0.711 | 0.00 | 0.01 | 0.02 |
| 0.721 | 0.721 | 0.00 | 0.01 | 0.01 |
| 0.731 | 0.731 | 0.00 | 0.02 | 0.03 |
| 0.741 | 0.741 | 0.01 | 0.02 | 0.04 |
| 0.751 | 0.751 | 0.01 | 0.03 | 0.05 |
| 0.761 | 0.761 | 0.02 | 0.07 | 0.03 |
| 0.771 | 0.771 | 0.03 | 0.12 | 0.04 |
| 0.781 | 0.781 | 0.04 | 0.07 | 0.02 |
| 0.791 | 0.791 | 0.04 | 0.05 | 0.02 |
| 0.801 | 0.801 | 0.05 | 0.09 | 0.02 |
| 0.811 | 0.811 | 0.06 | 0.10 | 0.00 |
| 0.821 | 0.821 | 0.07 | 0.07 | 0.01 |
| 0.831 | 0.831 | 0.08 | 0.12 | 0.00 |
| 0.841 | 0.841 | 0.07 | 0.07 | 0.01 |
| 0.851 | 0.851 | 0.11 | 0.11 | 0.03 |
| 0.861 | 0.861 | 0.11 | 0.05 | 0.05 |
| 0.871 | 0.871 | 0.12 | 0.09 | 0.02 |
| 0.881 | 0.881 | 0.18 | 0.18 | 0.05 |
| 0.891 | 0.891 | 0.20 | 0.17 | 0.11 |
| 0.901 | 0.901 | 0.32 | 0.18 | 0.49 |

Table 4.3: p-value for K, C and L between HE and P subjects

Our experiments indicated that $K$ and $C$ are getting lower and $L$ is getting higher while moving from healthy to schizophrenics, in the whole threshold range. This is clear when observing the aforementioned tables and the plots created by our tool (Figure 4.1). Instead of studying each measure independently, we attempt to quantify their interaction.

Towards this direction we concentrated in different values of $T$, where the values of $K$ and $C$ of schizophrenic patients were equal to those of healthy subjects. For the above values of $T$, the respective values of $L$ for patients were much greater than those for controls. This can be seen by comparing Table 4.4 and Table 4.5, where three selected different cases are presented. This was a first indication for the connection between $K$, $C$ and $L$. In order

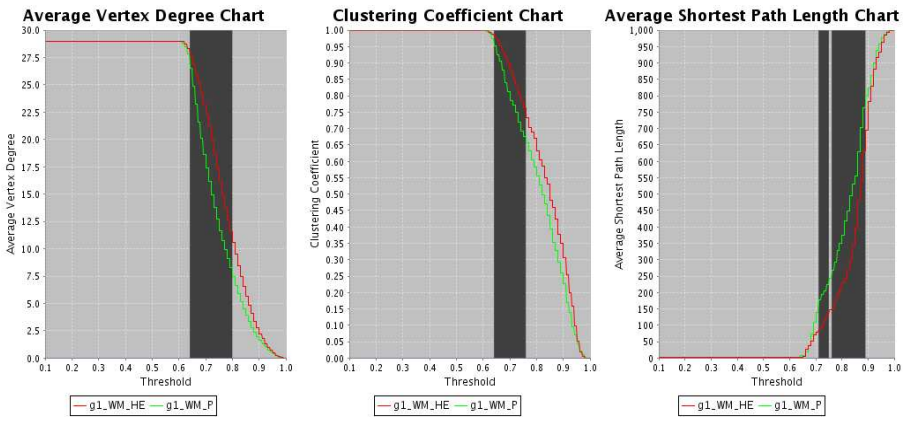



Figure 4.1: K, C and L plots for *gamma*1 band during WM

to make sure that the observed differences were important we conducted a t-test. The results of the statistical test are shown in Table 4.6. Here it is clear that for the selected thresholds the values of $K$ and $C$ are statistically equal, as $p \geq 0.05$, while the respective values of $L$ are statistically different, as $p \leq 0.05$.

| HEALTHY EDUCATED | | | |
|---|---|---|---|
| **T** | **K** | **C** | **L** |
| 0.791 | $11.56 \pm 1.22$ | $0.67 \pm 0.03$ | $210.75 \pm 32.57$ |
| 0.801 | $10.54 \pm 1.19$ | $0.63 \pm 0.03$ | $230.08 \pm 35.12$ |
| 0.831 | $7.46 \pm 1.00$ | $0.55 \pm 0.03$ | $294.93 \pm 40.91$ |

Table 4.4: Healthy Educated having comparable K and C with Patients

| PATIENT | | | |
|---|---|---|---|
| **T** | **K** | **C** | **L** |
| 0.781 | $9.09 \pm 1.04$ | $0.61 \pm 0.03$ | $327.30 \pm 45.81$ |
| 0.791 | $8.25 \pm 1.00$ | $0.58 \pm 0.03$ | $349.62 \pm 45.45$ |
| 0.811 | $6.66 \pm 0.91$ | $0.53 \pm 0.03$ | $418.53 \pm 45.09$ |

Table 4.5: Patients having comparable K and C with Healthy Educated

Additionally, we investigated the opposite direction. That is we tried to find values of $T$ where the values of $L$ for the patients and the healthy were equal. We hypothised that the respective values of $K$ and $C$ would be different. Again three selected cases are shown in Table 4.7 and Table 4.8. Additionally, we conducted a t-test to make sure that the results were

| | | p-value | | |
|---|---|---|---|---|
| $\mathbf{T}_{Healthy}$ | $\mathbf{T}_{Patient}$ | **K** | **C** | **L** |
| 0.791 | 0.781 | 0.13 | 0.15 | 0.05 |
| 0.801 | 0.791 | 0.15 | 0.26 | 0.04 |
| 0.831 | 0.811 | 0.56 | 0.68 | 0.05 |

Table 4.6: p-value for K, C and L with comparable K and C

statistically independent. As can be seen in Table 4.9 our assumption was correct, as for $L$ the $p$-values were greater than 0.05 while for $K$ and $C$ $p$-values were less than 0.05, making it clear that there is an interdependence between $K$, $C$ and $L$.

| **HEALTHY EDUCATED** | | | |
|---|---|---|---|
| **T** | **K** | **C** | **L** |
| 0.651 | 26.94 ± 0.27 | 0.97 ± 0.01 | 4.40 ± 3.33 |
| 0.661 | 26.10 ± 0.28 | 0.95 ± 0.02 | 27.60 ± 8.83 |
| 0.671 | 25.25 ± 0.36 | 0.94 ± 0.01 | 37.50 ± 10.10 |

Table 4.7: Healthy Educated having comparable L with Patients

| **PATIENT** | | | |
|---|---|---|---|
| **T** | **K** | **C** | **L** |
| 0.651 | 24.86 ± 0.72 | 0.92 ± 0.01 | 7.80 ± 4.58 |
| 0.661 | 23.19 ± 0.96 | 0.90 ± 0.02 | 17.62 ± 8.01 |
| 0.671 | 21.58 ± 1.11 | 0.88 ± 0.02 | 37.09 ± 12.78 |

Table 4.8: Patients having comparable L with Healthy Educated

| | | p-value | | |
|---|---|---|---|---|
| $\mathbf{T}_{Healthy}$ | $\mathbf{T}_{Patient}$ | **K** | **C** | **L** |
| 0.651 | 0.651 | 0.01 | 0.01 | 0.55 |
| 0.661 | 0.661 | 0.01 | 0.02 | 0.41 |
| 0.671 | 0.671 | 0.01 | 0.02 | 0.98 |

Table 4.9: p-value for K, C and L with comparable L

The physical meaning of this maneuver addresses the following two questions: Assuming both healthy and schizophrenic populations have the same average degree and clustering coefficient, is the network proportionally effi-

cient? According to Table 4.6 the answer is no, which means that for the above values of $T$ the respective values of $K$ and $C$ of the patients are statistically equal compared to those of healthy while the values of $L$ of the patients are statistically different and specifically much greater than those of healthy. On the other hand assuming both healthy and schizophrenic populations have the same average shortest path length, is the network proportionally efficient? According to Table 4.9 the answer is also no, which means that for the above values of $T$ the respective values of $L$ of the patients are statistically equal compared to those of healthy while the values of $K$ and $C$ of the patients are statistically different and specifically much smaller than those of healthy. These syllogisms lead to the suggestion that schizophrenic patients need significantly more direct node (channel) connections in order to perform the same WM task.

## 4.3 Local Properties of Brain Networks

In the previous sections we analysed a graph's properties using some global metrics. Now, we try a different angle and we move to local network metrics. As mentioned in Chapter 1 there is a debate about the kind of model that would best describe brain networks. To this end we investigate many such models (Section 3.4). In the following sections we describe the metrics and models used to accomplish this and some interesting results are presented.

In order to focus on the local structure of a network we use two metrics trying to investigate how they can help us reveal the true model behind brain networks. The first one is the graphlet frequency and the second is the network graphlet degree distribution agreement. Let us first define these two metrics.

The number of different connected networks on $n$ nodes increases exponentially with $n$. For $n = 3,\ 4\ and\ 5$, there are 2, 6 *and* 21 different connected networks on $n$ nodes, respectively. To avoid terminology confusing network motifs with network subgraphs (motifs are special types of subgraphs), we use the term graphlet to denote a connected network with a small number of nodes. All $3-5$-node graphlets are presented in Figure 4.2. We use the graphlet frequency, i.e. the number of occurrences of a graphlet in a network, as a new network parameter.

We generalize the notion of the degree distribution as follows. The degree distribution measures, for each value of $k$, the number of nodes of degree $k$. In other words, for each value of $k$, it gives the number of nodes "touching" $k$ edges. Note that an edge is the only graphlet with two nodes; henceforth, we call this graphlet $G_0$. Thus, the degree distribution measures the following: how many nodes "touch" one $G_0$, how many nodes "touch" two $G_0$s, ..., how many nodes "touch" $k$ $G_0$s. Note that there is nothing special about

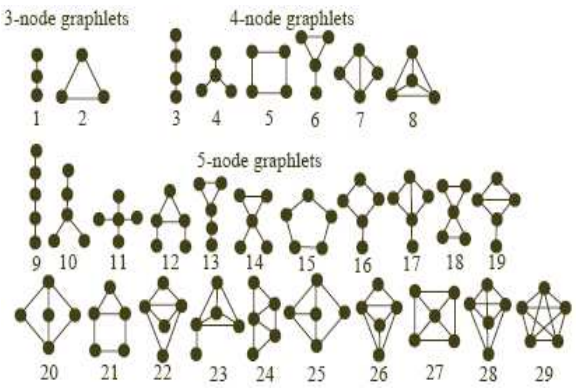



Figure 4.2: All 3-node, 4-node and 5-node connected networks (graphlets), ordered within groups from the least to the most dense with respect to the number of edges when compared to the maximum possible number of edges in the graphlet; they are numbered from 1 to 29

graphlet $G_0$ and that there is no reason not to apply the same measurement to other graphlets. Thus, in addition to applying this measurement to an edge, i.e., graphlet $G_0$, as in the degree distribution, we apply it to the twenty-nine graphlets $G_1, G_2, \ldots, G_{29}$.

When we apply this measurement to graphlets $G_1, G_2, \ldots, G_{29}$, we need to take care of certain topological issues. Lets take graphlet $G_1$ for instance. We ask how many nodes touch a $G_1$; however, note that it is topologically relevant to distinguish between nodes touching a G1 at an end or at the middle node. This is because nodes of a $G_1$ belong to one automorphism orbit, while the mid-node of a $G_1$ belongs to another. In Figure 4.3, we illustrate the partition of nodes of graphlets $G_1, G_2, \ldots, G_{29}$ into automorphism orbits (or just orbits for brevity); henceforth, we number the 73 different orbits of graphlets $G_1, G_2, \ldots, G_{29}$ from 0 to 72 (see [62] for more information). Analogous to the degree distribution, for each of these 73 automorphism orbits, we count the number of nodes touching a particular graphlet at a node belonging to a particular orbit. In this way, we obtain 73 distributions analogous to the degree distribution (actually, the degree distribution is the distribution for the $0^{th}$ orbit, i.e., for graphlet $G_0$). Thus, the degree distribution, which has been considered to be a global network property, is one in the spectrum of 73 "graphlet degree distributions (GDDs)" measuring local structural properties of a network. Note that GDD is measuring local structure, since it is based on small local network neighborhoods.

Having computed the 73 graphlet degree distributions for each network we wish to compare we are ready to measure the GDD agreement between them. Let $G$ a network (i.e. a graph). For a particular orbit $j$, let $d_G^j(k)$ be the sample distribution of the number od nodes in $G$ touching the appro-

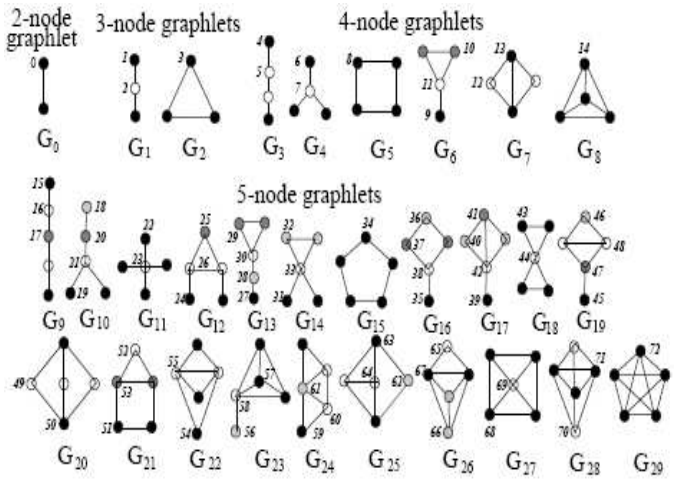



Figure 4.3: Automorphism orbits $0, 1, 2, \ldots, 72$ for the thirty 2, 3, 4, and 5−node graphlets $G_0, G_1, \ldots, G_{29}$. In a graphlet $G_i$, $i \in \{0, 1, \ldots, 29\}$, nodes belonging to the same orbit are of the same shade

priate graphlet $k$ times. That is $d_G^j$ represents the $j^{th}$ GDD. We define as the "normalized distribution" $N_G^j(k)$ as the fraction of the total area under the curve, over the entire GDD, devoted to degree $k$. Thus

$$N_G^j(k) = \frac{d_G^j(k)}{\sum_{k=1}^{\infty} d_G^j(k)} \,, \tag{4.1}$$

where formally the upper limit of the sum is unbounded but in practice it is finite due to the finite size of the graph. Finally, for two networks $G$ and $H$ and a particular orbit $j$, we define the "distance" $D^j(G, H)$ between their normalized $j^{th}$ distributions as

$$D^j(G, H) = \left( \sum_{k=1}^{\infty} [N_G^j(k) - N_H^j(k)]^2 \right)^{\frac{1}{2}} \,, \tag{4.2}$$

where again in practice the upper limit of the sum is finite due to the finite sample. The distance is between 0 and 1, where 0 means that $G$ and $H$ have identical $j^{th}$ GDDs and 1 means that their $j^{th}$ GDDs are far away. Next, we reverse $D^j(G, H)$ to obtain the $j^{th}$ GDD agreement:

$$A^j(G, H) = 1 - D^j(G, H) \,, \tag{4.3}$$

for $j \in \{0, 1, \ldots, 72\}$. Finally, the agreement between two networks $G$ and $H$ is the arithmetic mean of $A^j(G, H)$ over all $j$, i.e.

$$A(G, H) = \frac{1}{73} \sum_{j=0}^{72} A^j(G, H) \tag{4.4}$$

## 4.4 Local Properties Experimental Results

We have conducted experiments measuring the graphlet frequency for the schizophrenia and epilepsy data. Thus, we had 56.000 graphs, one for each group, band, state and threshold. For each graph of the real data we computed the same property for graphs that were generated according to the random (ER), generalized random (GR), scale-free (SF) and geometric random (GEO) models. For each model we generated 5 different graphs, having an extra number of 280.000 graphs which is an enormous amount of data to be analysed. We hypothized that our graphs should be better modeled with the GEO model. To this end we used the geometricness of a graph (Section 3.1) to help us identify the threshold regions that the graphs should be more geometric. If our intuition was correct the graphlet frequency metric should be approximated better from the GEO model.

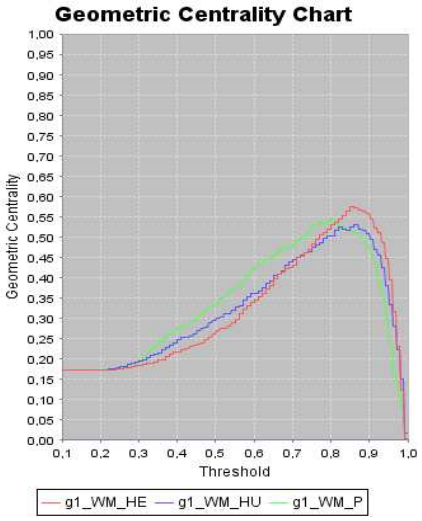


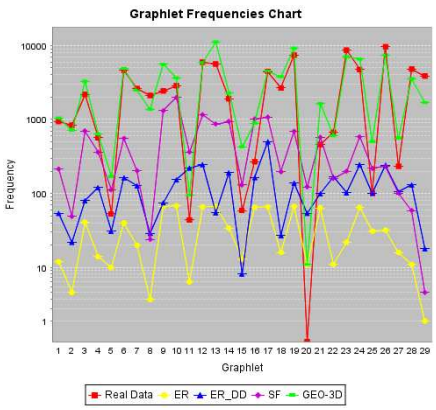

Figure 4.4: Geometricness for the three schizophrenic groups

We examined the way the fit between the plots of graphlet frequency of the real data and the four models altered as we increased the threshold. We observed that in the threshold regions where the real data graphs were too dense ($T \in [0 - 0.3]$) or too sparce ($T \in [0.8 - 1.0]$) the fit was poor. On the other hand when $T \in [0.3 - 0.8]$ the fit was good and in some cases very accurate.

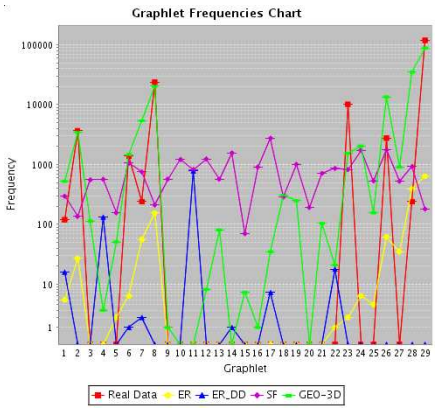
We also observed the way the geometricness of the graphs changed as we increased the threshold. We also found that there are three distinct threshold regions. The first is when the graph is too dense and the geometricness remains stable, the second when the graph is becoming sparce and the value of the metric is rapidly decreasing from its maximum value and the third in

the intermidiate case when the geometricness is increasing until it reaches a maximum. To our surpise the threshold region where the geometricness is increasing is approximately the same with the region where we had good fit between the plots of graphlet frequency of the real data and the GEO model.

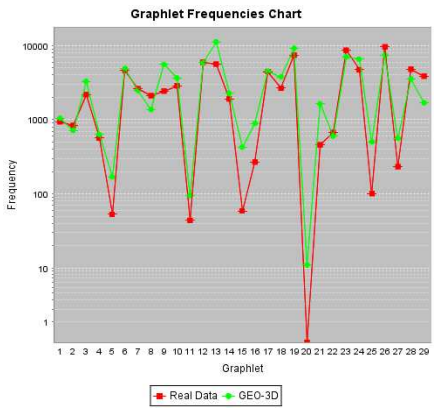
As an example we illustrate the case of healthy educated group in *gamma*1 band during WM and $T = 0.3$, where geometricness remains steady, and $T = 0.6$, where the geometricness is increasing (Figure 4.4). When we observe the graphlet frequency plot of the real data we notice that none of the ER, GR and SF models can fit it. Figure 4.5(b) and Figure 4.5(a) support this claim. On the other hand, as can be seen better in Figure 4.5(d) and Figure 4.5(c), the GEO model produces a poor fit when $T = 0.3$ but a very good fit when $T = 0.6$. This is true in the whole threshold range where geometricness is increasing, i.e. when $T \in [0.4 - 0.85]$.



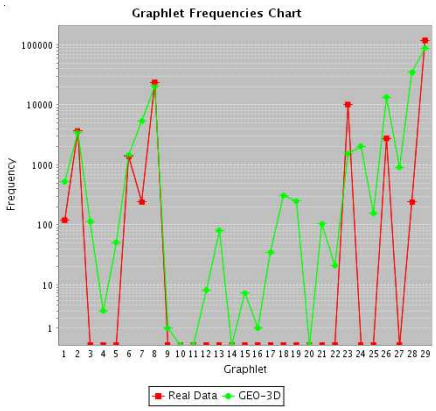

(a) Real Data vs all models ($T = 0.6$)




(b) Real Data vs all models ($T = 0.3$)




(c) Real Data vs GEO model ($T = 0.6$)




(d) Real Data vs GEO model ($T = 0.3$)

Figure 4.5: Graphlet frequency screenshot

# Chapter 5

# Tool Overview

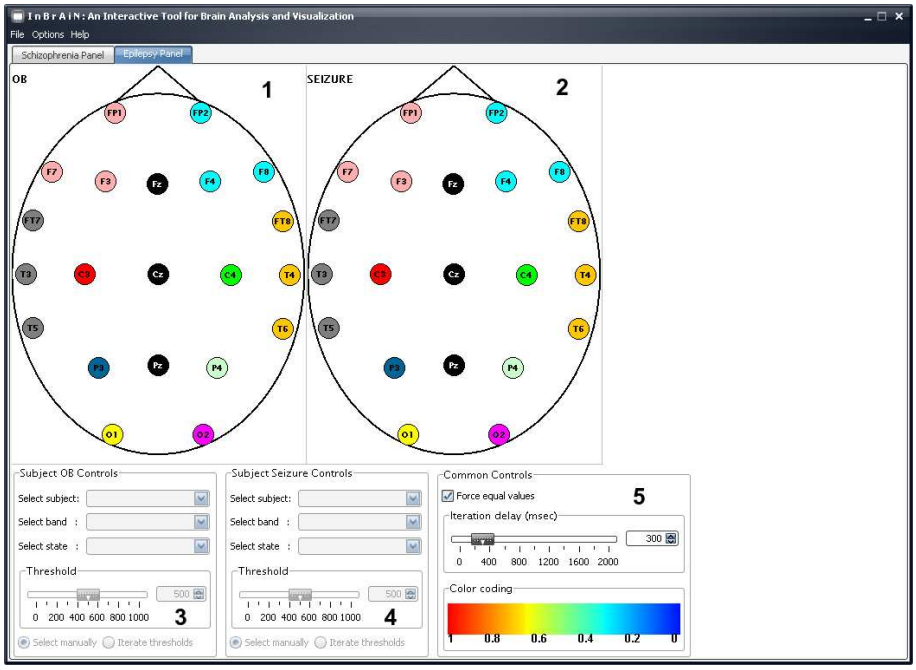
In this chapter we describe ways to visualize and analyze the graphs derived from EEG signals, by the technique mentioned in Section 4.1, using several graph drawing techniques and incorporate them smoothly into an easy-to-use framework in order to reveal and evaluate important properties of brain networks. First, we develop a static method which helps the doctors understand the inter-connectedness of the electrodes. Additionally, we invoke some well known graph algorithms in anticipation of a better, compared to the static method's, visualization outcome. These include force-directed and circular drawing algorithms.

## 5.1 Static Visualization Method

As described in Section 4.1, 30 electrodes for schizophrenia and 21 for epilepsy were used during the experiments. In order to visualize the topology of the emerged network we create a static framework where each electrode is depicted by a node placed in a position similar to the actual electrode's position on the human cortex. Thus, we manage to depict the brain network. This resulted in 1000 graphs to depict for each of the 40 healthy (educated and uneducated) and 20 patient subjects, for each of the two states, that is during Rest and WM, and for each of seven frequency bands, in terms of functional uniformity, acquired during the experiments. For the epilepsy case we have respectively one group consisting of 10 subjects, two states, namely OB and Seizure, and the same seven bands, namely $delta[0.5-4Hz]$, $theta[4-8Hz]$, $alpha1[8-10Hz]$, $alpha2[10-13Hz]$, $beta[13-30Hz]$, $gamma1[30-45Hz]$, $gamma2[45-90Hz]$. Figure 5.1 shows two screenshots of the visualizations for the schizophrenia and epilepsy case.

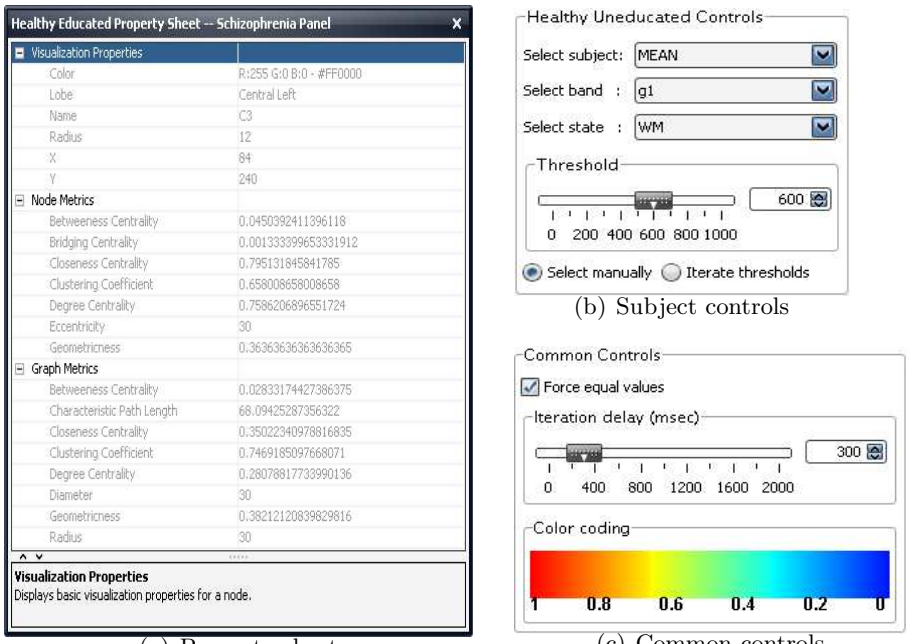
At this point some extra information about the two panels are needed. First, in Figure 5.1(a) we can see three subpanels numbered 1, 2 and 3. They correspond to the static visualization of the three different schizophrenic groups, namely HE, HU and P, respectively. In each of these panels a 2D


(a) Schizophrenia panel




(b) Epilepsy panel

Figure 5.1: Static cortex coordinates screenshot




(a) Property sheet

(b) Subject controls

(c) Common controls

Figure 5.2: Means to interact with the framework

representation of the human cortex is drawn as an oval shape. An arrow indicates the place where the subjects' eyes are. Inside this oval shape 30 circles exist that correspond to the electrodes placed on the subjects' head during the experiments. Each circle is named after the electrode it represents and is colored with a color indicating the lobe, that is the same area in the brain, it belongs.

Below them three other smaller panels exist numbered 4, 5 and 6. They provide means to controls various aspects of the graphs visualized for the three different schizophrenic groups (Figure 5.2(b) shows a detailed view of the HU controls). For example you can decide which subject, or the mean case for the group if you wish, band, state and specific threshold value to visualize. This is accomplished from the respective drop-down menus and slider present in each control panel. Additionally, you can choose to manually alter the threshold value or let the system automatically iterate all threshold values. Finally, a panel numbered 7 exists. From there we can control how fast the threshold iteration might be by specifying the miliseconds the system should wait until it increases the threshold value. Any change made to one group control panel can be applied to all three by selecting the appropriate check-box shown in Figure 5.2(c).

We need a way to visually differentiate the edges between two nodes based on the threshold. To this end we introduce a color coding shown

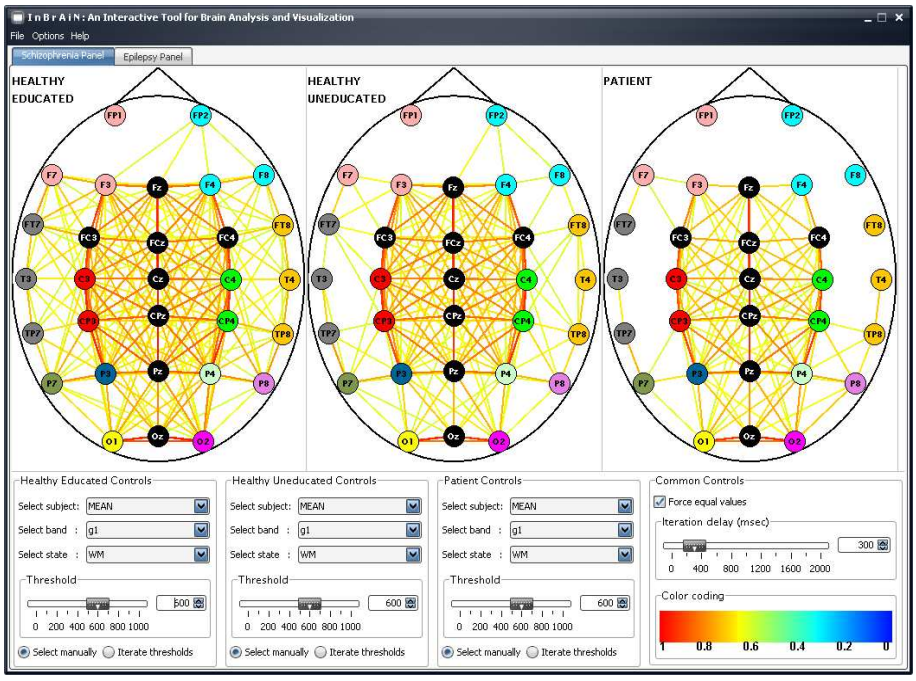



Figure 5.3: Schizophrenia screenshot

in Fig. 5.2(c). An edge's color turns to red if the nodes it connects are highly correlated. On the other hand the edge's color moves to blue. As an additional aid the more correlated two nodes are, the thicker the edge that connects them. So, when one iterates through the thresholds the most correlated nodes can easily be spoted as the edges that connect them are thicker and more red. In order to help this differentiation the more correlated two nodes are the later we draw the edge that connects them making it more visible than edges that correspond to lower correlations. Finally, many visualzation and graph properties of a specific node and the hole graph can be computed and presented to the user as shown in Figure 5.2(a). For an example case see Figure 5.3 where we visualize the mean case of each group, at $gamma1$ band during WM and $T = 600$.

Things are the same for the epilepsy case. As shown in Figure 5.1(b) there are two panels numbered 1 and 2 that correspond to a subject's OB and Seizure state. Additionally, there are two panels numbered 3 and 4 that control which subject, band, state and threshold value to visualize and a panel numbered 5 that controls the common functionalities. From now on we will only speek about the schizophrenia case as the same things hold for the epilepsy case as well.

This visualization is very useful to the doctors as they can see how the inter-connectedness in the brain changes and identify the critical thresholds

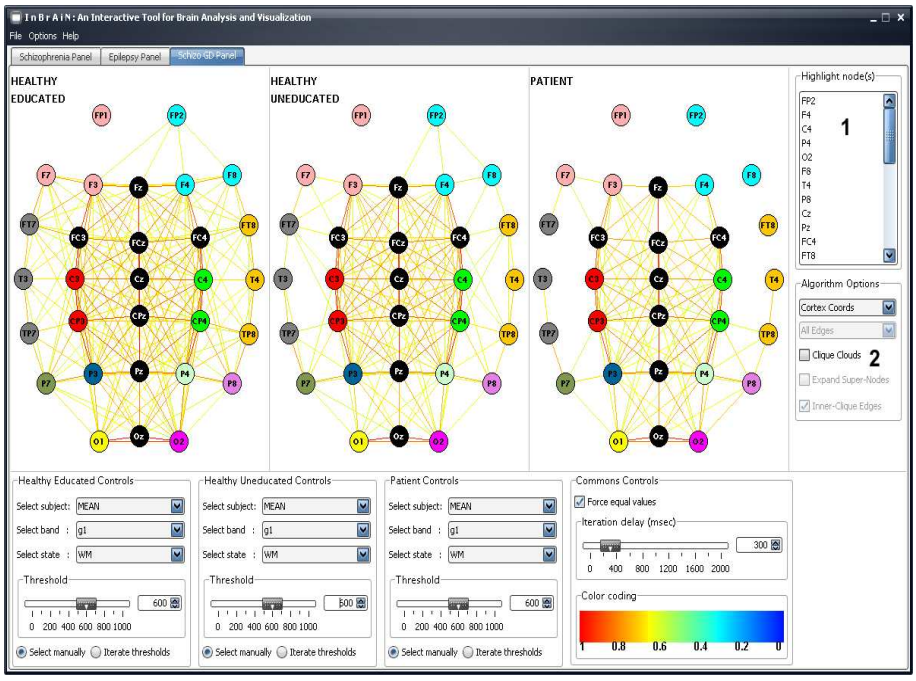


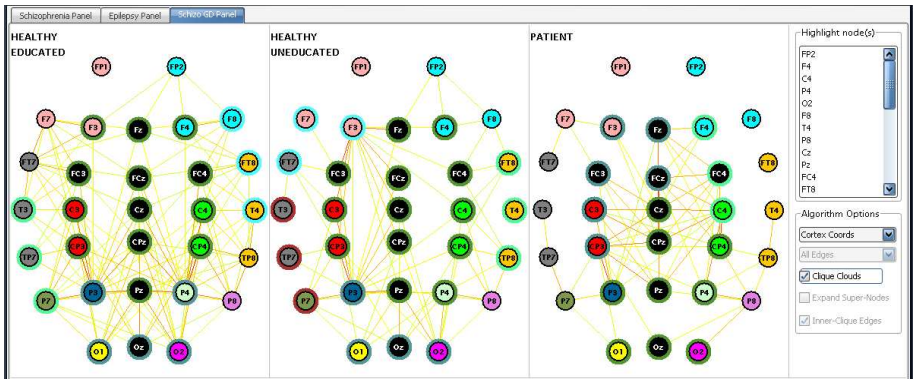

Figure 5.4: Schizophrenia graph drawing static screenshot

where this happens. Furthermore, they are able to verify already known properties and perhaps discover new ideas concerning the reasons for some disorders.

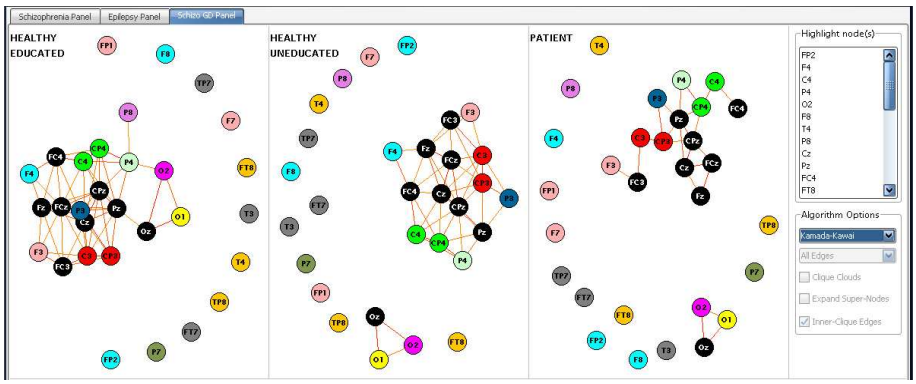
The proposed graphical framework is particularly useful as compared to the current clinical practice where the doctor can only observe the EEG between two specific nodes and at best have a few such EEGs on the screen. This is a severe drawback as they have to manually search for some prominent pair of nodes that according to the literature are held responsible for a specific disorder. Furthermore they can not focus at a specific threshold and see what the brain network looks like.

## 5.2 Graph Drawing Methods

Our next thought is to use some of the best known graph drawing techniques hoping for a clearer visualization outcome that could be of greater assistance. An example case of the static visualization method is shown in Figure 5.4. There are two extra panels numbered 1 and 2. The former can be used to quickly determine the position of a specific node or set of nodes. This is needed as the position of the nodes in all graph drawing methods is not fixed as is in the static case. The selected node is highlighted making it




(a) Node disjoint cliques "clouds"




(b) Kamada and Kawai algorithm

Figure 5.5: Schizophrenia graph drawing screenshots

easy to spot. The latter panel provides some functionality according to the visualization algorithm chosen. From here you can find all the node disjoint cliques. Around the nodes that belong to the same clique a small "cloud" of the same color is drawn and the edges that connect these nodes will be discarded (Figure 5.5(a)). Of cource you may change back to the normal view any time.

### 5.2.1 Force Directed Algorithms

First of all, we also transfer the static visualization method in this concept after embellishing it with some extra functionality. For example, you are given the opportunity to move the nodes from their fixed location if this is required. In the case where the graph is very dence you might want to see what connections are active only for a specific node or small set of nodes. This can be done by just clicking on the specisic node or nodes and only the

adjacent edges will be drawn.

The previous approach is very useful for doctors as they can see the topology of their patients' brains and be able to interact in some degree. Our next thought is to use some of the best known force-directed methods hoping for a clearer visualization outcome that could be of greater assistance. First of all, we use the Fruchterman and Reingold method [28], which is a simple model based on a combination of springs and electrical forces. Next we utilize the Kamada and Kawai method [41, 40], which is a more complex method that attempts to draw graphs such that the Euclidean distance between two vertices is near to the number of edges on the shortest graph-theoretic path between the vertices. Furthermore, the method of Eades is used [17]. Figure 5.5(b) shows an example case where we visualize the mean case of each group, at $gamma1$ band during WM and $T = 600$ using the Kamada and Kawai method.

The main drawback of using these methods is that the mapping of the nodes and the exact position of the electrodes on the brain is lost. In order to overcome this flaw, we color nodes that belong to the same lobe, which is a specific area of the human cortex, with the same color. We also introduce a tooltip with useful information about each node when one pauses over a specific node. Finally, you can instantly spot a specific node by just selecting it in the list present at the top right corner of the panel. What remains to be clarified is whether this approach could be potentially helpful to the doctors. The only knowledge that could be extracted from all the spring-embedder methods is the number of connected components and a nice placement of them.

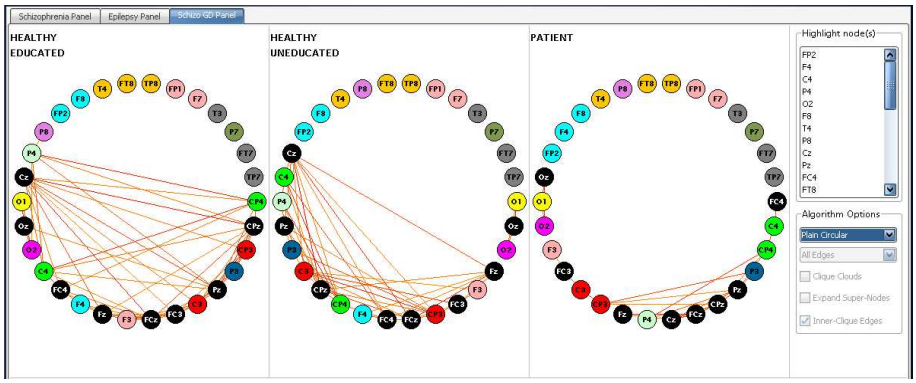
### 5.2.2 Circular Drawing Algorithms

A circular graph drawing is a visualization of a graph with the following charateristics:

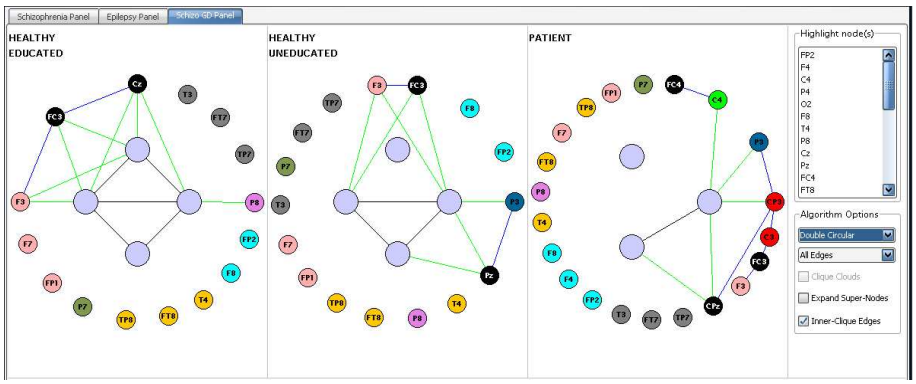
1. The graph is partitioned into clusters
2. The nodes of each cluster are placed on the circumference of an embedding circle
3. Each edge is drawn as a straight line segment

The problem of minimizing the number of crossings in a drawing is the well-known NP-Complete crossing number problem [29]. The more restricted problem of finding a minimum crossing embedding such that all the nodes are placed onto the circumference of a circle and all edges are represented with straight lines is also NP-Complete as proven in [68].

As described in [80, 71], a linear time technique, CIRCULAR, was introduced to produce circular graph drawings of biconnected graphs on a single embedding circle. In order to produce the drawings with fewer crossings,




(a) Circular algorithm




(b) Double Circular algorithm

Figure 5.6: Schizophrenia graph drawing circular screenshots

the authors presented an algorithm which tends to place edges toward the outside of the embedding circle while nodes are placed near their neighbors. The worst-case time requirement of CIRCULAR is O(m), where m is the number of edges. An important property of this technique is the guarantee that it will find a zero-crossing drawing for a given biconnected graph in case one exists.

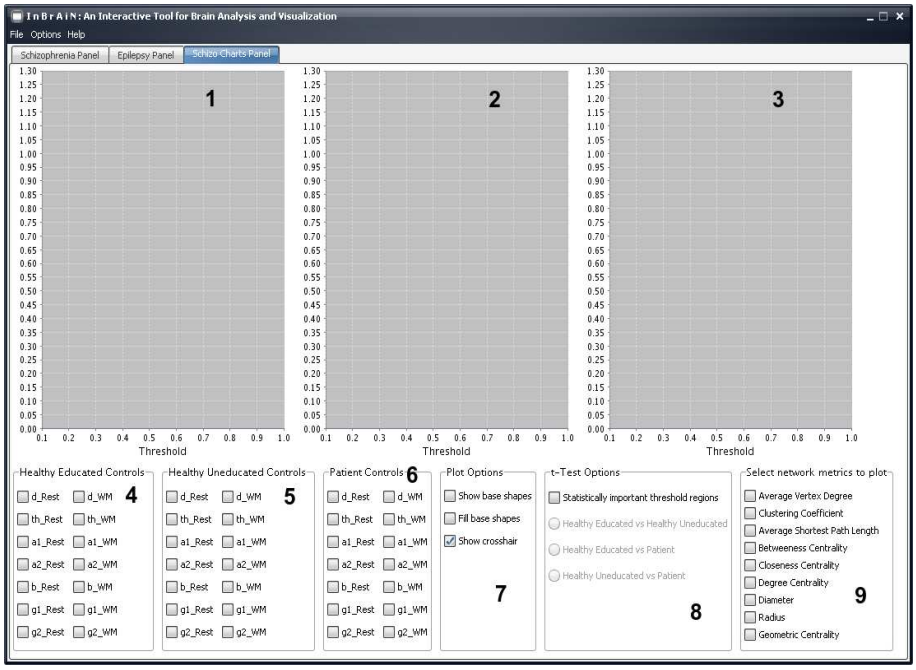
In [80, 72] another linear time algorithm, CIRCULAR-Nonbiconnected, was introduced for producing circular drawings of nonbiconnected graphs on a single embedding circle. Given a nonbiconnected graph G, it was first decomposed into biconnected components. In this technique, the layout of the resulting block-cutpoint tree on a circle was first produced and then the one for each biconnected component with a variant of CIRCULAR. The worst-case time requirement for CIRCULAR-Nonbiconnected is O(m) if we use a variant of CIRCULAR to layout each biconnected component. The resulting drawings have the property that the nodes of each biconnected

component appear consecutively. Furthermore, the order of the biconnected components on the embedding circle are placed according to a layout of the accompanying block-cutpoint tree and therefore the biconnectivity structure of a graph is displayed even though all of the nodes appear on a single circle.

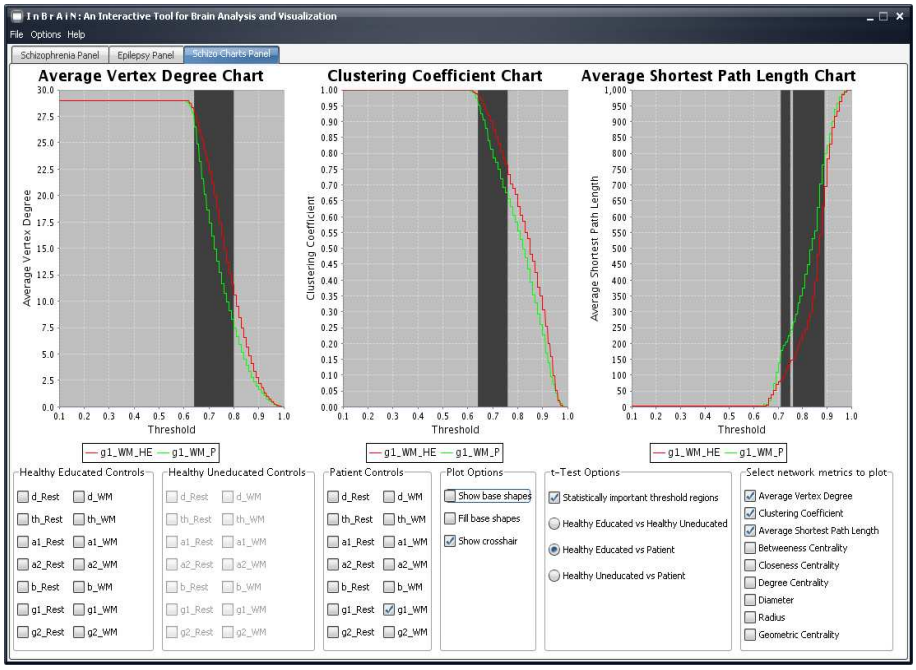
We incorporate the last algorithm in our framework. Since our networks are clustered and these clusters are important in identifying the regions of the brain that are active at the same time while performing a specific task. We implement two variants of this technique. In the first variant (Simple-Circular) we place the nodes according to the position assigned to them by the CIRCULAR-Nonbiconnected algorithm (Fig. 5.6(a)). In the second variant (Double-Circular) node-disjoint cliques in our graph are first identified and placed in the circumference of an inner cycle. These cliques can be displayed as super-nodes or can be decomposed to the vertices that form the clique, which are placed in a cycle with center the clique's position. Additionally, by pausing on a super-node one gets a useful tooltip that gives information about the nodes that form the specific node and the lobe each node belongs to. To find the cliques' positions we apply the CIRCULAR-Nonbiconnected algorithm in an effort to minimize the crossings among these important clusters. This leads to a clearer drawing that helps us understand how tightly the connected areas of the brain, that are active while performing a specific WM task, interact with other tightly connected areas. By doing that we are able to determine a specific node or set of nodes that are responsible for the poor connectivity or the disconnection of certain areas of the brain with each other. The remaining nodes, that do not belong to a clique, are placed in the circumference of an outer cycle. In order to maintain a clear drawing we place the nodes that are adjacent to some cliques in their mean angle (Fig. 5.6(b)).

## 5.3 Analysis Features

As described before there are many graph properties that could help us discriminate between different kind of networks and many of them can be plotted using this framework (Figure 5.7(a)). Here you can see the plots of three different metrics simultaniously in panels 1, 2 and 3. You can control which metrics to plot from the check-buttons in rectangle 9. For each metric you choose you can specify the group, band and condition to plot. This is easily accomplished by selecting the appropriate check-button in rectangles 4, 5 and 6 that are the controls for the HE, HU and P groups respectively. You can change the way each data point of each plot appears by using the chack-buttons of rectangle 7. Specifically, you can show a shape at each point or even fill it with the plots color. Additionally, you can invoke a crosshair in all plots. You can alter many of the chart properties such as the title and line-color by right-cliching on a chart. This will cause a property sheet to




(a) Charts panel




(b) Statistical important threshold regions

Figure 5.7: Schizophrenia charts screenshots

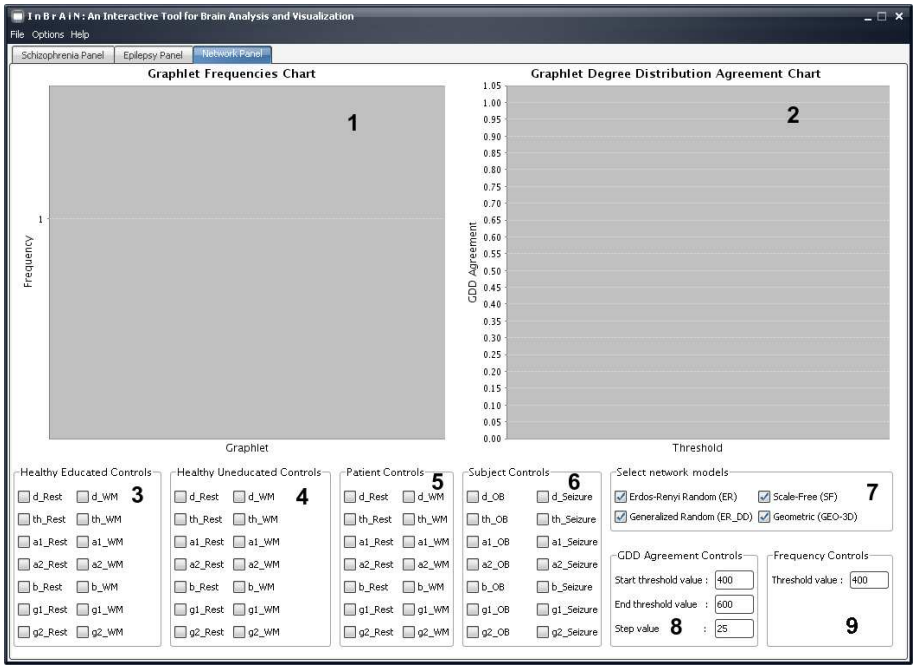
appear that will help you modify or inspect some of the chart properties.

Furthermore, by performing a statistical test, like t-test, to the input data, you can identify the threshold regions, where a statistically important difference exists. This is indicated by a vertical black marker as shown in Fig. 5.7(b).

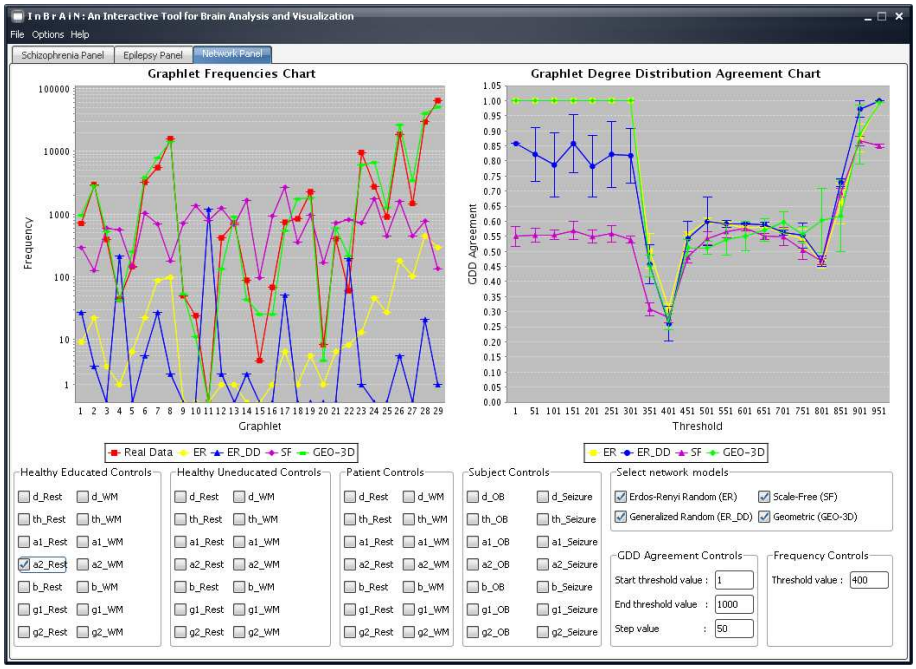
The last feature is particularly important. The statistically important threshold regions only offer information about the overall behavior of the network. For example in the case where the graph of the HE group has greater C compared to the graph of the P group for some band, state and threshold, the former graph has better local organization than the latter. However, we do not get any details as to where this better organization lies in the brain, which nodes make this happen, in what way these nodes are connected to the rest of the graph and generally what the two networks look like. In order to get this information and understand its impact to the topology of the graph, one should be able to see a static or a graph drawing visualization of the graph at these important threshold regions. This is accomplished by clicking on the markers. Getting to the visualization framework, it automatically focuses on the specified threshold. That reveals the exact interconnectedness which results in the aforementioned difference, making it easier for the doctors to discover where the problem might be.

Additionally, many local network properties can be visualized using the tool (Figure 5.8(a)). Here you can plot the graphlet frequency, panel 1, and the network GDD agreement, panel 2, for the real graph and the graphs produced from four different models such as Erdös-Rényi, Generalized Random, Scale-Free and Geometric Random. From panels 3, 4, 5, 6 and 7 you can choose the group, band, state and model. Furthermore, the threshold and the threshold region for which the frequency and the agreement would be plotted are controled from panels 9 and 8. In Figure 5.8(b) you can see the graphlet frequancy plot for the Healthy educated group in $alpha1$ band during rest for $T = 400$ and the GDD agreement plot for $T \in [1, \ldots, 1000]$.

This feature is very useful as it provides the means to investigate the way brain networks change, as the threshold alters. In the meantime you are able to focus on the local structure of the network rather than global one. This is very enticing as you have another option to study the exact interconnectedness of brain networks. In combination with the other visualization feautures it could reveal unknown properties.




(a) Initial network panel




(b) Network panel

Figure 5.8: Network panel screenshot

## 5.4 Designing the User Interface

The basic principles that can be applied to design an interactive application in order to improve its usability can be organized in three main categories [15] :

**Learnability** − the ease with which new users can begin effective interaction and achieve maximal performance.

**Flexibility** − the multiplicity of ways in which the user and system exchange information.

**Robustness** − the level of support provided to the user in determining successful achievement and assessment of goals.

In the following, we will subdivide these main categories into more specific principles.

### 5.4.1 Learnability

Learnability concerns the features on the interactive system that allow novice users to understand how to use it initially and how to attain a maximal level of performance. Table 5.4.1 contains a summary of the specific principles that support learnability.

| Principle | Definition |
|---|---|
| Predictability | Support for the user to determine the effect of future action based on past interaction history |
| Synthesizability | Support of the user to access the effect of past operations on the current state |
| Familiarity | The extent to which the knowledge and experience of a user in other real-world or computer-based domains can be applied when interacting with the new system |
| Generalizability | Support of the user to extend knowledge of specific interaction within and a across applications to other similar situations |
| Consistency | Likeness in input-output behavior arising from similar situations or similar task objectives |

Table 5.1: Summary of principles affecting learnability

### 5.4.2 Flexibility

Flexibility refers to the multiplicity of ways the end-user and the system exchange information. We identify several that contribute to the flexibility of interaction and these are summarized in Table 5.4.2.

| Principle | Definition |
|---|---|
| Dialog initiative | Allowing user freedom from artificial constraints on the input dialog imposed by the system |
| Multi-threading | Ability of the system to support user interaction pertaining to more than one task at a time |
| Task migratability | The ability to pass control for the execution of a given task so that it becomes either internalized by the user or the system or shared between them |
| Substitutivity | Allowing equivalent values of input and output to be arbitrarily substituted for each other |
| Customizability | Modifiability of the user interface by the user or the system |

Table 5.2: Summary of principles affecting flexibility

### 5.4.3 Robustness

In a work or task domain, a user is engaged with a computer in order to achieve some set of goals. The robustness of that interaction covers features that support the successful achievement and assessment of the goals. A summary of the principles that support robustness is presented in Table 5.4.3.

| Principle | Definition |
|---|---|
| Observability | Ability of the user to evaluate the internal state of the system from its perceivable representation |
| Recoverability | Ability of the user to take corrective action once an error has been recognised |
| Responsiveness | How the user perceives the rate of communication with the system |
| Task conformance | The degree to which the system services support all of the tasks the user wishes to perform and in the way that the user understands them |

Table 5.3: Summary of principles affecting robustness

### 5.4.4 Heuristic Evaluation

A heuristic is a guideline or general principle or rule of thumb that can guide a design decision or be used to critique a decision that has already been made. To aid the process of discovering usability problems, a set of 10 heuristics were developed by Nielsen [58] and are presented below. In order

to support the usability of the application during its design these heuristics were followed.

1. **Visibility of system status** The system should always keep users informed about what is going on, through appropriate feedback within reasonable time.

2. **Match between system and the real world** The system should speak the users' language, with words, phrases and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order.

3. **User control and freedom** Users often choose system functions by mistake and will need a clearly marked "emergency exit" to leave the unwanted state without having to go through an extended dialogue. Support undo and redo.

4. **Consistency and standards** Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform conventions.

5. **Error prevention** Even better than good error messages is a careful design which prevents a problem from occurring in the first place. Either eliminate error-prone conditions or check for them and present users with a confirmation option before they commit to the action.

6. **Recognition rather than recall** Minimize the user's memory load by making objects, actions, and options visible. The user should not have to remember information from one part of the dialogue to another. Instructions for use of the system should be visible or easily retrievable whenever appropriate.

7. **Flexibility and efficiency of use** Accelerators – unseen by the novice user – may often speed up the interaction for the expert user such that the system can cater to both inexperienced and experienced users. Allow users to tailor frequent actions.

8. **Aesthetic and minimalist design** Dialogues should not contain information which is irrelevant or rarely needed. Every extra unit of information in a dialogue competes with the relevant units of information and diminishes their relative visibility.

9. **Help users recognize, diagnose, and recover from errors** Error messages should be expressed in plain language (no codes), precisely indicate the problem, and constructively suggest a solution.

10. **Help and documentation** Even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user's task, list concrete steps to be carried out, and not be too large.

# Chapter 6

# Conclusions and Future Work

An insight into the EEG correlation networks formed in the human brain is given using the visualization methods mentioned in the previous section, for the first time. The static method was a good initial attempt. It gave the doctors the opportunity to see how all the nodes (electrodes) correlate with each other in a visual, quick and easy to use way compared to the manual and time consuming techniques used in the past. It would be interesting to see if a 3-dimensional version of the static method could further assist the doctors.

The visualizations yielded from the force directed methods still need to be evaluated. Doctors were able to identify the structure of the underlying network, as they revealed the connected components that formed the graphs and made a clearer drawing of them possible.

However, the two circular variants were more useful. The first variant showed with more detail the structure of the network, as doctors could easily identify the biconnected components of the graph. The second variant was of greater help as it revealed the important areas of the brain, which are co-activated. Additionally, it showed the way these areas are linked to each other and gave a hint as to where to search for the disorders present in schizophrenia. An interesting modification of the second variant would be to place in the inner cycle different kind of nodes rather than the node disjoint cliques. These nodes could be the most highly ranked nodes according to some metric or could even be user defined taking advantage of the doctors' expertise.

Finally, the plotting framework gave the doctors the opportunity to see how important metrics alter and how this influences the network's topology as one can change the focus from the graph's overall state to the detailed node interconnectedness. They were able to verify the assymetry in the function of the frontal lobe as well as identify prominent disturbances dur-

ing WM for the connections of the frontal and temporal lobes. Furtermore, the extensive experiments on local network properties revealed the potential existence of a new model that may describe brain functional networks, namely the geometric random model.

Future work includes:

- Further analysis to verify that brain networks are geometric
- New metrics/measures that could help us identify the control from patient graphs
- Find properties that would best describe brain networks
- A new visualization algorithm that would take into consideration the new measures and properties of brain networks

# Bibliography


[1] P. Abry. "Multirésolutions, algorithmes de décomposition, invariance d'échelles". Diderot Editeur, Paris, 1997.

[2] W. Aiello, F. Chung, and L. Lu. *"A random graph model for massive graphs"*. *In Proceedings of the thirty-second annual ACM symposium on Theory of computing*, pages 171–180, 2000.

[3] R. Albert, I. Albert, and G. L. Nakarado. *"Structural vulnerability of the north american power grid"*. *Physical Review E*, 69:025103, 2004.

[4] J. Arnhold, P. Grassberger, K. Lehnertz, and C. E. Elger. *"A robust method for detecting interdependencies: application to intracranially recorded EEG"*. *Physica D*, 134:419–430, 1999.

[5] A. L. Barabási and R. Albert. *"Emergence of scaling in random networks"*. *Science*, 286(5439):509–12, 1999.

[6] A. Barrat, M. Barthélemy, R. Pastor-Satorras, and A. Vespignani. *"The architecture of complex weighted networks"*. *Proceedings of the National Academy of Science USA*, 101(11):3747–3752, 2004.

[7] D. S. Bassett and E. Bullmore. *"Small-World Brain Networks"*. *The Neuroscientist*, 12(6):512–523, 2006.

[8] E. A. Bender and E. R. Canfield. *"The asymptotic number of labeled graphs with given degree sequences"*. *Journal of Combinatorial Theory A*, 24:296–307, 1978.

[9] G. E. Berrios and R. Porter. "A history of clinical psychiatry : the origin and history of psychiatric disorders". London; New Brunswick, NJ, 1999.

[10] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D. U. Hwang. *"Complex networks: structure and dynamics"*. *Phys. Rep.*, 424:175–308, 2006.

[11] T. H. Bullock, M. C. McClune, J. Z. Achimowicz, V. J. Iragui-Madoz, R. B. Duckrow, and S. S. Spencer. *"EEG coherence has structure in the millimeter domain: subdural and hippocampal recordings from epileptic patients". Electroencephalogr. Clin. Neurophysiol.*, 95:161–77, 1995.

[12] F. Chung and L. Lu. *"The average distances in random graphs with given expected degrees". Proceedings of the National Academy of Science USA*, 99:15879–15882, 2002.

[13] R. Cohen and S. Havlin. *"Scale-free networks are ultrasmall". Physical Review Letters*, 90(058701), 2003.

[14] I. Daubechies. *"The Wavelet transform time-frequency localization and signal analysis". IEEE Trans. Inform. Theory*, 36:961–1004, 1990.

[15] A. Dix, J. E. Finlay, G. D. Abowd, and R. Beale. "Human–Computer Interaction". Prentice Hall, 2004.

[16] R. J. Porter E. S. Goldensohn and P.A. Schwartzkroin. *"The American Epilepsy Society: An Historic Perspective on 50 Years of Advances in Research". Epilepsia*, 38(1):124–150, 1997.

[17] P. Eades. *"A Heuristic for Graph Drawing". Congr. Numer.*, 42:149–160, 1984.

[18] V. M. Eguíluz, D. R. Chialvo, G. A. Cecchi, M. Baliki, and A. V. Apkarian. *Scale-free brain functional networks.* 2005.

[19] P. Erdös and A. Rényi. *"On random graphs". Publicationes Mathematicae*, 6:290–297, 1959.

[20] P. Erdös and A. Rényi. *"On the evolution of random graphs". Publ. Math. Inst. Hung. Acad. Sci.*, 5:17–61, 1960.

[21] P. Erdös and A. Rényi. *"On the strength of connectedness of a random graph". Acta Mathematica Scientia Hungary*, 12:261–267, 1961.

[22] B. J. Fisch, B. J. Fisch, and R. Spehlman. "Fisch and Spehlmann's EEG primer : basic principles of digital and analog EEG". Elsevier, Amsterdam; New York, 1999.

[23] L. C. Freeman. *"A set of measures of centrality based on betweenness". Sociometry*, 40:35–41, 1977.

[24] L. C. Freeman. *"Centrality in social networks: Conceptual clarification". Social Networks*, 1:215–239, 1979.

[25] W. J. Freeman, R. Kozma, and P. J. Werbos. *"Biocomplexity: adaptive behavior in complex stochastic dynamical systems". Biosystems*, 59:109–23, 2001.

[26] N. E. Friedkin. *"Theoretical foundations for centrality measures"*. *American Journal of Sociology*, 96(6):1478–1504, 1991.

[27] K. J. Friston, C. D. Frith, P. F. Liddle, and R. S. J. Frackowiak. *"Functional connectivity: the principal component analysis of large (PET) data sets"*. *J. Cereb. Blood Flow Metab.*, 13:5–14, 1993.

[28] T. Fruchterman and E. Reingold. *"Graph Drawing by Force-Directed Placement"*. *Softw.-Pract.Exp.*, 21(11):1129–1164, 1991.

[29] M. Garey and D. Johnson. "Computers and Intractability: A Guide to the Theory of NP-Completeness". Freeman, 1979.

[30] M. T. Gastner and M. E. J. Newman. *"The spatial structure of networks"*. *The European Physical Journal B*, 49:247, 2006.

[31] J. W. Gibbs. *"Fourier Series"*. *Nature*, 59, 1898.

[32] C. M. Gray, P. Konig, A. K. Engel, and W. Singer. *"Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties"*. *Nature*, 338:334–337, 1989.

[33] C. M. Gray and W. Singer. *"Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex"*. *Proc Natl Acad Sci U S A*, 86:1698–1702, 1989.

[34] A. Grinsted, J. C. Moore, and S. Jevrejeva. *"Application of the cross wavelet transform and wavelet coherence to geophysical time series"*. *Nonlinear Processes in Geophysics*, 11:561–566, 2004.

[35] R. Guimerà, S. Mossa, A. Turtschi, and L. A. N. Amaral. *"The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles"*. *Proceedings of the National Academy of Science USA*, 102:7794–7799, 2005.

[36] Y. Hayashi. *"A review of recent studies of geographical scale-free networks"*. *physics/0512011*, 2005.

[37] H. H. Jasper. *"The 10-20 electrode system of the International Federation in Electroencephalography and Clinical Neurophysiology"*. *EEG Journal*, 10:370–375, 1958.

[38] E. R. John. *"A field theory of consciousness"*. *Conscious Cogn.*, 9:184–213, 2001.

[39] M. Kaiser and C. C. Hilgetag. *"Spatial growth in real-time networks"*. *Phys. Rev. E*, 69:036103, 2004.

[40] T. Kamada. *"Visualizing Abstract Objects and Relations". World Scientific Series in Computer Science*, 1989.

[41] T. Kamada and S. Kawai. *"An Algorithm for Drawing General Undirected Graphs". Inform. Process. Lett.*, 31:7–15, 1989.

[42] R. Kinney, P. Crucitti, R. Albert, and V. Latora. *"Modeling cascading failures in the north american power grid". European Physical Journal B*, 46:101–107, 2005.

[43] T. Koenig, L. Prichep, D. Lehmann, P. V. Sosa, E. Braeker, H. Kleinlogel, R. Isenhart, and et al. *"Millisecond by millisecond, year by year: normative EEG microstates and developmental stages". Neuroimage*, 16:41–8, 2002.

[44] V. Latora and M. Marchiori. *"Is the boston subway a small-world network?". Physica A*, 314:109–113, 2002.

[45] M. N. Livanov. "Spatial organization of cerebral processes". Wiley, New York, 1977.

[46] S. Milgran. *"The small world problem". Psychology Today*, 1(1):60–67, 1967.

[47] R. Milo, N. Kashtan, S. Itzkovitz, M. E. J. Newman, and U. Alon. *"On the uniform generation of random graphs with prescribed degree sequences". cond-mat/0312028*, 2003.

[48] M. Molloy and B. Reed. *"A critical point of random graphs with a given degree sequence". Random Structures and Algorithms*, 6:161–180, 1995.

[49] M. Molloy and B. Reed. *"The size of the largest component of a random graph on a fixed degree sequence". Combinatorics, Probability and Computing*, 7:295–306, 1998.

[50] R. Monasson. *"Diffusion, localization and dispersion relations on "small-world" lattices". European Physical Journal B*, 12(555), 1999.

[51] F. Mormann, K. Lehnertz, P. David, and C. Elger. *"Mean phase coherence as a measure for phase synchronization and its application to the EEG of epilepsy patients". Physica D*, 144:358–369, 2000.

[52] M. E. J. Newman. "Random graphs as models of networks". In Handbook of Graphs and Networks, (Bornholdt, S. & Schuster, H. G., eds), Wiley-VHC, Berlin, 2002.

[53] M. E. J. Newman. *"Structure and function of complex networks". SIAM Review*, 45(2):167–256, 2003.

[54] M. E. J. Newman, S. H. Strogatz, and D. J. Watts. *"Random graphs with arbitrary degree distributions and their applications"*. *Physical Review E*, 64:026118–1, 2001.

[55] M. E. J. Newman and D. J. Watts. *"Renormalization group analysis of the small-world network model"*. *Physical Review Letters A*, 263:341–346, 1999.

[56] M. E. J. Newman, D. J. Watts, and S. H. Strogatz. *"Random graphs with arbitrary degree distributions and their applications"*. *Proceedings of the National Academy of Science USA*, 99:2566–2572, 2002.

[57] E. Niedermeyer and F. H. Lopes da Silva. "Electroencephalography : basic principles, clinical applications, and related fields". Williams and Wilkins, Baltimore, 1999.

[58] J. Nielsen. "Heuristic evaluation. In Usability Inspection Methods". John Wiley, New York, 1994.

[59] P. L. Nunez. *"Toward a quantitative description of large-scale neocortical dynamic function and EEG"*. *Behav. Brain Sci.*, 23:371–437, 2000.

[60] Commission on Classification and Terminology ILAE. *"Proposal for Revised Clinical and Electroencephalographic Classification of Epileptic Seizures"*. *Epilepsia*, 22:489–501, 1981.

[61] A. V. Oppenheim, A. S. Willsky, and S. H. Nawab. "Signals and systems". Prentice Hall, Upper Saddle River, N.J., 1997.

[62] N. Pržulj. *"Biological Nework Comparison Using Graphlet Degree Distribution"*. *Bioinformatics*, 2006.

[63] R. Q. Quiroga, A. Kraskov, T. Kreuz, and P. Grassberger. *"Performance of different synchronization measures in real data: a case study on electroencephalographic signals"*. *Phys. Rev. E*, 65(041903), 2002.

[64] A. Rapoport. *"Nets with distance bias"*. *Bulletin of Mathematical Biophysics*, 13:85–91, 1951.

[65] A. Rapoport. *"Spread of information through a population with sociostructural bias: I. Assumption of transitivity"*. *Bulletin of Mathematical Biophysics*, 15:523–533, 1953.

[66] A. Rapoport. *"Contribution to the theory of random and biased nets"*. *Bulletin of Mathematical Biophysics*, 19:257–277, 1957.

[67] A. J. Rowan and E. Tolunsky. "A primer of EEG : with a mini-atlas". Butterworth-Heinemann, Philadelphia, PA, 2003.

[68] K. Nakajima S. Masuda, T. Kashiwabara and T. Fujisawa. *"On the NP-Completeness of a Computer Network Layout Problem". Proc. IEEE 1987 International Symposium on Circuits and Systems, Philadelphia, PA*, pages 292–295, 1987.

[69] G. Sabidussi. *"The centrality index of a graph". Psychometrika*, 31:581–603, 1966.

[70] C. E. Shannon. "Communication in the Presence of Noise. Proc. of the IRE". Proc. of the IRE, 194.

[71] J. M. Six and I. G. Tollis. *"Circular Drawings of Biconnected Graphs". Proc. of ALENEX '99, LNCS 1619, Springer-Verlag*, pages 57–73, 1999.

[72] J. M. Six and I. G. Tollis. *"Circular Drawings of Telecommunication Networks". Advances in Informatics, Selected Papers from HCI '99, D. I. Fotiadis and S. D. Nikolopoulos, Eds., World Scientific*, pages 313–323, 2000.

[73] O. Sporns. *"Network analysis, complexity and brain function". Complexity*, 8:56–60, 2002.

[74] C. J. Stam and B. W. van Dijk. *"Synchronization likelihood: an unbiased measure of generalized synchronization in multivariate data sets". Physica D*, 163:236–251, 2002.

[75] F. Takens, D. A. Rand, and L. S. Young. *"Detecting strange attractors in turbulence". Dynamical systems and Turbulence, Lecture Notes in Mathematics*, 898:366–381, 1981.

[76] R. W. Thatcher, P. J. Krause, and M. Hrybyk. *"Cortico-cortical associations and EEG coherence: a two-compartmental model". Electroencephalogr. Clin. Neurophysiol.*, 64:123–43, 1986.

[77] J. Theiler. *"Spurious dimension from correlation algorithms applied to limited time-series data". Phys. Rev. A*, 34(2427), 1986.

[78] C. Torrence and G. P. Compo. *"A practical Guide to Wavelet Analysis". Bull. Am. Meteorol. Soc.*, 79:61–78, 1998.

[79] M. Tudor, L. Tudor, and K. I. Tudor. *"Hans Berger (1873-1941)– the history of electroencephalography". Acta Med Croatica*, 59:307–313, 2005.

[80] J. M. Six (Urquhart). *"Vistool: A Tool For Visualizing Graphs". Ph.D. Thesis, The University of Texas at Dallas*, 2000.

[81] S. Wasserman and K. Faust. *"Social Network Analysis: Methods and Applications". Cambridge University Press*, 1994.

[82] D. J. Watts. "Small worlds : the dynamics of networks between order and randomness". Princeton University Press, 1999.

[83] D. J. Watts. "Six Degrees. The Science of a Connected Age". W.W. Norton & Company, 2003.

[84] D. J. Watts and S. H. Strogatz. *"Collective dynamics of 'small-world' networks"*. *Nature*, 393:440–442, 1998.

# Appendix A

# Computing Correlations

## A.1 Mean Squared Coherence (MSC)

Let us suppose we have two simultaneously measured discrete time series $x_n$ and $y_n$, $n = 1 \dots N$. The most commonly used linear synchronization method is the cross-correlation function ($C_{xy}$) defined as:

$$C_{xy}(\tau) = \frac{1}{N-\tau} \sum_{i=1}^{N-\tau} \left( \frac{x_i - \overline{x}}{\sigma_x} \right) \left( \frac{y_{i+\tau} - \overline{y}}{\sigma_y} \right) \tag{A.1}$$

where $\overline{x}$ and $\sigma_x$ denote mean and variance, while $\tau$ is the time lag. MSC or simply coherence is the cross spectral density function $S_{xy}$, which is simply derived via the FFT of Eq. A.1, normalized by their individual autospectral density functions. However, due to finite size of neural data one is able to actually estimate the true spectrum, known as periodogram, using smoothing techniques (e.g. Welch's method). Thus, MSC is calculated as:

$$\gamma_{xy}(f) = \frac{|\langle S_{xy}(f) \rangle|^2}{|\langle S_{xx}(f) \rangle| |\langle S_{yy}(f) \rangle|} \tag{A.2}$$

Where $\langle \cdot \rangle$ indicates window averaging in the case of Welch's method. The estimated MSC for a given frequency $f$ ranges between 0 (no coupling) and 1 (maximum linear interdependence).

## A.2 The Continuous Wavelet Transform (CWT)

Over the past decade, the WT has been developed into an important tool for time series analysis that contains nonstationary power (such as the EEG signal) at many different frequencies [14]. The CWT of a discrete sequence $x_n$ with time spacing $\delta t$ and $N$ data points ($n = 0 \dots N-1$) is defined as the convolution of $x_n$ with consecutive scaled and translated versions of the wavelet function $\psi_0(\eta)$:

$$W_n^X(s) = \sqrt{\delta t/s} \sum_{n'=0}^{N-1} x_{n'} \psi_0 * [(n'-n)\delta t/s] \tag{A.3}$$

$$\psi_0(\eta) = \pi^{-1/4} e^{i\omega_0 \eta} e^{-\eta^2/2} \tag{A.4}$$

where $\eta$ and $\omega_0 = 6$ is a non-dimensional "time" parameter and frequency, respectively. $\psi_0(\eta)$ describes the most commonly used wavelet type for spectral analyses: the normalized complex Morlet wavelet (Eq. A.4). The power spectrum of the WT is defined by the square of coefficients (Eq. A.3) of the wavelet series as $\left|W_n^X(s)\right|^2$. The notion of scale s is introduced as an alternative to frequency [1]. Thus, we may define frequency bands of interest, such as gamma band, capable of encapsulating the different functional frequencies of the brain.

## A.3 Wavelet Coherence

In a similar way to the definition of coherence, given two time series $X$ and $Y$, with wavelet transforms $W_n^X(s)$ and $W_n^Y(s)$, one can initially define the cross-wavelet spectrum as $W_n^{XY}(s) = W_n^X(s) W_n^{Y*}(s)$ , where $*$ denotes the complex conjugate. The cross-wavelet power is given by $|W_n^{XY}(s)|$ . If one closely resembles Eq. A.2 then the $WC$, $R_n^2$, of two signals may be defined as:

$$R_n^2(s) = \frac{\left|S(s^{-1} W_n^{XY}(s))\right|^2}{S\left(s^{-1}\left|W_n^X(s)\right|^2\right) \cdot S\left(s^{-1}\left|W_n^Y(s)\right|^2\right)} \tag{A.5}$$

where $S$ is a smoothing operator in time $S_t$ and scale $S_s$ such as $S(W) = S_s(S_t(W_n(s)))$ which for the Morlet wavelet is given by a Gaussian and a boxcar filter of width equal to 0.6, (the scale-decorrelation length) respectively [78, 34]:

$$S_t(W,t) = W_n(s) * c_1^{-t^2/2s^2} \tag{A.6}$$

$$S_s(W,s) = W_n(s) * c_2 \prod (0.6s) \tag{A.7}$$

where $c_1$ and $c_2$ are normalization constants and $\prod$ is the rectangle function.

## A.4 Robust state-space GS method (RSS-GS)

Alternatively, one may measure how neighborhoods (i.e., recurrences) in one attractor maps into the other. This idea turned out to be the most

robust and reliable way of assessing the extent of GS [4, 63]. First, we reconstruct delay vectors [75] out of our time series; $x_n = (x_n, \ldots, x_{n-(m-1)\tau})$ and $y_n = (y_n, \ldots, y_{n-(m-1)\tau})$, where $n = 1 \ldots N$, and $m$, $\tau$ are the embedding dimension and time lag, respectively. Let $r_{n,j}$ and $s_{n,j}$, $j = 1, \ldots, k$, denote the time indices of the $k$ nearest neighbors of $x_n$ and $y_n$, respectively. For each $x_n$ the squared mean Euclidean distance to its $k$ neighbors is defined as:

$$R_n^{(k)}(X) = \frac{1}{k} \sum_{j=1}^{k} (x_n - x_{r_{n,j}})^2 \tag{A.8}$$

The Y-conditioned squared mean Euclidean distance $R_n^{(k)}(X|Y)$ is defined by replacing the nearest neighbors by the equal time partners of the closest neighbors of $y_n$.

If the set of reconstructed vectors (point cloud $x_n$) has an average squared radius $R(X) = (1/N) \sum_{n=1}^{N} R_n^{(N-1)}(X)$, then $R_n^{(k)}(X|Y) \approx R_n^{(k)}(X) \ll R(X)$ if the systems are strongly correlated, while $R_n^{(k)}(X|Y) \approx R(X) \gg R_n^{(k)}(X)$ if they are independent. Hence, an interdependence measure is defined as [4]:

$$S^{(k)}(X|Y) = \frac{1}{N} \sum_{n=1}^{N} \frac{R_n^{(k)}(X)}{R_n^{(k)}(X|Y)} \tag{A.9}$$

Since $R_n^{(k)}(X|Y) \gg R_n^{(k)}(X)$ by construction, it is clear that $S$ ranges between 0 (indicating independence) and 1 (indicating maximum synchronization). Another normalized and more robust version of $S$ maybe defined as [63] and is the one actually used in this study:

$$N^{(k)}(X|Y) = \frac{1}{N} \sum_{n=1}^{N} \frac{R_n(X) - R_n^{(k)}(X|Y)}{R_n(X)} \tag{A.10}$$

## A.5 Phase Locking Value (PLV)

One of the mostly used phase synchronization measures is the PLV approach. It assumes that two dynamic systems may have their phases synchronized even if their amplitudes are zero correlated [51]. The PS is defined as the locking of the phases associated to each signal, such as:

$$|n\phi_x(t) - m\phi_y(t)| = const \tag{A.11}$$

However, in this case the phase locking ratio of $n : m = 1 : 1$, since both signals arise from the same physiological system (i.e., the brain).

In order to estimate the instantaneous phase of our signal, we transform it using the Hilbert transform (HT), whereby the analytical signal $H(t)$ is computed as:

$$H(t) = x(t) + i\widetilde{x}(t) \tag{A.12}$$

where $\widetilde{x}(t)$ is the HT of $x(t)$, defined as:

$$\widetilde{x}(t) = \frac{1}{\pi} PV \int_{-\infty}^{+\infty} \frac{x(t')}{t - t'} dt' \tag{A.13}$$

where PV denotes the Gauchy principal value. The analytical signal phase is defined as:

$$\phi(t) = \arctan \frac{\widetilde{x}(t)}{x(t)} \tag{A.14}$$

Therefore for the two signals $x(t)$, $y(t)$ of equal time length with instantaneous phases $\phi_x(t)$, $\phi_y(t)$ respectively the PLV bivariate metric is defined given by:

$$PLV = \left| \frac{1}{N} \sum_{j=0}^{N-1} e^{i(\phi_x(j\Delta t) - \phi_y(j\Delta t))} \right| \tag{A.15}$$

where $\Delta t$ is the sampling period and $N$ is the sample number of each signal. PLV takes values within the $[0, 1]$ space, where 1 indicates perfect phase synchronization and 0 indicates lack of synchronization.

## A.6 Synchronization Likelihood (SL)

Finally, the last measure (SL) used is an unbiased normalized synchronization estimator, closely related to the previous idea and to represent a normalized version of mutual information [74].

Supposing that $x_n$, $x_v$ and $y_n$, $y_v$ be the time delay vectors, SL actually expresses the chance that if the distance between $x_n$ and $x_v$ is very small, the distance between the corresponding vectors $y_n$ and $y_v$ in the state space will also be very small. For this, we need a small critical distance $\varepsilon_x$, such that when the distance between $x_n$ and $x_v$ is smaller than $\varepsilon_x$, $x$ will be considered to be in the same state at times $n$ and $v$. $\varepsilon_x$ is chosen such that the likelihood of two randomly chosen vectors from $x$ (or $y$) will be closer than $\varepsilon_x$ (or $\varepsilon_y$) equals a small fixed number $p_{ref}$, which is the same for $x$ and $y$, but $\varepsilon_x$ need not be equal to $\varepsilon_y$. Now SL between $x$ and $y$ at time $n$ is defined as follows:

$$SL_n = \frac{1}{N'} \sum_{\substack{v=1 \\ w_1 < |n-v| < w_2}}^{N} \theta(\epsilon_{y,n} - |y_n - y_v|)\theta(\varepsilon_{x,n} - |x_n - x_v|) \qquad \text{(A.16)}$$

Here, $N' = 2(w_2 - w_1 - 1)P_{ref}$ , $|\cdot|$ is the Euclidean distance and $\theta$ is the Heaviside step function, $\theta(x) = 0$ if $x \leq 0$ and $\theta(x) = 1$ otherwise. The value of $w_1$ is window equal to the Theiler correction for autocorrelation effects and $w_2$ is a window that sharpens the time resolution of the synchronization measure and is chosen such that $w_1 \ll w_2 \ll N$ [77]. When no synchronization exists between $x$ and $y$, $SL_n$ will be equal to the likelihood that random vectors $y_n$ and $y_v$ are closer than $\varepsilon_y$; thus $SL_n = p_{ref}$. In the case of complete synchronization $SL_n = 1$. Intermediate coupling is reflected by $p_{ref} < SL_n < 1$. Finally, SL is defined as the time average of the $SL_n$ values.