UNICAMP

UNIVERSIDADE ESTADUAL DE CAMPINAS

Instituto de Física "Gleb Wataghin"

CARLOS ALBERTO STEFANO FILHO

Performance analysis of graph metrics for assessing hand motor imagery tasks from electroencephalography data

Análise de desempenho de métricas de grafos para reconhecimento de tarefas de imaginação motora das mãos a partir de dados de eletroencefalografia

Campinas

2016

CARLOS ALBERTO STEFANO FILHO

PERFORMANCE ANALYSIS OF GRAPH METRCIS FOR ASSESSING HAND MOTOR IMAGERY TASKS FROM ELECTROENCEPHALOGRAPHY DATA

ANÁLISE DE DESEMPENHO DE MÉTRICAS DE GRAFOS PARA RECONHECIMENTO DE TAREFAS DE IMAGINAÇÃO MOTORA DAS MÃOS A PARTIR DE DADOS DE ELETROENCEFALOGRAFIA

Dissertation presented to the "Gleb Wataghin" Physics Institute of the University of Campinas in partial fulfillment of the requirements for the degree of Master in Physics.

Dissertação apresentada ao Instituto de Física "Gleb Wataghin" da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Mestre em Física.

Supervisor/Orientadora: Prof. Dr. GABRIELA CASTELLANO

Co-supervisor/Co-orientador: Prof. Dr. ROMIS RIBEIRO DE FAISSOL ATTUX

ESTE EXEMPLAR CORRESPONDE À VERSÃO FINAL DA DISSERTAÇÃO DEFENDIDA PELO ALUNO CARLOS ALBERTO STEFANO FILHO, E ORIENTADA PELA PROFA. DRA. GABRIELA CASTELLANO.

CAMPINAS

2016

Agência(s) de fomento e nº(s) de processo(s): CNPq, 165742/2014-3; CAPES, 1423625/2014

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Física Gleb Wataghin
Lucimeire de Oliveira Silva da Rocha - CRB 8/9174

St32p
Stefano Filho, Carlos Alberto, 1991-
Performance analysis of graph metrics for assessing hand motor imagery tasks from electroencephalography data / Carlos Alberto Stefano Filho. – Campinas, SP : [s.n.], 2016.

Orientador: Gabriela Castellano.
Coorientador: Romis Ribeiro de Faissol Attux.
Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de Física Gleb Wataghin.

1. Interfaces cérebro-computador. 2. Teorias dos grafos. 3. Eletroencefalografia. 4. Processamento de sinais. 5. Imaginação motora. I. Castellano, Gabriela,1970-. II. Attux, Romis Ribeiro de Faissol,1978-. III. Universidade Estadual de Campinas. Instituto de Física Gleb Wataghin. IV. Título.

Informações para Biblioteca Digital

Título em outro idioma: Análise de desempenho de métricas de grafos para reconhecimento de tarefas de imaginação motora das mãos a partir de dados de eletroencefalografia
Palavras-chave em inglês:
Brain-computer interfaces
Graph theory
Electroencephalography
Signal processing
Motor imagery
Área de concentração: Física
Titulação: Mestre em Física
Banca examinadora:
Gabriela Castellano
Diogo Coutinho Soriano
Rickson Coelho Mesquita
Data de defesa: 07-07-2016
Programa de Pós-Graduação: Física

FOLHA DE APROVAÇÃO

MEMBROS DA COMISSÃO JULGADORA DA DISSERTAÇÃO DE MESTRADO DE **CARLOS ALBERTO STEFANO FILHO - RA 101795** APRESENTADA E APROVADA AO INSTITUTO DE FÍSICA "GLEB WATAGHIN", DA UNIVERSIDADE ESTADUAL DE CAMPINAS, EM 07 / 07 / 2016.

COMISSÃO JULGADORA:

- Profa. Dra. Gabriela Castellano – Orientadora –DRCC/IFGW/UNICAMP
- Prof. Dr. Diogo Coutinho Soriano – UFABC
- Prof. Dr. Rickson Coelho Mesquita – DRCC/IFGW/UNICAMP

OBS.: Informo que as assinaturas dos respectivos professores membros da banca constam na ata de defesa já juntada no processo vida acadêmica do aluno.

CAMPINAS

2016

AGRADECIMENTOS

Primeiramente, agradeço à minha família, em especial, aos meus pais, Carlos e Rosana, e ao meu irmão, Rafael, pelo apoio durante os dois anos de mestrado, sempre atenciosos e dispostos a ouvir sobre o andamento do projeto.

Agradeço aos amigos de Sorocaba, dos tempos de colégio e que perduram até hoje; seja por discussões que mais efetivamente contribuíram para o desenvolvimento deste trabalho, ou simplesmente pelas horas descontraídas capazes de me manter longe das preocupações do projeto, tornando o período de seu desenvolvimento mais divertido: Athena, Bila, Blendon, Charlie, Cristian, Elaine, Fabi, Kath, Lucas, Júlia, Mafy, Mari, Paulo, Ricardo, Rodrigo, Sílvio e Taisa (em ordem alfabética). Também merecem agradecimentos os amigos de Campinas, tanto os da Física: Aline, Hítalo, Lucas, Marcelli e Welder; quanto os de fora dela: Buh, Paula, Patrícia, Tato e tantos outros que aqui não exponho para evitar demasiada extensão da lista.

Deixo um agradecimento particular ao Luccas, sempre companheiro, e que, principalmente ao final do trabalho, ouvia meus anseios praticamente diariamente, nunca deixando que faltasse o apoio incondicional que tanto me confortou.

Agradeço, também, aos colegas dos grupos de Neurofísica e de BCI, por todas as discussões e sugestões mais que proveitosas para o desenvolvimento deste trabalho.

Gostaria de agradecer enormemente à minha orientadora, Gabriela, por toda a paciência e dedicação nos últimos dois anos, e pelo incentivo durante os bons e maus tempos deste curto período de mestrado. Agradeço pelo exemplo profissional e por compartilhar todo seu conhecimento e experiência na área.

Agradeço, também, ao prof. Romis, meu coorientador, sempre disposto a me auxiliar quando foi preciso. Certamente tornou meu estudo acerca de BCIs e classificadores indubitavelmente muito mais simples.

Deixo, ainda, um agradecimento especial à recentemente falecida professora de Física Dolores Coutinho, dos tempos de colégio. Grande fonte de inspiração para a minha escolha de carreira, de quem sempre levarei muito

carinho e aprendizado de vida. Agradeço por todas as palavras de incentivo e apoio em uma decisão que mudaria o rumo da minha vida.

A todos, meus mais sinceros agradecimentos.

RESUMO

Interfaces cérebro-computador (BCIs, *brain-computer interfaces*) são sistemas cuja finalidade é fornecer um canal de comunicação direto entre o cérebro e um dispositivo externo, como um computador, uma prótese ou uma cadeira de rodas. Por não utilizarem as vias fisiológicas convencionais, BCIs podem constituir importantes tecnologias assistivas para pessoas que sofreram algum tipo de lesão e, por isso, tiveram sua interação com o ambiente externo comprometida. Os sinais cerebrais a serem extraídos para utilização nestes sistemas devem ser gerados mediante estratégias específicas. Nesta dissertação, trabalhamos com a estratégia de imaginação motora (MI, *motor imagery*), e extraímos a resposta cerebral correspondente a partir de dados de eletroencefalografia (EEG). Os objetivos do trabalho foram caracterizar as redes cerebrais funcionais oriundas das tarefas de MI das mãos e explorar a viabilidade de utilizar métricas da teoria de grafos para a classificação dos padrões mentais, gerados por esta estratégia, de usuários de um sistema BCI. Para isto, fez-se a hipótese de que as alterações no espectro de frequências dos sinais de eletroencefalografia devidas à MI das mãos deveria, de alguma forma, se refletir nos grafos construídos para representar as interações cerebrais corticais durante estas tarefas. Em termos de classificação, diferentes conjuntos de pares de eletrodos foram testados, assim como diferentes classificadores (análise de discriminantes lineares – LDA, máquina de vetores de suporte – SVM – linear e polinomial). Os três classificadores testados tiveram desempenho similar na maioria dos casos. A taxa média de classificação para todos os voluntários considerando a melhor combinação de eletrodos e classificador foi de 78%, sendo que alguns voluntários tiveram taxas de acerto individuais de até 92%. Ainda assim, a metodologia empregada até o momento possui várias limitações, sendo a principal como encontrar os pares ótimos de eletrodos, que variam entre voluntários e aquisições; além do problema da realização *online* da análise.

ABSTRACT

Brain-computer interfaces (BCIs) are systems that aim to provide a direct communication channel between the brain and an external device, such as a computer, a prosthesis or a wheelchair. Since BCIs do not use the conventional physiological pathways, they can constitute important assistive technologies for people with lesions that compromised their interaction with the external environment. Brain signals to be extracted for these systems must be generated according to specific strategies. In this dissertation, we worked with the motor imagery (MI) strategy, and we extracted the corresponding cerebral response from electroencephalography (EEG) data. Our goals were to characterize the functional brain networks originating from hands' MI and investigate the feasibility of using metrics from graph theory for the classification of mental patterns, generated by this strategy, of BCI users. We hypothesized that frequency alterations in the EEG spectra due to MI should reflect themselves, in some manner, in the graphs representing cortical interactions during these tasks. For data classification, different sets of electrode pairs were tested, as well as different classifiers (linear discriminant analysis – LDA, and both linear and polynomial support vector machines – SVMs). All three classifiers tested performed similarly in most cases. The mean classification rate over subjects, considering the best electrode set and classifier, was 78%, while some subjects achieved individual hit rates of up to 92%. Still, the employed methodology has yet some limitations, being the main one how to find the optimum electrode pairs' sets, which vary among subjects and among acquisitions; in addition to the problem of performing an online analysis.

FIGURE LISTING

TABLES LISTING

LIST OF ACRONYMS

BC	Betweenness centrality
BCI	Brain-computer interface
CAR	Common average removal
CC	Clustering coefficient
CNS	Central nervous system
DC	Degree centrality
EC	Eigenvector centrality
EEG	Electroencephalography
EPSP	Excitatory post-synaptic potential
ERD	Event related desynchronization
ERP	Event related potential
ERS	Event related synchronization
fMRI	Functional magnetic resonance imaging
ICA	Independent component analysis
IPSP	Inhibitory post-synaptic potential
LDA	Linear discriminant analysis
LSLDA	Least-squares based linear discriminant analysis
LSVM	Linear kernel based support vector machine
MI	Motor imagery
NIRS	Near-infrared spectroscopy
PSD	Power-spectral density
PSVM	Polynomial kernel based support vector machine
SDP	Second degree polynomial
SL	Surface Laplacian
SNR	Signal-to-noise ratio
SSVEP	Steady state visually evoked potential
STP	Single-term exponential
SVM	Support vector machine
TTE	Two-term exponential
VEP	Visually evoked potential

CONTENTS

Chapter 1

Introduction

The central nervous system's (CNS) role in how our environment is perceived by us is doubtlessly fundamental. It is responsible for processing and interpreting external stimulations such as light, sound, pain, heat and so on. There are, however, several disabilities that can affect its proper function and, therefore, decrease the quality of life of people that suffer from them.

Data from 2010 of the Brazilian Institute of Statistical Geography (IBGE – Instituto Brasileiro de Geografia Estatística) reveal that over 45 million people suffer from a type of disability [IBGE, censo demográfico 2010], accounting for 23.9% of the total Brazilian population. Specifically, visual (18.8% of the total population), motor (6.9%), hearing (5.1%) and mental/intellectual (1.4%) disabilities were pointed out in the study. Motor disabilities (as much as some other types of disabilities) can be caused by a variety of factors, such as traumatic injuries and diseases.

In the case of motor disabilities, traumatic injuries include damage done to the spinal cord of the CNS, which can result in limps paralysis. Paralysis denomination varies if only the legs or both legs and arms are affected. The first case is denoted by paraplegia, whilst the latter is called quadriplegia. Injuries to the spinal cord can be caused, for instance, by car accidents, violence acts and falls. Some diseases that can cause motor disabilities are cerebral palsy, muscular dystrophy, multiple sclerosis, spina bifida, amyotrophic lateral sclerosis, arthritis and stroke.

Depending on the extent of the disability, there are measures that can be implemented to increase the life quality of people affected by it. In the case of motor disabilities, an increasingly studied assistive technology over the last years is the brain-computer interface (BCI). BCIs are systems that enable communication with external devices through brain signals, without using the conventional physiological pathways. Thus, these systems do not require any physical movement, and, in theory, even people with severe disabilities should be able to control them, even locked-in patients [Kubler et al., 2001], the ones who have lost all their mobility, making it very hard to communicate with the external environment.

BCIs aim to provide a direct communication channel between the brain and an external device. Applications vary widely, and these systems can be used, for instance, to move a cursor in a computer screen [Cheng et al., 2004], control spelling devices with virtual keyboards [Obermaier et al., 2003], moving a wheelchair [Iturrate et al., 2009], monitoring emotional states of patients [Widge et al., 2014] and rehabilitation applications [Holper et al., 2010].

As ideal as they may sound, a lot of research is still needed to be done in order to optimize the performance of BCIs. Fundamental steps of BCIs include signal acquisition, processing and classification for the desired application. Given the complexity of the brain, this is not expected to be a trivial task. In fact, each operational step of a BCI is very demanding and can constitute a research topic by itself. Regarding brain signal acquisition, electroencephalography (EEG) has been the most used technique.

In EEG-BCIs, there are two main strategies currently used to generate mental patterns for identifying the user's intent: evoked potentials and imagery (see Chapter 2). Regarding imagery (more specifically, motor imagery), there is no optimum way established in the literature to proceed with such systems. Traditional analysis is restricted to applying different classification approaches to the spectral power density of the EEG brain signals. More recently, however, the concept of brain connectivity and how different brain regions interact with each other have been studied as potential and novel applications for BCIs [Gupta and Falk, 2015; Demuru et al., 2013; Elasuty and Eldawlatly, 2015; Salazar-Varas and Gutiérrez, 2015; Ghosh et al., 2015; Asensio-Cubero et al., 2016].

In this work, we focused on exploring approaches for feature extraction in motor imagery EEG-BCIs using metrics from graphs theory. We characterized graphs topology for each motor imagery task studied, extracted different metrics from these graphs and, then, we applied commonly used classifier algorithms in the BCI research area to discriminate between the tasks. Thus, the two main goals of this work were:

- To characterize EEG hand motor imagery response using graph features;
- To explore the feasibility of using graph metrics as features for a BCI classifier.

This thesis is divided as follows. Chapter 2 focuses on BCIs, reviewing the most recurring types in the literature and their applications. Also, a description of the technique used to extract the brain signals, EEG, is made. Some basic anatomical aspects are briefly reviewed, and methodological aspects, from how data is acquired to the technique limitations, are described. Chapter 3 presents basic concepts from graph theory, and how it can be used for the investigation of the proposed problem. Chapter 4 explains the methodology used, and Chapter 5 describes and discusses the obtained results. Finally, Chapter 6 presents conclusions and future perspectives. Also, a few interesting but not strictly necessary discussions and deductions, and sets of results, are presented in Appendixes A to D.

Chapter 2

EEG based Brain-Computer Interfaces (EEG-BCIs)

Brain-computer interfaces (BCIs) are systems that allow direct communication to external devices from brain signals, without using the conventional physiological pathways, such as talking or muscle movement. BCIs are an alternative form of communication and interaction for people affected by specific brain conditions or spinal cord injury preventing them to move. Currently, EEG is the most used technique for BCI signal acquisition, although some other techniques are also being studied, such as functional magnetic resonance imaging (fMRI) [Berman et al., 2011; Hermes et al., 2011; Halder et al., 2013] and near-infrared spectroscopy (NIRS) [Sitaram et al., 2007; Kanoh et al., 2009]. Each technique has its limitations, and different combinations of these techniques for BCI applications are also under study. This chapter focuses on EEG-BCI applications, it discusses brain electrical potentials and thought strategies commonly used in BCI, particularly, motor imagery (MI). Also, an overview of the EEG technique is given and processing methods commonly used in BCI signal processing are featured.

2.1. Electroencephalography

The electroencephalography (EEG) technique is used to measure the brain’s electrical activity. Its first recordings were done by Richard Caton (1842 – 1926), a British scientist, in 1875 [Finger, 2001]. In his studies, Caton used a galvanometer, an instrument that could measure small currents and potential differences between two points. He observed that positioning electrodes on two distinct points of the brain’s external surface yielded a current reading in the device. The first person, however, to perform an electroencephalogram exam per se was a German neurologist, Hans Berger (1873 – 1941). Presently, EEG may not be the most used technique, and many others are commonly practiced in diagnostic medicine, such as magnetic resonance imaging (MRI), computed tomography (CT) or positron emission tomography (PET). However, when it comes to measuring the brain electrical activity directly, EEG is the technique of choice. Among the main clinical uses of EEG, are searching for the focus of an epileptic crisis and evaluating sleep disorders. EEG has also been widely used to extract brain signals for Brain-Computer Interfaces (BCIs). In this section, some basic knowledge necessary for understanding the EEG signal generation and other relevant aspects of this technique are covered.

2.1.1. Biophysical aspects of the EEG signal generation

Basic anatomic aspects

The brain is an important part of the central nervous system (CNS), responsible for integrating and processing information. It has different functional regions; that is, distinct areas of the brain are commonly associated with different functions. One or more areas, however, can interact to yield some specific function. A common basic division when studying the brain comes in lobes: frontal lobe, temporal lobe, parietal lobe and occipital lobe (Figure 2.1).

The frontal lobe is the largest and most anterior part of each hemisphere [Ribas, 2010], and is responsible mainly for short-term memory, planning of future actions and motor control. The parietal lobe is concerned with sensory strip location, reading, writing, calculations and somatic sensation. The occipital lobe is largely associated with vision, and, the temporal lobe, with hearing, learning, memory and emotion. These latter three functions are due to the temporal lobe’s deep structures, such as the hippocampus and amygdaloid nuclei [Kandel et al., 2013].

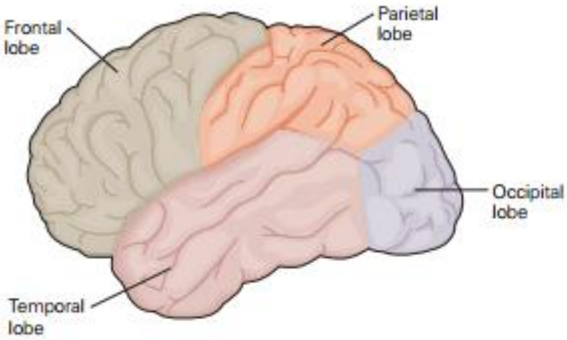


Figure 2.1. Brain lobe division. Extracted from [Kandel et al., 2013].

The nervous system contains two important types of cells: the neurons and the glial cells (or glia) [Kandel et al., 2013]. The human nervous system has about 86 billion neurons [Herculano-Houzel, 2012]. They can have many different morphologies, and their classification can depend on neurotransmitters used, electrophysiological properties or direction of neural impulse propagation.

Regardless of a neuron’s type, each one has four well-defined regions: the cell body (or soma), the dendrites, the axon and presynaptic terminals [Kandel et al., 2013]. The cell’s nucleus is located in the soma. Both dendrites and axon originate from the cell body. Dendrites are structures carrying electrical impulses from other neurons into the soma, and axons are the physical paths in which these impulses are carried from the soma to other neurons. Transmission or reception of neuronal impulses occur at the synaptic terminals. Figure 2.2 illustrates these regions for different types of neurons.

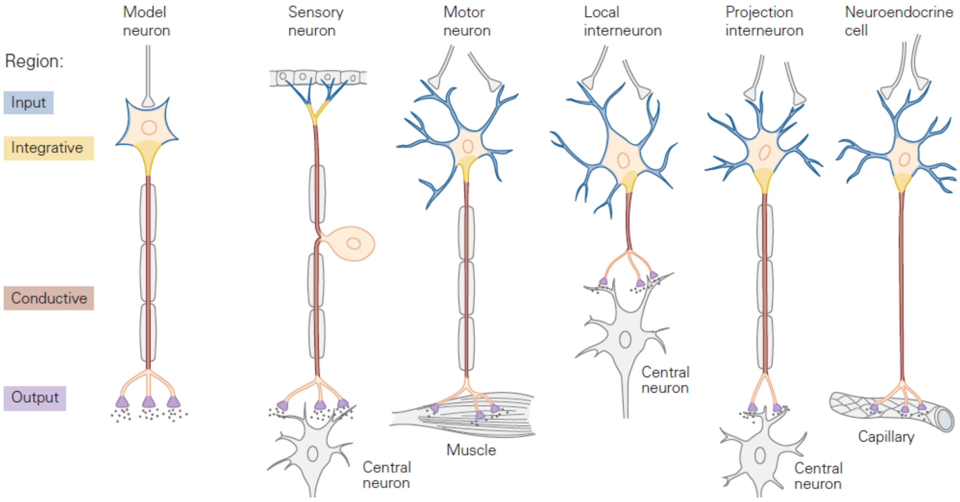


Figure 2.2. Different types of neurons and their components. **Adapted from [Kandel et al., 2013].**

In Figure 2.2, the ‘input’ site represents neuronal impulses that arrive from post-synaptic terminals at the dendrites, which are carried to the soma (integrative region), where they will be summed. The axon is represented by the ‘conductive’ region, and the pre-synaptic terminals are the ‘output’ region, where electrical impulses will be transmitted to other neurons.

Glial cells support neurons, commonly surrounding all of their different parts. Glia classification is usually done according to morphological and molecular criteria. All types of glia, however, have some common characteristics, such as being associated with neurons, having a different structure from neurons themselves and being lineally related to them [Saham, 2005].

Glia function and its contribution to the EEG signal have been matters under discussion. Some authors believe that the EEG measure is due mainly to neurons, with just a modest contribution from glial cells [Silva, 2010]. Their role, however, in generating extracellular potentials measurable with EEG has also been discussed [Speckmann et al., 2005; Somjen et al., 1979; Kuffler et al., 1966]. In [Saham, 2005], the authors review a variety of glial cells studies regarding some important aspects such as synaptogenesis, regulation of synaptic activity, neuronal conduction, neuronal migration and reciprocal control of cell survival between glia and neurons; but a decisive conclusion was not reached. More recent works have started to point out functional aspects of glial cells as contributors to the EEG signal. Astrocytes (a type of glial cell) have been associated with the maintenance of oscillations in the gamma range (about 25 – 80 Hz), through vesicular release in these cells. Also, TeNT (tetanus neurotoxin) expression in astrocytes has been reported to reduce the gamma band power density in EEG, in vivo. The gamma frequency range has been associated with several cognition functions, such as learning and memory, suggesting astrocytes may play a major role in these tasks [Lee et al., 2014]. This type of cells has also been related to important roles regarding neuronal circuits, indicating possible causal factors that regulate synchronized activation of neuronal ensembles [Poskanzer and Yuste, 2011].

Basically, the generation of the extracellular fields (which can be measured with EEG) lies in the cellular membrane structure and the ions that surround it. This membrane consists of a lipid bilayer with proteins inserted into it (Figure 2.3). While some of these proteins have purely structural function, others constitute channels available to the passage of certain types of ions, of which the most involved into the production and propagation of the electrical impulse are Na^+, K^+, Ca^{+2} and Cl^-.

A difference in ion concentration between intra and extracellular media results in an electrical potential gradient, which makes the ions diffuse through their specific protein channels (Figure 2.4). The so-called membrane potential is dictated by the ion concentration configuration between both media.

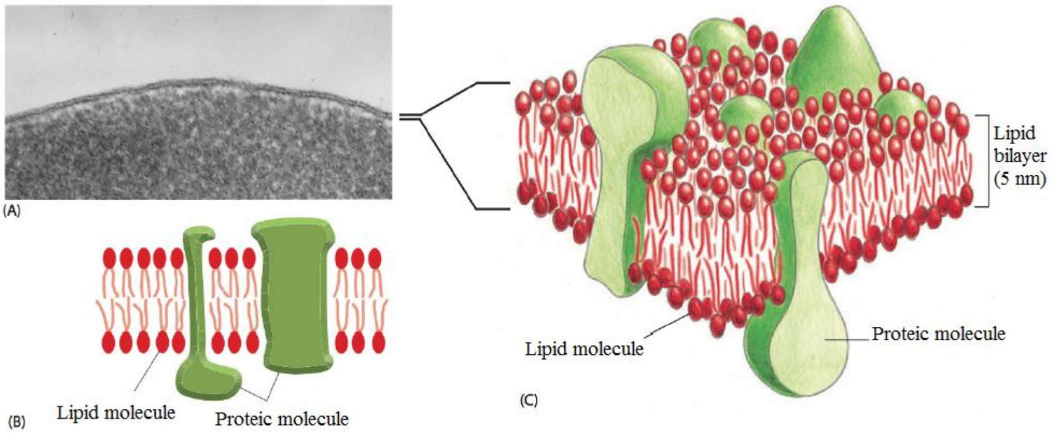


Figure 2.3. Cellular membrane structure. Adapted from [Alberts et al., 2010].

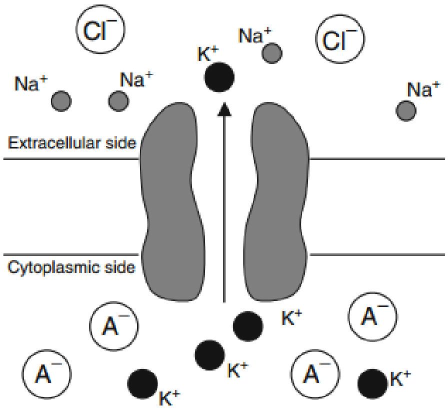


Figure 2.4. Illustration of ion transportation through protein channel. Extracted from [Ermentrout and Terman, 2010].

The membrane potential can be maintained either by diffusive passive ion transport, in which ion concentration gradients are dictated by the selective permeability of the protein channels, or active transport of ions, when the use of external energy is necessary, which in many cases comes from ATP hydrolysis. Regarding active transporters, the Na^+ - K^+ pump is maybe the most important one, and it functions by pumping three Na^+ ions out of the cell to every two K^+ ions that are pumped in for every pump turnover. Another relevant active transporter is the Na^+ - Ca^{+2} exchanger, which functions driving three Na^+ inward and one Ca^{+2} outward. This transporter is not ATP-based directly, since the Na^+ ions being driven down their concentration gradient is what provides the energy source in this case. This necessary Na^+ concentration gradient, however, is maintained by the ATP-based Na^+ - K^+ pump.

There are four physical laws that dictate ion movement in the biological membrane and, therefore, its electrical potential: particle diffusion caused by concentration differences, drift of ions caused by electric potential gradients, the relationship between the diffusion coefficient (D) and the drift mobility (μ), and the principle of separation of charges in biological systems. This last one states that, in a given volume, the total charges of cations and anions are approximately the same [Johnston and Wu, 1996]. The cell membrane, due to its charge separation, constitutes an exception to this rule. A mathematical description of this potential from these phenomena can be found in Appendix A.

The neuron as an electric dipole

When neurotransmitters arise from the synapses to the dendrites, they cause the post-synaptic terminal channels of the next neuron to open up. If the stimulus is excitatory, the membrane will depolarize; that is, its potential will increase to a more positive value, and the potential will be called an excitatory post-synaptic potential (EPSP). If, however, the stimulus is inhibitory, the membrane will hyperpolarize to a more negative potential, responding to an inhibitory post-synaptic potential (IPSP).

In the case of an IPSP, for instance, two situations could occur: the inflow of a negative current to the intracellular medium or an outflow of positive current to the extracellular medium (either way, the membrane potential hyperpolarizes). For an EPSP, the depolarization can occur due to the inflow of positive current or the outflow of negative current. Regardless of the stimulus type (IPSP or EPSP), the electrical flow causes potential changes along the cell, and the creation of regions of charge sink (current flowing into the cell) and source (current flowing out of the cell), inducing the neuron to behave like an electric dipole (Figure 2.5). The EEG signal is often attributed to this synaptic activity, being measured as the summation of electrical fields generated by the interaction of the dipoles. Non-synaptic activity can also contribute to the EEG signal, although to a lower extent [Olejniczak, 2006]. Neuronal responses such as the action potential (Appendix A) are considered too fast to significantly impact the EEG signals.

Since EEG electrodes are much larger than a single neuron, what they measure is an average signal of a population of brain cells. Besides, a single electrode measure is influenced, although to different degrees, by populations of neurons located not necessarily just below it. To yield a visible EEG measure, a combination of approximately 108 neurons in a minimal cortical area of 6 cm² are needed [Olejniczak, 2006]. Each neuron generates a relatively small amplitude signal and, taking into account the attenuation occurring at the skull and scalp, EEG mostly measures the neuronal activity of pyramidal cells located on the cortex surface [Sanei and Chambers, 2007]. Note that the signal recorded by a specific electrode does not detect electrical activity exclusively from the site upon which it is located. In fact, as dipoles, all cortex neurons contribute to the signal, although some more than others, given the characteristics of the electrical field of a dipole. Thus, it is necessary to keep in mind that the signal of an EEG electrode can be highly affected by various neurons populations. An elaborated analysis should take this into account.

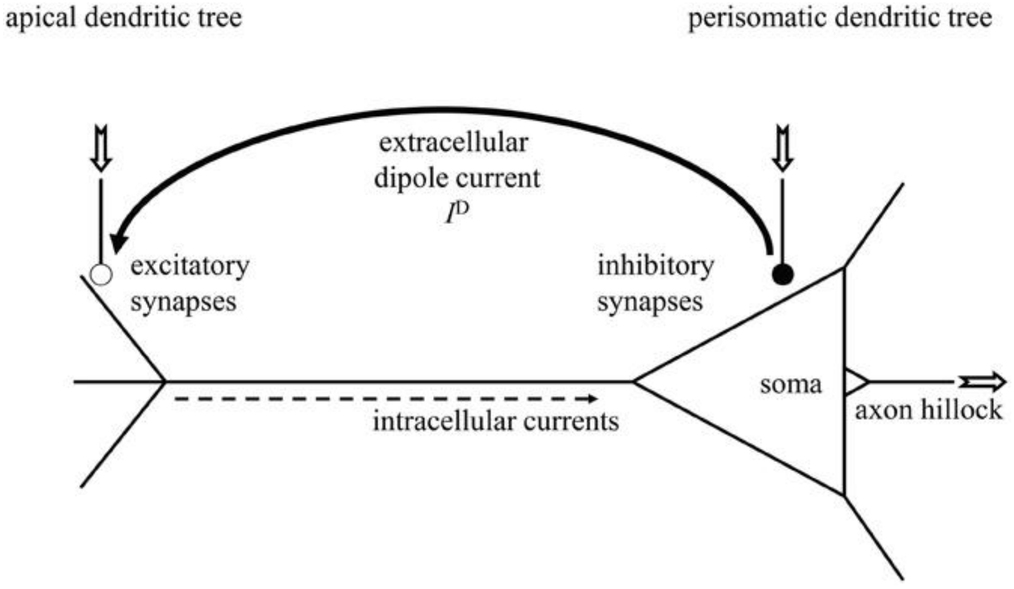


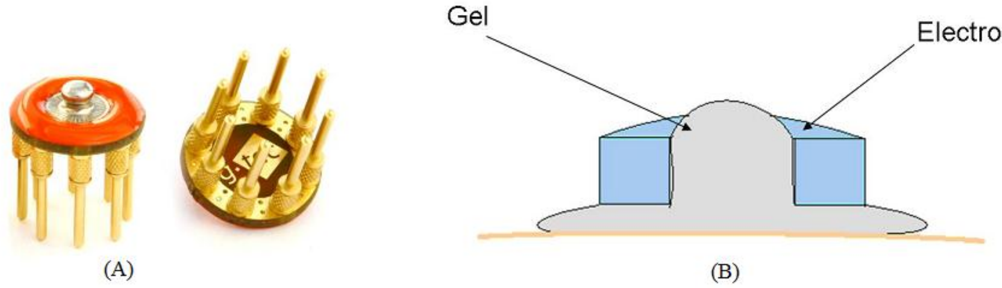
Figure 2.5. Neuron as an electric dipole illustration. In this example, inhibitory synapses function as charge source, and excitatory synapses, as charge sink. Extracted from [Graben and Rodrigues, 2013]

2.3.2. EEG instrumentation

Electrodes

Electrodes are the EEG sensors responsible for measuring variations in the brain electrical potential. They are positioned on the subject's scalp, and their measure provides data from a population of neurons noninvasively. The measured signal has a very low amplitude and suffers from a great attenuation, which makes the acquisition step very challenging and relevant, greatly affecting data quality. EEG electrodes are made of metal and can have different shapes, such as discs, needles or cups. In addition, they can be dry or gel (or paste) based electrodes (Figure 2.6). For dry electrodes, there exists direct contact between scalp and sensor, whilst for gel based electrodes this contact is intermediated by a layer of an electrolyte material.

The electrode material plays an important role in EEG signal acquisition, and not only the materials' electrical conductivity should be taken into account. Electrodes are made of metal, which discharges ions into solution when in contact with the electrolyte. Therefore, an ion-electron exchange occurs between electrode and electrolyte, resulting in the creation of two layers in the space between these two components: one of ions, and another one of electrons. The formation of each one of these layers occurs at different rates, which depend on the materials used. The rate difference causes a voltage to appear, termed the *half-cell* potential [Usakli, 2009], which can interfere with the signal quality.



(A)

(B)

Figure 2.6. Illustration of (A) dry and (B) gel based electrodes. Figure (A) extracted from http://www.gtec.at. and (B) from http://www.unicog.org/pmwiki/uploads/Main/EEGElectrode.jpg.

With some electrodes, there is a free exchange of charge across this double layer, and they are termed to be nonpolarized (or reversible) electrodes. They are usually indicated by the symbol Ag-AgCl, due to its constitution: the metal Ag electrode is coated by an AgCl gel layer. These electrodes are very common, due to their low impedance, low susceptibility to motion artifacts and capacity to record very slow potential changes [Beltramini, 2014; Picton et al., 2000].

When the transfer of charge is very small across the double layer, the electrode is said to be polarized (or nonreversible). In this case, the device functions as a capacitor, cutting off high frequencies and DC voltages [Usakli, 2009].

The impedance of the electrode tissue interface depends on several factors, such as skin preparation, existing hair in between and electrolyte temperature [Usakli, 2009]. Since the skin is filled with sweat glands, it can generate ionic potentials. Besides, its external layer has insulator properties. Therefore, it is common to use a kind of paste to remove this outer layer.

Gel based electrodes are commonly present in caps with pre-defined electrode positioning (Figure 2.7). When working with this type of device, it is necessary to clean the scalp's skin before gel application, in order to remove superficial dust and oil. In addition, filling the space inside the electrode with gel is a demanding task and, in some applications, 100 or more electrodes may be necessary. Had caps not been invented, the electrodes would need to be positioned one at a time (which may actually be the case for some applications).

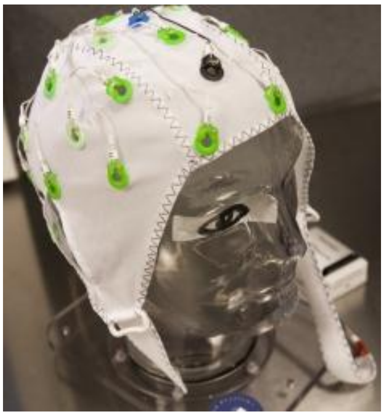


Figure 2.7. EEG positioning cap with gel based electrodes. Extracted from [Aalto university website].

The type of electrode to use will depend upon the application sought. Dry electrodes are a more recent technological development and more easily set, since they have direct contact via mechanical pressure. Gel-based electrodes, on the other hand, can take a reasonable time to prepare for acquisition (in some cases, a few hours). A few studies show that there are optimum conditions within which these electrodes can replace gel-based ones in specific applications [Mihajlovic and Molina, Fiedler et al., 2014].

When choosing which set to use, then, the sought result should be kept in consideration, such as specifications for each type, including signal-to-noise ratio and electrode impedance. Also, when possible, acquiring and comparing data with both types of electrodes should be a good manner to determine if gel-based electrodes could be replaced by dry ones. However, such a study is very demanding and not always possible to do.

Electrodes positioning systems

As an attempt to give EEG acquisitions some common basis, positioning systems have been developed (Figure 2.8). In positioning systems, letters indicate anatomical regions: frontal (F), occipital (O), central (C), temporal (T), parietal (P) and ears (A). Central electrodes are always accompanied by the letter ‘z’. Numbers indicate to which hemisphere the position belongs to: even numbers are located on the right hemisphere, and, odd numbers, on the left one. Nasion and inion are common reference points. A system’s name indicates how distances between adjacent electrodes are calculated. In the 10-20 system, for instance, the nasion-inion and left-right ear distances are measured. Electrodes next to these reference points are placed at 10% of the distance measured. Intermediate electrodes are placed at 20% of that distance (Figure 2.9).

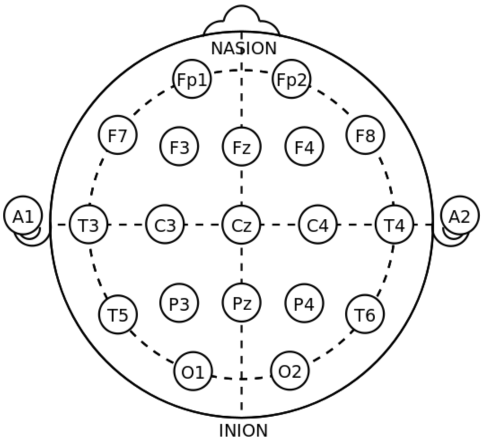


Figure 2.8. EEG positioning system example. 10-20. Extracted from [Wikimedia commons website].

There are other positioning systems, such as the 10-10 or 10-5 systems. They differ only in the distance used to place electrodes from one another (and, consequently, in the total number of electrodes each system comprises).

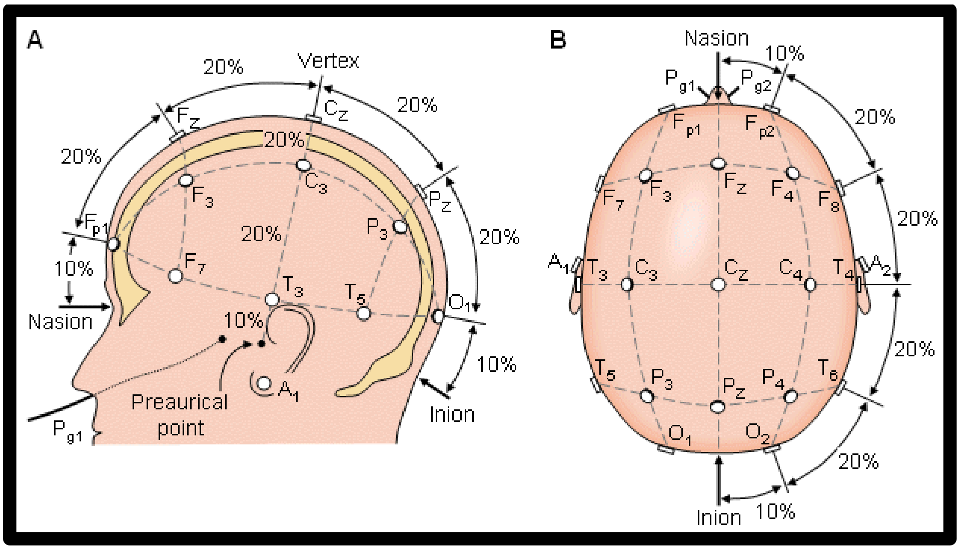


Figure 2.9. 10-20 system positioning. Taken from http://www.bem.fi/book/13/13.htm.

Noise

EEG recordings can provide very good insight about the situation under study. The signal is, however, very noisy, making its analysis difficult, tricky and challenging in many occasions. In general, there are four strategies used to reduce noise: elimination of noise sources, averaging, rejection of noisy data and noise removal [Repvos, 2010]. Also, there are two types of noise that are usually present in EEG data: external and physiological noise.

External sources include those of electromagnetic fields, such as TVs, computers, AC power lines, mobile phones, notebooks and so on. Noise that comes from these sources can easily be controlled by simply avoiding letting any device of this kind near the EEG sensors. When possible, replacing devices working on AC with others that work on DC is also helpful. In addition, isolating the EEG recording room from any external electromagnetic source, such as in a Faraday cage, should provide an effective minimization of electromagnetic external noise. This approach, however, has a very high building cost [Repvos, 2010].

Physiological noise sources occur in many forms, and they include interference from the heart’s electrical activity, blinking artifacts, motion artifacts, breathing artifacts and the skin potentials. Motion artifacts can be avoided by asking the patient not to move (of course, this is not possible if one is performing a motor study, or one in which some kind of movement is necessary). Blinking artifacts are a little trickier, in the sense that, being an automatic behavior, trying to interrupt blinking might induce artifacts in the signal [Repvos, 2010]. Skin potentials exist because of the presence of sweat glands' ionic potentials and the outer insulator layer of the skin. An abrasive cream can be applied to remove this layer, and it is a good way to reduce these potentials. Also, skin potentials can vary during acquisition depending on the temperature or the individual's stress level, factors that can lead to sweat.

Averaging in order to reduce noise is used when the noise is assumed to be random, with zero mean and constant variance. Being this the case, averaging over the experiment time should cancel the noise out. This, however, is only useful when a stable and reproducible signal is sought, such as is the case for an evoked response in an event related potential (ERP) [Repvos, 2010].

Rejection of noisy data consists simply of eliminating parts of the data made unusable by noise. One way to do so is by simple visual inspection: some artifacts really stand out to the eye (Figure 2.10), being easily spotted. Besides, apparent artifacts usually show up in more than one channel. Some EEG analysis software tools perform automatic artifact detection. However, a good knowledge of the artifact sought is important to correctly set the parameters of the software's algorithm.

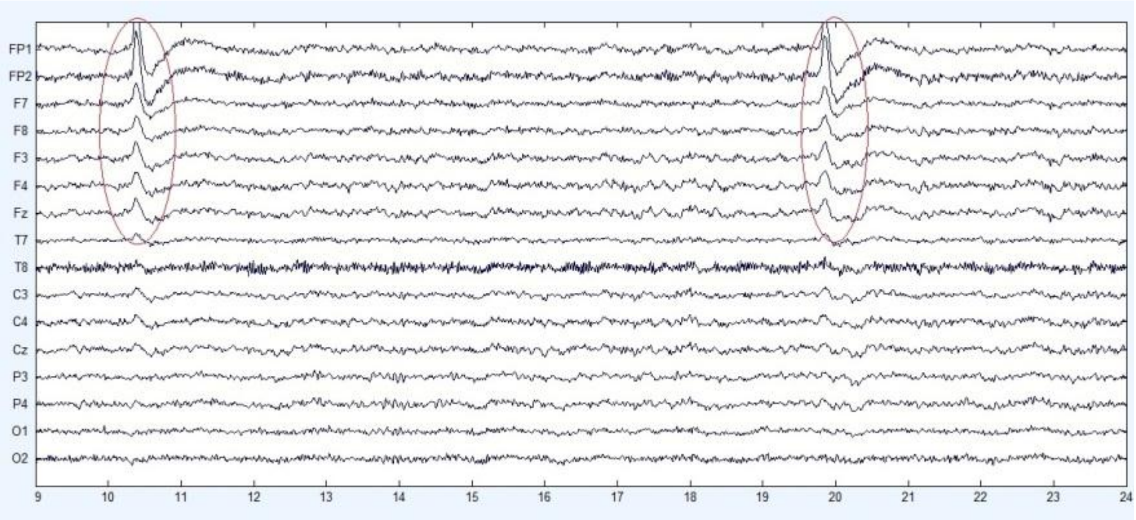


Figure 2.10. Example of EEG blinking artifact. It is really noticeable that the artifact stands out, having a larger amplitude than the rest of the data. Taken from http://www.intechopen.com/.

The removal of noise usually includes some kind of filtering or subtracting of a noisy component from the signal by previous application of some sort of mathematical regression.

With filtering, it is possible to remove specific frequencies (or frequency range) that are a priori known to be mainly noise; for instance, the noise introduced by the electrical power line (50 Hz or 60 Hz, depending on the location). Other types of filters that are not in the frequency domain can also be used, such as the CAR (Common Average Filter) spatial filter. This filter averages the data from all channels and subtracts this value from every channel, in order to eliminate common features that are present in all channels at the same time.

In the case of subtraction of a noisy component, this can be done by modeling the component (for example, the cardiac rhythm) and directly subtracting it from the signal. For instance, independent component analysis (ICA) could be used to estimate the sources of the signal, identify purely noisy components, remove them from the original signal and reconstruct it to obtain a signal that is “free” of that particular noise source [Jung et al., 2000; Repvos, 2010].

Amplifier

The EEG amplifier is a differential amplifier. It amplifies differences between two input signals, rejecting any common voltage between them [Laplante, 2005]. This means that any common artifact present in the two input signals should be suppressed; which is particularly good for EEG, since environmental electromagnetic interference, for instance, usually has a much larger amplitude than the physiological response.

The signal measured by the amplifier has two important bases: the ground electrode and the reference electrode. The ground electrode is simply the one that is chosen to be at zero potential (that is, 0 V). The reference electrode should, ideally, be positioned on an electrically neutral site. In practice this is not always possible, and the signal measured by each electrode represents the activity upon that specific electrode and upon the reference electrode [Beltramini, 2014]. Common locations for placing reference electrodes include the mastoids, ear lobes and the nose. Another alternative consists of using the signal average of all electrodes as the reference. This latter option is more precise when there is a large number of electrodes, since it is the situation that results in good scalp coverage.

The amplifier's ability to suppress common voltages between the two input signals is characterized by what is called "common mode rejection" (CMR). The larger the CMR is, the higher the signal-to-noise ratio (SNR). In CMR, the input signals must be treated equivalently, and, to accomplish this, the ratio of the electrode impedance to the amplifier's input must be considered [Kappenman and Luck, 2010].

Electrode impedances are largely dictated by skin properties. An increase in electrodes impedances lead to a CMR decrease, accompanied by a decrease in SNR. Therefore, in recordings using high impedance electrodes, more trials may be needed to perform signal averaging in order to increase the SNR, which could be very time demanding and not feasible in some cases. If all other acquisition parameters are equal, low impedance electrodes should give a more statistically significant recording.

As already stated, the use of abrasive creams and proper skin cleaning can reduce the skin's influence to the electrodes impedance. For many electrodes systems, however, reducing each electrode's impedance is a very demanding task. Besides, abrasion of the skin may cause bleeding into the electrodes, which will have to be disinfected, another onerous activity. An alternative to work with a high impedance electrode system dealing with the SNR reduction is to use high entrance impedance amplifiers [Kappenman and Luck, 2010]. Nonetheless, using a high entrance impedance amplifier does not account for a problem generated by the skin potentials artifacts.

Any differences in the skin's conductance under the electrodes leads to a different voltage offset for each electrode, creating an electrical potential between them. If the electrical properties of the skin in the sites under each electrode are different, this potential will vary over time. Two main ways to reduce skin potential artifacts are: to keep the recording environment cool and dry, which should avoid changes in the sweat level; and the abrasion of the skin [Kappenman and Luck, 2010].

The entrance voltage in the amplifier can be approximated by [Beltramini, 2014]:

$$V_{\mathrm{diff}} \approx \frac{\mathrm{V_A}}{1+\frac{\mathrm{Z_A}+\mathrm{Z_B}}{\mathrm{Z_{diff}}}+\frac{\mathrm{Z_A}}{\mathrm{Z_{cm}}}} - \frac{\mathrm{V_B}}{1+\frac{\mathrm{Z_A}+\mathrm{Z_B}}{\mathrm{Z_{diff}}}+\frac{\mathrm{Z_B}}{\mathrm{Z_{cm}}}}. \qquad \textbf{(2.1)}$$

In (2.1), $\mathrm{V_A}$ and $\mathrm{V_B}$ stand for the measured potential in electrodes "A" and "B", respectively, in reference to the ground electrode. $\mathrm{Z_A}$ and $\mathrm{Z_B}$ are these electrodes' impedances, $\mathrm{Z_{diff}}$ represents the entrance differential impedance and $\mathrm{Z_{cm}}$ is the common mode entrance impedance, that is, when $\mathrm{V_A}$ and $\mathrm{V_B}$ are equal. This approximation stands when $\mathrm{Z_{mc}} >> \mathrm{Z_A}$, $\mathrm{Z_{mc}} >> \mathrm{Z_B}$ and $\mathrm{Z_{diff}} >> \mathrm{Z_A} + \mathrm{Z_B}$ [Beltramini, 2014]. Clinically, standards for EEG digital recording state that electrodes entrance impedance should be less than 5 kΩ; for the amplifier, the minimum entrance impedance should be of 100 MΩ [Nuwer et al., 1998].

When the denominators of (2.1) approach 1, V_{diff} is simply V_A - V_B, and this constitutes the most accurate situation possible. This occurs when Z_{diff} and Z_{cm} are much larger than Z_A and Z_B; that is, the lower the electrodes impedances are when compared to Z_{diff} and Z_{cm}, the closer to the best scenario situation one would be.

In regarding pre-processing of the data, to properly visualize and analyze the EEG signals, it is necessary that they undergo a series of steps. Raw EEG data has a very low amplitude, of the order of microvolts, and contains five components: desired biopotential, undesired biopotentials, the power supply interference of 50/60 Hz (and its harmonics), interference signals due to the tissue/electrode interface and noise [Sanei, 2007; Teplan, 2002]. The signals must be amplified before the ADC (analog-to-digital converter) and filtered, mainly to reduce noise.

Usually, high-pass filters have a cut-off frequency of about 0.5 Hz, to eliminate the slow component of breathing. To eliminate the 50/60 Hz power supply noise, commonly a notch filter with a null frequency of 50/60 Hz is used [Sanei, 2007]. It is important to make sure that the amplifier and filters do not distort the signal.

2.3.3. EEG frequency bands

EEG signal amplitudes usually range from 2μV to some hundreds of microvolts. Commonly, frequencies from 1 to 100 Hz are analyzed, and they are often divided for study purposes (Table 2.1). EEG sampling frequencies (which should not be confused with the measured frequencies) can range from 100 Hz to 5 kHz, in some cases.

Table 2.1. EEG frequency bands. There exist some variations upon these definitions in the literature. This table was based upon the indications of [Vidal, 1973].

Band name	Frequency range (Hz)
Delta	1-4
Theta	4-7
Alpha	7-13
Beta	13-30
Gamma	30 - 100

Every band is also associated to some brain state. The delta band is commonly linked to deep sleep stages of normal adults or people under anesthesia influence; the theta band is associated with sleeping in adults and is very present in infants and children; the alpha band is present when a person is not focusing in a specific task; the beta band is associated with problem solving, states of attention, cognitive functions and active thought; the gamma band is known to be linked to stress or anxiety states. Some authors do not even dissociate the beta and gamma bands, putting them under the same 'beta band' label. Other subdivisions such as alpha 1, alpha 2, beta 1, beta 2 and so on are also possible, depending on the application and area of study.

Another important band that is not listed in the table (because it is contained within the alpha band) is the mu (μ) rhythm. It is defined by activity extracted from 8 to 13 Hz over the sensorimotor cortex during waking neural activity [Volkmar, 2013]. It is well established that this rhythm reflects the synchronized activity of large groupings of pyramidal neurons in the brain's motor cortex [Pfurtshceller et al., 1997; Volkmar, 2013]. The decrease in the power of this band and the gradual desynchronization of neural activity is known to constitute this band attenuation. According to [Volkmar, 2013], μ band attenuation has been observed *"during both execution and observation of actions falling within one's behavioral repertoire"*. The mu rhythm is particularly relevant in motor imagery based brain-computer interfaces, and will be discussed ahead.

2.3.4. General remarks about EEG

EEG is the technique of choice when it comes to measuring the brain's electrical activity, being widely used in monitoring activity in coma patients, produce biofeedback, locate epilepsy's seizures origin, test drug effects, investigate sleep disorder, among others. Given the technique's portability, non-invasiveness and low-cost, it has been widely used in BCI (brain-computer interface) research. It provides relatively high sampling rates, reaching 5 kHz in some cases, which enables a very good temporal resolution to measure electrophysiological phenomena. Although EEG has poor spatial resolution, it is still able to monitor localized activity from populations of neurons. It, however, records a signal that is very noisy and highly influenced by external electromagnetic sources, requiring pre-processing steps able to compensate for this.

Nonetheless, implementing all needed steps, EEG can provide very useful information and insight into electrophysiological phenomena occurring within the brain, constituting an efficient tool for studying this fascinating organ.

2.2. What is a BCI?

Brain-computer interfaces (BCIs) are systems capable of acquiring a user’s brain signals, processing them and interpreting them in order to interact with an external device, for example a computer, a wheelchair, or a prosthesis of some kind. Figure 2.11 shows a general scheme of a BCI: brain signals are generated by the system's user, according to a pre-defined strategy; they are measured by a technique; then, the signals must be processed to reduce noise and artifacts; and some characteristic of the signal is extracted, usually called the signal "features". Feature extraction is a key step in BCIs success, since this step basically sends to the classifier algorithm what characterizes the desired and expected brain response. The classifier algorithm has the role of associating each feature set to different signals, that are, finally, translated to specific commands to be sent to the external device that the user wishes to control.

BCIs depend upon the user's intentional control, since they should associate a response to a mental task performed by him/her. [Pfurtscheller et al., 2005]. This means that the user generates mental patterns, according to a thought strategy, which the system associates with some event [Graimann et al., 2010]. The used strategy imposes limits to the BCI hardware and software, especially in the manner neural signals are processed. As with any human activity, it requires practice to improve, and the training time depends on the strategy itself. In addition, the system itself should adapt to its user, being able to respond to variations between different users and to the change of their strategies in time. Depending on the type of brain signals generated, ideally, a feedback can be given to the user in order to let them know if the chosen strategy is giving the expected results. If these results are not being achieved, he/she can try to modulate his/her brain signals and obtain better results.

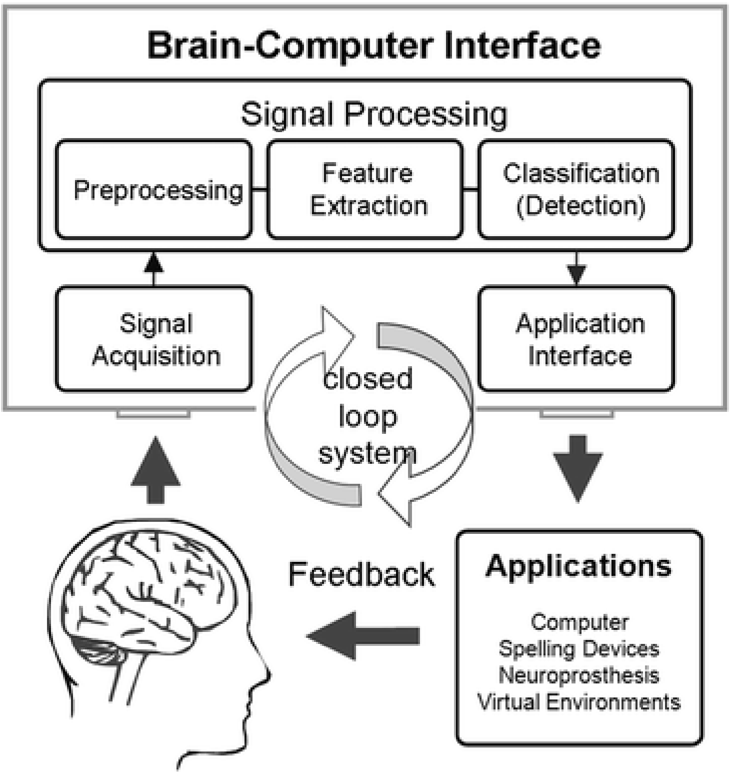


Figure 2.11. A general scheme of a BCI. Extracted from [Pfurtscheller et al., 2005].

Initially, BCIs were developed to improve the life quality of patients with serious motor deficits. In particular, the ones in the called "locked-in" state; a state in which they have lost all their mobility and have absolutely no manner of communication with the external environment [Wolpaw et al., 2002; Kubler et al., 2001]. More recently, however, BCIs have been investigated for other uses, such as monitoring the emotional state of patients [Widge et al. 2014] and soldiers [Miranda et al. 2014], monitoring the wakefulness state of drivers [Chuang et al. 2010], videogames [Bos et al. 2010] etc. BCIs can also provide insight into functional mechanisms of the brain, since understanding how responses to be used in such systems are generated is a crucial step of a BCI.

BCIs can rely on endogenous or exogenous tasks. The first ones are independent of external stimulation, and are based on mental tasks that can generate brain signals simply by enough concentration, such as motor imagery. This can be, however, a very tiring approach, as focusing in these types of task is not so simple. The latter ones are based on evoked responses due to external stimulation, such as visual stimuli. Their behavior is currently more well-established than signals generated by endogenous tasks, which constitutes an advantage. Nonetheless, their dependence on external stimulation is their main disadvantage, since additional external devices are needed to generate stimulation. Regardless of what type of task a BCI relies on, each one has its own advantages and disadvantages, which should be considered for each specific case.

Currently, there are many techniques for measuring brain signals, from non-invasive ones, such as functional magnetic resonance imaging (fMRI), near-infrared spectroscopy (NIRS), EEG and magnetoencephalography (MEG); semi-invasive techniques, such as electrocorticography (ECoG); to totally invasive ones, such as neural probes. For use in BCIs, each technique has advantages and disadvantages. The main advantage of the non-invasive methods is the fact that they are, precisely, non-invasive. This, however, comes at the cost of more recording noise and harder to decode data. Still in the non-invasive techniques, fMRI has the best spatial resolution but a high cost, while EEG has better temporal resolution, worse spatial resolution but lower cost. In addition, EEG equipment is portable. With all this considered, EEG has been the most used technique in BCI studies [Panoulas et al., 2010], and was also adopted in the present work. In the following section, we describe the main strategies used for generating brain signals in an EEG-based BCI.

2.3. Main strategies to generate input signals for EEG-BCI

In EEG-BCI studies, brain signals are usually triggered by a specific external stimulus; the EEG response is then called an "event related potential", or ERP. An ERP is, therefore, a very small voltage that is generated in the brain in response to specific events or stimuli [Sur and Sinha, 2009]. The main types of ERPs used for EEG-BCIs will be discussed in this section.

2.3.1. P300

P300 is an ERP, so called due to the fact that its positive amplitude peak occurs at about 300 ms after initial stimulation (Figure 2.12). It is a response located at the parietocentral scalp area (FCz, Cz and Pz) [Al-ani and Trad, 2010; Picton, 1992; Pritchard, 1981]. The first report of the P300 wave dates from 50 years ago [Sutton et al., 1965]. In this study, the authors found results that suggested an inverse relationship between the P300 amplitude and the stimulus probability [Pritchard, 1981]; that is, the lower the occurrence probability of a stimulus, the higher the evoked P300 amplitude. The P300 wave is associated with decision-making, being a brain response to an event with low occurrence probability and to a task that requires the subject's attention. It can be evoked by auditory, visual or somatosensory stimuli that are particularly significant and infrequent.

In Figure 2.12, N1, P2 and N2 indicate components of the waveform that are sensory evoked responses which do not correspond to the task responsible for evoking the P300. For instance, let P3 be evoked by a specific target event, "T", and N1, N2 and P2 be components present in a standard stimulation, "S". Every time a new stimulus is the same as the standard stimulus "S", P300 is not evoked, and only N1, N2 and P2 are the present components. If, however, the new stimulus is different from "S" (that is, "T"), and the subject is focused on the task, the P3 response is evoked in addition to N1, N2 and P2. The scheme in Figure 2.13 illustrates this process. This is known as the "context updating theory of P300" [Polich, 2003].

In BCI applications, the P300 wave is associated with the peak picking (PP) technique [Al-ani and Trad, 2010; Meinicke et al., 2003; Garrett et al., 2003; Bayliss et al., 2004; Bayliss and Inverso, 2005; Samili Khorshidi et al., 2007; Hoffman et al., 2008]. PP is an algorithm designed to detect the P3 component in the waveform "using the difference between the minimum and maximum amplitude in a trial" [Al-ani and Trad, 2010]. To do so, the difference between the maximum and minimum points of the recorded P300 data is evaluated within a time window. If this difference exceeds a threshold value, the algorithm response is a detected P3 component.

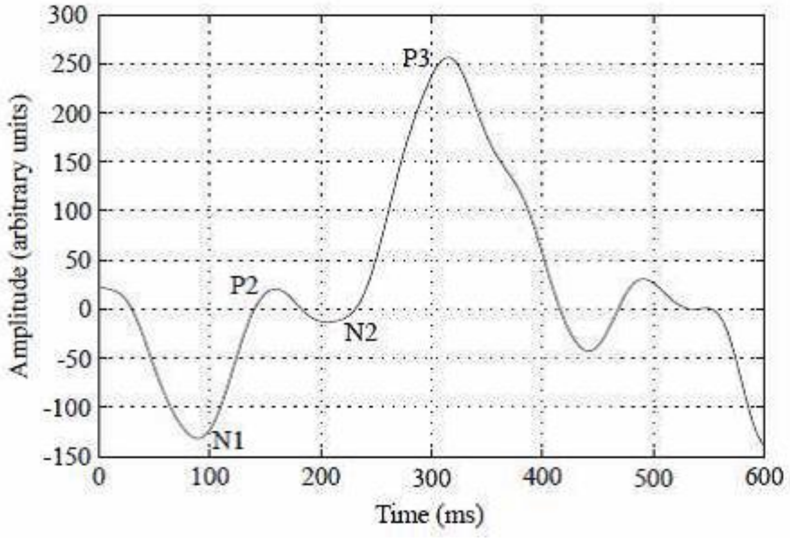


Figure 2.12. P300 potential illustration. The P300 component of the illustrated waveform is labeled as 'P3'. Extracted from [Al-ani and Trad, 2010].

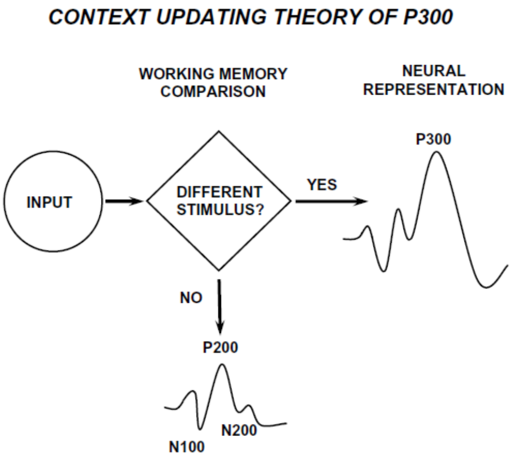


Figure 2.13. Context updating theory of P300 scheme. The P3 potential is evoked only in response to the target stimulus "T". Extracted from [Polich, 2007].

One of the main applications of P300 in BCIs has been the P300-BCI speller, which displays a matrix containing letters and numbers. Depending on the system, each letter flashes one at a time, or an entire row or column of the matrix flashes at a time. The principle is that a P3 component should be evoked every time the intended letter flashes (Figure 2.14). When this happens, the system detects the P300 wave and is able to identify the selected row, column or letter. Eventually, the user can spell entire sentences and communicate with the outside environment. BCIs of this type have been reported in the literature with accuracy rates 95%, showing it is a promising application [Donchin et al., 2000; Farwell and Donchin, 1988; Kaper et al., 2004; Krusienski et al., 2006; Rivet et al., 2009; Sellers et al., 2006].

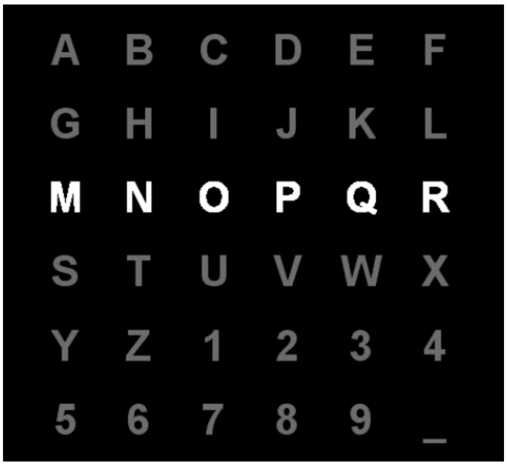


Figure 2.14. Character matrix displayed for P300-BCI application. Extracted from [Rivert et al., 2009].

Other applications involve other types of decision-making processes, such as discriminating between daily tasks [Hoffman et al., 2008], or giving commands to a wheelchair [Iturrate et al., 2009].

Given its nature and detectability, P300 is a well-established response that has been demonstrated to be very useful and successful in BCI applications.

2.3.2. SSVEP

A visually evoked potential (VEP) is an ERP that occurs in response to visual stimulation. The steady state visually evoked potential (SSVEP) is a special case of VEP, and it constitutes a response to stimuli modulated in a specific frequency range, from 3.5 to 75 Hz [Panoulas et al., 2010]. The SSVEP response has the same fundamental frequency as that from the source that originated it. For instance, if a person is stimulated by a 14 Hz blinking LED, the corresponding SSVEP response will have its fundamental frequency set at 14 Hz. It usually also includes the fundamental frequency's harmonics series (Figure 2.15). This type of response is more pronounced at the occipital lobe, since it is the region of the brain directly associated with visual processing.

It is important to note that SSVEPs and VEPs are different, and what distinguishes them is the repetition rate of the stimulation. An isolated visual stimulation will not produce an SSVEP. This response will only occur if the stimulus is presented repetitively, from a rate of 5 to 6 Hz or greater, originating an oscillatory response in the brain - the SSVEP [Al-ani et al., 2010].

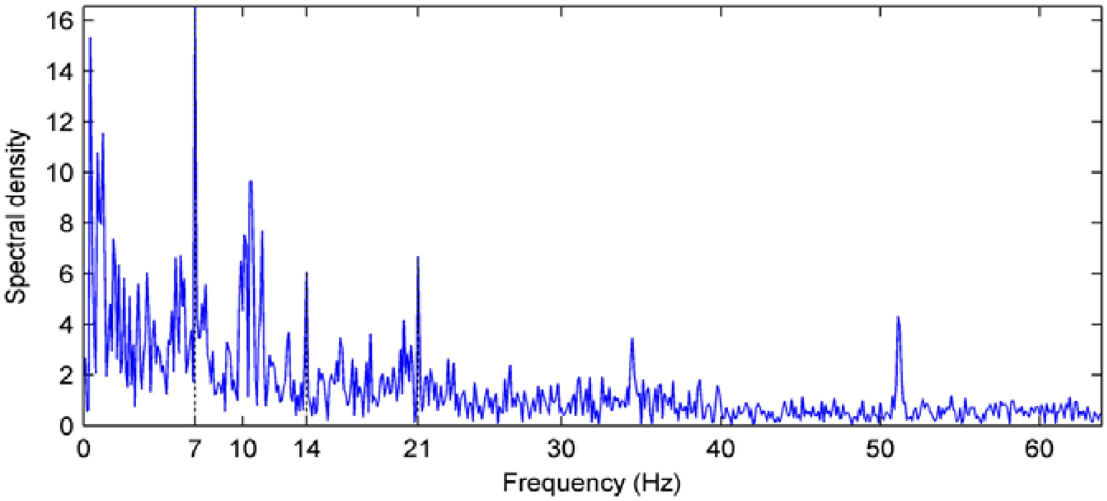


Figure 2.15. SSVEP response example. It is possible to note the higher peak at the fundamental frequency (7 Hz) and smaller peaks at the harmonic frequencies (14 Hz, 21 Hz, 35 Hz). Extracted from [Panoulas et al., 2010].

If two or more frequencies are presented as stimuli, the one that the BCI's user is focusing on is the prevalent one in the SSVEP response [Panoulas et al., 2010]. Therefore, the use of SSVEP in BCIs commonly involves presenting different flashing stimuli on a screen, and associating a specific command to each stimulus. The user's choice can then be identified by analyzing the peaks of the signal in the frequency-domain (Figure 2.16).

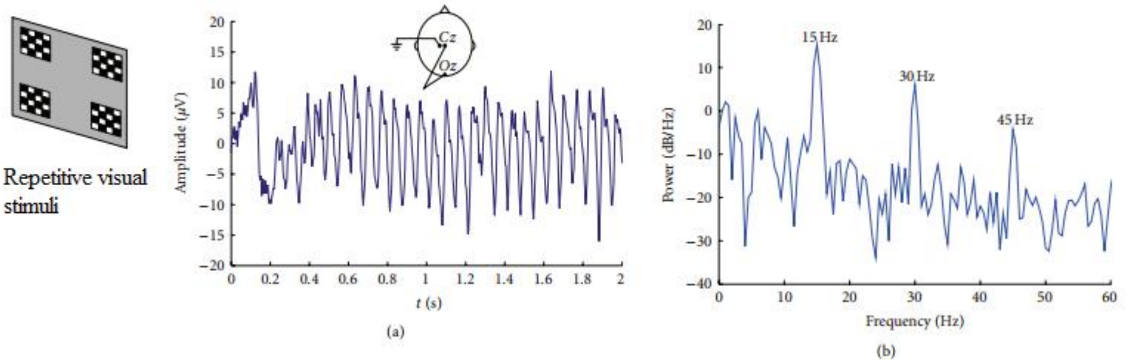


Figure 2.16. SSVEP-BCI scheme. Left: choices are presented to the user, flashing at different frequencies. Center: the EEG signal is measured at the occipital cortex. Right: The Fourier Transform of the EEG signal shows peaks at the frequency corresponding to the stimulus the user was looking at, accompanied by its harmonics. Adapted from [Zhu et al., 2010].

SSVEP-BCIs have the advantage of generating a robust response without requiring much training from the user. However, they do require their users to have good voluntary eye movement control [Panoulas et al., 2010]. Overall, given the nature of the SSVEP response, it seems a very feasible tool for BCI use. In [Guger et al., 2012], the authors explore how feasible an SSVEP-BCI was to 53 volunteers, obtaining an average accuracy of 95.5%, with no subject obtaining accuracies below 60%. They concluded that SSVEP-BCIs are systems that can provide very good classification rates with small training sections.

2.3.3. Motor Imagery (MI)

Some changes in EEG data occur due to events that cause an increase or decrease in the spectral power of certain frequency bands, such as motor behavior, sensory stimulation and mental imagery [Pfurtscheller et al., 2001]. These changes are often viewed with aid of the Fourier Transform in the frequency domain of the signal. When an increase happens, the event is called an event-related synchronization (ERS), and, when a decrease occurs, it is an event-related desynchronization (ERD). The reason for this to happen is considered to be due to the synchronization degree between neuronal populations [Panoulas et al., 2010]; that is, neurons firing synchronously should yield an ERS, and neurons firing asynchronously should yield an ERD.

ERS and ERD can occur simultaneously [Pfurtscheller et al., 2001]. At the same EEG electrode, for instance, at the same time, a certain frequency band can undergo an ERS, while another undergoes an ERD. In addition, at different scalp locations, the same frequency band can undergo an ERD at some region and an ERS at another one.

MI can be seen as a mental rehearsal of the motor act itself, without movement realization, and that triggers responses at similar areas to those for motor tasks (sensorimotor areas) [Pfurtscheller et al., 2001]. Hand movement is very well located in the motor cortex, and it is currently well established that MI of the hands results in an ERD of the μ band and central β rhythm at the contra-lateral hemisphere, and an ERS within the γ band at the ipsi-lateral hemisphere [Al-ani et al., 2010; Pfurtscheller and Aranibar, 1977; Pfurtscheller, 1999c; Neuper and Pfurtscheller, 1999a].

MI patterns are a little distinct than overt movement. During movement execution, the ERD found has a bilateral behavior [Pfurtscheller et al., 2001]. Therefore, ERDs of the contra-lateral hemisphere for μ and β rhythms are more evident during MI than during movement execution.

Figure 2.17 illustrates the contralateral ERD during hand MI for distinct subjects at different frequency ranges. Also, some ERS can be seen on the ipsilateral hemisphere for some subjects and frequencies.

MI-BCIs have an important advantage over P300-BCIs and SSVEP-BCIs, since they do not need external stimulation. MI-BCI users can, theoretically, imagine whatever they desire at any time, making it possible to give the system its input whenever it would be needed. In practice, however, it is not that simple. Although MI has been considered to have this characteristic response, imagining movements requires great focus, attention and effort, which the BCI user usually must train to develop [Slenes et al., 2013; McFarland et al., 1998; Hwang et al., 2009]. In addition, the brain response to MI usually has a large inter-variability and intra-variability, making the task of the classifier algorithm very demanding. On the other hand, MI-BCIs have been successful in aiding patients toward their motor rehabilitation, since they activate similar areas to the overt movement realization [Sharma et al., 2009; Wang et al., 2010].

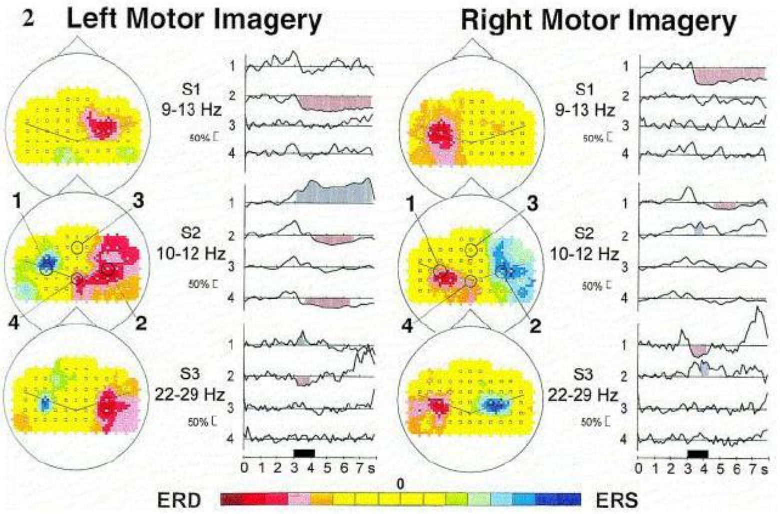


Figure 2.17. EEG “activations” when subjects imagine right or left hand movement. ERDs (red) and ERSs (blue) are indicated in a color scale for all three subjects (S1, S2 and S3). Note that MI of a specific hand causes ERDs on the contralateral hemisphere to which this task is done; while ERSs usually occur on the ipsilateral hemisphere. In addition, time courses of recorded signals are shown on the right to each scalp map, to each scalp location denoted by ‘1’, ‘2’, ‘3’ or ‘4’. Extracted from [Pfurtscheller and Neuper, 1997].

In MI-BCI applications, it is usual to search for distinguishable patterns between two responses. For instance, to use left and right hand imaginations and attribute a command to each one of the hands. In a typical protocol, data from both hand MIs are acquired separately. Data are pre-processed and the feature extraction is made. It is very common for the feature to be the power-spectral density (PSD) of specific frequency bands. In theory, ERS and ERD should be noticeable as changes in the PSD. Left hand MI, for instance, should cause the contra-lateral μ rhythm PSD to decrease, due to the occurring ERD. Distinct values of PSD at specific locations for each task can, then, be used to train a classifier algorithm. More recently, feature extraction of MI using graphs has also been studied, and will be discussed later on.

MI, despite its challenges, presents itself as a very interesting strategy, mainly because it does not require external stimulation and seems to be the a “natural” approach. It has been subject of intense research in the BCI area, with various studies trying to overcome its challenges

2.4. Common signal processing techniques in EEG-BCIs

In this section, some common preprocessing methods in BCI research and aspects regarding feature extraction that are relevant to the understanding of this work will be discussed.

2.4.1. Temporal Filtering

Usually for BCI applications, the information sought is located in well-specified frequency bands of interest. In this context, frequency filtering refers to filtering the signal in order to leave only specific frequency bands. If the alpha band is sought, for example, a band-pass filter between 8 to 12 Hz can be used; low-pass frequency filters, on the other hand, can eliminate sources of (high-frequency) electromagnetic noise; and so on.

2.4.2. Spatial Filtering

The application of spatial filters in BCI aims mainly the enhancement of the recorded EEG signal by increasing its signal-to-noise ratio. Two spatial filters are highlighted here: common average referencing and surface Laplacian.

A common average referencing (CAR) spatial filter averages the EEG signal across all recording electrodes, and subtracts this result, pointwise, at each electrode. This is expected to eliminate common artifacts arising in all channels at the same time. Differences in results of distinct EEG studies have been attributed, partly, to the difference on referencing the recording signal [Hagemman et al., 2001]. Assuming that the influence of the reference electrode is of equal magnitude in all EEG recording channels, then, a CAR filter should eliminate this effect over the recorded data [Stanny, 1989; Al-ani et al., 2010]. Mathematically:

$$V_{CAR}^{i} = V_{R}^{i} - \frac{1}{n}\sum_{j=1}^{n} V_{R}^{j}, \qquad \textbf{(2.2)}$$

in which n is the total number of electrodes, V_{R}^{i} is the voltage time series recorded at electrode i and V_{CAR}^{i} is the voltage time series at the same *i* electrode after the filter application.

The surface Laplacian (SL) consists of the instantaneous second derivative of the electrical potential spatial distribution. This filter enhances the high-frequency spatial components, and can achieve high spatial resolution by using a large number of electrodes [Al-ani et al., 2010; Panoulas et al., 2010]. This procedure accentuates localized activity and attenuates more diffuse activity [McFarland et al., 1997]. A SL calculation, as defined by [McFarland et al., 1997], uses a finite difference method to approximate the second derivative by subtracting the mean activity of the surrounding electrodes of some electrode *i* of interest as follows:

$$V_{SL}^{i} = V_{R}^{i} - \sum_{j \in S_i} g_{ij} V_{R}^{j}, \tag{2.3}$$

with

$$g_{ij} = \frac{\frac{1}{d_{ij}}}{\sum_{j \in S_i} \frac{1}{d_{ij}}}. \tag{2.4}$$

S_i represents the set of electrodes that surround the i^{th} electrode under consideration. d_{ij} is the distance between electrodes i and j., that can be taken to be nearest-neighbors or not (Figure 2.18).

CAR and SL are amongst the most used space filters in EEG-BCI research, with increasingly popularity [Al-ani et al., 2010]. They are relatively simple approaches that can be used to enhance the EEG signal-to-noise ratio.

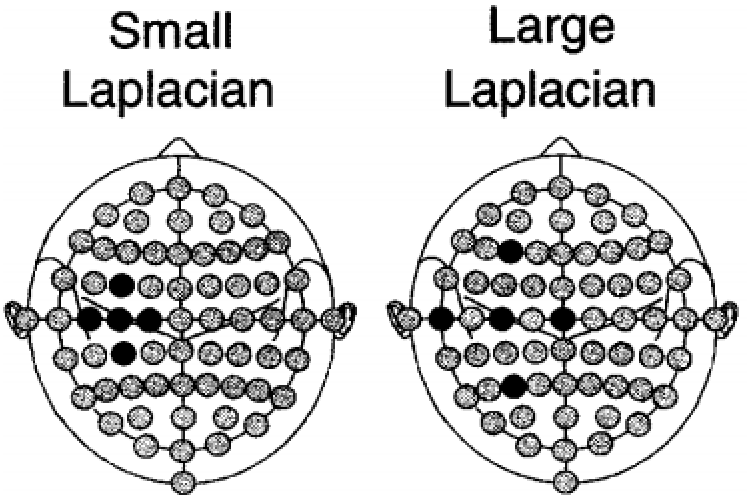


Figure 2.18. Examples using nearest-neighbors' electrodes (small Laplacian) and next-nearest-neighbors (large Laplacian). What should change in the SL filter equation is just the distance between the electrodes. Channels to be included in the filter's range are black. Adapted from [McFarland et al., 1997].

2.5. Classification approaches

The classification step is a crucial part of a BCI's scheme (see Figure 2.1), as it is the phase where the user's intent is recognized and transformed into a command to be sent to the desired application. In this section, two classification approaches commonly used in BCI applications are discussed: linear discriminant analysis and support vector machines.

2.5.1. Linear Discriminant Analysis

Due to its robustness and simplicity, linear discriminant analysis (LDA) is used for all types of BCI. In MI-BCI, many works use LDA providing, as input, the PSD of specific frequency bands (usually the µ rhythm), as discussed above. The mathematical formulation of the problem is stated below.

Given an input data vector **x**, whose components are x_1, x_2,...,x_n, (the so called “feature vector”), a linear classifier has the following structure:

$$y = \boldsymbol{w}^T\boldsymbol{x} + x_0, \tag{2.5}$$

in which (.)^T indicates the transpose of a matrix (or vector), **w** is a weight vector, to be determined, and x_0 is an offset value. Expression (2.5) can be thought of as the projection into a line, by means of **w**, of the data vector **x**. If ||**w**|| = 1, the *y* value is just the projection of the corresponding *x* into a line in the direction of **w**. The values of the components of **w** are not important, since they are scaling factors for y; the direction of **w**, however, matters. Simply choosing to project the input data in lines having different directions can have great implication on the separation efficiency (Figure 2.19). However, if the input data from the chosen feature are highly overlapping, even the best choice of **w** may not be able to satisfactorily separate the classes [Duda et al., 2001].

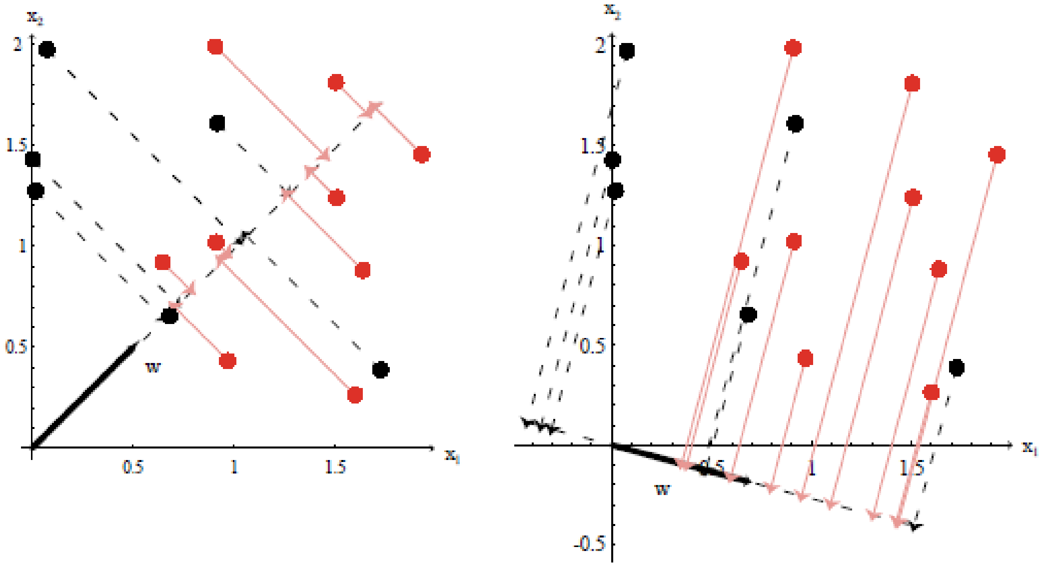


Figure 2.19. Two different projection examples for the same original input data. Projection on the right shows better separation. Extracted from [Duda et al., 2001].

In general, for linear classifiers, distinction between two classes is done from a hyperplane of the form $\boldsymbol{w}^T\boldsymbol{x} - \lambda = 0$ (see Figure 2.20), where λ is a decision threshold. If it is sought to discriminate between two classes, C$^+$ and C$^-$, then:

$$\begin{cases} y = \boldsymbol{w}^T\boldsymbol{x} \geq \lambda \rightarrow \text{the algorithm decides for } \mathrm{C}^+ \\ y = \boldsymbol{w}^T\boldsymbol{x} < \lambda \rightarrow \text{the algorithm decides for } \mathrm{C}^- \end{cases}. \quad \textbf{(2.6)}$$

In MI-BCI research, searching for differences between recorded MI tasks is basically a pattern recognition problem. This type of situation is usually not simple, and its output can be affected by a variety of factors, such as the feature's choice and noise. In some cases, prior knowledge of the problem under study can be of great help in defining which features should be chosen. Ideally, it is sought to use a feature that is both simple and robust, in the sense of not being easily affected by noise. In some cases, however, this is not possible, and prior knowledge and context become even more important.

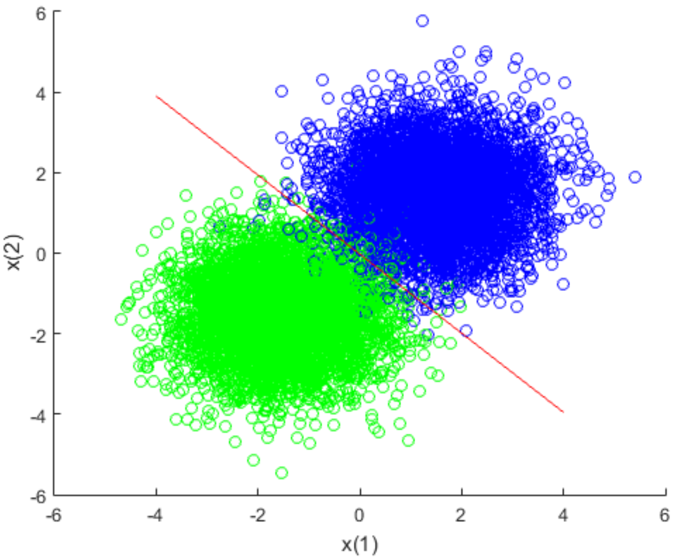


Figure 2.20. Example of two classes (blue and green) optimal separation using the least-square LDA method. The red line is the hyperplane that best separates the two clusters. Data were generated randomly.

Many criteria can be used for choosing the parameters and decision threshold of the problem. In this thesis, to determine the ***w*** vector, we applied a least-squares based methodology. The least-squares method is commonly applied to many situations of data regression and classification.

In the case of data classification, the idea is to find the optimum hyperplane that best separates two data classes; for instance, classification of new data between two classes – right and left hands MI. By labeling each class, for instance, by “1” and “-1”, the closer the classifier output (y) is to the corresponding class' label, the more accurate the classifier is. Therefore, defining $\boldsymbol{l}$ as the labels' vector, then:

$$l_i = \begin{cases} 1, \text{if } x_i \epsilon C^+ \\ -1, \text{if } x_i \epsilon C^- \end{cases}. \tag{2.7}$$

Note that each element of $\boldsymbol{l}$ need not be "1" or "-1", this is merely an example. Generally, any two constants would suffice.

In the most general case, let $\boldsymbol{X}$ be an $n \times d$ matrix, as follows (note that x_{i0} elements are analogous to the offset value x_0 of (2.5)), for which each i^{th} row is a vector $\boldsymbol{x}_i^T$:

$$\boldsymbol{X} = \begin{pmatrix} x_{10} & x_{11} & \cdots & x_{1d-1} & x_{1d} \\ \vdots & & \ddots & & \vdots \\ x_{n0} & x_{n1} & \cdots & x_{nd-1} & x_{nd} \end{pmatrix}. \tag{2.8}$$

The labels' vector $\boldsymbol{l}$ and the attribute matrix $\boldsymbol{X}$ are related by $\boldsymbol{w}$:

$$\boldsymbol{l} = \boldsymbol{wX}. \tag{2.9}$$

It is sought to minimize the quadratic norm of the residual value between the classifier output and the actual class label. Let this quantity be identified by the cost function $J(\boldsymbol{w})$. Then:

$$J(\boldsymbol{w}) = ||\boldsymbol{Xw} - \boldsymbol{l}||^2. \tag{2.10}$$

Alternatively, Equation (2.10) can also be written explicitly in terms of the matrix elements as:

$$J(\boldsymbol{w}) = \sum_{i=1}^{n} (\boldsymbol{w}^T \boldsymbol{x_i} - l_i)^2. \tag{2.11}$$

To minimize $J(\boldsymbol{w})$, its first derivative at a given point must equal zero, and its second derivative at that same point must be positive. Therefore:

$$\frac{dJ(\boldsymbol{w})}{d\boldsymbol{w}} = 0 \rightarrow 2 \sum_{i=1}^{n} (\boldsymbol{w}^T \boldsymbol{x_i} - l_i)\boldsymbol{x_i} = 0. \tag{2.12}$$

Expression (2.12) can also be written as:

$$2\,\boldsymbol{X}^T(\boldsymbol{X}\boldsymbol{w}-\boldsymbol{l})=0. \tag{2.13}$$

Then:

$$\boldsymbol{X}^T\boldsymbol{X}\boldsymbol{w}=\boldsymbol{X}^T\boldsymbol{l}. \tag{2.14}$$

Matrix $\boldsymbol{X}^T\boldsymbol{X}$ is a square d-by-d matrix and often nonsingular [Duda et al., 2001]. Therefore, the solution for $\boldsymbol{w}$ from Equation (2.14) is:

$$\boldsymbol{w}=(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{l}\;=\;\boldsymbol{X}^{\dagger}\boldsymbol{l}. \tag{2.15}$$

In (2.15), $\boldsymbol{X}^{\dagger}\overset{\text{def}}{=}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T$ is known as the pseudoinverse of matrix $\boldsymbol{X}$. Note that when the number of samples available for classification and the number of features are equal, that is, $n = d$, then the pseudoinverse is square, and coincides with the matrix's regular inverse [Duda et al., 2001].

In addition to providing a relatively simple mathematical formulation of the problem, LDA classifiers present the advantage of being of low computational cost when compared to more elaborated methods. Besides, they can be very robust given the right type of input features.

2.5.2. Support Vector Machines

Support vector machines (SVMs) also seek to discriminate between classes by finding the hyperplane that best separates them. To do so, however, only the support vectors from each data class are used (providing the name for the method). Support vectors are the ones that lie on the margin of each data class (Figure 2.21). The hyperplane to be found must maximize the distance between the hyperplane itself and the support vectors.

Firstly, let us consider a two linearly separable classes classification problem. As previously, let $\boldsymbol{x}_i$ denote the input data, l_i be the label for each $\boldsymbol{x}_i$ and $\boldsymbol{w}$ be a weight vector. Given a parameter b, we can write

$$\begin{cases}\boldsymbol{w}\boldsymbol{x}_i+b\geq+1\rightarrow\;l_i\;=+1\\ \boldsymbol{w}\boldsymbol{x}_i+b\leq-1\rightarrow\;l_i\;=-1\end{cases}. \tag{2.16}$$

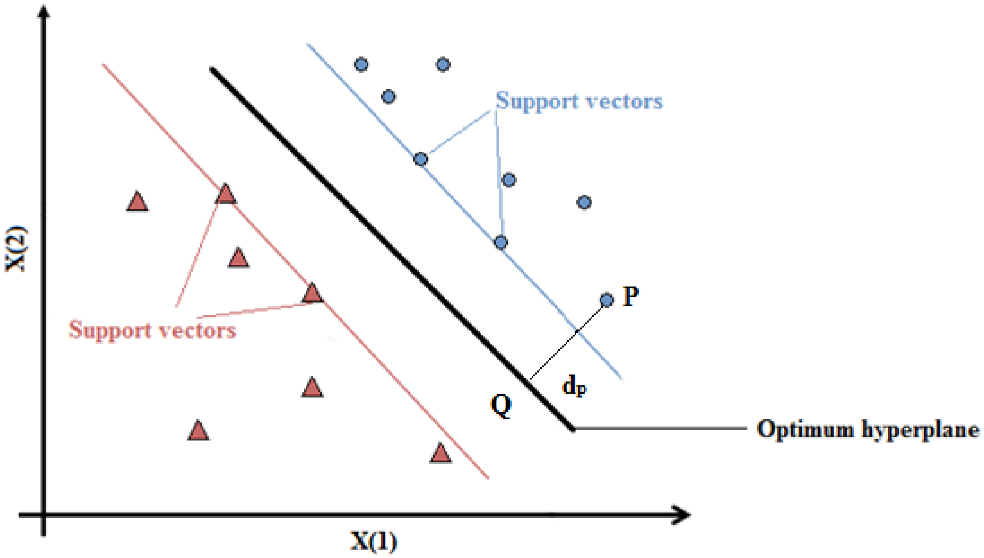


Figure 2.21. SVM optimum hyperplane to separate margins between two classes.

Basically, (2.16) states that $\boldsymbol{wx}_i + b$ belongs to either one of the classes, depending on whether it is regarding a threshold indicating the margins of the classes. These inequalities can be summarized into just one expression:

$$l_i(\boldsymbol{wx}_i + b) - 1 \geq 0, \qquad \textbf{(2.17)}$$

which is valid for all i.

If data are linearly separable, then the equality situation in (2.16) occurs for data lying on the margins – that is, the support vectors. Then:

$$l_i(\boldsymbol{wx}_i + b) = 1. \qquad \textbf{(2.18)}$$

Let us take the equation of the hyperplane to be

$$f(\boldsymbol{x}) = \boldsymbol{w}^T\boldsymbol{x} + b = 0. \qquad \textbf{(2.19)}$$

Since the separating hyperplane is the decision boundary, any point on it must obey Equation (2.19). Note that $\boldsymbol{w}$ is orthogonal to the separating hyperplane.

Now, let d_P be the distance between an arbitrary point P and the separating hyperplane, and Q be the projection of P onto the hyperplane (see Figure 2.21). If $\boldsymbol{x}_P$ represents the data point on P, then, point Q is given by:

$$\boldsymbol{x}_Q = \boldsymbol{x}_P - \frac{d_P}{y_P}\frac{\boldsymbol{w}}{||\boldsymbol{w}||}. \tag{2.20}$$

y_P merely scales the expression according to whether the positive or negative labeled class is under analysis. Since Q lies on the decision boundary, it must satisfy Equation (2.19). Then, putting (2.20) into (2.19):

$$\boldsymbol{w}^T\left(\boldsymbol{x}_P - \frac{d_P}{y_P}\frac{\boldsymbol{w}}{||\boldsymbol{w}||}\right) + b = 0. \tag{2.21}$$

Proceeding with the calculations:

$$d_P = y_P\left(\frac{\boldsymbol{w}^T\boldsymbol{x}_P + b}{||\boldsymbol{w}||}\right). \tag{2.22}$$

Given all training set points $\boldsymbol{x}_P$ with their respective labels y_P, the geometric margin of parameters (w, λ) with respect to this set is given by the smallest possible d_P value. Maximizing the distance to this margin is a manner to find the optimal boundary, related to the classification boundary. This problem is analogous to minimizing the norm of $\boldsymbol{w}$, subject to constraints (2.17).

The solution for this type of problem leads to an optimization problem of quadratic order. Its solutions may be found from the following Lagrangian [Leite, 2016]:

$$L(\boldsymbol{w}, \lambda, \mathrm{b}) = \frac{1}{2}\boldsymbol{w}^T\boldsymbol{w} - \sum_{i=1}^{n}\lambda_i[y_i(\boldsymbol{w}^T\boldsymbol{x} + b) - 1]\,; \tag{2.23}$$

where λ represents the Lagrange multipliers and n is the number of samples available for training the classifier. To solve this problem, we want to minimize L with respect to $\boldsymbol{w}$ and b, and maximize it with respect to Lagrange's multipliers. This means that:

$$\frac{\partial L(\boldsymbol{w}, \lambda, b)}{\partial \mathbf{w}} = \boldsymbol{w} - \sum_{i=1}^{n}\lambda_i y_i \boldsymbol{x}_i = 0; \tag{2.24}$$

$$\frac{\partial L(\boldsymbol{w}, \lambda, b)}{\partial b} = \sum_{i=1}^{n}\lambda_i y_i = 0. \tag{2.25}$$

From (2.26):

$$\boldsymbol{w}_{optimum} = \sum_{i=1}^{n} \lambda_i y_i \boldsymbol{x}_i. \tag{2.26}$$

Expression (2.28) shows that the optimum weight vector is a linear combination of the input data, for which the λ_i are non-zero. Now, putting (2.26) and (2.27) into (2.25) yields:

$$L(\boldsymbol{w}, \lambda, \mathrm{b}) = \sum_{i=1}^{n} \lambda_i - \frac{1}{2} \boldsymbol{w}_{optimum}^{T} \boldsymbol{w}_{optimum}. \tag{2.27}$$

Thus, the task is to maximize L with respect to the Lagrange multipliers in order to find the maximum margin subject to constraints (2.25) and (2.17).

In the cases for which data classes are not linearly separable in the original space, it is possible to choose a function $\phi(\boldsymbol{x})$ that maps the data points to a higher dimension space, in which classes are linearly separable (Figure 2.22).

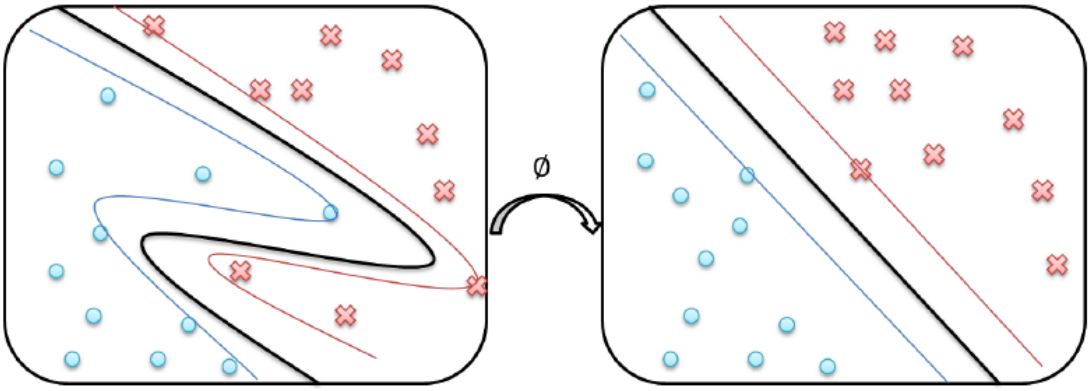


Figure 2.22. Mapping of input dataset x by function ϕ. ϕ(x) takes the feature points to a new space of higher dimensionality, where classes are linearly separable. Extracted from [Leite, 2016].

The mathematical formulation for this problem is very much alike to what was done previously. This time, however, the classifier output is written in terms of $\phi(\boldsymbol{x})$:

$$f(\boldsymbol{x}) = \boldsymbol{w}^T \phi(\boldsymbol{x}) + b. \tag{2.28}$$

Thus, for this case, the optimum $\boldsymbol{w}$ is now taken to be

$$\boldsymbol{w}_{optimum} = \sum_{i=1}^{n} \lambda_i y_i \phi(\boldsymbol{x}_i). \tag{2.29}$$

Using (2.29) into (2.28) yields:

$$f(\boldsymbol{x}) = \sum_{i=1}^{n} \lambda_i y_i \phi^T(\boldsymbol{x}_i)\,\phi(\boldsymbol{x}) + b. \tag{2.30}$$

Note that the classifier's output depends on a scalar product involving the function $\phi(\boldsymbol{x})$. Therefore, it is not necessary to know exactly how $\phi(\boldsymbol{x})$ operates over all data points; instead, the classifier can operate indirectly via the inner product of the function $\phi(\boldsymbol{x})$. In this context, it becomes useful to define a kernel function $K(\boldsymbol{u},\boldsymbol{v})$ as:

$$K(\boldsymbol{u},\boldsymbol{v}) = \phi^T(\boldsymbol{u})\phi(\boldsymbol{v}). \tag{2.31}$$

It can be ensured that $K(\boldsymbol{u},\boldsymbol{v})$ indeed defines this scalar product, given that it satisfies the following conditions (known as Mercer's conditions) [Cortes and Vapnik, 1995]:

$$\iint K(\boldsymbol{u},\boldsymbol{v})g(\boldsymbol{u})g(\boldsymbol{v})d\boldsymbol{u}d\boldsymbol{v} > 0; \tag{2.32}$$

this must be satisfied for all g, given that:

$$\int g^2(\boldsymbol{u})d\boldsymbol{u} < \infty. \tag{2.33}$$

A variety of kernel functions can be used. In this work, we used linear and polynomial kernels. A polynomial kernel of degree m is given by (2.34).

$$K(\boldsymbol{u},\boldsymbol{v}) = (\alpha\boldsymbol{u}^T\boldsymbol{v} + c)^m. \tag{2.34}$$

Adjustable parameters are the polynomial degree m, the constant c and the slope α. A linear kernel can be obtained simply by setting $m = 1$.

SVMs present the possibility of working with non-linear approaches and, even though the fitting is based only on the support vectors (which are, usually, in smaller number than the total input samples for the classifier training), the algorithm is very robust. Both linear and polynomial SVM were tested in this work, and their performances were compared to the least-squares based LDA.

2.6. Final remarks regarding MI EEG-BCIs

EEG-BCIs have been shown to be a viable alternative for direct communication from brain signals. However, their main problem is to identify specific patterns within the recorded data that provide good classification rates. This poses other questions that must be answered to satisfactorily deal with this problem: what is the best manner to perform data acquisition? What would be the best processing steps for a specific situation (if such one does exist)? What classifier should be used, and with what feature? Answering some of these questions requires prior knowledge of the problem under study.

In the case of MI-BCIs, during data acquisition, the subject must be provided with enough time to satisfactorily perform the proposed tasks. A lot of attention is necessary during trials to ensure no artifact will compromise data analysis. For MI, subjects tend to report having a hard time imagining their movements, suggesting that this task is not as trivial as it may seem.

Processing steps should be able to enhance the EEG SNR. EEG being a technique that is highly affected by noise, it can be very difficult to identify specific mental patterns if the signal has not been properly treated. Also, distinct combinations of processing steps used, method used for feature selection and classification algorithm can affect the classifier’s results. Also, high inter-subject variability suggests that BCIs’ performance can be enhanced if they are adapted to each individual user.

As discussed, there are many factors that can affect a BCI’s performance. A considerable amount of research is being done in each one of a BCI’s crucial steps: data acquisition, signal processing and data classification. Specifically, for MI-BCIs, still no optimum way of going through these steps has been established in the literature. In particular, the approach presented here explores non-conventional methods for feature extraction through metrics from graphs theory. In the next chapter, a description of graph theory is given.

Chapter 3

Graph Theory

The main purpose of this work was to investigate the feasibility of using graph features obtained from MI tasks to distinguish between left and right hand motor imagery. Therefore, a basic notion of graph theory and some common metrics associated with it is necessary for fully understanding what was done. Graph theory is a subject about which whole books could be written, which is obviously not the purpose of this work. The aim of this chapter, then, is to summarize the main properties that have been commonly used in BCI studies with graphs. The chapter starts by shortly defining graphs, explaining common metrics and, finally, indicating how these have been used in BCI applications.

.

3.1. Basics of Graph Theory

Graphs theory is a mathematical field that allows description of phenomena from many areas, such as social networks, communications, neural networks and so on. Basically, a graph consists of an ensemble of elements that can interact with each other, pair-wise. Each of these elements is denominated a "node", and the interaction between them is termed a "link". If two nodes do not interact, it is said there is no link between them. Nodes can also be called "vertices" or "points", and links are sometimes termed "edges".

In the context of using graphs to model brain regions' interactions by means of EEG data, two main approaches are largely used. One option is to consider each individual EEG electrode as a node, and to establish links between these nodes by some similarity measure using the signals' features, such as amplitude and phase. Another approach is to determine the EEG signal's sources and, then, to consider these sources as the graph's nodes. Their links can be determined in the same manner as previously stated. Regardless of the approach, the idea of applying graphs to EEG data is that they provide a model of interaction between distinct neuronal populations.

A more formal definition of graphs requires a mathematical formulation. A graph is a pair $G = (\mathrm{V}, \mathrm{E})$ of sets where $V = \{\mathrm{v}_1, \mathrm{v}_{2,\dots,}\mathrm{v}_\mathrm{i}\}$ represents its vertices, and $E = \{(\mathrm{e}_1, \mathrm{e}_2), \dots, (\mathrm{e}_\mathrm{i}, \mathrm{e}_\mathrm{j})\}$ contains the edges. The notation $(\mathrm{e}_\mathrm{i}, \mathrm{e}_\mathrm{j})$ indicates that there is an existing link between nodes *i* and *j*. Figure 4.1 is an example of a simple graph, with $V = \{1, 2, 3, 4, 5\}$ and $E = \{(1,2), (2,1), (2,3), (3,2), (2,5), (5,2), (4,5), (5,4)\}$.

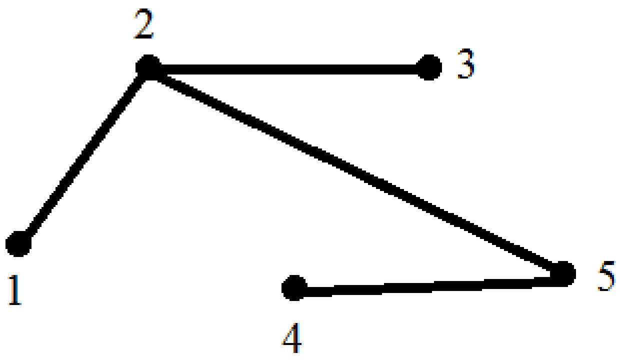
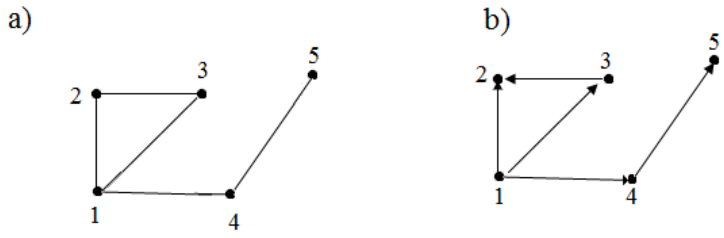


Figure 3.1. Illustrative example of a graph with 5 nodes. Lines connecting nodes represent the links between them.

The mathematical representation of a graph is its adjacency matrix ($\boldsymbol{A}$). The matrix elements, a_{ij}, represent the connections between nodes *i* and *j*. For an undirected graph, the interaction between any two nodes is symmetric, which means that $a_{ij} = a_{ji}$. On the other hand, in a directed graph, in which interaction between two nodes has a directionality (such as an effect of causality or transfer of information), it is possible to have $a_{ij} \neq a_{ji}$. In this case, the link is taken to be from node *i* to node *j*. Figure 4.2 shows an example, with binary undirected (left) and directed (right) graphs, and their corresponding adjacency matrices.



$$A = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \qquad A = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Figure 3.2. Example of undirected graph (a) and directed graph (b). Both graphs have the same nodes; however, their adjacency matrices are significantly different.

Graphs can also be classified as "weighted" or "unweighted". The latter type are binary graphs, i.e., links between two nodes can either exist ($a_{ij} = 1$) or not ($a_{ij} = 0$). A weighted graph can have links with different values (weights), that are not necessarily zero or one. In this case, it is usual to indicate the elements of the adjacency matrix by w_{ij}, which represent the values of the links.

Weighted and unweighted graphs can be built from the same basic principle. For example, suppose it is sought to construct an adjacency matrix that describes the interaction between different brain regions by measuring its electrical activity using EEG electrodes. A statistical correlation method could be used to compare the signal's time series, and to assign a correlation coefficient to each pair *i*-*j* of electrodes, c_{ij}. Therefore, each c_{ij} is an element of the adjacency matrix that describes this weighted graph. If it is desired to work with an unweighted graph, a threshold can be defined, to set every c_{ij} to 1, if its value exceeds the threshold, or 0, otherwise.

3.2. Common Graph Metrics

When working with small graphs, simple visualization of the graph may provide all the information needed from it. Frequently, however, graphs are built with a large number of nodes, making its analysis via simple visualization unfeasible. In such a case, graph metrics can help to extract information from the system represented by the graph. In this section, some common graph metrics often used are presented.

3.2.1. Degree

The size of a graph is given by its total number of edges, and a graph's order is indicated by its total number of nodes [Bessa et al., 2010]. The degree property takes into consideration how many links are adjacent to each node. For an undirected graph, the degree of node *i* (k_i) is given by:

$$k_{\mathrm{i}} = \sum_{\mathrm{i}=1}^{\mathrm{N}} a_{ij}, \tag{3.1}$$

in which *N* is the total number of nodes. The degree is intuitively extended to weighted graphs by just replacing a_{ij} by w_{ij}.

For directed graphs, it is possible that a node may only have links going inward or outward (Figure 3.2). It is necessary, then, to define two distinct types of degree: the in-degree $\left(k_i^{in}\right)$, being the number of links a node receives, and the out-degree (k_i^{out}), which is the number of links exiting the node.

$$k_i^{in} = \sum_{\mathrm{i}=1}^{\mathrm{N}} a_{ij}\,; \tag{3.2}$$

$$k_i^{out} = \sum_{\mathrm{j}=1}^{\mathrm{N}} a_{ji}. \tag{3.3}$$

The mean degree of a graph is simply the arithmetic mean of the degree. For an undirected graph, this means that:

$$\langle k \rangle = \frac{1}{\mathrm{N}} \sum_{\mathrm{j}=1}^{\mathrm{N}} k_i. \tag{3.4}$$

The notation $\langle x \rangle$ was used to express the mean value of the variable x. In this type of graph, in summing $\sum_{\mathrm{j}=1}^{\mathrm{N}} k_i$, it is accounted for the total number of edges (E), twice. Then:

$$\langle k \rangle = \frac{2\mathrm{E}}{\mathrm{N}}, \tag{3.5}$$

for an undirected graph.

In the case of directed graphs, the edges start and end in specific nodes, and the double-accounting that previously occurred for undirected graphs is no longer present. Then:

$$\langle k_i^{out} \rangle = \sum_{\mathrm{i}=1}^{\mathrm{N}} k_i^{out}\,; \tag{3.6}$$

$$\langle k_i^{in} \rangle = \sum_{\mathrm{i}=1}^{\mathrm{N}} k_i^{in}\,; \tag{3.7}$$

$$\langle k_i^{in} \rangle = \langle k_i^{out} \rangle \;=\; \frac{\mathrm{E}}{\mathrm{N}}. \tag{3.8}$$

An important property of a graph directly associated with the degree is the "degree distribution". It indicates the probability $P(k_i)$ of node i having the degree k. The degree distribution can be best seen in the form of a histogram (Figure 3.3).

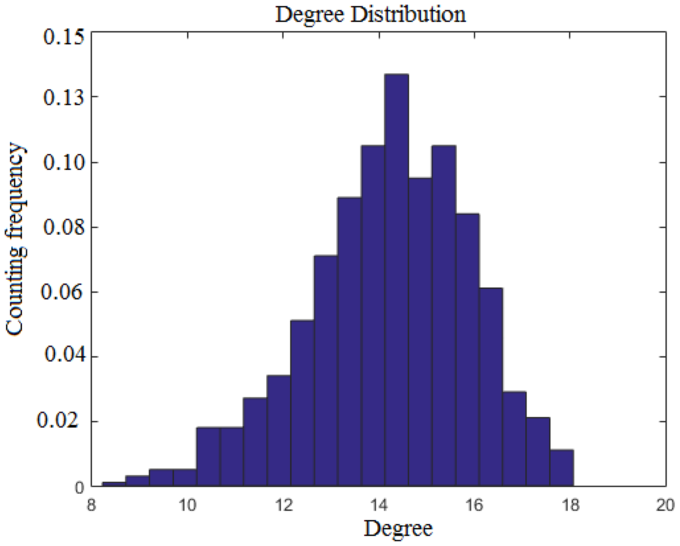


Figure 3.3. Graph's degree distribution example.

3.2.2. Clustering Coefficient

The clustering coefficient (CC) is a metric associated with the degree, that contains information on how the nodes are organized. It indicates the probability that two nodes that are adjacent (linked) to a certain node *i* are also adjacent to each other. Therefore, the clustering coefficient (C) can be expressed as:

$$C_i = \frac{\text{number of closed triplets connected to node 'i'}}{\text{number of triplets of connected vertices, centered on node 'i'}}. \tag{3.9}$$

Figure 3.4 shows different cases of clustering coefficient values for the same node. In (a), all neighbors to node *i* are also neighbors to each other, accounting for a CC of 1. In (c), these neighbors are not connected to each other, resulting in a CC of 0. In (b), an intermediate situation is shown.

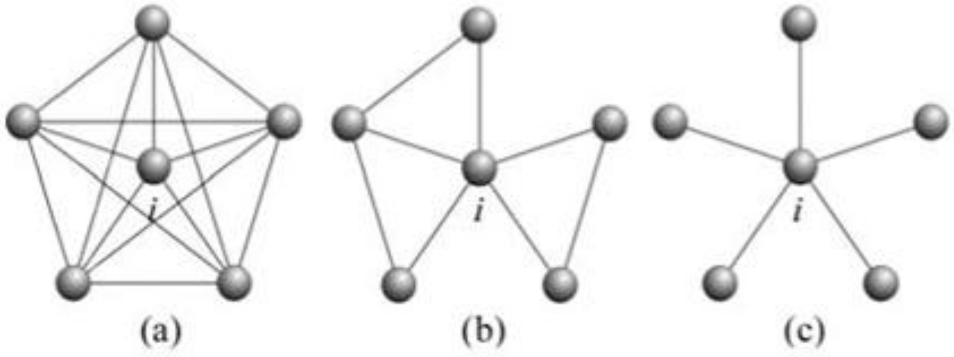


Figure 3.4. Clustering coefficients for the central node of a star graph. Three situations are shown: an entirely connected graph (a), with a CC for node *i* of 1; the opposite situation, for which CC for the same node is 0 (c); and an intermediate case (b). Extracted from [Costa et al., 2008].

Let k_i be the degree of node *i* for an unweighted graph. This means that this node has k_i neighbors. Then, the maximum possible number of edges (E^i_{max}) occurs when all neighbors of *i* are connected to *i* and to each other, simultaneously, which translates into the denominator of (3.9). In this situation:

$$\mathrm{E}^{\mathrm{i}}_{\mathrm{max}} = \binom{k_i}{2}; \tag{3.10}$$

$$\mathrm{E}^{\mathrm{i}}_{\mathrm{max}} = \frac{\mathrm{k_i}!}{(\mathrm{k_i}-2)!\,2!} = \frac{\mathrm{k_i}(\mathrm{k_i}-1)(\mathrm{k_i}-2)!}{2(\mathrm{k_i}-2)!}; \tag{3.11}$$

$$\mathrm{E}^{\mathrm{i}}_{\mathrm{max}} = \frac{\mathrm{k_i}(\mathrm{k_i} - 1)}{2}. \tag{3.12}$$

Let the numerator of (3.9) be expressed as *b_i*. It can be seen as the product of elements of the adjacency matrix, since a contribution for the cluster coefficient will only exist if there is a closed triplet at node *i*; that is, if two nodes (*j* and *l*), adjacent to *i*, have non-zero matrix elements. Then, to compute b_i, it must be taken into consideration three matrix elements, analyzing if a link exists between *i* and *j*, *i* and *l*, and *j* and *l*, simultaneously. Mathematically:

$$b_i = \sum_j \sum_l a_{ij} a_{il} a_{jl}. \tag{3.13}$$

Therefore, the expression for the clustering coefficient can be rewritten to take the following form:

$$C_i = \frac{\mathrm{b_i}}{\mathrm{E}^{\mathrm{i}}_{\mathrm{max}}}. \tag{3.14}$$

Substituting (3.12) and (3.13) into (3.14) yields

$$C_i = \frac{2 \sum_j \sum_l a_{ij} a_{il} a_{jl}}{\mathrm{k_i}(\mathrm{k_i} - 1)}. \tag{3.15}$$

The average clustering coefficient is calculated simply by averaging C_i over the total number of nodes, N.

$$\langle C \rangle = \frac{1}{\mathrm{N}} \sum_{\mathrm{i}=1}^{\mathrm{N}} C_i. \tag{3.16}$$

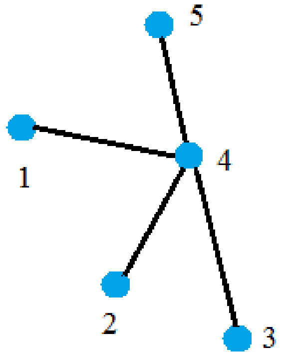
The extension of the calculation of the clustering coefficient for weighted graphs is not so straightforward as for the degree. Antoniou and Tsompa reviewed common calculations used to extend this concept to weighted graphs. They showed that all definitions of the clustering coefficient for a weighted network reduced to the definition (3.15) by replacing the adjacency matrix elements by their respective weights, w_{ij} [Antoniou and Tsompa, 2008].

3.2.3. Characteristic Path Length

It may be sought how far apart two nodes are from each other. A path between two nodes *i* and *j* of a graph can be defined as a sequence of *n* vertices, $v_1, \ldots, v_n$, in which: $v_1 = i$ and $v_n = j$; and for which any *l* defined in the interval $1 \leq l \leq n - 1$, there is a link between v_l and v_{l+1}; and there is no vertices or links repetition within this sequence [Bessa et al., 2010].

The path length between *i* and *j* refers to the number of edges connecting these two nodes corresponding to this path. The distance between these nodes is given by the shortest path length between them (l_{ij}). When *i* and *j* are adjacent to each other, $l_{ij} = 1$.

Using this concept, it is possible to define the graph distance matrix (L): a square matrix in which each element l_{ij} indicates the distance between nodes *i* and *j* (Figure 3.5).



$$L = \begin{bmatrix} 0 & 2 & 2 & 1 & 2 \\ 2 & 0 & 2 & 1 & 2 \\ 2 & 2 & 0 & 1 & 2 \\ 1 & 1 & 1 & 0 & 1 \\ 2 & 2 & 2 & 0 & 2 \end{bmatrix}$$

Figure 3.5. The graph's distance matrix (L) contains the minimum path length between nodes 'i' and 'j'. For instance: from node '1', it is necessary to undergo two edges to reach node '5', and only one link, to node '4'.

For every node *i* it is possible to define its minimum path length by averaging all minimum distances to every possible reachable node from *i*:

$$\langle l_i \rangle = \frac{1}{\mathrm{N}-1} \sum_{\mathrm{j}=1}^{\mathrm{N}-1} l_{ij}. \tag{3.17}$$

The network's average minimum path is the mean of every l_i:

$$\langle l \rangle = \frac{1}{\mathrm{N}} \sum_{\mathrm{i}=1}^{\mathrm{N}} l_i. \tag{3.18}$$

As with the degree, the minimum path can directly be extended to weighted networks. Two different generalizations for this case are present in the literature, usually used in communication and transportation networks [Antoniou and Tsompa, 2008]. To define a new notation, let d_{ij} be the equivalent of l_{ij}, for weighted graphs. Then, regarding communication networks, the shortest path length between nodes *i* and *j* becomes:

$$d_{ij} = min \sum_{i,j} \frac{1}{w_{ij}}. \quad \textbf{(3.19)}$$

Under this context, a large connection weight indicates great efficiency on communication between the two nodes i and j.

For transportation networks, the definition is:

$$d_{ij} = min \sum_{i,j} w_{ij}. \quad \textbf{(3.20)}$$

For both types of networks (communication or transportation), average values of the shortest path length are direct generalizations of (3.17) and (3.18).

3.2.4. Centrality measures

Differently than from the previous metrics, centrality measures can only occur as local properties, as they are related to the relative importance of a node within a network. Since they can be defined under specific criteria, different types of centrality measures have been established in the literature [Newman, 2005]. In this thesis, three types of centrality were studied: degree centrality (DC), betweenness centrality (BC) and eigenvector centrality (EC).

As will be explained further (Chapter 4), working with weighted graphs was preferable. In our case, then, the DC coincides with the strength (or weighted degree) of a node. If $\boldsymbol{W}$ is a weighted adjacency matrix whose elements w_{ij} represents every connection weight between any pair of nodes *i* and *j*, then, the DC for node *i* is given by [Zhang et al., 2012; Barrat et al., 2004]:

$$DC_i = \sum_{i,j} w_{ij}. \quad \textbf{(3.21)}$$

Note that the DC is a very simple measure, in the sense that only information about the weight of the connections is taken into consideration. A node can, however, possesses low connection weights and, even so, be important to the network regarding some other criterion. Consider, for instance, Figure 3.6, below. To simplify our discussion, let us assume that all connections have equal weight (this, however, need not to be the case). Node A, then, will present low DC when compared to nodes like B. On the other hand, this node is also the only one that connects the blue and red parts of the graph, being vital to perform communication between the two areas. Therefore, in the sense of information flow, for instance, node A would be one of the most (if not the most) central nodes of the graph. The BC metric refers to this idea: node A would have a high BC value.

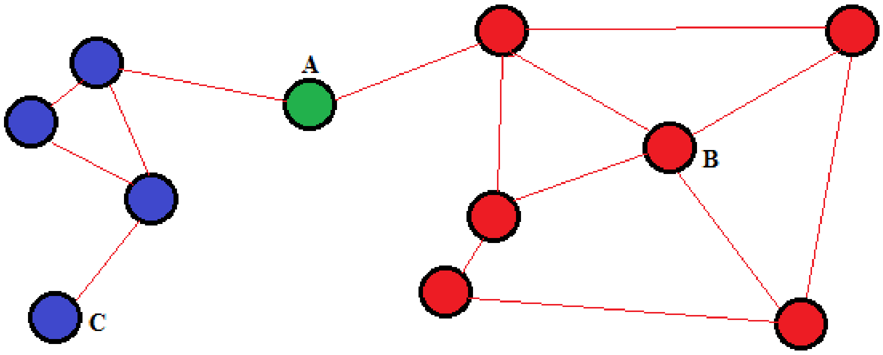


Figure 3.6. A graph illustration for centrality measures comparison. Although node A has low DC values, it is vital for other interpretations, such as information flux through the red and blue parts of the graph.

Continuing with the example in Figure 3.6, node B is highly connected and can, also, intermediate flux (of information, people etc.) between other nodes, mainly within the red ones. It could, then, show high DC and BC values. Node C has only one connection and does not connect any other two nodes of the graph. It would, then, present low DC and zero BC. Figure 3.6, then, is an example that an important node regarding one centrality metric is not, necessarily, central under another criterion as well.

BC, then, accounts for how important a node is to intermediate interaction between any other two nodes of a graph. In more technical terms, BC for node *i* is related to how many shortest paths between any other nodes *j* and *k* (l_{jk}) necessarily pass through *i* ($l_{jk}(i)$). Mathematically [Monteiro, 2014]:

$$BC_i = \sum_{i \neq j \neq \mathrm{k}} \frac{l_{jk}(i)}{l_{jk}}. \qquad \textbf{(3.22)}$$

BC values can be normalized within the range (0,1) simply by the multiplication by the factor $\frac{2}{(N-1)(N-2)}$, *N* being the number of nodes. BC measure frequently assumes that network flow preferentially happens through the graph's shortest paths.

EC for node *i* is related to the i$^{\text{th}}$ component of the eigenvector associated with the largest eigenvalue of the adjacency matrix [Monteiro, 2004]. This measures account for the quality of the connections, in the sense that nodes with high EC values tend to connect with other nodes that, also, have high ECs [Newman, 2008]. The following paragraphs aim to provide a deeper insight into the idea EC is based on.

The EC is also called the Gould's index of accessibility, since Gould was the first person to address this problem, for geographical applications [Gould, 1967]. The idea behind this measure is that the accessibility of a node should depend somehow on the other nodes . This can be seen as a consequence of the measure representing how accessible a node can be from other nodes. As a mathematical formulation, let x_i determine the importance of node i. Dependence on the importance of other neighbor nodes j implies that, at the simplest scenario, x_i is proportional to a linear combination of the importance of the nodes it is connected to. That is:

$$x_i = \mathrm{c} \sum_j a_{ji} x_j. \tag{3.23}$$

In (3.23), for simplicity's sake, let us consider a_{ij} as each element of a binary and symmetrical adjacency matrix, $\boldsymbol{A}$, of a graph. c is a proportionality constant, independent of i. Rewriting (3.23) in matrix notation gives:

$$\boldsymbol{x} = \mathrm{c}\boldsymbol{A}^T\boldsymbol{x}; \tag{3.24}$$

with $\boldsymbol{x}$ being a vector comprising each node's importance measure x_i. Dividing both sides of (3.24) by c:

$$c^{-\mathbf{1}}\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{x}. \tag{3.25}$$

Note that (3.25) is an eigenvector equation for matrix $\boldsymbol{A}^T$ (which is the same as $\boldsymbol{A}$, since it was assumed to be a symmetric matrix), with eigenvalues c^{-1}. This equation also gives an idea of why eigenvalues and eigenvectors may be associated to a notion of importance of a node, although it is not a formal description. The more detailed description given below is based on [Straffin, 1980], a work the reader is referred to, if he/she wishes a more complete formulation.

Instead of working with the matrix $\boldsymbol{A}$, let us define a new matrix $\boldsymbol{B}$, such that

$$\boldsymbol{B} = \boldsymbol{A} + \boldsymbol{I}, \tag{3.26}$$

with $\boldsymbol{I}$ being an $n \times n$ identity matrix. Note that $\boldsymbol{B}$ has the same eigenvectors as $\boldsymbol{A}$, but with eigenvalues that are larger by one unity. Also, note that each element of $\boldsymbol{A}^{k}$, that is, the adjacency matrix up to power k, represents the number of manners to travel between any two nodes. From the above definition of $\boldsymbol{B}$, the elements of $\boldsymbol{B}^{k}$ are related to a very similar measure: the number of manners to travel between any two nodes by paths of length k [Straffin, 1980].

Another feature used for justification of the EC comes from the Perron-Frobenius theorem, which states that, if $\boldsymbol{M}$ is a non-negative $n \times n$ matrix, then there is an eigenvalue λ_1 such that:

- λ_1 is a positive and real value, being a simple root of the characteristic equation;
- for any eigenvalue λ_i, with $i \neq 1$, $\lambda_1 > |\lambda_i|$;
- λ_1 has a unique eigenvector $\mathbf{v}_1$, up to constant multiples, and this eigenvector has all positive components.

Since matrix $\boldsymbol{B}$ is primitive [Strafiin, 1980], Perron-Frobenius' theorem applies. The eigenvectors of $\boldsymbol{B}$, $\boldsymbol{v}_i$, can constitute a basis for which any vector $\boldsymbol{y}$ that is non-orthogonal to them can be written as:

$$\boldsymbol{y} = \sum_i a_i \boldsymbol{v_i}, \tag{3.27}$$

with $\boldsymbol{v}_1 \neq 0$.

In the basis of its eigenvectors, $\boldsymbol{B}^{k}$ can be diagonalized to its eigenvalues. Then:

$$\boldsymbol{B}^k \boldsymbol{y} = \sum_i \lambda_i^k a_i \boldsymbol{v}_i \ . \tag{3.28}$$

Dividing both terms of (3.28) by λ_1:

$$\frac{\boldsymbol{B}^k \boldsymbol{y}}{\lambda_1^k} = \frac{1}{\lambda_1^k} \sum_i \lambda_i^k a_i \boldsymbol{v}_i \ . \tag{3.29}$$

Now, since λ_1 is the largest eigenvalue, then:

$$\lim_{k \to \infty} \frac{\boldsymbol{B}^k \boldsymbol{y}}{\lambda_1^k} = a_1 \boldsymbol{v}_1. \tag{3.30}$$

Note that each i^{th} element of $\boldsymbol{v}_1$ corresponds to the node i, and is seen as a measure of accessibility of that node, providing the notion for the EC.

3.3. Graphs in MI-BCI applications

The most common way to search for the expected response of a motor imagery task is to analyze the signal in the frequency domain. A decrease in the amplitude of the µ rhythm is usually expected on the motor cortex in the contralateral hemisphere regarding the hand MI.

One of the main problems with any type of BCI is the large inter and intra-variability. The system must be capable of adapting to the user, and adaptation to the system by the user must also occur. An issue that arises from this is finding reproducible results. In this context, establishing “good” features as input to the classifier is essential. Theoretically, power spectra of specific frequency bands should be able to do this task; however, MI-BCIs have been demonstrated to be very complex, in the sense that MI is a task most people are not used to. In addition, there is not yet an optimum well established way for extracting features in MI-BCIs (and maybe this cannot be accomplished).

New approaches for feature extraction have been studied for MI-BCIs using graph theory by modeling the situation using a mathematical graph, and extracting metrics and characteristics that could be used in a classifier algorithm. It is important to note that a graph is a mathematical tool that does not, necessarily, describe an existing network (although many networks can, in fact, be modeled very accurately by graphs). As with any attempt to describe natural phenomena, it is only a model, that may work well, but not necessarily accounts for the reality; it is merely a way of perceiving and describing it. Regardless, many studies in neuroscience currently make use of graph theory to model brain interactions, and data analysis usually undergoes a few common steps shown in Figure 3.7.

Nodes determination of the graph to be constructed varies according to the technique used to extract the brain signals. FMRI (functional magnetic resonance imaging) and PET (positron emission tomography), for instance, are techniques denominated as "voxel-based" modalities, while EEG, MEG (magnetoencephalography) and fNIRS (functional near-infrared spectroscopy) are "sensor-based" modalities. For these last techniques, there are mainly two ways to determine graphs nodes, as already stated in Section 3.1: either the nodes are directly assigned to sensors, or to their reconstructed sources [Fallani et al., 2014].

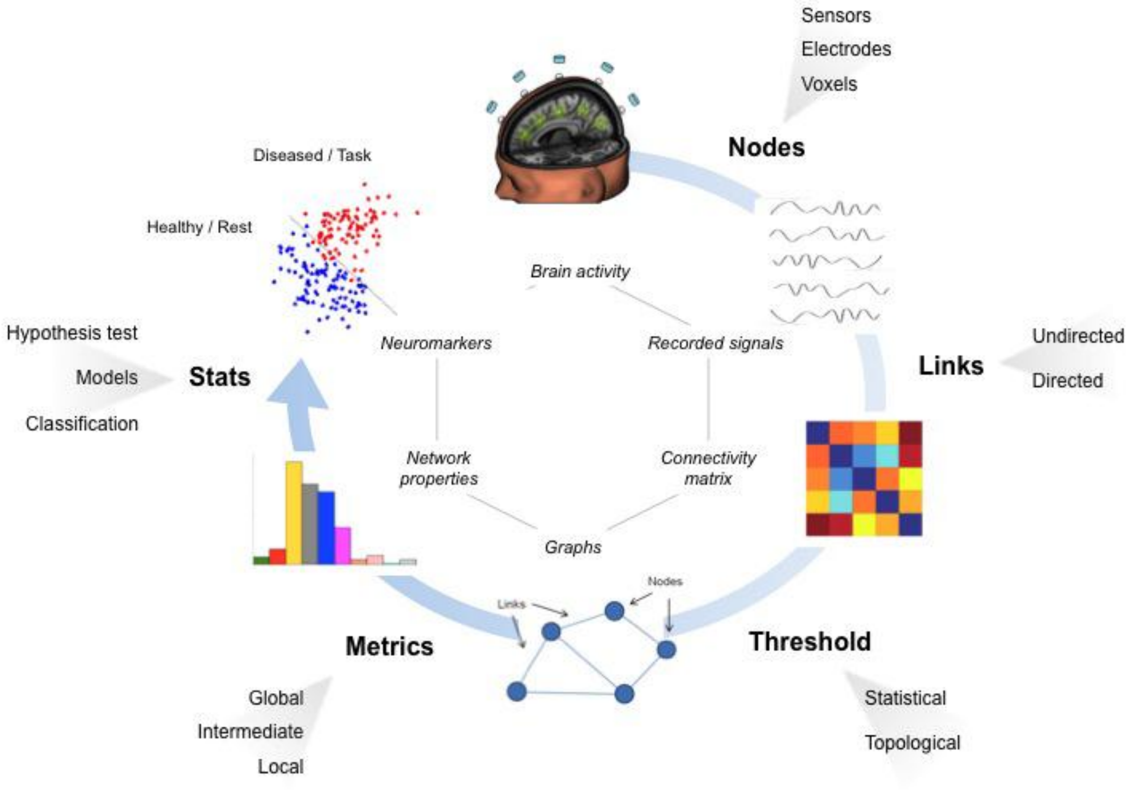


Figure 3.7. Usual steps to build graphs from recorded brain signals. Nodes indicate specific brain sites, which may vary according to the technique used. Links between them are estimated through some correlation method. In this example, there is a definition of a threshold value to determine which links should be maintained in the graph. As final steps, graph metrics are extracted to characterize and classify different tasks or populations. Extracted from [Fallani et al., 2014].

To establish the links, regardless of being a directed or undirected graph, weighted or unweighted graph, some statistical method to analyze the similarity between two time series should be used. In general, these methods take the N nodes, pairwise, attributing a similarity coefficient to them according to comparisons made between their time series, aiming to establish how similar they are to each other. By the end of this step, an N×N matrix containing all the similarity coefficients is obtained, sometimes called the correlation matrix, which originates the graph adjacency matrix. If it is sought to work with an unweighted graph, a threshold should be defined, in order to keep only the most significant links of the correlation matrix for the adjacency matrix. For a weighted graph, the common procedure is to normalize each element of the correlation matrix by dividing its value by the maximum value of the matrix, resulting in the weighted adjacency matrix (or simply weights matrix).

In EEG-BCI and connectivity studies, a variety of correlation methods and applications are reported in the literature, such as Pearson's correlation, to evaluate affective states [Gupta and Falk, 2015]; phase locking value, to develop new tools for processing and analyzing EEG data [Shamas et al., 2015]; signal coherence, for multi-class BCI [Salazar-Varas and Gutiérrez, 2015]; directed transfer function [Ghosh et al., 2015]; and graph lifting transform, for developing an online BCI game [Asensio-Cubero et al., 2016]. Regardless of the method, the basic idea is the same: estimate the relationship between two nodes to determine if there is a link between them (in the case of unweighted graphs) or how significant is the link between them (in the case of weighted graphs).

Network studies have been an increasing field in neuroscience, accompanied by a fast development of new methods to be applied. Graph metrics have been widely used to characterize normal or abnormal brain states, or to try to determine a specific brain task. Graph applications have provided a unique opportunity for better understanding the brain [Fallani et al., 2014]. However, it should be noted that lack of physiological knowledge may result in purely mathematical measures, with questionable clinical value. In fact, the use of models growing in complexity can assist in developing ideas that are closer to the real functional mechanisms of the brain. For instance, it is well known that not all neurons are the same, nor are the manners through which they interact to one another; thus, models that take this into account (considering, for example, non-uniform nodes and edges) should be able to provide more accurate insights on the brain given, of course, the correct physiological interpretation.

In general, graphs in neuroscience seem like a very reasonable approach, being new analytical tools capable of providing a more complex understanding of the brain. However, according to [Fallani et al., 2014], the reason why network neuroscience has gained increasing popularity is, also, the reason why "the risk of a rush towards its frenetic and counterproductive application becomes more and more concrete", and therefore, they should be treated carefully.

Chapter 4

Materials and Methods

4.1. Data acquisition

This study is a subproject of a larger project, “DesTINe – Desenvolvimento de Tecnologias de Informação para Neurologia”, which was approved by the Ethics Committee of UNICAMP in 2010 (registration number at CEP-UNICAMP 791/2010, C.A.A.E. number 0617014600010). All subjects signed a written informed consent previous to data acquisition.

EEG data from 8 healthy subjects (mean age 24 ± 4 years, 7 men) were acquired at a 5000 Hz sampling rate using the BrainAmp amplifier (BrainProducts, Germany) with 64 Ag/Cl pre-defined electrodes of the BrainCap MR (10/10 system). In this system, the electrical contact between scalp and electrode is made through a contacting AgCl gel. The ensemble was connected to a computer equipped with the BrainVision Recorder software. The cap, amplifiers ensemble and the computer can be seen in Figure 4.1.

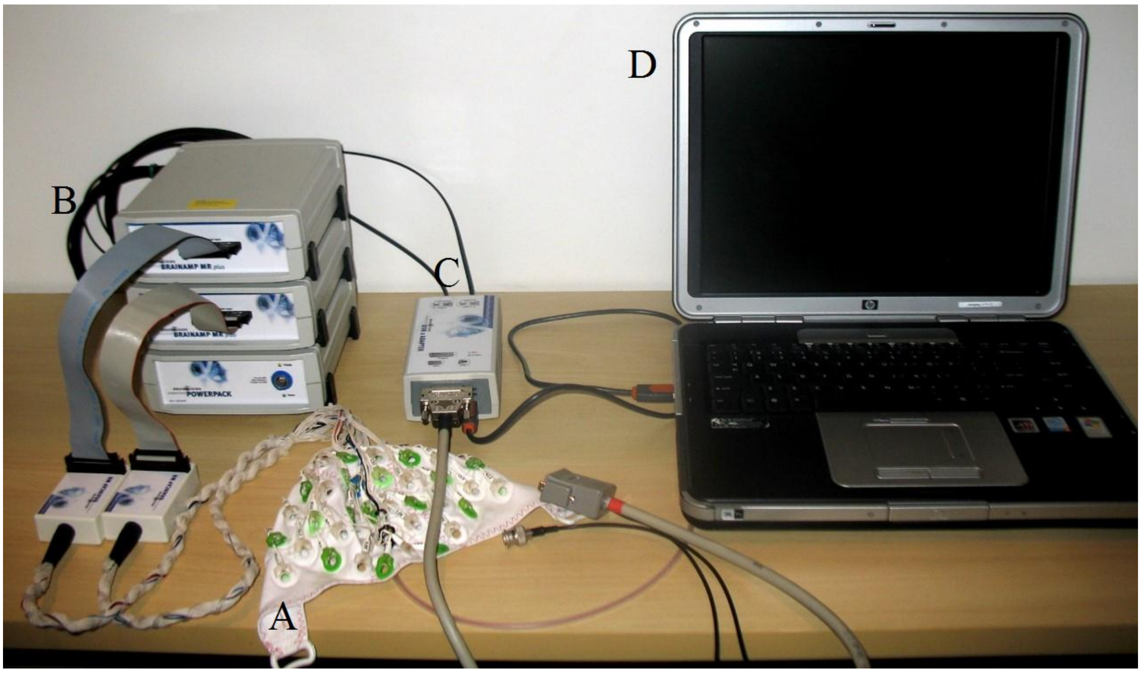


Figure 4.1. Experimental apparatus. The signals are extracted using the EEG cap (A) and passed to the amplifiers (B). Each amplifier supports 32 channels, accounting for the total of 64 channels. Signals undergo through the device in (C) to reach the computer (D).

The experiment was designed in alternating rest and task blocks, each block with duration of 10 s (Figure 4.2). The task involved either hand motor imagery (two acquisitions) or hand movement (two acquisitions), alternating left and right hands with the rest blocks. Total duration of each acquisition was of 170 s.

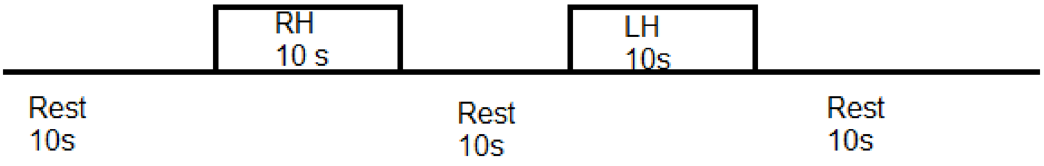


Figure 4.2. Experimental paradigm. RH and LH stand for “right hand” and “left hand”, respectively.

Subjects were placed in front of a computer screen, sitting comfortably. A chronometer was shown on the screen, so they could keep track of each block’s duration. They were instructed when to start and stop tasks, and movement acquisitions made sure they properly understood the experimental protocol. To keep track of possible sources of artifacts, every time the subject made a noticeable movement, notes were taken. Also, movement acquisitions supposedly ensure that subjects properly understood instructions given and would, hopefully, execute the protocol correctly. When necessary, acquisitions were redone.

4.2. Data preprocessing

Data preprocessing included frequency filtering in specific bands of interest and downsampling to 256 Hz in *EEGLAB* [http://sccn.ucsd.edu/eeglab], a *MATLAB* suite. Also, data were filtered in two bands of interest: μ (7 – 13 Hz) and β (13 – 30 Hz). CAR (Common Average Removal) spatial filtering was also used. Downsampling was performed because the MATLAB platform was not able to deal with the amount of data generated by a 5 kHz sampling rate. Nevertheless, since the maximum frequency that was worked with was of 30 Hz, this should not pose a problem.

It was hypothesized that, submitted to the same stimuli during each 10 s block, graphs representing the cerebral cortex during MI tasks should not suffer great variations. Nonetheless, to compensate for sudden and unexpected variations, every three one-second blocks of EEG data were averaged to produce a smoothed 1 s data sample, from which the graph features were extracted (Figure 4.3). Therefore, from the original 4 blocks of 10 s each, the analysis was made with four new blocks of 8 smoothed “one-second” blocks. Considering that the system's response tie was sought to be of 1 s, 32 adjacency matrices were built, one for each of the 32 s available for analysis (4 blocks x 8 smoothed seconds).

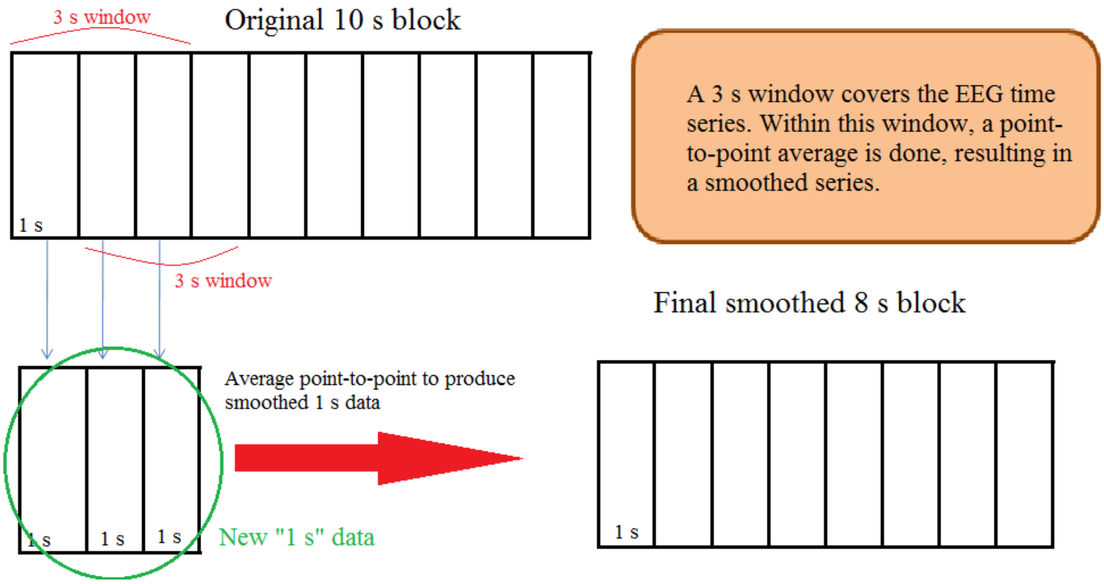


Figure 4.3. Data smoothing illustration.

4.3. Brain connectivity

In studying the brain, there are three main commonly used terms to define different kinds of connectivity: anatomical connectivity, which indicates a physical existing connection between brain areas; functional connectivity, in which usually time series correlations are used to determine how two brain sites interact; and effective connectivity, which indicates how a signal from one region of the brain influences other areas. In this work, the concept of functional connectivity was used.

To estimate connectivity, different methods to determine the similarity between EEG time series were used, namely, the motifs method and Pearson’s correlation. These will be discussed below. These similarity measures were combined with graph theory in order to build a representative brain network of the tasks under analysis. The use of networks (graphs), built based on signals extracted by techniques such as EEG or fMRI, to represent the brain, has been very common in recent years, showing the relevance of the area for studying the brain [Friston, 2011; Marzetti et al. 2008; Sakkalis, 2011; Belmonte et al. 2004; Supekar et al. 2008; Rosário, 2013].

4.4. The motifs method

Basically, the motifs method is a qualitative approach that translates the EEG time series into new ones (the motifs series), from which similarity coefficients can be calculated. This is accomplished by identifying how the time series varies and, then, by labeling every possible behavior that could happen. Analyses based on different numbers of points will yield distinct pattern possibilities (which constitute the *motifs*). [Olofsen et al., 2008; Rosário et al., 2015]. This is illustrated in Figure 4.4.

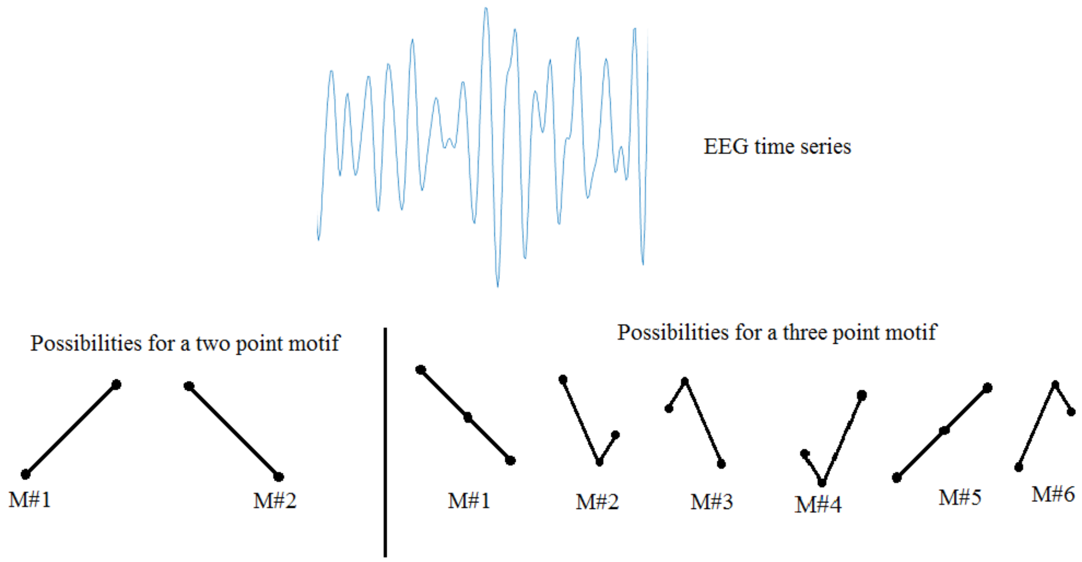


Figure 4.4. Possibilities for EEG patterns with two and three point motifs.

Figure 4.4 shows examples of how points of an EEG time series could be translated into motifs. A time window containing the number of points of the desired motifs must be determined. This time window is slid over the series, labeling each window with one of the possible patterns, therefore translating the EEG time series into the new motif series. Note that, in doing so, information on how great is the variation between two time points is irrelevant, and just the qualitative general form of the signal is taken into account. This procedure ends up having a smoothing effect on the signal, and therefore the length of the motifs series is always smaller than that of the original time series. For an ‘n’ point motif, the length L_M of the motif series, in terms of the original series length, L_0, is

$$L_M = L_0 - (n - 1). \tag{4.1}$$

The number n introduced above is also called the motif's degree. It determines the length of the window to cover the time series. Figure 4.5 illustrates this process to a time series, using n = 3.

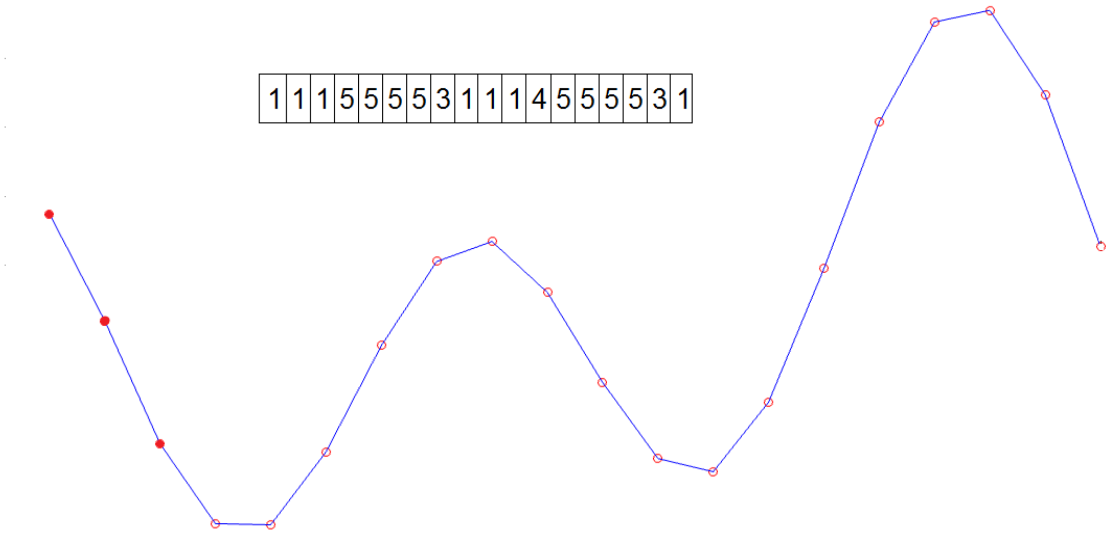


Figure 4.5. Example of motifs translation for an EEG time series with 20 points. n = 3. Every 3 data points are identified as one of the possible labels of Figure 4.4.

Using the motifs approach, the similarity between two electrodes time series will be given by the similarity between their motifs series, which is computed as follows. Let *J* be the variable containing the number of times the same motif has been encountered when comparing two series. For distinct values of lag **λ**, then, *J* of a specified point *i* of the series can be defined as follows:

$$J_i^{\lambda} = \begin{cases} 1, if\ M\sharp x_i = M\sharp y_{i+\lambda} \\ 0, otherwise \end{cases}. \tag{4.2}$$

.

In doing so, it is possible to find the total number of motif coincidences given a certain lag value by simply summing J_i^{λ} over *i*.

A coefficient c_{xy} can then be defined as

$$c_{xy} = max\left(\sum_{i=1}^{L_M} J_i^{\lambda_0}, \sum_{i=1}^{L_M} J_i^{\lambda_1}, \dots, \sum_{i=1}^{L_M} J_i^{\lambda_N}\right). \tag{4.3}$$

.

c_{xy} stands for the greatest number of times the same motif was found in both time series, considering every different lag value accounted for. When λ=0, it makes no difference which series is used as reference for the comparison. For other lag values, however, this order is important. Therefore, the definition of another coefficient, c_{yx}, is justified. For c_{xy}, the motif is firstly identified in series X and afterwards in series Y; and vice-versa for c_{yx}.

With the c_{xy} and c_{yx} coefficients, another quantity, the synchronization degree (Q), can be defined as:

$$Q_{xy} = \frac{\max\{c_{xy}, c_{yx}\}}{L_M}. \tag{4.4}$$

Note that, if λ is zero, then c_{xy} and c_{yx} are equal, since comparisons will be done equally for the two cases. This, however, needs not to be true for larger lag values.

Another relevant index indicates the synchronization direction (Equation (4.5), q). It can be seen as an indicative of which coefficient between c_{xy} or c_{yx} is larger. In the context of directed graphs, it can be used to indicate the direction of the edge between two nodes.

$$q_{xy} = sign(c_{xy} - c_{yx}). \tag{4.5}$$

Sign(x) designates the sign function, which can be defined as:

$$sign(x) = \begin{cases} -1, if\ x < 0 \\ 0, if\ x = 0 \\ 1, if\ x > 0 \end{cases}. \tag{4.6}$$

Table 4.1 summarizes the motifs properties and quantities described above.

Table 4.1. Summary of motifs properties and quantities.

Property	Description
Motif degree (n)	Number of points to be used in the motif.
Lag (λ)	Indicates the delay from which motifs series are compared.
c coefficient	Contains information about how similar two motif series are.
Q	Normalized from 0 to 1, it is analogue to c.
q	Indicates if there is a preferential direction for the synchronization between two time series.

We chose to work with the motifs method because, besides its low computational cost, it provides a qualitative approach to analyze a very noisy signal. Thus, we believed it could provide a more truthful similarity value between the electrodes time series. In our analysis (Chapter 5), we also noticed that it could provide adjacency matrix calculations faster than more traditional methods in the field, such as Pearson's Correlation, described in the next section. For an online BCI application, this time difference can be extremely significant: while graphs calculation using Pearson's correlation often took about two minutes, it could be done in a matter of few seconds with the motifs method.

4.5. Pearson's Correlation

Pearson's correlation between variables 'x' and 'y' is defined as [Ross, 2010]:

$$\rho_{xy} = \frac{cov(x,y)}{\sqrt{var(x)var(y)}}.; \qquad \textbf{(4.7)}$$

in which 'cov(x,y)' is the covariance of 'x' and 'y', and 'var(x)' denotes the variance of the series represented by 'x'.

Pearson's correlation was mostly used as a form of control, by comparing results obtained with the less popular method of the motifs. Note, however, that this type of correlation, when applied directly to the EEG electrodes, can induce auto-correlations, since the activity that influences the recorded signal on a specific electrode can also affect the recorded signal at electrodes positioned on other sites of the scalp.

4.6. Graphs construction and metrics calculations

To build graphs from adjacency matrices, it is first necessary to define whether these will be weighted or non-weighted. If they are non-weighted, it is necessary to stipulate a threshold value. Electrodes will be considered to be correlated if their similarity value (using either the motifs or the Pearson's correlation method) is above this threshold, in which case a value of '1' is attributed to that link; otherwise, the pair of electrodes will be considered independent, and a value of '0' is attributed to the link, meaning no 'connection' between them. This approach, however, proved to be very dependent on the threshold value, posing the question of how to determine the best value for it. Therefore, the results that will be shown and discussed regard weighted adjacency matrices.

To analyze the data, two independent graphs were constructed, one for each cerebral hemisphere. Therefore, data from the 64 channels were split to form two separate networks. Central electrodes, except for Cz and Fz, were left out of the analysis. Below, Figure 4.6 shows which electrodes were used for each graph's construction. Blue electrodes denote nodes from the left hemisphere, and red electrodes are the ones constituting the right hemisphere. Purple electrodes belong to both graphs, although they do not connect them. Therefore, the left (right) hemisphere is composed of electrodes FP1 (FP2), AF3 (AF4), AF7 (AF8), Fz (Fz), F1 (F2), F3 (F4), F5 (F6), F7 (F8), FC1 (FC2), FC3 (FC4), FC5 (FC6), FT7 (FT8), FT9 (FT10), Cz (Cz), C1 (C2), C3 (C4), C5 (C6), T7 (T8), CP1 (CP2), CP3 (CP4), CP5 (CP6), TP7 (TP8), TP9 (TP10), P1 (P2), P3 (P4), P5 (P6), P7 (P8), PO3 (PO4), PO7 (PO8) and O1 (O2). Initial analyses with no hemisphere separation were also performed; however, they are not displayed in this work. The hemispheres' split was the approach of choice mainly because, as a first step, this work's interest was aimed at localized properties. This split, however, can imply in loss of considerable inter-hemisphere information.

Using the motifs methods, for constructing the graphs, each element of the adjacency matrix (w_{ij}) was considered to be the synchronization degree defined in Equation (5.4). Also, only a lag $\lambda=0$ was used in the analysis. When Pearson's correlation was used, each element of the adjacency matrix was taken to be the correlation index (if non-negative) or zero (if negative).

The weighted degree or strength (also the degree centrality, for this case) (s_i) of the node 'i' was calculated by simply summing the value of all the links (a_{ij}) connected to that node:

$$s_i = \sum_j w_{ij}. \tag{4.8}$$

The calculation for the clustering coefficient of a weighted network is not so well-established, since it does not follow from a direct generalization from unweighted graphs properties [Antoniou et al., 2009]. In this work, it was chosen to use equation (7) as defined in [Antoniou et al., 2009] (and reproduced by Equation (4.9), below).

$$C_i = \frac{\sum_j \sum_k w_{ij} w_{jk} w_{ki}}{\left(\sum_j w_{ij}\right)^2 - \sum_j w_{ij}^2}. \tag{4.9}$$

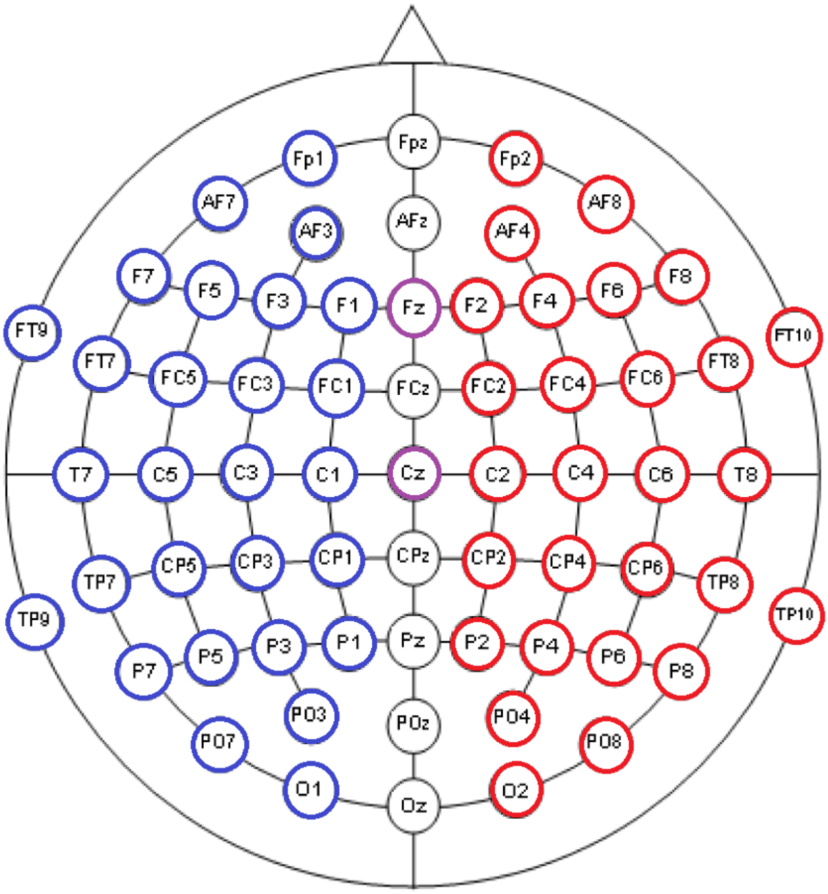


Figure 4.6. Electrodes used for graphs' construction. Blue electrodes belong to the left hemisphere graph; red electrodes belong to the right hemisphere graph; purple electrodes belong to both graphs.

The characteristic path length was calculated from averaging shortest distances between all graph nodes. Shortest paths for a weighted graph follows from a direct generalization of the binary case [Antoniou et al., 2009], and were calculated according to Dijkstra's algorithm. In using this library, distance between two nodes are larger if the weight of the connection between them is lower. The characteristic path length (l) is then given by (4.10):

$$l = \frac{1}{N(N-1)} \sum_{i \neq j} d_{ij}. \tag{4.10}$$

N stands for the number of nodes in the graph, and d_{ij} is the shortest path between nodes i and j. d_{ij} is associated with elements of the inverse of the adjacency matrix.

Betweenness centrality (BC) was calculated as stated by (4.11):

$$C_b^i = \sum_{j \neq i \neq k} \frac{l_{jk}(i)}{l_{jk}}. \tag{4.11}$$

Eigenvector centrality (EC) was calculated according to its description in section 3.3.4.

4.7. Data classification

Data classification followed two approaches: a linear least-squares based discriminant analysis (LSLDA) and support vector machines (SVM) (both explained in Chapter 2). Also, two kernel functions for the SVM were tested, namely, linear and polynomial (third degree) ones. Both SVM algorithms were used as predefined MATLAB functions. Thus, adjustable kernel parameters were set according to the platform's default settings.

As inputs to the classifier, effects of using each cerebral hemisphere separately and their combinations were tested. Combinations were done using the difference of a metric's value on a specific node located on the right hemisphere minus its value on the correspondent contralateral node; for instance: the clustering coefficient on C4 minus its value on C3, and so on. In the following sections, this input approach is denoted as "difference". The idea of this "difference" input is try to couple metrics values from both hemispheres, studying how it can affect classification results by comparing to the case when only one of the hemispheres was used for data classification. Also, other combinations, such as the "sum" or "ratio" of values were tested, not yielding better results than the ones for the "difference" scenario.

In addition, we investigated the dependence on the classification results with the electrodes used for classification. As a first approach, all graph nodes were used for classification. On a second approach, all combinations using the minimal number of electrodes possible were tested, and offline analysis displays the best one to be used for classification purposes.

All classification tests were conducted in the leave-one-out fashion. That is, from "T" total available samples, "T-1" were used to train the classifier, whilst the remaining one was used to test its outcome response.

Chapter 5

Results and Discussion

While searching for differences in left and right hands MI networks properties, two methods were used to construct the graph’s adjacency matrices: motifs and Pearson’s correlation. The latter was mostly used as a form of control, presenting very similar results to the motifs’ method. In this chapter, we only present results from the motifs’ method.

In the first subsection of this chapter, five metrics for characterizing the graphs under each MI task were analyzed: strength (4.8), clustering coefficient (4.9), characteristic path length (4.10), betweenness centrality (4.11) and eigenvector centrality. These results are shown in scalp topographic maps, as a colormap, with the mu band being shown on the top row and, the beta band, on the bottom one. Note that, due to the interpolation used to create such figures, non-zero values may appear in sites in-between electrodes, even though, strictly, they do not constitute graph nodes. In some cases, results for different subjects were qualitatively very similar, and it was chosen not to show them all. Therefore, a group average behavior is presented. Individual results for the strength can be found in Appendix B.

The second subsection contains classification tests using the methods described in Chapters 2 and 4. The effects of using all nodes and specific ones within the graph for predicting MI classes were studied and compared.

Also, it is important to emphasize that the terms “electrodes” and “nodes” are often used here as synonyms (although, strictly, they are not), simply because each graph node was constituted by an electrode. In addition, it is necessary to distinguish between “central electrodes/nodes” and “central line electrodes/nodes”. The first refers to electrodes labeled by the “C” letter, and that lie on either of the scalp hemispheres. The latter regards the electrodes positioned on the scalp’s central line: Cz and Fz.

5.1. Graphs' Topology

5.1.1. Strength (S; also the degree centrality)

The strength (s_i) for each node *i* was calculated as described by (4.1). An average over all experimental blocks and subjects was made, and results are shown in Figure 5.1.

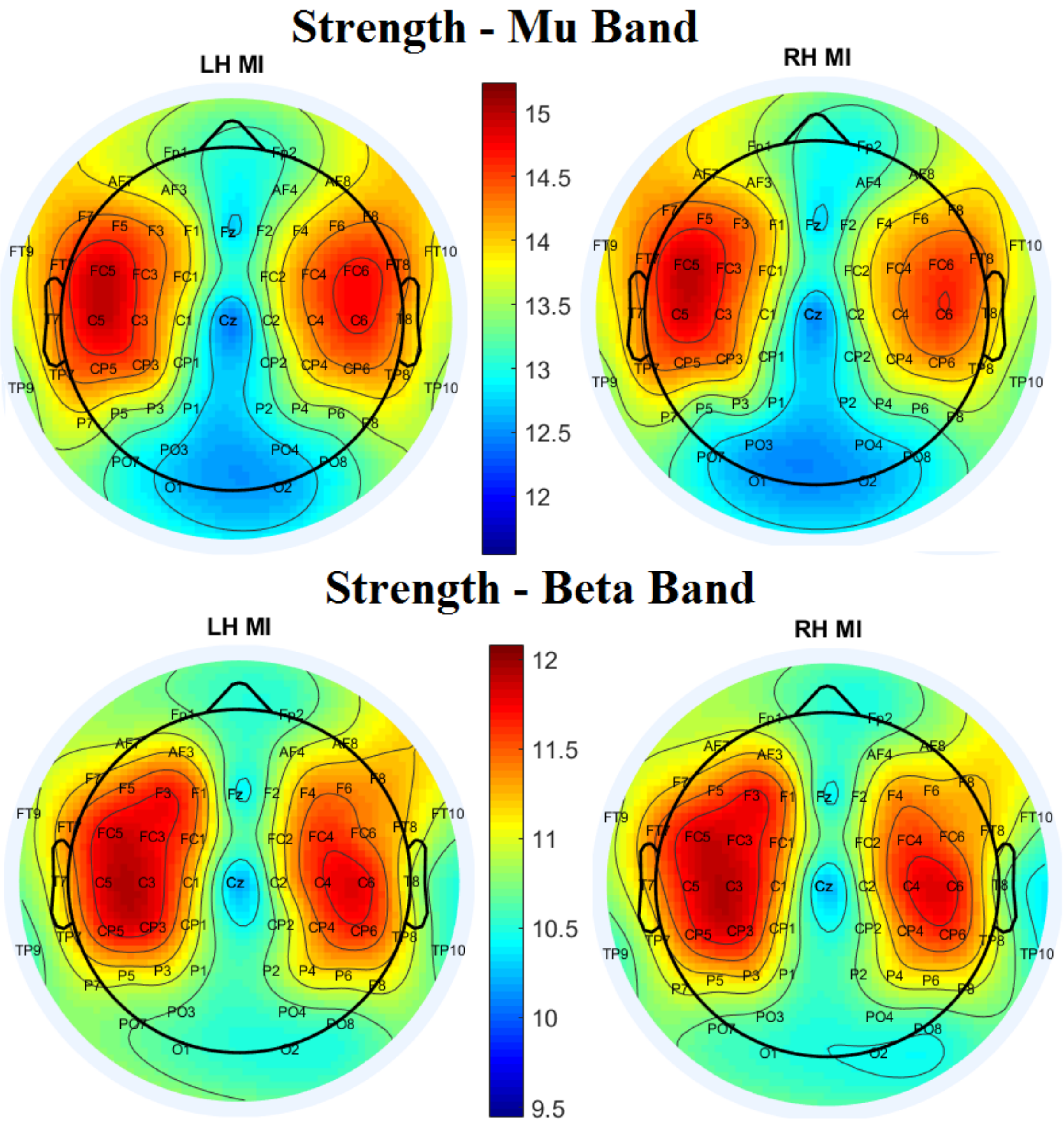


Figure 5.1. Strength values distribution over the scalp. Left: left hand (LH) MI; right: right hand (RH) MI. Top row: mu band; bottom row: beta band.

It was hypothesized that the well-established response for MI in the frequency domain should also manifest itself in the built graphs. Nodes with the largest strength are usually C3/C4, FC5/FC6 and C5/C6. Electrodes FC5/FC6 are positioned on the primary motor cortex (PMC), an area related to movement but not to MI, although some subjects may experience activations in these areas depending on their thinking strategy [Velásquez-Martínez et al., 2013]. C3/C4 and C5/C6 are commonly appearing electrodes in MI studies; they are positioned on the motor cortex. In fact, a study has found, by means of the Fourier coefficients, that opposing differences related to MI's ERDs were maximum around the C3/C4 regions [Haufe et al., 2010] (within the mu band). The expected response for MI is an ERD in the contralateral hemisphere. This means that neurons fire in less synchrony, which should result in less similarity between time series.

The top row of Figure 5.1 shows the results for the strength for the mu band. However, this property, that in our case, gets larger as more synchronous the time series are – is larger for these motor areas (around C3/C4). Nonetheless, note that this metric is calculated considering the graph as a whole, and not just these specific regions. Thus, nodes C3/C4, FC5/FC6 and C5/C6 having the largest strength values simply means that, on the whole graph perspective, their time series are slightly more similar to the series of all other nodes, not necessarily reflecting the strictly local expected behavior. Therefore, the strength may not be the best metric when it comes to reflect the expected ERDs due to MI. Smallest strength values were usually present on occipital and occipitoparietal electrodes, such as O1/O2, P1/P2 and PO3/PO4, and central line electrodes (Fz and Cz). This means that these areas are related in a less synchronous way to the rest of the cortex regions during the MI tasks.

An almost identical pattern was obtained for the beta band (Figure 5.1, bottom row). The main difference is that strength values are smaller for this band. Note that both MI tasks generate very similar behaviors for how strength values are distributed over the scalp.

Again, largest strengths can be found on frontal, central and central-parietal areas, while the lowest values lie among electrodes on the scalp central line (Cz and Fz), frontal-parietal, parietal-occipital and occipital regions.

Looking at both rows in Figure 5.1, the similar behavior between both bands is evident. Basically, points of maxima and minima remain at the same electrodes (or regions around them), although the beta band presents smaller values. For the beta band, nodes located on occipital and parietal-occipital sites showed slightly larger strengths, relative to the rest of the graph.

In order to quantitatively analyze possible differences, values for all subjects were averaged over both acquisitions (Table 5.1).

Table 5.1. Strength mean values ± standard deviation for each graph during each MI task for all subjects, averaged over both acquisitions.

		Left hemisphere graph		**Right hemisphere graph**	
Band	**Subject**	Left hand MI	Right hand MI	Left hand MI	Right hand MI
	1	14.1	14.1	13.9	13.9
	2	14.2	14.2	14.4	14.3
	3	13.1	13.1	12.8	12.9
μ	4	14.1	14.1	13.5	13.5
	5	13.8	13.6	14.0	13.7
	6	13.8	13.7	13.7	13.3
	7	13.4	13.4	13.6	13.6
	8	13.8	13.6	14.1	13.6
	Mean	**13.8±0.4**	**13.8±0.4**	**13.8±0.5**	**13.6±0.4**
	1	11.3	11.2	11.0	10.8
	2	12.1	12.0	11.8	11.7
	3	10.7	10.7	10.1	10.1
β	4	11.4	11.5	11.0	11.0
	5	10.9	10.8	10.9	10.9
	6	11.1	11.2	11.2	11.1
	7	10.5	10.5	11.1	11.2
	8	11.3	11.5	11.2	11.2
	Mean	**11.2±0.5**	**11.2±0.5**	**11.0±0.5**	**11.0±0.5**

For the same frequency band, it is possible to note that, for the same subject, all values are very close to each other, presenting very little variation between RH and LH MI tasks. Differences between both graphs are also minor. In fact, the relatively large standard deviation value makes it practically impossible to distinguish between MI classes by simply analyzing the graph's mean strength. Also, average values on the right hemisphere are usually slightly smaller.

Regarding discrimination between MI tasks, these results suggest that using only the average strength of a graph may not be the best approach. In fact, it is hard to notice a behavioral pattern for all subjects (if such one does exist) from Table 5.1.

5.1.2. Clustering Coefficient (CC)

Figure 5.2 shows clustering coefficient values distribution over the scalp, for the two frequency bands.

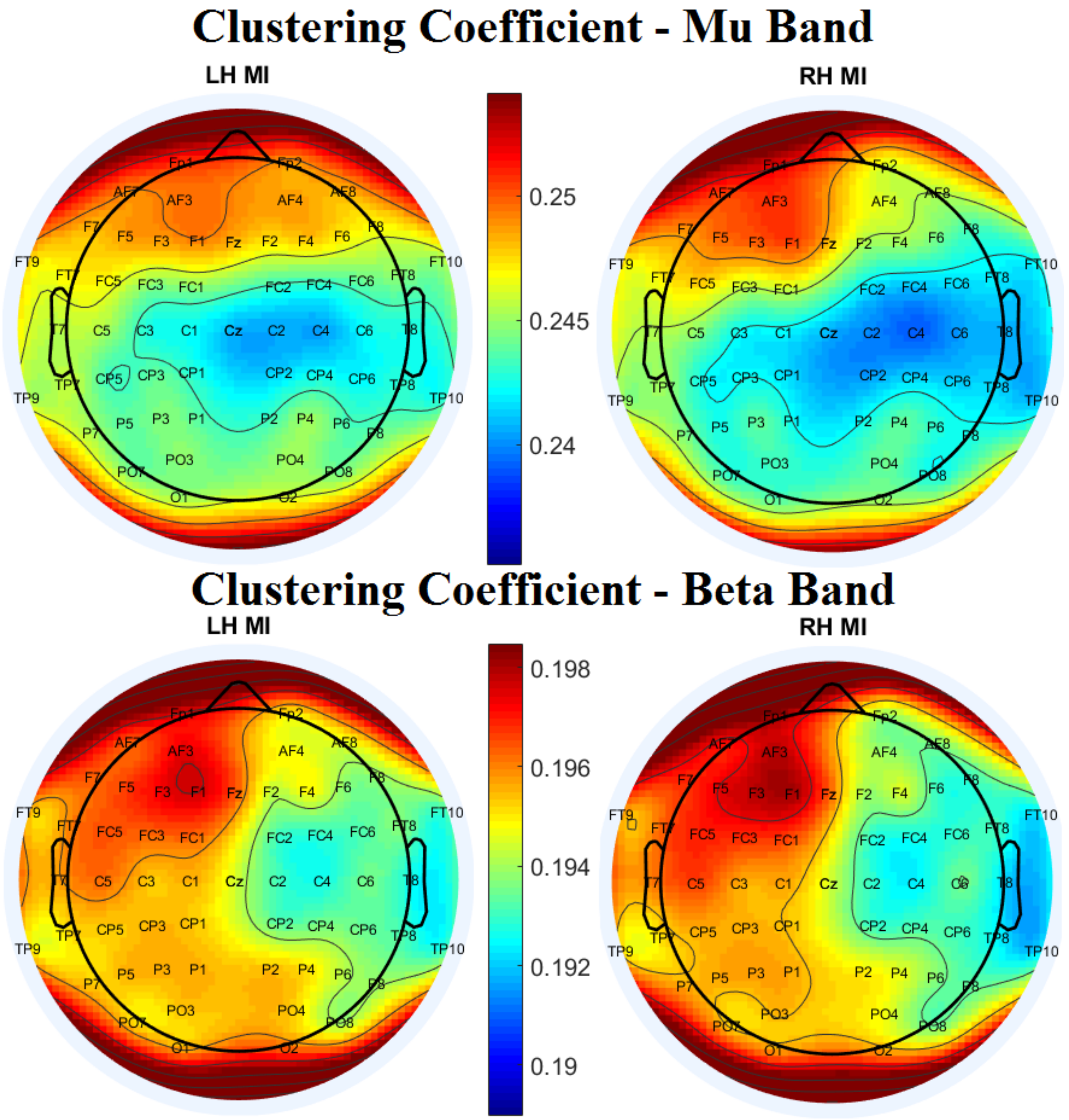


Figure 5.2. Clustering coefficient values distribution over the scalp. Left: left hand (LH) MI; right: right hand (RH) MI. Top row: mu band; bottom row: beta band.

For the mu band, for both hemispheres, MI of the contralateral hand produces slightly larger CC values to that hemisphere (with just one exception for the left hemisphere – FT9). Largest CC values were present in frontal-related nodes, regardless of the MI task.

Regarding the beta band, CC values are usually larger during the contralateral hand MI for both hemispheres. Largest CCs are found in the frontal nodes for the left hemisphere, and in the parietal and occipital areas, for the right hemisphere. As with the strength, values are lower for the beta band.

Electrodes such as C3/C4, C5/C6 and FC5/FC6, that previously presented the largest strengths, now lie among the ones with the lowest clustering coefficients. This indicates that the regions activated in MI, although possessing connections with more weight (in the sense of generating more similar EEG patterns), present slightly less tendency to forming clusters than the rest of the graph. CC emphasizes interactions between a node's nearest-neighbors and, therefore, it is a better indicative of local interaction than the strength. Thus, we believe that the decrease in the CC values of these motor electrodes are more directly related to the expected ERDs caused by MI than the results found for the strength.

Comparing clustering coefficient results for both bands (Figure 5.2), we see that regardless of the frequency band, similarly to the strength, distributions of values along electrodes are very similar to one another: in this case, largest and lowest values are present approximately at the same nodes. This suggests that, even though MI is reflected differently throughout the brain for different frequency bands, which can be seen from the clear decreases in the strength and clustering coefficients when comparing one band to another, there seems to be a more general state that dictates the very similar behavior observed for both mu and beta bands in distinct regions (electrodes), something that may be related to the manner the brain itself is structured.

Graphs mean clustering coefficient values were evaluated during each motor imagery task (Table 5.2). For the same band, mean values are all very close to each other, and the uncertainty range causes them to overlap. Analysis of the mean CC suggests that the contra-laterality observed for most nodes individually is not reproduced in most subjects as a global graph behavior. This and the referred overlapping indicate that the mean CC alone does not seem to be a promising feature in assessing MI tasks.

Table 5.2. Graphs' mean clustering coefficient values during each motor imagery task, for all subjects. Values are averaged over both acquisitions for each subject.

		Left hemisphere graph		**Right hemisphere graph**	
Band	**Subject**	LH MI	RH MI	LH MI	RH MI
	1	0.249	0.249	0.247	0.246
	2	0.254	0.258	0.258	0.259
	3	0.231	0.230	0.225	0.226
μ	4	0.251	0.250	0.240	0.239
	5	0.243	0.240	0.247	0.242
	6	0.244	0.248	0.248	0.240
	7	0.238	0.236	0.239	0.239
	8	0.245	0.256	0.252	0.250
	Mean	**0.244±0.007**	**0.246±0.009**	**0.245±0.009**	**0.243±0.009**
	1	0.198	0.197	0.193	0.189
	2	0.213	0.213	0.210	0.209
	3	0.188	0.187	0.178	0.176
β	4	0.200	0.201	0.192	0.192
	5	0.191	0.190	0.190	0.190
	6	0.196	0.197	0.197	0.196
	7	0.184	0.184	0.194	0.195
	8	0.199	0.203	0.197	0.198
	Mean	**0.196±0.009**	**0.197±0.009**	**0.194±0.009**	**0.193±0.009**

5.1.3. Characteristic Path Length (L)

Figure 5.3 shows each node minimum path length (l_i) for the both bands, calculated according to Equations (3.17) and (3.18). Again, a very similar pattern is found for both cerebral hemispheres during the studied MI tasks.

For the mu band, on the left hemisphere, all frontal related nodes, except for FC1, FC5 and FT7 present larger l_i's values during LH MI periods. Although the difference is very little, this can be seen by the lower extent of blue on this hemisphere during this particular task. This is also valid for central electrodes Cz, C1 and C3. Parietal-central, temporal-parietal and parietal regions exhibit a mixed behavior. All occipital related areas showed larger values during RH

MI. On the right hemisphere, patterns between both MI tasks are almost identical; a few differences can be observed on nodes such as AF4, AF8 and parietal nodes, regions where characteristic path lengths were slightly larger during RH MI.

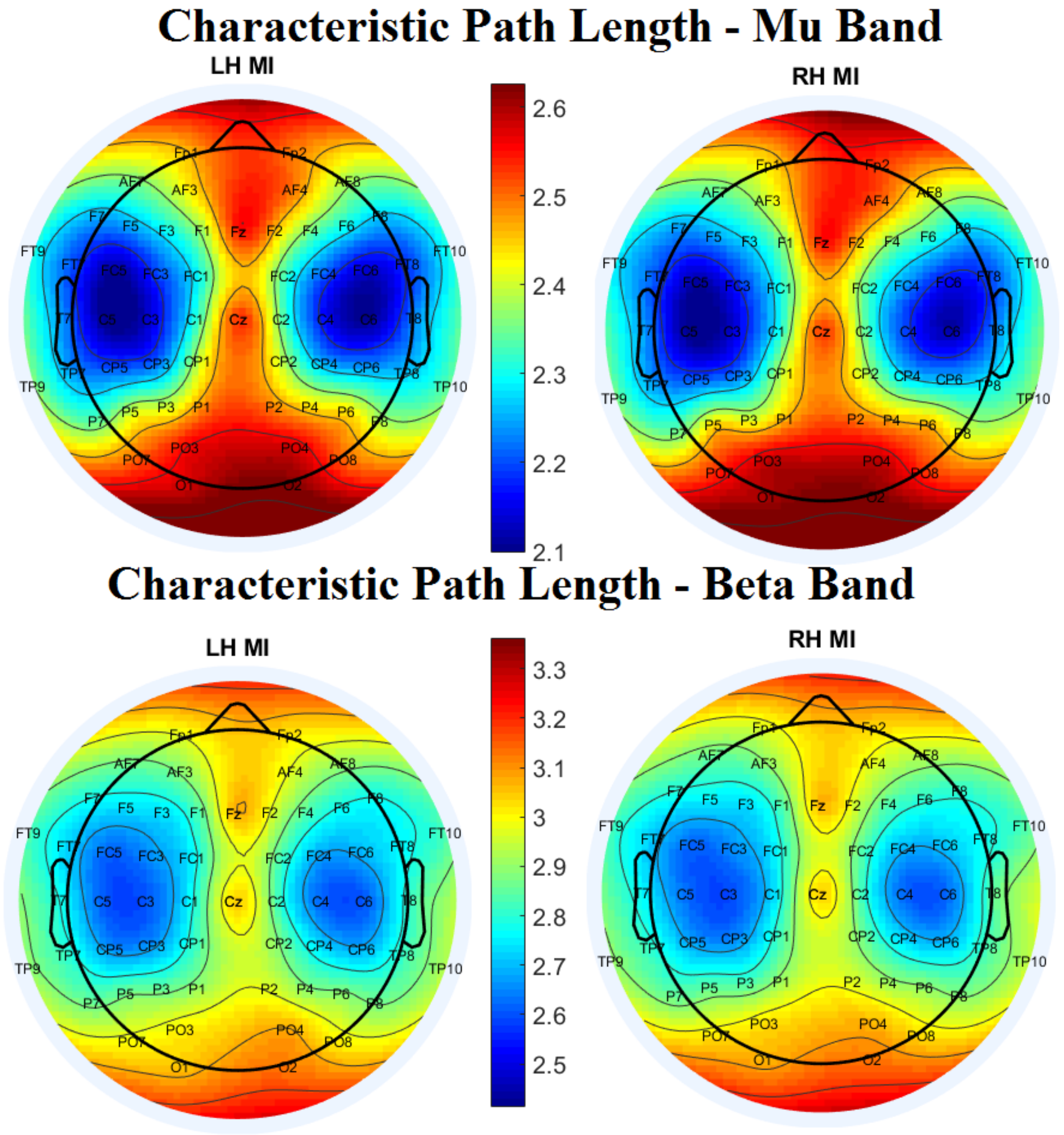


Figure 5.3. Characteristic path length values distribution over the scalp. Left: left hand (LH) MI; right: right hand (RH) MI. Top row: mu band; bottom row: beta band.

Still regarding the mu band, on both hemispheres, central line electrodes lie amongst the ones with the largest path length values, along with nodes related to the occipital region. Lowest values are presented by electrodes C5 and C6.

For the beta band, for both hemispheres and tasks, highest path lengths are displayed by electrodes on the occipital related regions. This could be associated to the fact that these electrodes previously presented low strength and clustering coefficient values when compared to the rest of the graph. In terms of our analysis, this suggests that their connections had less weight, meaning that more cost is required to travel between them. Besides, they were in less connected regions, regarding clustering formation. Therefore, less possibilities of minimum path lengths existed for reaching the rest of the graph from these nodes, which could explain the highest observed node characteristic length values for these areas. In an analogous manner, the lowest node path length present in motor regions, such as FC3/FC4, FC5/FC6, C1/C2, C3/C4 and C5/C6, can be related to how these areas showed higher strengths and clustering coefficients. From Figure 5.3, little to no difference can be spotted when comparing the two different MI tasks. Regions of minima and maxima are practically the same.

Again, regarding comparison between the two frequency bands, the general behavior across them is very similar, in the sense that nodes presenting maximum and minimum values remain the same. Also, there is a clear distinction between both bands' values. This time, however, differently than what occurred for the strength and the clustering coefficient, values are larger in the beta band. Smaller strengths and clustering coefficients may indicate a less connected network, and with less tendency to form clusters, resulting in a decrease in the number of possibilities of minimum cost paths to travel from one node to another, which could reflect on the largest characteristic path length observed for these graphs. This may be related to the fact that the ERDs caused by MI are more pronounced in the mu band than in the beta band.

Table 5.3 presents characteristic path length values for both graphs during hand MI tasks. It can be seen that, for both bands, characteristic path length values were slightly larger, on average, over the right hemisphere. Individually, this was also true for most subjects, with exception of Subjects 2 and 7 (mu band) and Subjects 5 to 8 (beta band).

Table 5.4 summarizes average results obtained. All these results suggest that graphs global properties may not be the most significant manner to assess MI tasks. Local properties, that is, how the values were distributed across each graph node, however, seem to present a more feasible possibility, exhibiting a general behavior reproducible over all subjects. In this context, two centrality measures were explored (betweenness and eigenvector centrality), which will be discussed below.

Table 5.3. Graph's mean characteristic path length values during each motor imagery task, for all subjects. Values are averaged over both acquisitions.

		Left hemisphere graph		**Right hemisphere graph**	
Band	**Subject**	Left hand MI	Right hand MI	Left hand MI	Right hand MI
	1	2.27	2.26	2.29	2.29
	2	2.28	2.30	2.26	2.30
	3	2.42	2.42	2.46	2.44
	4	2.28	2.28	2.38	2.38
μ	5	2.29	2.34	2.35	2.32
	6	2.39	2.40	2.41	2.47
	7	2.37	2.36	2.35	2.34
	8	2.40	2.34	2.43	2.44
	Mean	**2.34±0.06**	**2.34±0.06**	**2.37±0.07**	**2.37±0.07**
	1	2.78	2.79	2.85	2.80
	2	2.62	2.65	2.74	2.77
	3	2.90	2.92	3.09	3.12
	4	2.72	2.73	2.85	2.83
β	5	2.82	2.86	2.85	2.83
	6	2.89	2.86	2.88	2.90
	7	2.98	2.98	2.80	2.78
	8	2.87	2.78	2.86	2.86
	Mean	**2.82±0.11**	**2.82±0.11**	**2.87±0.10**	**2.86±0.11**

Table 5.4. Summary of average basic metrics values. Results are shown for both graphs and during each MI task. < > indicate average value. S = strength; CC = clustering coefficient; L = characteristic path length.

Band	**Property**	**Left hemisphere**		**Right hemisphere**	
		LH	RH	LH	RH
	<S>	13.8 ± 0.4	13.8 ± 0.4	13.8 ± 0.5	13.6 ± 0.4
μ	<CC>	0.244 ± 0.007	0.246 ± 0.009	0.245 ± 0.009	0.243 ± 0.009
	<L>	2.34 ± 0.06	2.34 ± 0.06	2.37 ± 0.07	2.37 ± 0.07
	<S>	11.2 ± 0.5	11.2 ± 0.5	11.0 ± 0.5	11.0 ± 0.5
β	<CC>	0.196 ± 0.09	0.197 ± 0.09	0.194 ±0.09	0.193 ±0.09
	<L>	2.8 ± 0.1	2.8±0.1	2.9±0.1	2.9±0.1

5.1.4. Betweenness Centrality (BC)

BC results are shown in Figure 5.10, for both frequency bands. There are well defined regions regarding larger or smaller BC values. In some of them, BC increases for electrodes located farther from the scalp central line; that is the case for frontal, frontocentral (partially), central (partially), parietal (partially) and centroparietal areas. In other regions, an inverse pattern can be identified, such as in the frontocentral (partially), central (partially), and parietal (mostly) areas.

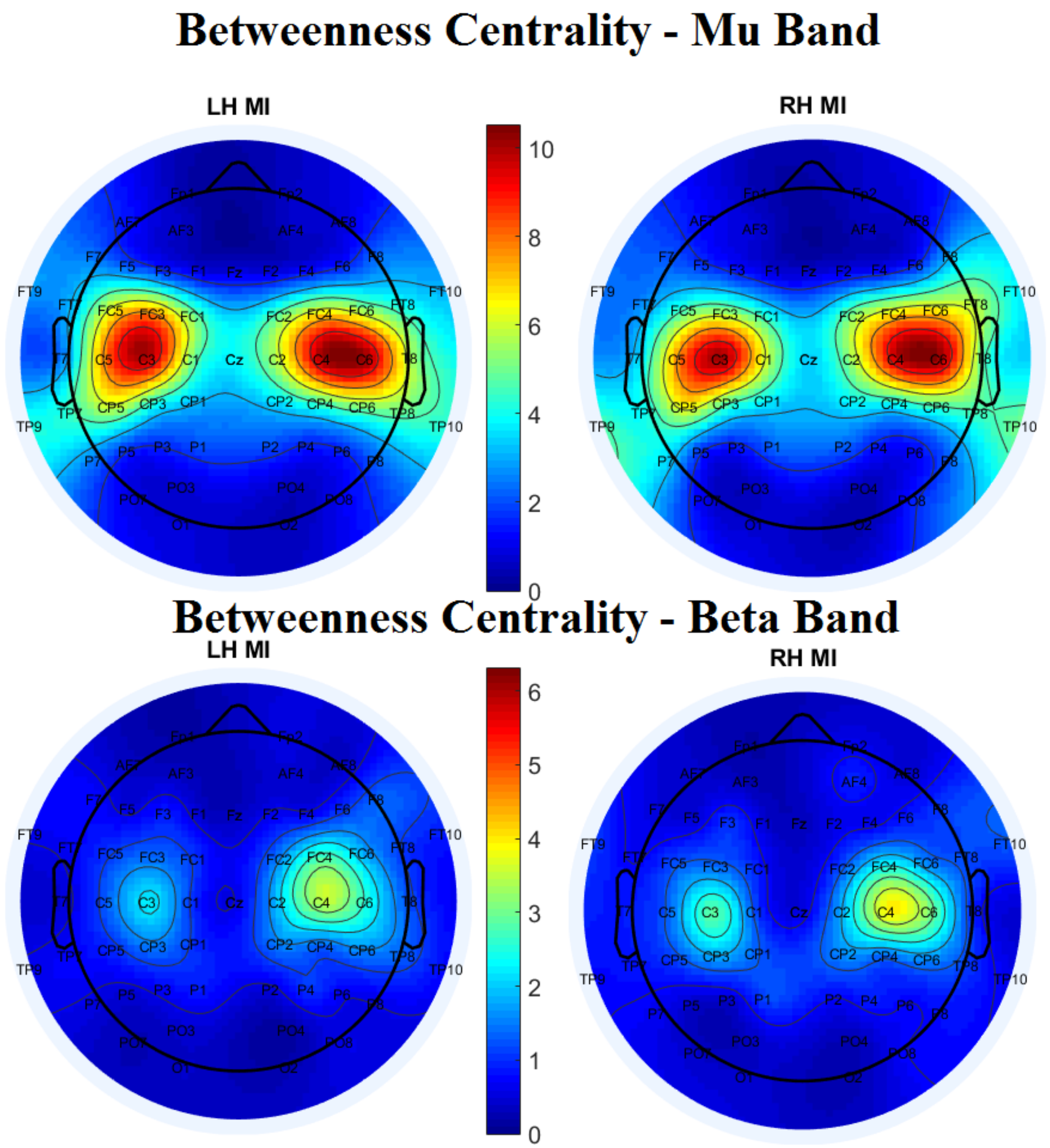


Figure 5.4. Betweenness centrality values distribution over the scalp. Left: left hand (LH) MI; right: right hand (RH) MI. Top row: mu band; bottom row: beta band.

Nodes that are amongst the ones with the lowest BC values are related to the occipital (PO3/PO4, PO7/PO8 and O1/O2), frontal-parietal, some frontal regions and intermediate areas in between these last two (FP1/FP2, AF3/AF4, AF7/AF8, Fz and F1/F2). Occipital-related nodes also displayed low degree centralities. Frontal-related electrodes, however, did not; which is an indicative that, although they may be regarded as relatively important nodes from the degree point of view, BC results suggest that, regarding an interpretation of information flow, they may be as unimportant to MI as the occipital electrodes. Nonetheless, these two centrality measures converge to the fact that motor cortex electrodes, such as FC5/FC6, C3/C4 and C5/C6, are the most central (or important) ones during MI tasks. This is actually what would be expected, since traditional frequency analysis of hand MI response is basically focused on these regions.

Note that the betweenness centrality is a measure related to how many minimum path lengths passing through a specific electrode are needed in order to travel between any two nodes of a graph. From Figures 5.3 and 5.4, it can be seen that there is a general tendency of nodes with larger characteristic path lengths to display smaller BC values. This is not unexpected, since a larger path length for a node means that, from a specific node, it is of greater cost to travel to any other node of the network. An interpretation of BC states that nodes with greater values control the network flow (information, energy, passengers etc.) [Monteiro, 2014]. Therefore, in our case, flow of information for motor imagery data would preferentially pass through nodes on the motor area (C3/C4, C5/C6 and FC3/FC4), which also would present smaller shortest path lengths, when compared to the rest of the graph.

In order to better visualize the relationship of BC and a node's path length, values were plotted into a scattering diagram (Figure 5.5). Each data point of the scattering plot corresponds to a pair of values (l_i, BC_i) of a node i. Regardless of the task or cerebral hemisphere, there seems to exist a general tendency for the BC to decrease its value as a node's path length increases. Three mathematical models were tested for fitting the data: a second degree polynomial (SDP; $f(x) = ax^2 + bx + c$), a single-term exponential (STE; $f(x) = Ae^{bx}$) and a two-term exponential (TTE; $f(x) = Ae^{bx} + Ce^{dx}$). The best fit was chosen according to the reduced chi-squared value. The closer the reduced chi-squared is to one, the better the model fits the data (Table 5.5).

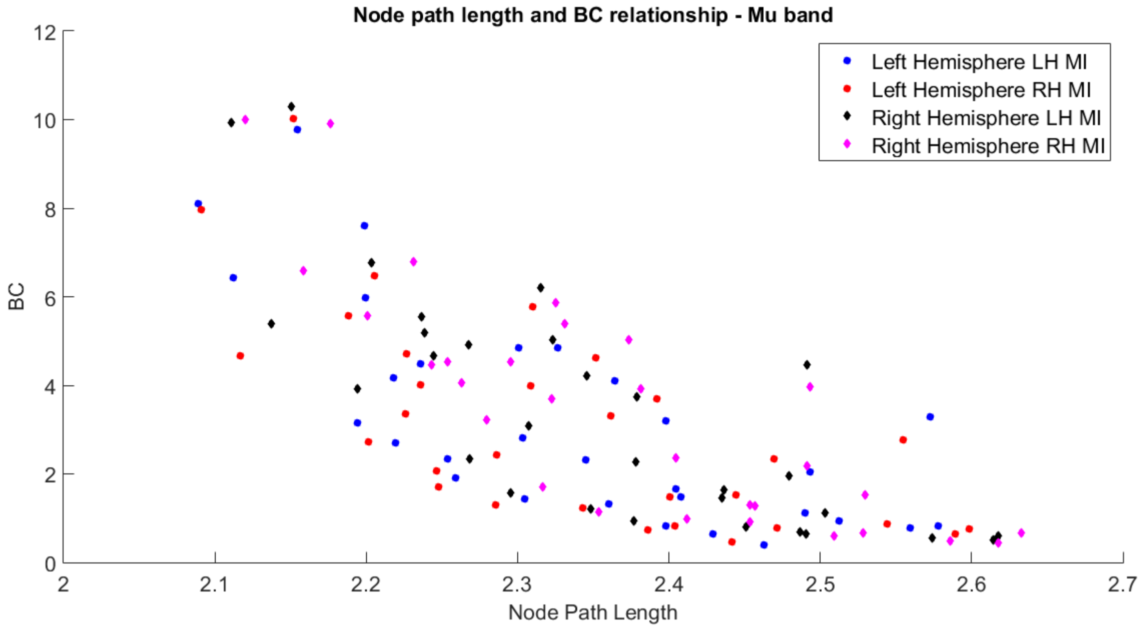


Figure 5.5. Relationship between BC and node path length (mu band). Both hemispheres and MI tasks are shown, in different colors and shapes. There are 30 data points (one for each electrode) regarding each entry of the graphic.

Table 5.5. Reduced chi-squared values for the three models tested (mu band). SDP: $f(x) = ax^2 + bx + c$; STE: $f(x) = Ae^{bx}$; TTE: $f(x) = Ae^{bx} + Ce^{dx}$.

Model	**Reduced chi-squared**			
	Left hemisphere		Right hemisphere	
	LH MI	RH MI	LH MI	RH MI
SDP	0.6931	0.8609	0.6913	0.5352
STE	0.7179	0.8586	0.6753	0.5196
TTE	0.7034	0.8553	0.6714	0.5196

Table 5.5 shows that all three models produced similar chi-squared values. Also, RH MI on the left hemisphere resulted in the best fitting. In addition, reduced chi-squared values were larger on a specific hemisphere during the contralateral hand MI, meaning that values for path length and BC are in more accordance with the three proposed models during these periods. In addition, LH MI shows little variation in fit quality comparing both hemispheres, while RH MI presents a more drastic difference, with reduced chi-squared values on the right hemisphere falling to almost half of their value on the left hemisphere. This, along with previous results, suggest that both cerebral hemispheres are not entirely symmetrical, at least regarding hand MI.

The observed behavior for the BC within the beta band was similar (bottom row, Figure 6.5). For both MI tasks, the same nodes presented more significate BC values (C3, C4, C6, FC4 – generally, electrodes from motor related areas), although they were slightly larger during the RH MI. Mainly all qualitative aspects occurring for the mu band were also present in the beta band. Again, the main difference is seen in the absolute values: they are smaller for the beta band for almost all electrodes.

A similar pattern was also obtained regarding the relationship between the BC and nodes path length values (Figure 5.6). Also for the beta band, it was found that the general tendency for BC values was to decrease, while a node's path length increased.

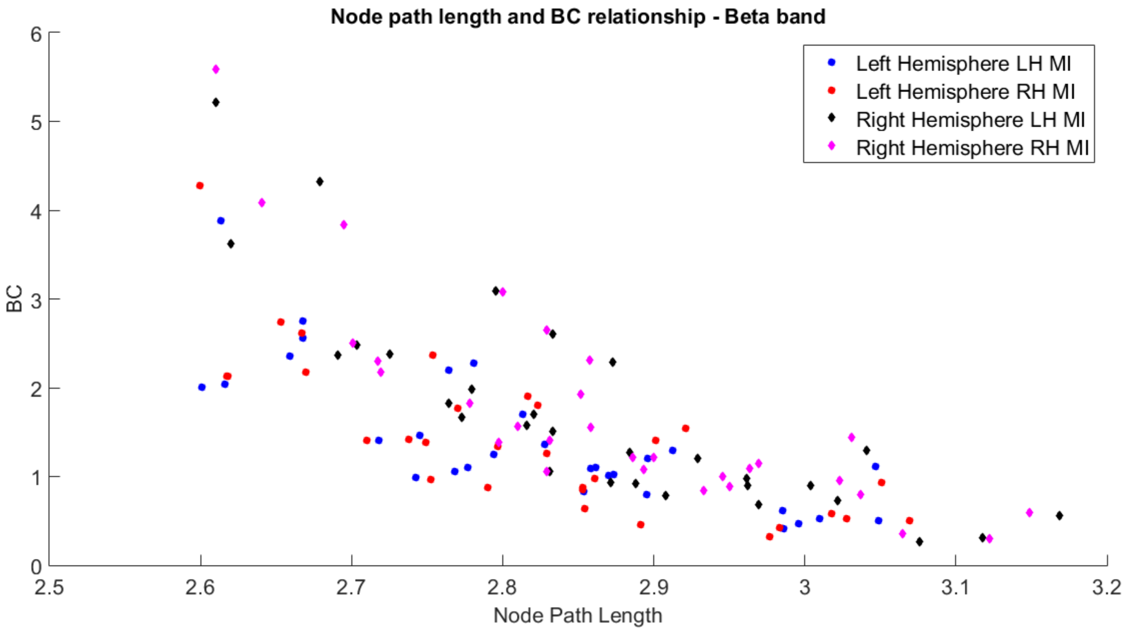


Figure 5.6. Relationship between BC and node path length (beta band). Both hemispheres and MI tasks are shown, in different colors and shapes. There are 30 data points (one for each electrode) regarding each set of data points.

5.1.5. Eigenvector Centrality (EC)

Average behavior for EC is shown in Figure 5.7.

For the mu band, on the left hemisphere, all frontal-related areas, except for the node located at FT7, showed higher EC values during the contralateral hand MI periods. For most central-parietal, parietal and occipital areas values are larger during imagery of the ipsilateral hand. On the right hemisphere, frontal-related electrodes, except for FC2, also present slightly larger ECs during the contralateral hand MI. For more posterior regions from Cz, most nodes show higher values during the ipsilateral hand MI.

Still regarding the mu band, on both hemispheres, there is a very similar qualitative behavior, just as was found for the other metrics. Again, nodes such as FC5/FC6 and C5/C6 showed the highest EC values, and centralities for the nodes positioned on the scalp central line (Fz and Cz) are amongst the lowest values. Note that, for both hemispheres, regardless of the MI task, the EC becomes greater as a node becomes farther apart from the medial fissure. The only two exceptions are the temporal-related areas (FTs and TPs electrodes).

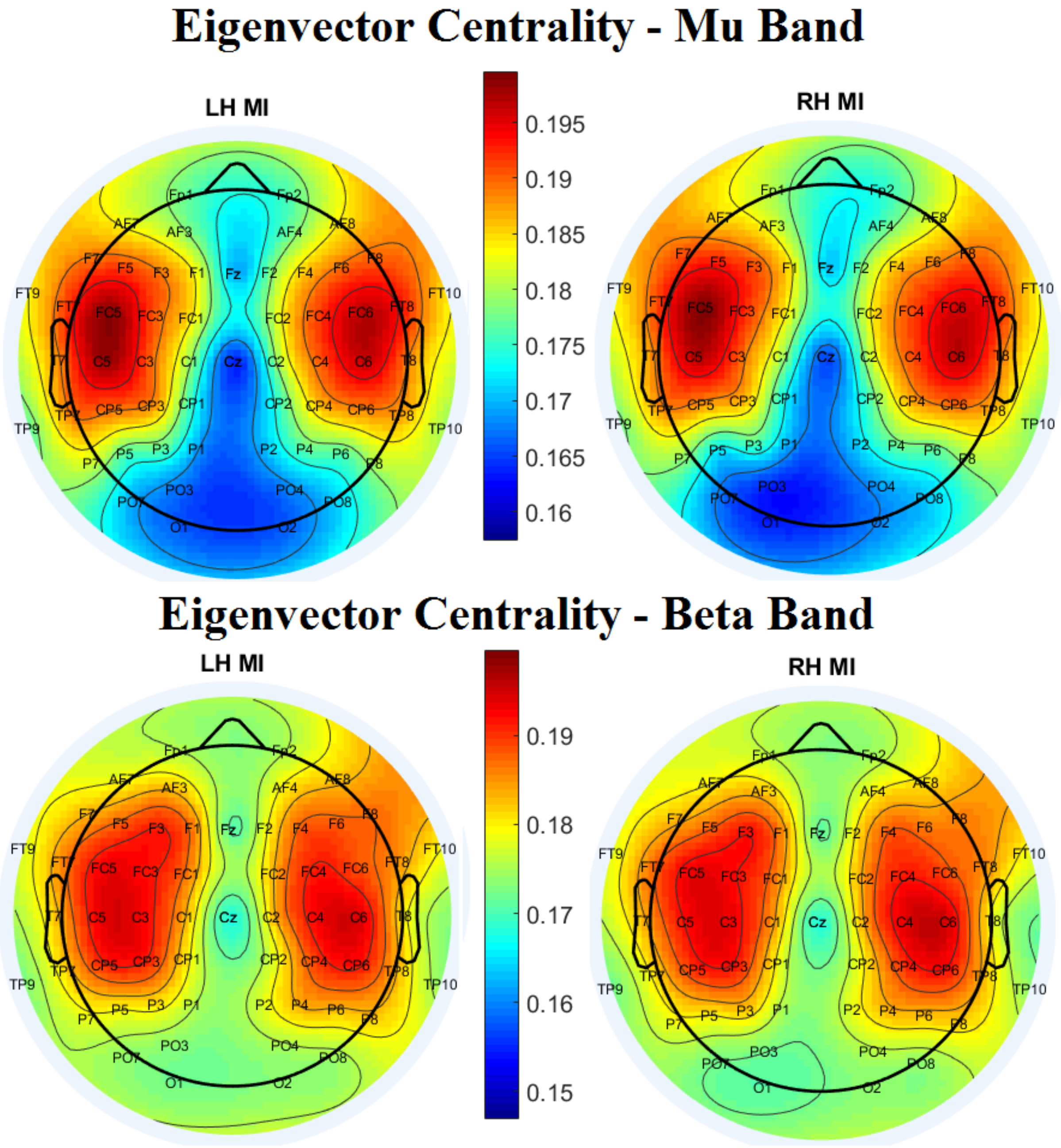


Figure 5.7. Eigenvector centrality values distribution over the scalp. Left: left hand (LH) MI; right: right hand (RH) MI. Top row: mu band; bottom row: beta band.

Again there exists a very similar pattern between both bands. This time, however, the difference in absolute values across them is not as significant as it was for the other metrics.

5.1.6. General remarks regarding the graphs' topology

From the metrics that were studied regarding graphs' topology evaluation, five were chosen to be displayed here: strength (which also corresponds, in our case, to the degree centrality), clustering coefficient, characteristic path length, betweenness centrality and eigenvector centrality. Actually, a few more measures were analyzed, such as strength distribution, entropy, energy (sum of the squares of the eigenvalues of the graph's Laplacian) and entropy of the discrete Fourier Transform of the Laplacian's eigenvalues. It was chosen not to show them, as they had little to add regarding discrimination between MI tasks – which is, in fact, the main goal of this work – and their study proved to be a little more complex, from a conceptual point of view.

The first three metrics – strength, clustering coefficient and characteristic path length – could be evaluated individually or as global graph properties. It was found that global properties made the distinction between MI tasks very unclear in some cases (if not virtually impossible), since all values were either equal to each other or highly overlapping due to the estimated standard deviation, which mainly reflected subject inter-variability and noise. When searching for distinguishable patterns between each MI task for each node (local properties), however, it was found that these three basic metrics provided similar and reproducible information, in the sense that they were able to generate similar patterns for all subjects.

On both hemispheres, the overall qualitative behavior was very similar during the two studied MI tasks. Note, however, that there was not, necessarily, a symmetry between both cerebral hemispheres. Also, simply glancing at the figures makes it very hard to find differences between MI conditions. Since overall graph properties were less useful than local metrics when trying to assess MI tasks, the study of other two centrality measures took place: betweenness and eigenvector centralities.

Betweenness centrality distribution over graphs' nodes presented well defined areas containing larger or smaller values. Nodes that presented the largest values were located on motor-related areas, such as FC5/FC6, C3/C4 and C5/C6. Lowest BCs were found in some frontal-related sites and parietal-occipital and occipital nodes. Also, there were regions (parietal) where BC values tended to become smaller as nodes became farther away from the medial fissure, and areas (frontal and center-parietal) where BC tended to increase as electrodes became farther away from this fissure. For BC, this means that nodes farther from the medial fissure can develop a greater or lower crucial role, in the sense of information flow, depending on the anatomic position they are on. However, a reason for why this may be is yet to be found.

Eigenvector centrality also showed defined regions for which values were higher or lower, although they are less easy to spot. Smallest values can be found for nodes located on the occipital area, while highest values lie amongst electrodes FC5/FC6, C3/C4 and C5/C6.

Note that three analyzed centrality measures – degree, betweenness and eigenvector – provide complementary information. The degree centrality regards the weight of each connection: the greater it is for a node, the more high-valued links it has attached to it. The betweenness centrality aims to measure how essential a node is in connecting any other two nodes of a graph, under the context of geodesic paths. Eigenvector centrality is related to the quality of the links, in the sense that nodes with high EC values tend to connect themselves to other nodes with, also, high EC values. Thus, nodes with high values of a given centrality do not, necessarily, display an equivalently large value for other centrality measure. For the degree centrality, for instance, frontal and occipital electrodes presented relatively large values when compared to the maximum obtained, which was dramatically reduced for the BC case (Figs. 5.1, 5.4 and 5.7).

For all centrality measures, motor-related electrodes (FC5/FC6, C3/C4 and C5/C6) presented the highest values, although it is for the BC that this behavior was most evident. Also, for all of them, lowest centrality values can be found on some frontal and occipital-related electrodes. This is expected to some extent, since a motor task was analyzed, which could reflect the greater involvement of motor cortex areas during hand motor imagery execution.

Nodes FC5/FC6, C3/C4 and C5/C6 also displayed the lowest characteristic path length values, indicating that their connections to the other nodes are of less cost and, therefore, traveling between them is preferable (which is reflected by their high BC).

In summary, it was found that, qualitatively, some graph metrics investigated showed a contra-laterality behavior regarding MI tasks at specific sites; which could be expected, given the nature of the MI response. However, classification of these tasks demands a more robust analysis. The results obtained when approaching this problem are shown in the next section.

5.2. Classification results

5.2.1. Strength, clustering coefficient and characteristic path length

Given their simplicity and recurring role in characterizing networks, the following metrics: strength, clustering coefficient and characteristic path length, were tested as inputs to the classifiers. Mean classification rates are shown in Table 5.6. Three classification methods were tested for each frequency band: a linear least-square based discriminant analysis (LSLDA), a linear support vector machine (LSVM) and a polynomial SVM (PSVM). Also, the effect of using each hemisphere individually for classification was studied ("Left hemisphere" and "Right hemisphere" columns). Finally, "Difference" refers to using the difference of a metric's value on the right hemisphere node minus its value on the correspondent contra-lateral node in the left hemisphere, as input to the classifier. Classification results for each subject, metric and frequency band can be found in Appendix C.

Results shown in Table 5.6 were obtained using all graph nodes for classification.

Table 5.6. Mean classification rates (strength, clustering coefficient and characteristic path length). Largest values obtained are highlighted in bold.

Band/Method	Feature	Classifier input		
		Left hemisphere (%)	Right hemisphere (%)	Difference (%)
μ/LSVM	Strength	64 ± 8	65 ± 9	67 ± 7
	Clustering Coefficient	54 ± 7	53 ± 8	54 ± 11
	Path Length	53 ± 7	52 ± 9	52 ± 9
μ/PSVM	Strength	51 ± 4	49 ± 3	61 ± 6
	Clustering Coefficient	54 ± 9	53 ± 8	54 ± 11
	Path Length	53 ± 10	53 ± 8	53 ± 6
μ/LSLDA	Strength	64 ± 8	64 ± 6	64 ± 7
	Clustering Coefficient	56 ± 9	59 ± 7	61 ± 9
	Path Length	53 ± 10	54 ± 6	53 ± 9
β/LSVM	Strength	66 ± 11	66 ± 10	67 ± 8
	Clustering Coefficient	52 ± 9	56 ± 9	53 ± 11
	Path Length	54 ± 10	54 ± 10	56 ± 14
β/PSVM	Strength	50 ± 5	52 ± 5	**68 ± 8**
	Clustering Coefficient	51 ± 10	53 ± 9	52 ± 11
	Path Length	52 ± 10	54 ± 9	53 ± 13
β/LSLDA	Strength	65 ± 11	64 ± 10	**68 ± 8**
	Clustering Coefficient	60 ± 9	61 ± 8	60 ± 8
	Path Length	53 ± 9	53 ± 10	54 ± 14

Analyzing each hemisphere individually, it can be noted that it makes little difference using any one of them for data classification. Highest variation comes from classifying the clustering coefficient within the mu band with a LSLDA: mean classification accuracy obtained using only the left hemisphere was 56%, while using the right hemisphere accounted for a slightly larger accuracy, of 59%. Also, strength classification provided the highest accuracies, except for PSVM classification within the mu band. CC and path length classification showed similar results.

Using the difference between corresponding right and left hemisphere values as input caused a huge increase in classification results in some cases. For instance, strength classification within the beta band using the LSVM method increased of 16 – 18%. This shows that combining metric values from corresponding nodes can drastically alter classification outcome in some cases. For most lines of the table, however, using the difference achieves similar classification rates to the ones obtained for each hemisphere individually. In summary, using the metric difference between corresponding nodes, strength classification provided the highest accuracies for all cases, while CC and path length showed similar (lower) results – the same qualitative result obtained using metrics for one hemisphere.

Overall, with a few exceptions, accuracies for the β-band are slightly larger. In fact, highest accuracies (68%) were obtained within this band, using PSVM and LSLDA methods, both for the strength as input. In this case, the linear method would be preferable, since it has considerably less computational cost. The LSVM resulted in a very similar accuracy (67%).

Nonetheless, note that most classification rates are associated with a relatively large standard deviation, which reflects the great inter-subject variability of the method under study. Some subjects were able to achieve 75% accuracy, while others were barely able to overcome 50% rates (which is practically a by-chance classification).

Results from Section 5.2.1 suggested that more significant differences may be found by searching for specific sites of the graph, rather than analyzing the global pattern. Therefore, since the LSLDA method provided the best results of Table 5.6 using the difference as the classifier input, this approach was used to test the effect of using specific electrode pairs for classification (Table 5.7). It is important to emphasize that Table 5.7 presents maximum classification accuracies obtained when searching for the optimum sets of nodes that best classify the data, for each metric. All possible combinations of up to three node pairs were tested, and results show the maximum rates obtained with a given number of pairs. This can be regarded as some type of limit, ideal accuracy rate. Also, values displayed are averaged across all subjects and acquisitions.

Table 5.7. Average maximum accuracies obtained when using optimal electrodes sets for classification. Highlighted values indicate the best results for each metric used as input.

Band/Number of electrodes pairs	Strength (%)	CC (%)	Path Length (%)
μ/One	62 ± 4	64 ± 5	60 ± 4
μ/Two	73 ± 2	70 ± 4	73 ± 3
μ/Three	76 ± 2	72 ± 6	77 ± 3
β/One	64 ± 5	61 ± 7	62 ± 6
β/Two	74 ± 5	72 ± 6	74 ± 5
β/Three	**78 ± 5**	77 ± 6	**78 ± 5**

For the three metrics, accuracy increases with the number of optimum electrode pairs used for data classification. However, as seen from Table 5.7, classification rates when using only one optimum set are comparable to the situation where all graph nodes are used, providing rates within the range 60-68% (Table 5.6). This suggests that there must exist a turning point where classification rate stops increasing with the number of optimum electrode sets used. Also, some nodes may possess irrelevant information for classification, which could actually cause classification rates to decrease significantly, as suggested from Tables 5.6 and 5.7.

For most cases, results for the beta band are slightly larger. Also, note that standard deviation values are significantly smaller than the ones found by the previous approach (Table 5.6), reflecting a more robust classification mean. Besides, different metric attained more similar values in Table 5.7.

Using three pairs of electrodes, the best result obtained was 92% (clustering coefficient, beta band, three pairs of electrodes, subject 1), and the worst was 64% (clustering coefficient, beta band, one pair of electrodes, subject 5).

The method of choosing only optimum node sets for every subject for classification rises, along with a significant increase in classification accuracy, a huge limitation: these optimum pairs vary across subjects, and even for the same subject, between different acquisitions made just minutes apart.

5.2.2. Centrality measures

The same procedure was done with the centrality measures (degree, betweenness and eigenvector). From the conclusions draw from Table 5.6, only the difference was used as input, in the same manner as it was previously done. Results are shown in Table 5.9.

Table 5.9. Mean classification accuracies ± standard deviation for the centrality measures. DC (degree centrality), BC (betweenness centrality), EC (eigenvector centrality).

Band/Method	**Centrality measure**		
	DC (%)	**BC (%)**	**EC (%)**
μ/LSLDA	64 ± 7	55 ± 6	63 ± 8
μ/LSVM	61 ± 6	51 ± 7	58 ± 7
μ/PSVM	67 ± 7	52 ± 7	57 ± 7
β/LSLDA	**68 ± 8**	54 ± 7	67 ± 7
β/LSVM	67 ± 8	56 ± 8	62 ± 6
β/PSVM	**68 ± 8**	54 ± 4	60 ± 5

The BC presented the lowest accuracy rates, with a maximum of 56%, barely above pure chance (50%). DC showed the largest accuracies, with a maximum value of 68%, and EC displayed values in between the other two centrality measures. Note, however, that in some cases, differences in classification accuracies rates are within the standard deviation ranges. This can impose a difficulty in establishing which metric would be the best to choose.

Again, as with Table 5.6, results for the β-band are slightly better, and best classification approaches were the LSLDA and PSVM methods. Within the μ-band, however, the PSVM classifier performed better (for the DC measure).

It can be seen that different metrics perform better with different classifying approaches for different frequency bands. Within the μ-band, DC provided better results when using a PSVM classifier, whilst for the other two measures, LSLDA would be the preferred method. Within the β-band, the PSVM resulted in higher accuracies for the BC, while the LSLDA was better for DC and EC.

Combinations of the three centrality measures were also tested, in order to check for possible improvements in classification results (Table 5.10). The combination that yielded better results consisted of degree and eigenvector centrality, which is expected, since these two provided the largest classification rates previously (Table 5.9).

Table 5.10. Mean classification accuracies ± standard deviation for centrality measures combinations. Bold values highlight the largest one obtained.

Band/Method	**DC+BC (%)**	**EC+BC (%)**	**DC+EC (%)**	**All (%)**
μ/LSLDA	55 ± 7	54 ± 7	56 ± 8	55 ± 8
μ/LSVM	55 ± 5	52 ± 6	67 ± 6	55 ± 5
μ/PSVM	52 ± 5	52 ± 8	61 ± 5	53 ± 5
β/LSLDA	52 ± 7	52 ± 6	52 ± 6	51 ± 6
β/LSVM	61 ± 6	56 ± 6	67 ± 6	61 ± 6
β/PSVM	56 ± 4	54 ± 3	**68 ± 7**	57 ± 4

Overall, combining centrality measures did not cause improvement. In some cases, it even worsened the results. For instance: from Table 5.9, the best classification rate obtained for the β/LSLDA band/method was 68%. This rate (for the same band and classifier) decreased to a maximum of 52% in Table 5.10. Thus, even though the three types of centralities provide complementary information, it can be seen that it does not mean that combining all this information will necessarily improve classification rates.

To explore really relevant classification information, as was done in Section 5.2.1, the effect of using specific pairs of electrodes and using them as input to the classifier was studied (Table 5.11). Also as done previously, the classifier used was the LSLDA since, when classifying these inputs separately, it was the one that gave maximum classification rates at the least computational cost.

Table 5.11. Average maximum classification results obtained using individual pairs of electrodes for classification (centrality measures). DC (degree centrality); BC (betweenness centrality); EC (eigenvector centrality).

Band/Number of electrodes sets	**DC (%)**	**BC (%)**	**EC (%)**
Mu/One	62 ± 4	55 ± 1	63 ± 4
Mu/Two	73 ± 2	69 ± 2	73 ± 4
Mu/Three	76 ± 2	73 ± 2	77 ± 2
Beta/One	64 ± 5	56 ± 2	64 ± 3
Beta/Two	74 ± 5	70 ± 3	73 ± 3
Beta/Three	**78 ± 5**	74 ± 3	76 ± 4

The type of centrality that provided best distinction between MI tasks was the degree centrality, reaching a maximum classification rate of 78%. Again, results for the beta band were slightly larger for almost all cases.

Similarly to what happened to the classification results for the strength, clustering coefficient and characteristic path length, classification accuracy increases with the number of optimum electrodes used for classification. Even the metric that performed worst (BC), provided accuracies comparable to the other two metrics (74%) with this approach.

5.2.3. Optimum pairs of electrodes for data classification

The most recurring optimum electrodes for classification are shown in Figures 5.17 and 5.18, in the form of histograms. All metrics were taken into consideration; that is, nodes that composed at least one optimum set of classification for at least one of the metrics were counted to produce these results. Note that there are two acquisitions for each subject and, therefore, the maximum number of times a pair of nodes can be present per metric is 2 x 8 = 16 times, for each number of pairs used for classification (that is, tests with one, two or three optimum pairs). Then, considering all possibilities of numbers of optimum pairs to use, this number raises up to 16 × 3 = 48 times. When considering the five studied metrics, the maximum number of times a pair of nodes can be present per band is 5 x 48 = 240 times. This value was used to normalize the counting number in the histograms of Figures 5.8 and 5.9. Also, only one of the pair constituents is shown in the horizontal axis. A detailed discrimination containing all optimum sets of electrodes pairs per subject, metric and band can be consulted in Appendix D.

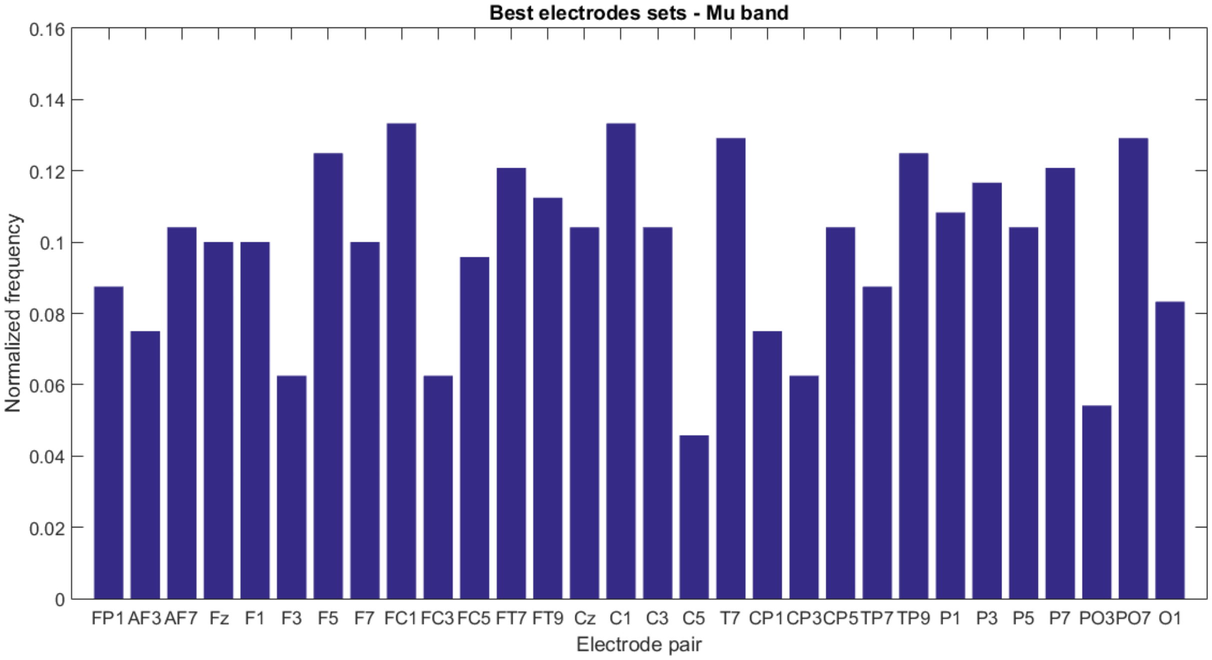


Figure 5.8. Normalized counting frequency of optimum electrodes for classification (mu band). Greatest number of occurrence of a pair was 32, for FC1/FC2 and C1/C2 pairs.

For the mu band, the most recurrent pair occurred at a maximum number of times of 32 (C1/C2 and FC1/FC2). Note, however, that this is relatively small, representing only about 13% of the achievable total of 240. The C5/C6 pair was the least recurrent, being present only at about 5% of times. Even though this pair is in the motor area, it presented less relevant information for classification than, for instance, the O1/O2 pair, located on the occipital cortex. Best pairs for classification were F5/F6, FC1/FC2, FT7/FT8, C1/C2, T7/T8, TP9/TP10 and PO7/PO8. Note that this distribution basically collects information from areas scattered all over the scalp: from frontal to occipital areas and from nearest and furthest regions from the scalp's central line.

For the beta band (Figure 5.18), the most recurrent pair was situated in the parietal cortex: P5/P6, and it occurred 38 times (about 16%). The least recurrent pair was T7/T8, with a relative frequency of about 6%. Best pairs for classification were AF3/AF4, AF7/AF8, F5/F6, FT9/FT10, C3/C4, P3/P5 and P5/P6.

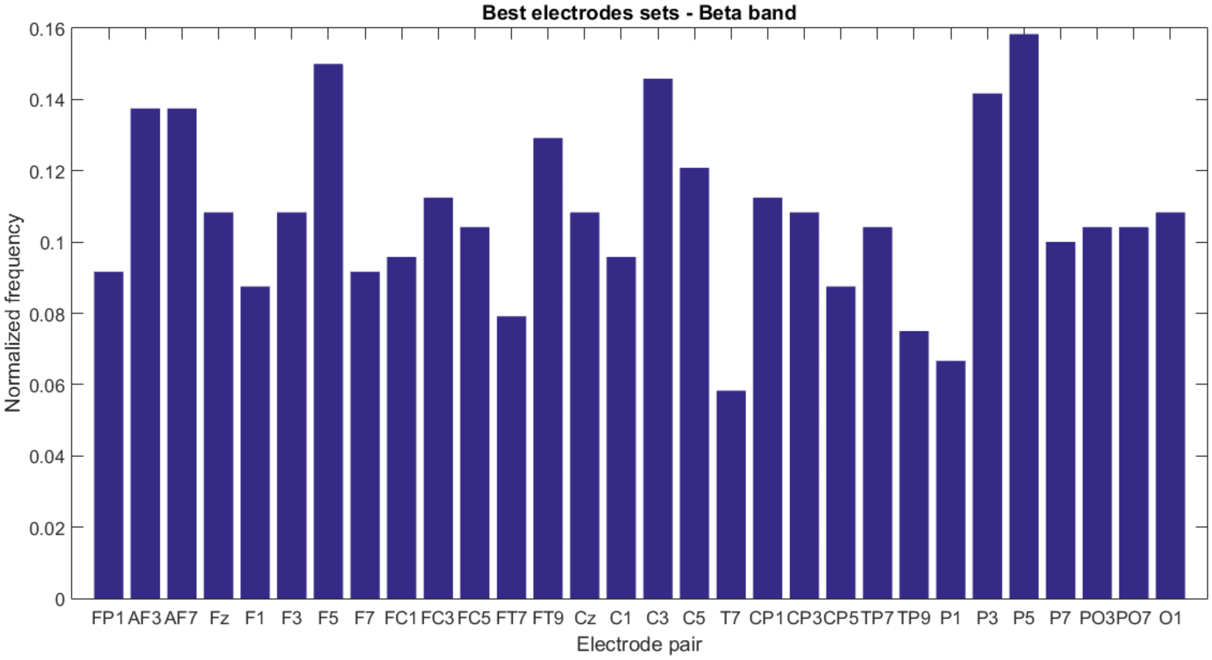


Figure 5.9. Normalized counting frequency of the optimum electrodes for classification (beta band). Greatest number of occurrence of a pair was 38, for the P/P6 pair.

Although the three centrality measures suggested that electrodes on motor areas ("FC" and "C" labeled nodes) were most central during the studied MI tasks, they do not necessarily are the best approach for data classification. For instance, the C5/C6 pair in the mu band and is amongst the least recurrent ones. This suggests that data classification is much more complex than a simple visual inspection, and that results over 70% (up to 80% to 90% in some cases; see Appendix C) can be obtained given the right electrode pairs combination and, classification algorithm.

5.2.4. General remarks regarding data classification

Within each band (mu and beta), three methods were tested for data classification: a linear least-squares based discriminant analysis (LSLDA), a linear support vector machine (LSVM) and a polynomial support vector machine (PSVM). For all measures, general findings can be drawn.

First of all, classification accuracy can be improved by working with optimal sets of electrodes for each subject. When doing so, classification accuracy increased by almost 20% in some cases. Nonetheless, this brings along several limitations; mainly, that these sets vary between subjects and between acquisitions for the same subject (Tables 5.7 and 5.11). Therefore, comprehending why such sets exist and how to determine them could play a major role in increasing the classifiers success rate. Note that these optimum pairs need not be in motor related areas. It is important to emphasize that these sets were found through an offline analysis, and this approach would limit the performance of an online BCI, if these limitations were not to be resolved.

By using just information from one optimum classification set it is possible to reproduce accuracy rates that are comparable (or even better, in some cases) to the ones when information from all nodes is used. We believe this is due to the fact that a considerable fraction of the electrodes does not contain relevant information regarding discrimination between MI tasks. In fact, the analysis of graph topology (Section 5.1) revealed that more significant differences were found when the graph was analyzed locally in specific regions, instead of being considered as a whole.

For all tests (different combinations of band, metrics, nodes and classifier), some subjects achieved higher accuracies than others. MI is a very complex and highly demanding task, thus, it is not expected that all users will produce distinguishable patterns in the same manner. When using three optimum classification pairs, for instance, some subjects achieved classification rates of 90%, while others provided results near 70% (see Appendix C). We believe that results could have been enhanced had subjects been submitted to MI training prior to data acquisition.

To the best of our knowledge, there are few works in the literature who attempted a similar approach to this particular classification problem, and none was able to produce better accuracy rates. Thus, our approach seems promising, although further studies are still necessary to overcome the limitations discussed above.

Chapter 6

Conclusions and future perspectives

In this work, we used metrics from graph theory in order to characterize the functional brain networks associated with motor imagery tasks of right and left hands, executed separately. We also attempted to use these metrics as input for a classifier in a BCI system.

To build the graphs, we used the motif method, since we believed a qualitative approach would be interesting to estimate similarities between the recorded EEG time series, given the highly noisy nature of the EEG signal. Besides, the motif method proved to be of very low computational cost. Nonetheless, more traditional analyses using Pearson's correlation were also done, for comparison with the results obtained with the motif method. For all cases tested, the qualitative behavior described in Chapter 5 was very similar in both approaches, with just slight differences in absolute values.

A first approach to build the graphs consisted of thresholding the connectivity matrix with different threshold values, each corresponding to a percentage of the maximum matrix element (from 10% to 90%). Results, however, varied a lot depending on the choice of threshold. Thus, due to the lack of a rigorous criterion to choose a threshold value for the connectivity matrix, and to avoid the risk of losing information, we chose to work with weighted graphs

Several graphs metrics not shown in this dissertation were tested: the graph's entropy, the Laplacian's energy, the entropy of the Fourier transform of the Laplacian's eigenvalues, and the degree distribution. A few parameters originating from the graph's minimum spanning tree were also analyzed. They did not provide much complementary information to what has been presented and, therefore, were not included in this dissertation. We chose to present what we considered to be the simplest and most recurring metrics regarding graphs studies, analyzing their feasibility for data classification: degree (also degree centrality), clustering coefficient, characteristic path length, betweenness centrality and eigenvector centrality.

In preprocessing steps, data were frequency filtered in the EEG specific bands, as described in Table 2.1. Motor imagery studies report mainly response in mu and beta bands and, therefore, these were the bands we worked with in this study. Additional bands were also tested, but no interesting result arose from them.

CAR filtering aimed to enhance the EEG SNR by removing artifacts arising at all channels at the same time. This can also reduce the correlation effect generated on different electrodes due to measuring signals arising from a same pool of neurons. In addition, we used a smoothing operation (Figure 4.2) to compensate for sudden and unexpected signal variations. This significantly increased classification results. Without it, most rates achieved when using all graph nodes barely surpassed 50 %.

The three centrality measures revealed that the most important nodes involved in the studied tasks lied on the motor area; namely, C3, C4, C5, C6, FC3 and FC4. This makes sense, given that these were motor tasks. Three distinct metrics were used because they provide complementary information. For instance: values for degree centrality were not very different between nodes, while for betweenness centrality, motor nodes presented higher values compared to other nodes. These three measures state that: the aforementioned nodes (C3, C4, C5, C6, FC3, FC4) on motor areas are the ones whose connections have the largest weights (degree centrality), the ones which are most important for information flow (betweenness centrality) and the ones that, also, tend to connect themselves to other central nodes (eigenvector centrality). Therefore, this confirms that motor areas are central during motor tasks.

Regarding data classification, we showed that relevant information is not present in all electrodes. In fact, the system’s performance can be optimized if it is known which are the best sets of nodes to be used. Results showed that using just one optimum pair produces similar rates than when all graph nodes are used. Besides, increasing number of electrodes to up to three pairs was accompanied by an increase in classification rate. Therefore, since using all electrodes did not provide the best rates, there must exist a turning point for which increasing the number of electrode pairs stops enhancing the classification rates.

Using these optimum sets, maximum average (over subjects) classification rate was of 78%, with some subjects achieving rates as high as 92% in some cases. This shows that differences between MI tasks, although hard to be spotted from the figures of Section 5.1, can exist in specific projections of attributes combinations. Thus, exploring combinations of metrics and values across frequency bands seems like a natural approach as a next step, along with the use of computational methods that can provide insights of optimum feature extraction techniques. This could also add to better understand the main limitation with our methodology; namely, the fact that the optimum electrode sets vary across subjects and acquisitions, making its definition complicated for an online application.

Between and within-subject variability is, indeed, a well-known variable when working with human data. In fact, we also found slight variations regarding which classifier algorithm and features performed best for each user. Besides, a BCI system and its user are in constant adaptation to one another. Finally, subjects tested here were not used to MI tasks. A study conducting MI training sessions between acquisitions could investigate if training could enhance the MI response and, therefore, the classification results.

From our results, some important questions arise:

- Can other metrics provide better classification?
- Can signal enhancing techniques (such as the surface Laplacian) improve classification results, if applied to the optimum classification electrodes?
- Is the difference between metric values on corresponding electrodes (in both hemispheres) really the best approach to provide features for the classifier?
- Is working with weighted and undirected graphs the best approach?
- Can new combinations of the studied metrics improve classification results?
- Can the use of longer data segments for classification enhance results?
- Can other classification algorithms provide better classification rates?
- Why are there optimum sets of electrodes for classification? What makes them be so?
- Up until to what point does using additional optimum electrode pairs for classification increase the accuracy rate?
- Can combinations of graph metrics with the more traditional frequency analysis improve classification results?
- Still using the motifs methods, considering different lag times can alter classification output?

The use of the human brain to directly control external devices raises a high degree of fascination and interest, and it allows application in many fields. So far, BCIs seem the way to do so. Any contribution that could enhance these systems performance, or even the ones that establish negative approaches should be taken into consideration. Our approach investigated a still novel topic allying BCI and graphs. Although the classification results obtained were not outstanding, important questions that could aid overcoming the method’s limitations were risen. Also, to the best of our knowledge, no work in the literature has obtained better classification results using graph features alone for MI-BCIs. We established that there are optimum sets of electrodes for each subject that should be searched for classification, since they provide better accuracies than using all nodes. Limitations still need to be investigated, as well as strategies to overcome them. We believe that more refined processing steps for finding the optimum electrode sets may increase our results above the 80% classification rate.

BCIs still have many obstacles to overcome in order to become well established communication systems. Advances in our understanding of the brain and engineering still need to be done. The general BCI concept is outstanding, but practical issues surely impose various limitations. The extent to which these systems can be applied, however, motivates their research, as challenging as it can be.

References

Aalto University website. http://ani.aalto.fi/en/ami_centre/facilities/monitoring_systems/eeg/. Accessed on 06/03/2016.

Al-ani T., Trad D. Signal Processing and Classification Approaches for Brain-Computer Interface, Intelligent and Biosensors, Vernon S. Somerset (Ed.), ISBN: 978-953-7619-58-9. *InTech*. DOI: 10.5772/7032 (2010).

Alberts B., Johnson A., Lewis J., Raff M., Roberts K., Walter P. Biologia Molecular da Célula. *Artmed*, 5th ed., Porto Alegre, Brazil (2010).

Antoniou I. E., Tsompa E. T. Statistical Analysis of Weighted Networks. *Discrete Dynamics in Nature and Society*, 2008, article ID 375452, 16 pages (2008).

Asensio-Cubero J., Gan J. Q., Palaniappan R. Multiresolution analysis over graphs for a motor imagery based online BCI game. *Computers in Biology and Medicine*, 68: 21-26 (2016).

Barrat A., Barthélemy M., Pastor-Satorras R., Vespignani A. The architecture of complex weighted networks. *PNAS*, 101(11): 3747 – 3752 (2012).

Bayliss J. D., Inverso S. A., Tentler A. Changing the P300 Brain-Computer Interface. *Journal of CyberPsychology & Behavior*, 7(6): 694-704 (2004).

Bayliss J. D., Inverso S. A. Automatic Error Correction Using P3 Response Verification for a Brain-Computer Interface. *Proc. 11th International Conference on Human- Computer Interaction*, July 22-27, Las Vegas, NV. Mahwah: Lawrence Relbaum Associates (2005).

Beltramini G.C., Análise temporal de correlatos hemodinâmicos associados à atividade epileptiforme através da técnica de EEG-RMf simultâneos. *UNICAMP, PhD theseis* (2014).

Berman B., Horovitz S., Venkataraman G., Hallett M. Self-modulation of primary motor cortex activity with motor and motor imagery tasks using real-time fMRI-based neurofeedback. *Neuroimage*, 59(2): 917 - 925 (2011).

Bessa A. D., Santos L. B. L., Martinez L. P. N. R., Costa M. C., Cardoso P. G. S. Introdução às Redes Complexas (2010). Avaliable at: http://www.dpi.inpe.br/geocxnets/wiki/lib/exe/fetch.php?media=wiki:introducao_as_redes_complexas.pdf. Accessed on 01/26/2016.

Birch G.E., Bozorgzadeh Z., Mason S.G. Initial On-Line Evaluations of the LF-ASD brain–computer interface with able-bodied and spinal-cord subjects using imagined voluntary motor potentials. *IEEE Transactions on Neural Systems Aand Rehabilitation Engineering*, 10(4): 219 – 224 (2002)

Blackwood D.H., Muir W.J. Cognitive brain potentials and their application. *The British Journal of Psychiatry*, 157(9): 96-101 (1990).

Bos D. P. O., Reuderink B., van de Laar B., Gürkök H., Mühl P., Nijholt A., Heylen D. Brain-Computer Interfacing and Games, chapter 10. In: Tan D.S., Nijholt A. Brain-Computer Interfaces - Applying our minds to Human-Computer Interaction. *Springer*. ISBN: 978-1-84996-271-1 (2010).

Cheng M., Jia W., Gao X., Gao S., Yang F. Mu rhythm-based cursor control: an offline analysis. *Clinical Neurophysiology*, 115: 745 – 751 (2004).

Cincotti F., Mattia D., Babiloni C., Carducci F., Salinari S., Bianchi L., Marciani M.G., Babiloni F. The Use of EEG Modifications Due to Motor Imagery for Brain–Computer Interfaces. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 11(2): 131-133 (2003).

Chuang C.H., Lai P.C., Ko L.K., Kuo B.C., Lin C. T. Driver’s Cognitive State Classification toward Brain Computer Interface via using a Generalized and Supervised Technology. *IEEE*. DOI: 10.1109/IJCNN.2010.5596835 (2010).

Cortech Solutions website. http://www.cortechsolutions.com/Products/EC/EC-AC/EC-AC-DR. Assessed on 01/07/2016.

Cortes C., Vapnik V. Support-Vector Networks. *Machine Learning*, 20: 273 – 297 (1995).

Costa L.F., Rodrigues F.A., Cristino A.S. Complex networks: The key to systems biology. *Genetics and Molecular Biology*, 31(3): 591 – 601 (2008).

Donchin E., Spencer K. M., Wijesinghe R. The Mental Prosthesis: Assessing the Speed of a P300-Based Brain–Computer Interface. *IEEE Transactions on Rehabilitation Engineering*, 8(2) (2000).

Duda R.O., Hart P.E., Stork D. G. Pattern Classification. *Wiley*, 2nd edition (2001).

EEGLAB website. http://sccn.ucsd.edu/eeglab/. Assessed on 12/28/2015.

Einstein A. Investigations on the theory of the Brownian motion. *Dover Publications, Inc.*, USA (1956).

Ermentrout G.B., Terman D.H. Mathematical Foundations of Neuroscience. *Springer*, New York (2010).

Fallani F.D. V., Richiardi J., Chavez M., Achard S. Graph analysis of functional brain networks: practical issues in translational neuroscience. *Philosophical Transactions of the Royal Society B: Biological Sciences* (2014). Avaliable at: http://arxiv.org/abs/1406.7391v1. Accessed on 01/26/2016.

Farwell L A., Donchin E. Talking of the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and clinical Neurophysiology*, 70: 510-523 (1988).

Fiedler P., Haueisen J., Jannek D., Griebel S., Zentner L., Vaz F., Fonseca C. Comparison of three types of dry electrodes for electroencephalography. *Acta Imeko*, 3(3): 33-37 (2014).

Finger S. Origins of neuroscience: a history of exploration into brain functions. *Oxford USA POD*. ISBN: 978-0-19-514694-3 (2001).

Garrett D., Peterson D. A., Anderson C. W., Thaut M. H. Comparison of linear, nonlinear, and feature selection methods for EEG signal classification. *IEEE Trans. On Rehabilitation Engineering*, 11: 141-144 (2003).

Ghosh P., Mazumder A., Bhattacharyya S., Tibarewala D. N., Hayashibe M. Functional Connectivity Analysis of Motor Imagery EEG signal for Brain-computer Interfacing Application. *7th Annual International IEEE EMBS Conference on Neural Engineering* (2015).

Gould P.R. On the Geographical Interpretation of Eigenvalues. *Transactions of the Institute of British Geographers*, 42: 53 – 86 (1967).

Graben P.B., Rodrigues S. A biophysical observation model for field potentials of networks of leaky integrate-and-fire neurons. *Frontiers in Computational Neuroscience*, 6(100) (2013).

Graimann B., Allison B. Z., Pfurtscheller G. Brain-computer interfaces: Revolutionizing human-computer interaction. *Springer* (2010).

Gross C. Early History of Neuroscience. In: Adelman G. Encyclopedia of Neuroscience. *Birkhauser*, 843 – 847 (1987).

Guger C., Allison B.Z., Großwindhager B., Pruckl R., Hintermuller C., Kapeller C., Bruckner M., Krausz G., Edlinger G. How Many People Could Use an SSVEP BCI? *Front. Neurosci.*, 6 (2012).

Gupta R., Falk T. H. Affective State Characterization based on Electroencephalography Graph-Theoretic Features. *7th International IEEE/EMBS Conference on Neural Engineering*, 577-580 (2015).

Gtec website. *http://www.gtec.at.*

Hagemann D., Naumann E., Thayer J. F. The quest for the EEG reference revisited: a glance from brain asymmetry research. *Psychophysiology*, 38(5): 847 - 857 (2001).

Halder S., Varkuti B., Bogdan M., Kubler A., Rosenstiel W., Sitaram R., Birbaumer N. Prediction of brain-computer interface aptitude from individual brain structure. *Frontiers in Human Neuroscience*, 7. Article 105 (2013).

Haufe S., Tomioka R., Dickhaus T., Sannelli C., Blankertz B., Nolte G., Muller K.R. Localization of class-related mu-rhythm desynchronization in motor imagery based Brain-Computer Interface sessions. *Annual International Conference of the IEEE Engineering in Medicine and Biology*, Buenos Aires: 5137-5140 (2010).

Herculano-Houzel, S. The remarkable, yet not extraordinary, human brain as a scaled-up primate brain and its associated cost. *PNAS*, 109(1): 10661 - 10668 (2012).

Hermes D., Vansteensel M.J., Albers A.M., Bleichner M.G., Benedictus M.R., Mendez Orellana C., Aarnoutse E.J., Ramsey N.F. Functional MRI-based identification of brain areas involved in motor imagery for implantable brain–computer interfaces. *Journal of Neural Engineering*, 8 (2011).

Hoffmann U., Vesin J. M., Ebrahimi T., Diserens K. An efficient P300-based brain-computer interface for disabled subjects. *Journal of Neuroscience Methods*, 167(1): 115-125 (2008).

Holper L., Muehlemann T., Scholkmann F., Eng K., Kiper D., Wolf M. Testing the potential of a virtual reality neurorehabilitation system during performance of observation, imagery and imitation of motor actions recorded by wireless functional nearinfrared spectroscopy (fNIRS). *Journal of Neuroengineering and Rehabilitation*, 7(57) (2010).

Hwang H.J., Kwon K., Im C.H. Nerofeedback-based motor imagery training for braincomputer interface (BCI). *Journal of Neuroscience Methods*, 179(1): 150-156 (2009).

IBGE, censo demográfico de 2010. Data avaliable at the institute’s website: http://www.ibge.gov.br/home/. Accessed at 05/29/2016.

Iturrate I., Antelis J. M., Kubler A., Minguez J. A Noninvasive Brain-Actuated Wheelchair Based on a P300 Neurophysiological Protocol and Automated Navigation. *IEEE Transactions on Robotics*, 25(3): 614 - 627 (2009).

Johnston D., Wu S.MS. Foundations of cellular neurophysiology, *the MIT press*, London, England (1995).

Jung T.P., Makeig S., Lee T.W., McKeown M.J., Brown G., Bell, A.J., Sejnowski T.J. Independent Component Analysis of Biomedical Signals. *The 2nd Int'l Workshop on Indeppendent Component Analysis and Signal Separation*: 633-44 (2000).

Kandel E.R., Schwarts G. H., Jessell T. M., Siegelbaum S.A., Hudspeth A.J. Principles of Neural Science, *McGrawHill Medical* (2013).

Kanoh S., Murayama Y., Miyamoto K., Yoshinobu T., Kawashima R. A NIRS-based Brain-Computer Interface System during Motor Imagery: System Development and Online Feedback Training. *31st Annual International Conference of the IEEE EMBS* (2009).

Kaper M., Meinicke P., Grossekathoefer U., Lingner T., Ritter H. BCI Competition 2003—Data Set IIb: Support Vector Machines for the P300 Speller Paradigm. *IEEE Transactions on Biomedical Engineering*, 51(6): 1073 - 1076 (2004).

Krusienski D. J., Sellers E. W., Cabestaing F., Bayoudh S., McFlarland D. J., Vaughan T. M., Wolpaw J. R. A Comparison of Classification Techniques for the P300 Speller. *Journal of Neural Engineering*, 3(4): 299-305 (2006).

Kubler A., Kotchoubey B., Kaiser J. Brain-Computer Communication: Unlocking the Locked in. *Psychological Bulletin*, 127(3): 358-375 (2001).

Kuffler S. W., Nicholls J. G. The physiology of neuroglial cells. *Erg Physiol.*, 57:1–90 (1966).

Laplante P.A. Comprehensive Dictionary of Electrical Engineering. *CRC Press*, 2nd ed., p. 190 (2005).

Lee H.S., Ghetti A., Pinto-Duarte A., Wang X., Dizewczapolski G., Galimi F., Huitron-Resendiz S., Piña-Crespo J.C., Roberts A.J., Verma I.M., Sejnowski T.J., Heinemann S.F. Astrocytes contribute to gamma oscillations and recognition memory. *PNAS*, 111(32): E3343 – E3352 (2014).

Leite S.N.C. Contribuições ao desenvolvimento de interfaces cérebro-computador baseadas em potenciais evocados visualmente em regime estacionário. *Unicamp*, PhD thesis (2016).

McFarland D. J., McCane L. M., David S. V., Wolpaw J. R. Spatial filter selection for EEG-based communication. *Electroencephalography and clinical Neurophysiology*, 103: 386 - 394 (1997).

McFarland D., McCane L., Wolpaw J. EEG-Based Communication and Control: Short-Term Role of Feedback. *IEEE Transactions on Rehabilitation Engineering*, 6(1) (1998).

Meinicke P., Kaper M., Hoppe F., Heumann M., Ritter H. Improving Transfer Rates in Brain Computer Interfacing: a Case Study. In: Becker S., Thrun S., Obermayer K. Advances in Neural Information Processing Systems, *MIT Press*, Cambridge (2003).

Mihajlovic V., Molina G. G. To what extent can dry and water-based EEG electrodes replace conductive gel ones? A Steady State Visual Evoked Potential Brain-Computer Interface Case Study. *ICBE Conference Paper* (2011).

Miranda R. A., Casebeer W. D., Hein A. M., Judy J. W., Krotkov E. P., Laabs T. L., Manzo J. E., Pankratz K. G., Pratt G. A., Sanchez J. C., Weber D. J., Wheeler T. L., Ling G. S. F. DARPA-funded efforts in the development of novel brain–computer interface technologies. *Journal of Neuroscience Methods*. DOI: 10.1016/j.jneumeth.2014.07.019 (2014).

Monteiro L.H.A. Sistemas Dinâmicos Complexos. *LF editorial*, 2nd ed. (2014).

Naatanen, R. Processing negativity: an evoked-potential reflection of selective attention. *American Psychological Association*, 92(3): 605-640. DOI: 10.1037/0033-2909.92.3.605 (1982).

Neuper, C., Pfurtscheller G. Motor imagery and ERD Related Desyncronization, *Handbook of Electroencepalography and Clinical Neurophysiology (Revised Edition)*, 6: 303 - 525 (1999a).

Newman M. E. J. A measure of betweenness centrality based on random walks. *Social Networks*, 27: 39 – 54 (2005).

Newman M. E. J. The mathematics of networks. In: Blume L. E., Durlauf S.N. The New Palgrave Encycplopedia of Economics. 2nd ed. (2008).

Nuwer M. R., Comi G., Emerson R., Fuglsang-Frederiksen A., Guérit J. M., Hinrichs H., Ikeda A., Luccas F. J. C., Rappelsberger P. IFCN standards for digital recording of clinical EEG. *Electroencephalography and Clinical Neurophysiology*, 106: 259 - 261 (1998).

Obermaier B., Müller G.M., Pfurtscheller G. “Virtual Keyboard” Controlled by Spontaneous EEG Activity. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 11(4): 422 - 426 (2003).

Olejniczak P. Neurophysiologic Basis of EEG. *Journal of Clinical Neurophysiology*, 23(3): 186 – 189 (2006).

Panoulas K. J., Hadjileontiadis L. J., Panas S.M. Brain-Computer Interface (BCI): Types, Processing Perspectives and Applications, Capítulo 14. In: Tsihrintzis GA, Jain LC. Multimedia Services in Intelligent Environments – Integrated Systems. *Springer*. ISBN-13: 978-3642133954 (2010).

Pfurtscheller G., Aranibar A. Event related cortical desynchronization detected by power measurements of scalp EEG. *Electroenceph Clin Neurophysiol*, 42: 817-826 (1977).

Pfurtscheller, G., Neuper, C., Andrew, C., & Edlinger, A. Foot and hand area mu rhythms. *International Journal of Psychophysiology*, 26:121–135 (1997).

Pfurtscheller G. EEG Event-related Desynchronisation (ERD) and Event-related Synchronisation (ERS). In Niedermeyer E., Lopes da Silva, F.H. (Eds.), Electroencephalography. Basic Principles, Clinical Applications, and Related Fields, *forth ed., Williams and Wilkins, Baltimore* (1999c).

Pfurtshceller G., Neuper C., Birbaumer N. Human Brain-Computer Interface, Chapter 14. In: Vaadia E, Riehle A. Motor cortex in voluntary movements: a distributed system for distributed functions. *CRC Press*. ISBN-13: 9780849312878 (2005).

Picton T.W. The P300 Wave Wave of the Human Event-Related Potential. *Journal of Clinical Neurophysiology*, 9(4): 456-479 (1992).

Picton T.W., Bentin S., Berg P., Donchin E., Hillyard S. A., Johnson R Jr., Miller G.A., Ritter W., Ruchkin D.S., Rugg M.D., Taylor M.J. Guidelines for using event-related potentials to study cognition: Recording standards and publication criteria. *Psychophysiology* 37(2): 127-152 (2000).

Polich J. Overview of P3a and P3b. In: Polich J. Detection of change: event-related potential and fMRI findings. *Boston, MA: Kluwer*, p. 83 - 98 (2003).

Polich J. Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, 118: 2128 - 2148 (2007).

Poskanzer K.E., Yuste R. Astrocytic regulation of cortical UP states. *PNAS*, 108(45): 18453 – 18458 (2011).

Pritchard W.S. Psychophysiology of P300. *Psychological Bulletin*, 89(3): 506-540 (1981).

Repvos G. Dealing with Noise in EEG Recording and Data Analysis. *Infor Med. Slov.*, 15(1): 18 – 25 (2010).

Ribas G.C. The cerebral sulci and gyri. *Neurosurg Focus*, 28(2): E2 (2010).

Rivet B., Souloumiac A., Attina V., Gibert G. xDAWN Algorithm to Enhance Evoked Potentials: Application to Brain Computer Interface. *IEEE Trans. Biomedical Eng.* (2009).

Rosário R., Cardoso P.T., Muñoz M.A., Montoya P., Miranda J.G.V. Motif-Synchronization: A new method for analysis of dynamic brain networks with EEG. *Physica A*, 439: 7 – 19 (2015).

Ross S. Probabilidade: um curso moderno com aplicações. *Bookman*, 8th ed. (2010).

Rudell A.P., Fox S.E. The propagation potential: An axonal response with implications for scalp-recorded EEG. *Biophysics Journal*, 60: 556 – 567 (1991).

Saham S. Glia–Neuron Interactions in Nervous System Function and Development. *Current topics in development Biology*, 69: 39 – 66 (2005).

Salazar-Varas R., Gutiérrez D. Feature Extraction for Multi-class BCI using EEG coherence. *7th Annual International IEEE EMBS Conference on Neural Engineering* (2015).

Salimi Khorshidi G., Jaafari A., Motie Nasrabadi A., Hashemi Golpayegani, M. Modifying the classic peak picking technique using a fuzzy multi agent to have an accurate P300-based BCI. *EUSFLAT Conf.*, 2:143-147 (2007).

Salmoni A., Schmidt A., Walter C. Knowledge of results and motor learning: A review and critical reappraisal. *Psychologic. Bull*., 95: 355-386 (1984).

Sanei S., Chambers J.A. EEG signal processing. *John Wiley & Sons LTD*, England (2007).

Sellers E. W., Krusienski D. J., McFarland D. J., Vaughan T. M., Wolpaw J. R. A P300 event-related potential brain–computer interface (BCI): The effects of matrix size and inter stimulus interval on performance. *Biological Psychology*, 73: 242 - 252 (2006).

Shamas M., Wendling F., Falou W. El, Hassan M. EEGNET: A Novel Tool for Processing and Mapping EEG Functional Networks. *7th International IEEE/EMBS Conference on Neural Engineering* (2015).

Sharma N., Baron J. C., Rowe J. B. Motor Imagery After Stroke: Relation Outcome to Motor Network Connectivity. *Ann. Neurol*., 66(5): 604-616 (2009).

Silva, F.L. EEG: Origin and Measurement. In: Mulert C., Lemieux L. EEG-fMRI, *Springer*, Berlin (2010).

Sitaram R., Zhang H., Guan C., Thulasidas M., Hoshi Y., Ishikawa A., Shimizu K., Birbaumer N. Temporal classification of multichannel near-infrared spectroscopy signals of motor imagery for developing a brain–computer interface. *NeuroImage*, 34: 1416 – 1427 (2007).

Slenes G. F., Beltramini G. C., Lima F. O., Li, L. M. The use of fMRI for the evaluation of the effect of training in motor imagery BCI users. *Neural Engineering (NER), 2013 6th International IEEE/EMBS Conference on, pp.686-690, 6-8 Nov. 2013* (2013).

Somjen G.G., Trachtenberg M. Neuroglia as generator of extracellular current. In: Speckmann E.J., Caspers H., Origin of CerebralField Potentials. *Stuttgart: Thieme*: 21–32 (1979).

Speckmann E. J., Elger C. E., Gorji A. Neurophysiologic Basis of EEG and DC Potentials. In: Niedermeyer E., Silva F.L. Electroencephalography: basic principles, clinical applications and related fields. *Lippincott Willians& Wilkins* (2005).

Stanny R. R. Mapping the event related Potentials of the brain: Theoretical issues, Technical considerations and computer programs. *Naval aerospace Medical Research Laboratory, Naval Air Station, Pensacola, Florida* (1989).

Straffin P.D. Linear Algebra in Geography: Eigenvectors of Networks. *Mathematics Magazine*, 53(5): 269 – 276 (1980).

Sur S., Sinha V.K. Event-related potential: An overview. *Industrial Psychiatry Journal*, 18(1): 70-73. DOI: 10.4103/0972-6748.57865 (2009).

Sutton S., Braren M., Zubin J., John E.R. Evoked potential correlates of stimulus uncertainty. *Science Magazine,* 150(3700): 1187-1188 (1965).

Teplan M. Fundamentals of EEG Measurement. *Measurement Science Review*, 2(2) (2002).

Unicog website. *http://www.unicog.org/pmwiki/uploads/Main/EEGElectrode.jpg.*

Usakli A.B. Improvement of EEG Signal Acquisition: An Electrical Aspect for State of the Art of Front End. *Computational Intelligence and Neuroscience*, 2010 (2010).

Velásquez-Martínez LF, Álvarez-Meza AM, Castellanos-Domínguez CG. Motor Imagery Classification for BCI Using Common Spatial Patterns and Feature Relevance Analysis. In Vicente J.M.F., Sánchez J.R.A., López F.P., Moreo F.J.T. Natural and Artificial Computation in Engineering and Medical Applications. *5th International Work-Conference on the Interplay Between Natural and Artificial Computation, IWINAC Mallorca, Spain, June 2013, Proceedings, Part II* (2013).

Vidal J.J. Toward Direct Brain-Computer Communication. *Annual Review of Biophysics and Bioengineering*, (2): 157-180 DOI:10.1146/annurev.bb.02.060173.001105 (1973).

Volkmar F. R. Encyclopedia of Autism Spectrum Disorders. *Springer*, New York (2013).

Wang W., Collinger J. L., Perez M. A., Tyler-Kabara E. C., Cohen L. G., Birbaumer N., Brose S. W., Schwartz A. B., Boninger M. L., Weber D. J. Neural Interface Technology for Rehabilitation: Exploiting and Promoting Neuroplasticity. *Phys. Med. Rehabil. Clin. N. Am.*, 21(1): 157-178 (2010).

Widge A. S., Dougherty D. D., Moritz C. T. Affective brain-computer interfaces as enabling technology for responsive psychiatric stimulation. *Brain-Computer Interfaces*, 1(2): 126-136 (2014).

Wikimedias commons website. *https://commons.wikimedia.org/wiki/File:21_electrodes_of_International_10-20_system_for_EEG.svg*.

Wolpaw J. R., Birbaumer N., McFarland D. J., Pfurtscheller G., Vaughan T. M. Brain-computer interfaces for communication and control. *Clinical Neurophysiology*, 113(6): 767-91 (2002).

Zhang Y., Zhang Z., Wei D., Deng Y. Centrality Measures in Weighted Networks Based on Amoeboid Algorithm. *Journal of Informatics & Computational Science*, 9(2): 369 – 376 (2012).

Zhu D., Bieger J., Molina G. G., Aarts R. M. A survey of stimulation methods used in SSVEP based BCIs. *Computational Intelligence and Neuroscience*, Article 702357, 12 pages (2010).

Appendix A

Some membrane electrophysiological aspects

A.1. Membrane resting potential derivation

Since the protein channels allow charge to pass through them in the form of an electric current and the lipid bilayer acts as an electrical insulator between intra and extracellular membrane sites, an electrical modeling of the cellular membrane as an electric circuit is in order. The resistive component represents the protein channels, and the capacitive component, the lipid bilayer. This is sketched in Figure A.1. In this figure, all possible types of ion channels that could compose the membrane structure are summarized into one resistance component (“R_m”, the membrane resistance), and in the membrane’s capacity, “C_m”. “I_m” is the current flowing through the membrane, which is divided into two parts: the ionic current (“I_i”) and the capacitive current (“I_c”). “V_m” is the membrane potential and “E_r” represents the equilibrium potential, occurring when the ion’s membrane potential balances the ion’s concentration gradient.

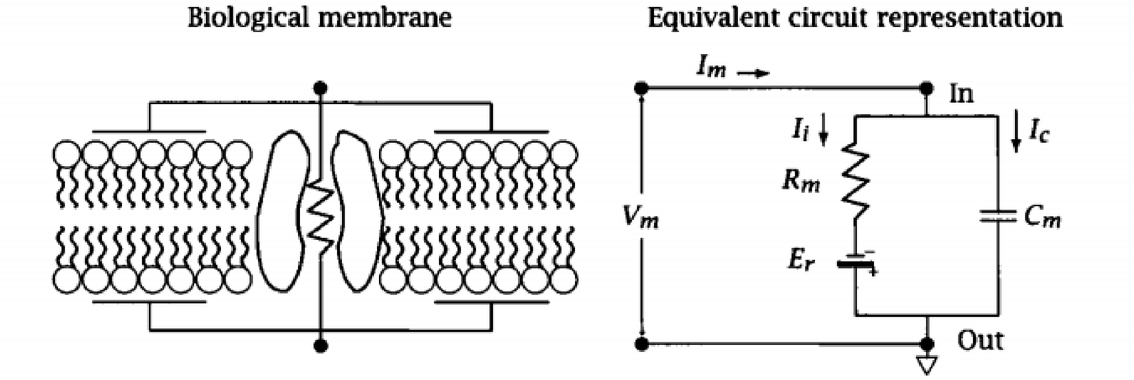


Figure A.1. Membrane equivalent circuit representation. **Extracted from [Johnston and Wu, 1995].**

Fick’s law of diffusion states that the diffusion flux (J_{diff}), in molecules/s.cm², is related to the diffusion coefficient and the rate to which the particle concentration ([C]) varies in a length dimension (x), by:

$$J_{\text{diff}} = -\mathrm{D}\frac{\partial[\mathrm{C}]}{\partial \mathrm{x}}. \qquad \textbf{(A.1)}$$

The negative sign of (A.1) indicates that the flow goes from higher to lower concentration. One can see that the process of diffusion, therefore, occurs down the concentration gradient and it is as larger as greater this gradient's value is.

Ohm's law for drift establishes that charged particles in a system interact with its inherent electrical field, and experience a drift flux (J_{drift}), in molecules/s.cm^2, that is proportional to the magnitude of the electrical field and the medium's electrical conductivity (σ):

$$J_{\text{drift}} = \sigma \mathrm{E}. \qquad \textbf{(A.2)}$$

Since the electrical filed is related to the electrical potential (V) by $E = -\frac{\partial \mathrm{V}}{\partial \mathrm{x}}$, and, rewriting σ in terms of the mobility coefficient and the particle concentration, then equation (A.2) takes the form

$$J_{\text{drift}} = -\mu[\mathrm{C}]\frac{\partial \mathrm{V}}{\partial \mathrm{x}}. \qquad \textbf{(A.3)}$$

As in (A.1), the drift current for positive charges takes place down its concentration gradient, and it is proportional to the gradient magnitude (this time expressed as an electrical potential gradient).

The relationship between drifting and diffusion is based on Einstein's assumptions in his 1905 work on Brownian motion [Einstein, 1956]. Assuming a random walk process for the small particles, the relationship between D and μ is taken to be [Johnston and Wu, 1996]

$$D = \frac{\mathrm{kT}}{\mathrm{q}}\mu. \qquad \textbf{(A.4)}$$

In (A.4), 'k' is the Boltzmann's constant, T is the temperature in Kelvin and 'q' stands for the molecule charge. In cells, ions movement is often influenced by both drift and diffusion processes. Since the resistance presented by the biological medium to these phenomena is the same, they become additive [Johnston and Wu, 1996]. Let 'J' be the total ion fluency that accounts for both drifting and diffusion, then

$$J = J_{\text{drift}} + J_{\text{diff}}. \qquad \textbf{(A.5)}$$

Taking (A.2) and (A.3) into (A.5) yields

$$J = -\mu[\mathrm{C}]\frac{\partial \mathrm{V}}{\partial \mathrm{x}} - \mathrm{D}\frac{\partial [\mathrm{C}]}{\partial \mathrm{x}}. \qquad \textbf{(A.6)}$$

From Thermodynamics and Electromagnetism, recalling that the constant of ideal gases (R) is $R = \mathrm{k}N_A$, with $\mathrm{N_A}$ being Avogadro's constant; and that Faraday's constant (F) is $F = \mathrm{e}N_A$, with 'e' denoting the electron's elementary charge, then it is possible to rewrite equation (A.4) as

$$D = \frac{\mathrm{RT}}{N_A \mathrm{Ze}}\mu. \qquad \textbf{(A.7)}$$

In (A.7), 'Z' stands for the ion valence. Substituting the elementary charge, 'e':

$$D = \frac{\mathrm{RT}N_A}{N_A \mathrm{ZF}}\mu = \frac{\mathrm{RT}}{\mathrm{ZF}}\mu. \qquad \textbf{(A.8)}$$

Taking μ in terms of D:

$$\mu = \frac{\mathrm{ZF}}{\mathrm{RT}}\mathrm{D}. \qquad \textbf{(A.9)}$$

Taking (A.9) into (A.6) yields

$$J = -\frac{ZF}{RT}D[\mathrm{C}]\frac{\partial \mathrm{V}}{\partial \mathrm{x}} - D\frac{\partial [\mathrm{C}]}{\partial \mathrm{x}}. \qquad \textbf{(A.10)}$$

Equation (A.10) expresses the total molar flux through the cell membrane. The electric current ($\mathrm{I_M}$) associated with it is nothing more than the molar flux multiplied by $\mathrm{N_A q}$, since this last term contains information about the total number of ions passing through the membrane and its charge. But $N_A q = N_A Ze = \mathrm{ZF}$, then:

$$I_M = \mathrm{ZFJ}. \qquad \textbf{(A.11)}$$

Taking (A.10) into (A.11):

$$I_M = -\mathrm{ZF}\left(\frac{ZF}{RT}D[\mathrm{C}]\frac{\partial \mathrm{V}}{\partial \mathrm{x}} + D\frac{\partial [\mathrm{C}]}{\partial \mathrm{x}}\right). \qquad \textbf{(A.12)}$$

Assuming a constant electrical field through the cell membrane, it is possible to make the approximation $\frac{\partial \mathrm{V}}{\partial \mathrm{x}} \sim \frac{V}{L}$, in which 'L' stands for the membrane length. In doing so, equation (A.12) becomes

$$I_M = -\mathrm{ZF}\left(\frac{ZF}{RT}D[\mathrm{C}]\frac{\mathrm{V}}{\mathrm{L}} + D\frac{\partial[\mathrm{C}]}{\partial\mathrm{x}}\right). \quad \textbf{(A.13)}$$

Rearranging its terms:

$$\frac{\partial[\mathrm{C}]}{\partial\mathrm{x}} = -\frac{I_M}{DZF} - \frac{ZF}{RT}\frac{\mathrm{V}}{\mathrm{L}}[\mathrm{C}]\,. \quad \textbf{(A.14)}$$

Defining a new variable $w = -\frac{I_M}{DZF} - \frac{ZF}{RT}\frac{\mathrm{V}}{\mathrm{L}}[\mathrm{C}]$, then $dw = -\frac{ZF}{RT}\frac{\mathrm{V}}{\mathrm{L}}\mathrm{d}[\mathrm{C}]$, and equation (A.14) takes the following form:

$$-\frac{RTL}{ZFV}\frac{\partial\mathrm{w}}{\partial\mathrm{x}} = w. \quad \textbf{(A.15)}$$

(A.15) is a first order ordinary differential equation that can be solved through variables separation and simple integration. The following steps are simply calculus and algebraic operations.

$$\int_{in}^{out}\frac{\partial\mathrm{w}}{\mathrm{w}} = \int_{in}^{out} -\frac{ZFV}{RTL}\partial\mathrm{x}. \quad \textbf{(A.16)}$$

'in' and 'out' indicate that the integration should be carried over considering the physical dimensions of the cell that delimits the intra and extracellular environment.

$$\ln(w)\,|_{in}^{out} = -\frac{ZFV}{RTL}L = -\frac{ZF}{RT}\mathrm{V}\,. \quad \textbf{(A.17)}$$

$$\ln\left(\frac{w_{out}}{w_{in}}\right) = -\frac{ZF}{RT}\mathrm{V}\,. \quad \textbf{(A.18)}$$

$$\frac{w_{out}}{w_{in}} = e^{-\frac{ZF}{RT}\mathrm{V}}. \quad \textbf{(A.19)}$$

Recalling the definition of the variable 'w':

$$-\frac{I_M}{DZF} - \frac{ZF}{RT}\frac{\mathrm{V}}{\mathrm{L}}[C]_{out} = -e^{-\frac{ZF}{RT}\mathrm{V}}\left(\frac{I_M}{DZF} + \frac{ZF}{RT}\frac{\mathrm{V}}{\mathrm{L}}[C]_{in}\right). \quad \textbf{(A.20)}$$

$$\frac{I_M}{DZF}\left(-1 + e^{-\frac{ZF}{RT}\mathrm{V}}\right) = \frac{ZF}{RT}\frac{\mathrm{V}}{\mathrm{L}}\left([C]_{out} - e^{-\frac{ZF}{RT}\mathrm{V}}[C]_{in}\right). \quad \textbf{(A.21)}$$

$$I_M = \frac{Z^2F^2DV}{RTL}\frac{\left([C]_{out} - e^{-\frac{ZF}{RT}\mathrm{V}}[C]_{in}\right)}{\left(-1 + e^{-\frac{ZF}{RT}\mathrm{V}}\right)}. \tag{A.22}$$

The quantity ‘D/L’ can be defined as the membrane permeability ‘P’. It indicates how easily an ion can diffuse through its protein channel. With this, equation (A.22) takes the form

$$I_M = \frac{Z^2F^2PV}{RT}\frac{\left([C]_{out} - e^{-\frac{ZF}{RT}\mathrm{V}}[C]_{in}\right)}{\left(-1 + e^{-\frac{ZF}{RT}\mathrm{V}}\right)}. \tag{A.23}$$

Equation (A.23) expresses the current through the membrane due to ion ‘C’. Considering contributions from more ions, the total current passing through the membrane should be the sum of each ion’s own individual current.

At rest, the net membrane current is zero, since the quantities of charge flowing inward and outward are equal. Then, setting I_M to zero in (A.23):

$$P[C]_{out} - e^{-\frac{ZF}{RT}\mathrm{V}}P[C]_{in} = 0. \tag{A.24}$$

$$e^{-\frac{ZF}{RT}\mathrm{V}} = \frac{P[C]_{out}}{P[C]_{in}}. \tag{A.25}$$

$$\mathrm{V} = -\frac{RT}{ZF}\ln\left(\frac{P[C]_{out}}{P[C]_{in}}\right). \tag{A.27}$$

$$\frac{ZF}{RT}\mathrm{V} = -\ln\left(\frac{P[C]_{out}}{P[C]_{in}}\right). \tag{A.26}$$

Equation (A.27) indicates the resting potential of a certain ion ‘C’. When considering more types of ions, the expression can be generalized to

$$\mathrm{V} = \frac{RT}{ZF}\ln\left(\frac{\sum_i P_i\,[i]_{in}}{\sum_i P_i\,[i]_{out}}\right). \tag{A.28}$$

Knowing the permeability values of the ions during a certain event (such as an action potential), one could estimate the membrane potential using (A.28). This equation is known as the Goldman-Hodgkin-Katz (GHK) equation.

A.2. Hodgkin and Huxley's model for the action potential

The membrane potential can be regarded simply as the potential gradient between the intra and extracellular media. In Hodgkin and Huxley's model for the giant squid axon's membrane, the so called gating particles are responsible for controlling the opening and closing of ionic channels, thus altering the membrane's permeability value to that ion. A specific channel can be associated with more than just one type of gate particle, and the combination of the two states – opened and closed – of each one of them, then, dictates the membrane's permeability to that ion. The membrane equivalent circuit can be regarded as the one shown in Figure A.2. R_{Na} and R_K indicates the resistances values associated with the sodium and potassium ion channels, respectively. Also, I_{Na} and I_K represent the currents that flow through each one of them. R_L and I_L are the leaking resistance and current, respectively. The leaking current origins from the passive transportation of ions through gate-free channels [Ermentrout and Terman, 2010]. Just as before, E_x is the resting potential for each ion 'x', and C_M stands for the membrane's capacitance.

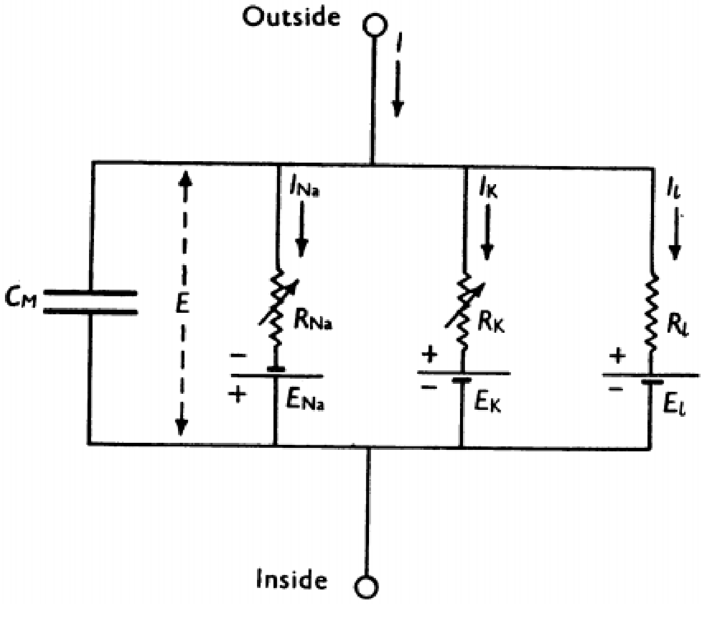


Figure A.2. Membrane equivalent circuit representation for Hodgkin and Huxley's model for the action potential. Extracted from [Hodgkin and Huxley, 1952].

In studying the mathematical membrane properties, it is usual to work with the channel's conductance (g), instead of its resistance (R), defined as:

$$g = \frac{1}{R}. \tag{A.29}$$

According to Kirchoff's law for electrical circuits, the sum of the electrical current within a closed loop must equal zero. Then:

$$I_C + I_{Na} + I_K + I_L = 0, \tag{A.30}$$

to which I_C is the current flowing through the circuit's capacitive component. (A.30) can be rewritten in terms of the ions conductance and the membrane capacity as:

$$C_M \frac{dV}{dt} + g_{Na}(V - E_{Na}) + g_K(V - E_K) + g_l(V - E_l) = 0. \tag{A.31}$$

In Hodgkin and Huxley's model, while g_l is considered to be constant, the sodium (g_{Na}) and potassium (g_K) conductance are potential-dependent, and their values would be controlled by the gating particles, since their states can alter the channel's permeability. Thus, denoting by g_i^{max} the maximum conductance value the channel for ion 'i' can assume, then:

$$g_{Na} = f_{Na}(V,t) g_{Na}^{máx}. \tag{A.32}$$

$$g_K = f_K(V,t) g_K^{máx}. \tag{A.33}$$

The $f_i(V,t)$ are merely functions dependent on the time and the membrane's potential. In this form, it becomes clear that the conductance values vary according to a given potential over time.

Ionic channels can be characterized according to two distinct states: open (O) or closed (C). The transition between these states occur due to conformational changes in the proteins caused by variations in the electrical field (and, subsequently, in the electrical potential) generated by the ions distributed around the cell's membrane. The rate to which this process develop is, then, dependent on the membrane potential. Figure A.3 illustrates these processes, in which 'a' and 'b' denote transition rates from state O to C, and C to O, respectively.

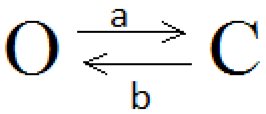


Figure A.3. Transition scheme between states O and C.

Assuming that ionic channels can only assume one between these two states, then, if a certain channel has probability ‘p’ of being in the O state, then, the probability of it being in the C state is ‘1-p’. Let N_O and N_C be the number of channels existing, respectively, in states O and C. Then, the rate according to which there is a change in the number of channels within a given state is:

$$\frac{dN_O}{dt} = bN_C - aN_O. \tag{A.34}$$

$$\frac{dN_F}{dt} = aN_A - bN_F. \tag{A.35}$$

Similar equations to (A.34) and (A.35) can be written in terms of a new quantity N = N_O + N_C by noting that the number of channels in a given state depends upon the fraction channels in the other state, the probability of their transition and the rate to which they do so. Mathematically:

$$\frac{N_O}{N} = b(1 - p). \tag{A.36}$$

$$\frac{N_C}{N} = ap. \tag{A.37}$$

The equations stated above dictate relations between fractions of open and closed channels, their transitional rates and the probability ‘p’. Note that, macroscopically, ‘p’ also represents the fraction of number of channels in the O state. Also, since ‘a’ and ‘b’ are potential dependent, alterations in the membrane potential cause the transition rates value to change accordingly. This means that the value of ‘p’ also varies over time, since it is subject to boundary conditions explicit in Equations (A.36) and (A.37). This variation is not instantaneous, and it can be expressed as:

$$\frac{dp}{dt} = \frac{N_A}{N} - \frac{N_F}{N}. \tag{A.38}$$

Putting (A.36) and (A.37) into (A.38) yields:

$$\frac{dp}{dt} = b(1-p) - ap. \tag{A.39}$$

When a stationary state is reached (that is, when transitions between states O and F are equal):

$$\frac{dp}{dt} = 0 => b(1-p_{ss}) - ap_{ss} = 0. \tag{A.40}$$

The index 'ss' is used here to indicate 'steady state'. Then:

$$b = p_{ss}(a+b); \tag{A.41}$$

and

$$p_{ss} = \frac{b}{a+b}. \tag{A.42}$$

Some algebraic steps applied to Equation (A.39) allows it to be rewritten in terms of the new quantity defined by (A.42). (A.39) can be also written as:

$$\frac{dp}{dt} = b - p(a+b). \tag{A.43}$$

Combining (A.42) and (A.43):

$$\frac{dp}{dt} = (a+b)p_{ee} - (a+b)p. \tag{A.44}$$

Or:

$$\frac{dp}{dt} = (a+b)(p_{ee} - p). \tag{A.45}$$

(A.45) is a first-order linear differential equation that can be solved by separation of variables:

$$\int \frac{dp}{p_{ee} - p} = \int (a+b)dt\,; \tag{A.46}$$

$$-\ln(p_{ee} - p) = (a+b)t + K. \tag{A.47}$$

'K' is a real constant. Proceeding with the calculations:

$$\ln(p_{ee} - p) = -[(a+b)t + K]; \qquad \textbf{(A.48)}$$

$$p_{ee} - p = e^{-(a+b)t-K}; \qquad \textbf{(A.49)}$$

$$\mathrm{p} = p_{ee} + e^{-K}e^{-(a+b)t}; \qquad \textbf{(A.50)}$$

$$p(t) = p_{ee} + He^{-(a+b)t}. \qquad \textbf{(A.51)}$$

In (A.51), 'H' denotes a new real constant. This equation expresses the variation of the probability 'p' over time. 'H' can be determined by means of some boundary conditions. Setting $p(0) = p_0$, then:

$$p_0 = p_{ee} + H => H = p_0 - p_{ee}. \qquad \textbf{(A.52)}$$

Using this value of 'H' in (A.51):

$$p(t) = p_{ee} + (p_0 - p_{ee})e^{-(a+b)t}; \qquad \textbf{(A.53)}$$

$$p(t) = p_0 + (p_0 - p_{ee})(1 - e^{-(a+b)t}). \qquad \textbf{(A.54)}$$

(A.54) describes how the probability of a channel being open over time varies. It is defined in terms of the transition rates 'a' and 'b', the initial value 'p_0' and the steady state value 'p_{ss}'.

In Hodgkin and Huxley's model for the action potential, sodium and potassium channels are dictated by more than just one gating particle. Assuming these particles to be independent from one another, the new probability 'P(t)' of the channel being open would be given by

$$P(t) = [p(t)]^x, \qquad \textbf{(A.55)}$$

in which 'x' is the number of needed gating particles of type 'p'. Note that 'p' is defined in (A.54) for the general case. Through empirical results, Hodgkin and Huxley described potassium channels as being mediated by four *n-type* gating particles, while sodium channels would be dictated by three *m-type* particles and one *h-type* particle [Hodgkin and Huxley, 1952]. *M* and *n* particles are activated by membrane depolarizations, while *h* is deactivated by this process [Johnston and Wu, 1995]. Mathematically, it is possible to write conductance values for each type of channel as function of the probabilities associated with these gating particles. The resulting expressions take forms similar to (A.32) and (A.33), just by replacing $f_{Na}(V,t)$ by m^{3}h, and $f_K(V,t)$ by n^4:

$$g_{Na} = n^4 g_{Na}^{\text{máx}}; \tag{A.56}$$

$$g_K = m^3 h g_K^{\text{máx}}. \tag{A.57}$$

Note that the proposed gating particles should behave like the 'p' variable. Thus, it is possible to write equations as (A.43) to any one of them:

$$\frac{dx}{dt} = b_x - x(a_x + b_x). \tag{A.58}$$

The variable 'x' can represent either *m*, *n* or *h*.

Each 'a' and 'b' value associated with its respective gating particle was obtained empirically by Hodgkin and Huxley, and are reproduced below:

$$\begin{gathered} b_n(V) = 0{,}01\frac{10-V}{e^{\frac{10-V}{10}}-1} \\ b_m(V) = 0{,}1\frac{25-V}{e^{\frac{25-V}{10}}-1}; \\ b_h(V) = 0{,}07e^{\frac{-V}{20}} \\ a_n(V) = 0{,}125e^{\frac{-V}{80}} \\ a_m(V) = 4e^{\frac{-V}{18}}; \\ a_h(V) = \frac{1}{1+e^{\frac{30-V}{10}}}. \end{gathered} \tag{A.59}$$

Note that expressions for the rates associated with the *h* particle are different than the ones found for the other particles, a consequence of the difference in these particle's nature [Johnston and Wu, 1995], as previously referred.

Putting it all together, the equation for the action potential generation under Hodgkin and Huxley's model actually involve the solution of a set of differential equations:

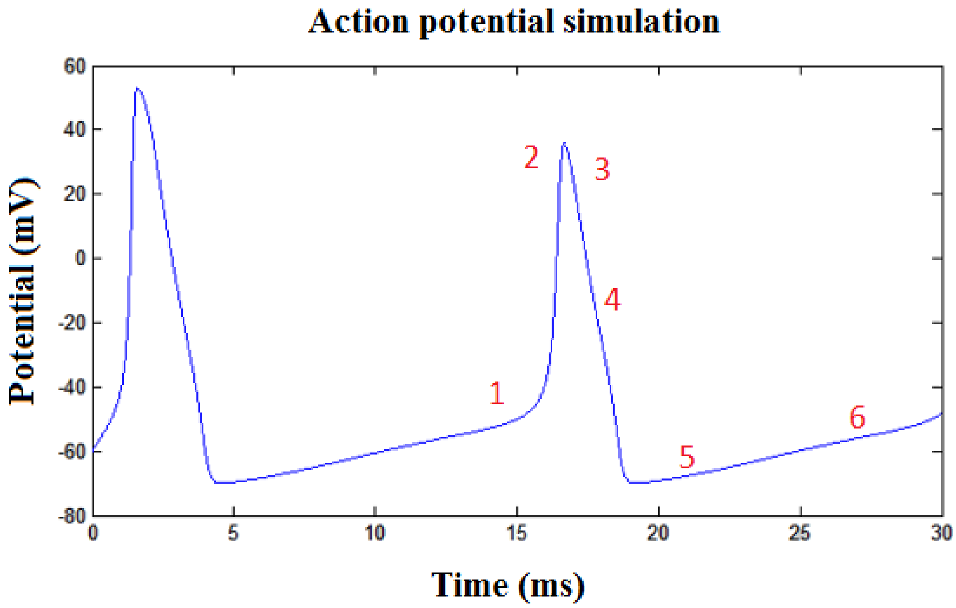
$$\begin{gathered} C_M\frac{dV}{dt} + g_{Na}^{\text{máx}} m^3 h(V - E_{Na}) + g_K^{\text{máx}} n^4(V - E_K) + g_l(V - E_l) = I_m \\ \frac{dn}{dt} = b_n(1-n) - a_n n \qquad b_n(V) = 0{,}01\frac{10-V}{e^{\frac{10-V}{10}}-1} \quad a_n(V) = 0{,}125e^{\frac{-V}{80}} \end{gathered} \tag{A.60}$$

$$\frac{dm}{dt} = b_m(1-m) - a_m m \qquad b_m(V) = 0{,}1\frac{25-V}{e^{\frac{25-V}{10}}-1} \qquad a_m(V) = 4e^{\frac{-V}{18}}$$
$$\frac{dh}{dt} = b_h(1-h) - a_h h \qquad b_h(V) = 0{,}07e^{\frac{-V}{20}} \qquad a_h(V) = \frac{1}{1+e^{\frac{30-V}{10}}}.$$

Figure A.4 displays a form obtained for the action potential when solving (A.60) using Euler's method [Howell, 2004] for numerical differential equations solving. The firing of this potential is an 'all-or-nothing' event; that is, it will occur only if the sum of the potentials received by a neuron exceeds some threshold value. The most important aspects of this phenomenon are marked in red in Figure A.4, and basically consist of the following steps:

- At '1', an incoming stimulation increases the sodium channel's conductance, making them more permeable to this ion, and depolarizing the cell membrane; thus, increasing its potential to more positive values. If this depolarization is sufficiently high, the firing of the action potential occurs;
- Sodium channels continue to open up, increasingly depolarizing the cell's membrane, up to a maximum value (region '2', in Figure A.4);
- During '3', potassium channels start to open up. Note that potassium and sodium channels are dictated by different gating particles, allowing them to present distinct properties. Thus, potassium channels take a longer time to open up than those associated with sodium, allowing membrane depolarization. The change in their conductance can be seen in Figure A.5;
- At '4', following influx of potassium ions due to the opening of their channels, the membrane repolarizes. Also, at this phase, *h* gating particles start to exercise its inactivation role to the sodium channel. This is the stage in which the membrane begins to return to its equilibrium state;
- At '5', the cell reaches a polarization stage in which its potential is beneath its resting value. The membrane is said to be hyperpolarized;
- Finally, during '6', the membrane returns to its equilibrium state, aided by the sodium-potassium bomb.
- Analysis of conductance values for the potassium and sodium ions during the action potential event can provide complementary information about it (Figure A.5). Note that sodium channels open up faster, being responsible for the cell membrane's depolarization. When sodium's conductance reaches its maximum value, almost all potassium channels are, also, open; leading to the beginning of the membrane's

repolarization. Sodium channels also close faster than potassium ones, which can be seen by the rate both curves decrease.



- **Figure A.4. Action potential form illustration.** This figure was generated by solving the set of differential equations in (A.60) using Euler's method. Some parameters were set as to match values used by Hodgkin and Huxley in their original paper [Hodgkin and Huxley, 1952]: I_m = 10 µA/cm²; C_m = 0.01 µA/cm²; Euler's method initial value for potential: -60 mV (membrane's resting potential value); temporal step of 0.01 ms.

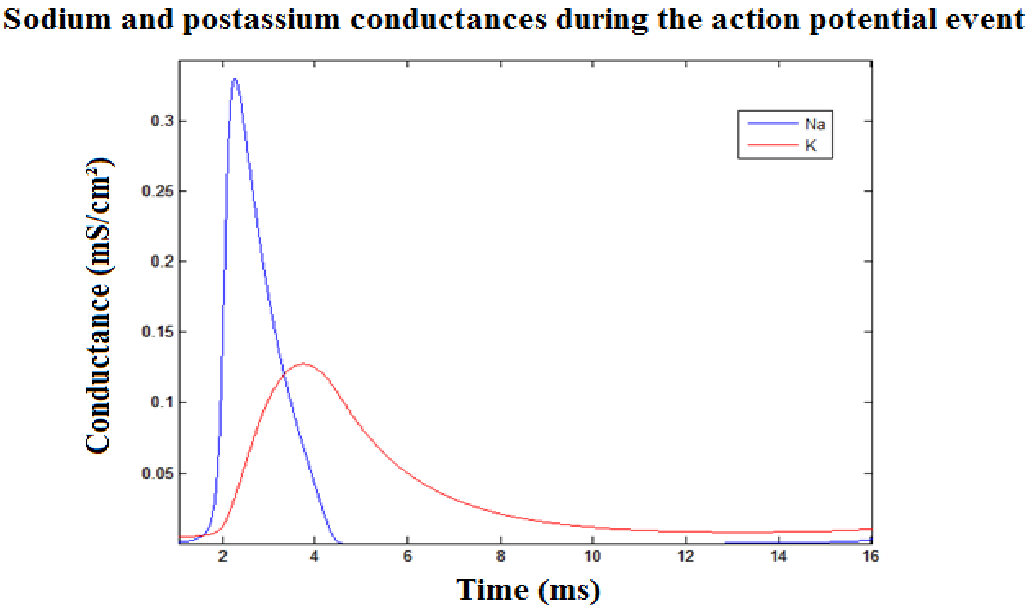


Figure A.5. Sodium and potassium conductance variation during the action potential firing. These curves were obtained using (A.56) and (A.57), and by solving expressions for the respective gating particles.

Note that the above description only covers the generation of the action potential. Its propagation through neuron’s axon involve modulations of its own. Some models, such as the so-called *core conductor model* [Johnston and Wu, 1995], treat axons as simple cylindrical conductors. Several other considerations can be made, which can substantially simplify the problem. Also, EEG signals are usually assumed to not be direct measures of the action potential per se; they are, on the other hand, considered to be mainly due to extracellular electrical fields, as descripted in Chapter 2. They, however, could not occur if not by previous action potentials.

Appendix B

Strength individual results

This appendix presents scalp distributions for the strength (S) values, for every subject. This metric was chosen because it was the one whose qualitative behavior could be best observed and reproduced along subjects (although subject 3 showed considerable variation).

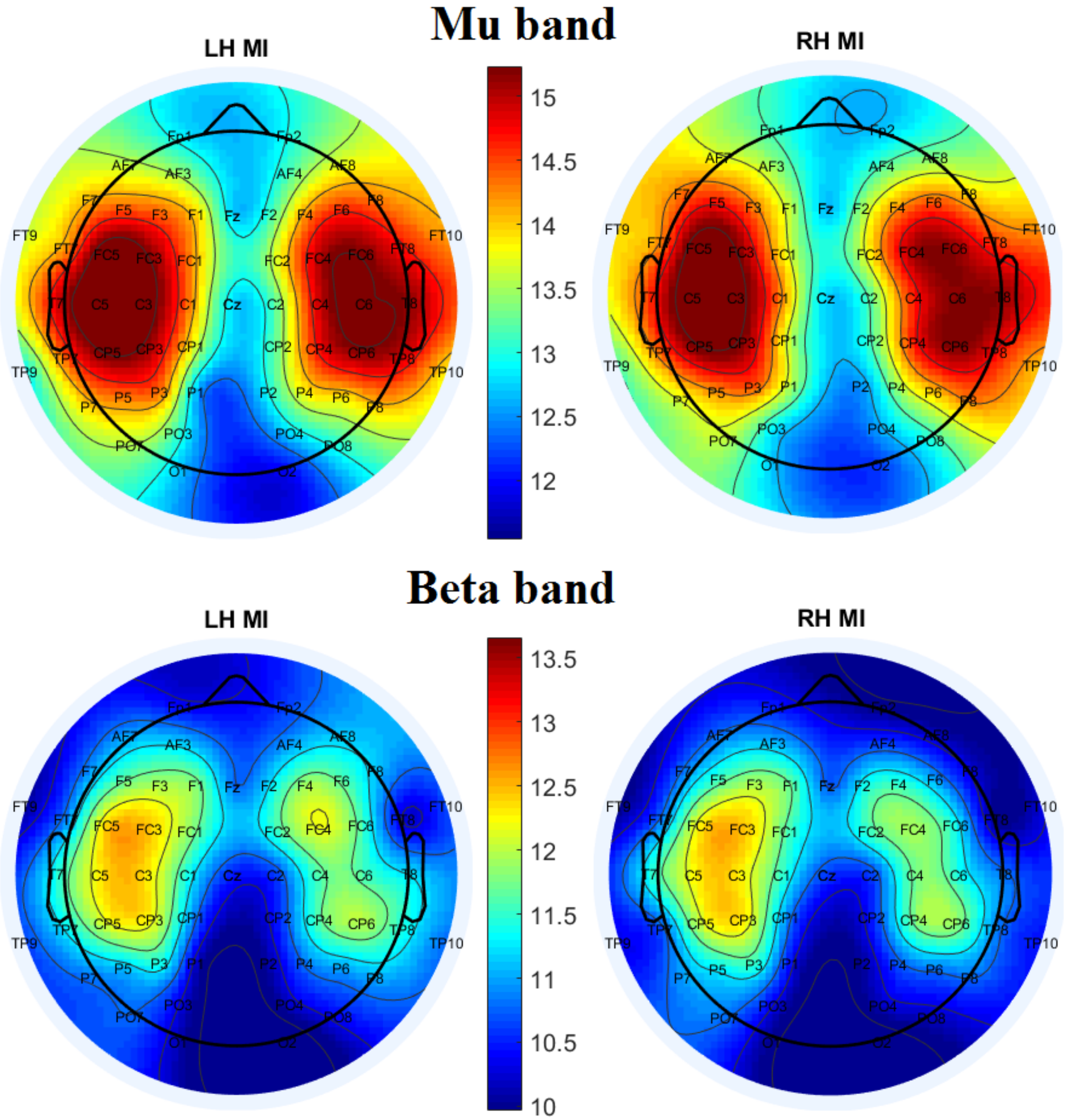


Figure B.1. Strength values distribution over the scalp (subject 1).

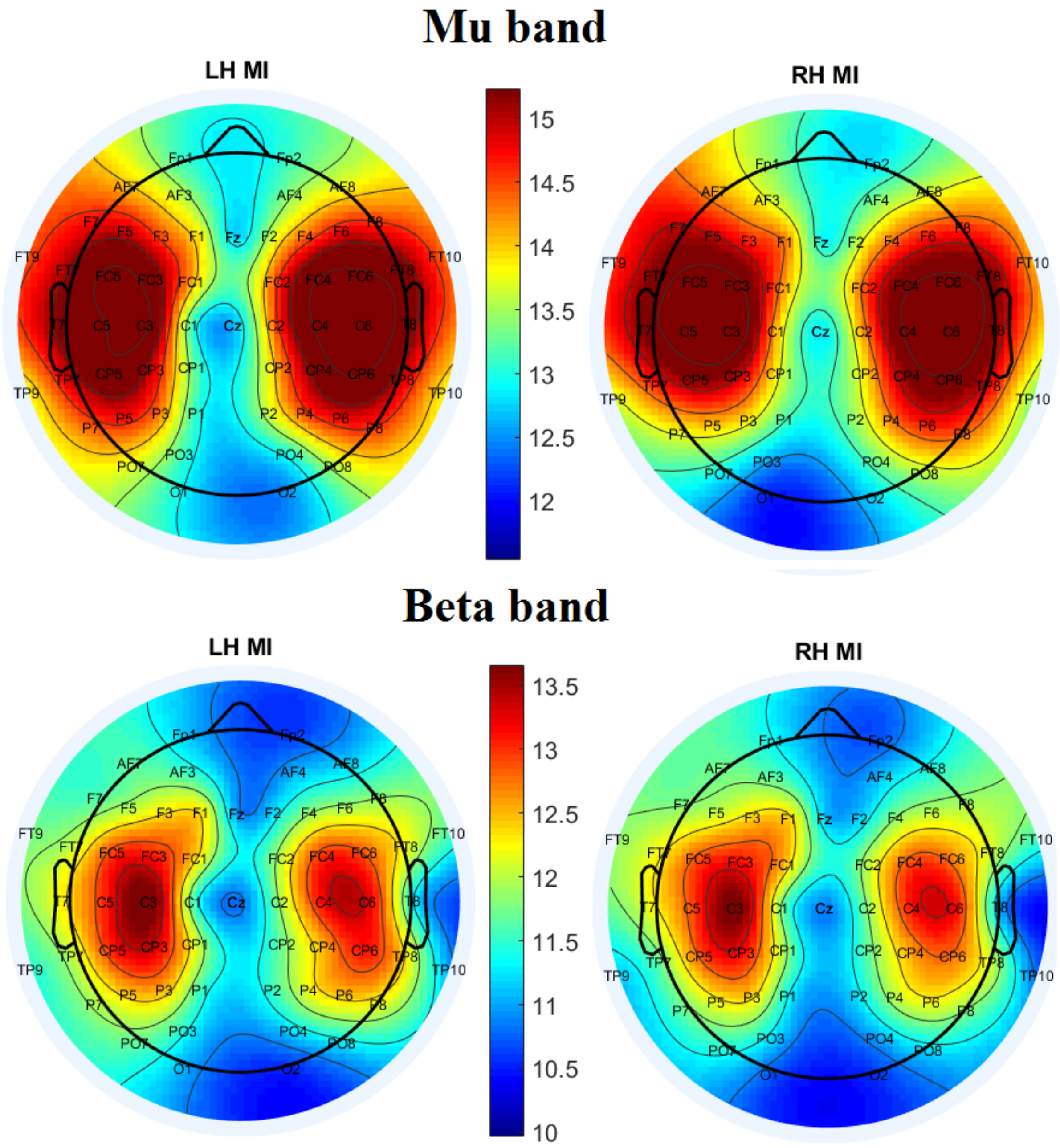


Figure B.2. Strength values distribution over the scalp (subject 2).

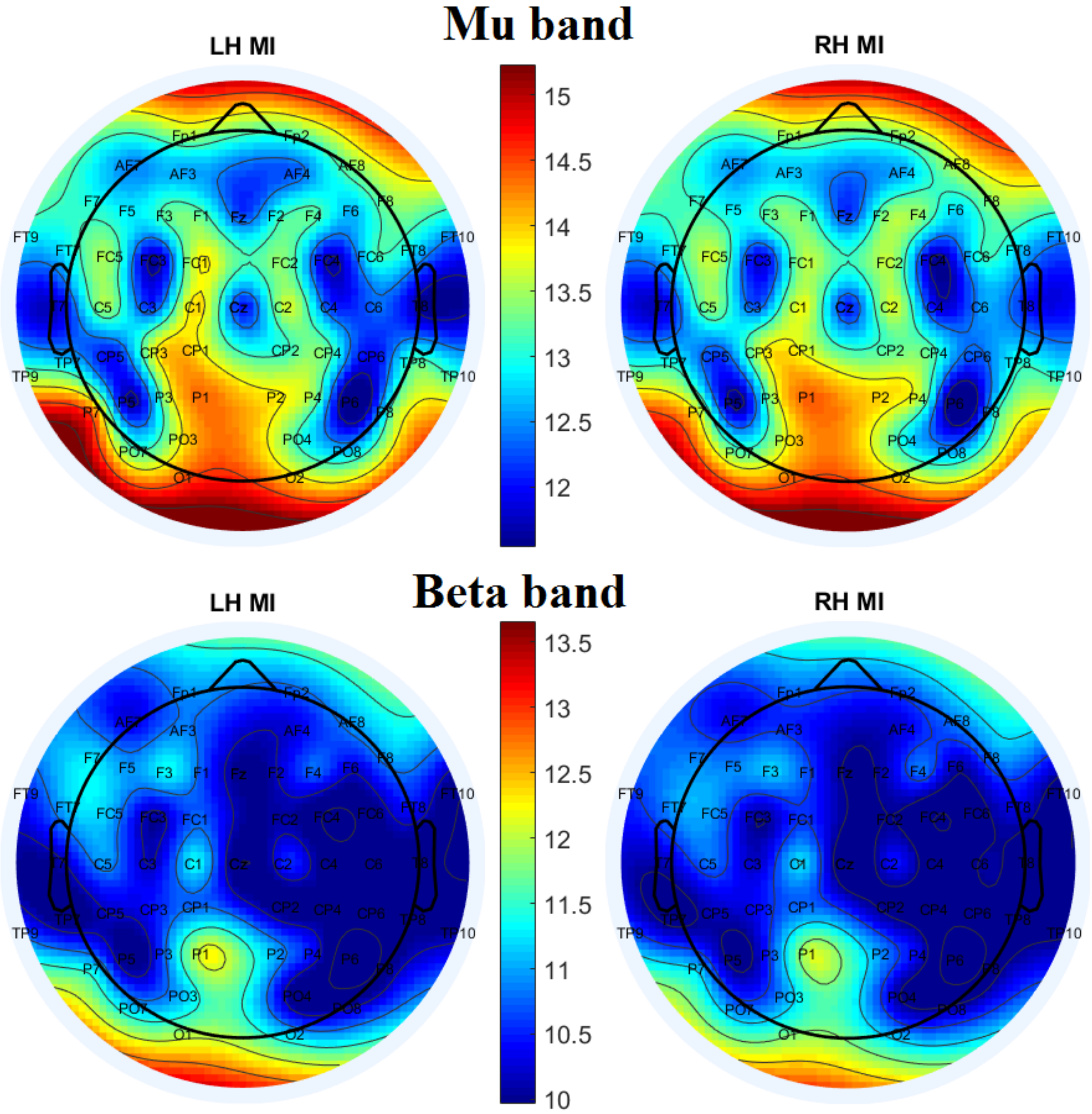


Figure B.3. Strength values distribution over the scalp (subject 3).

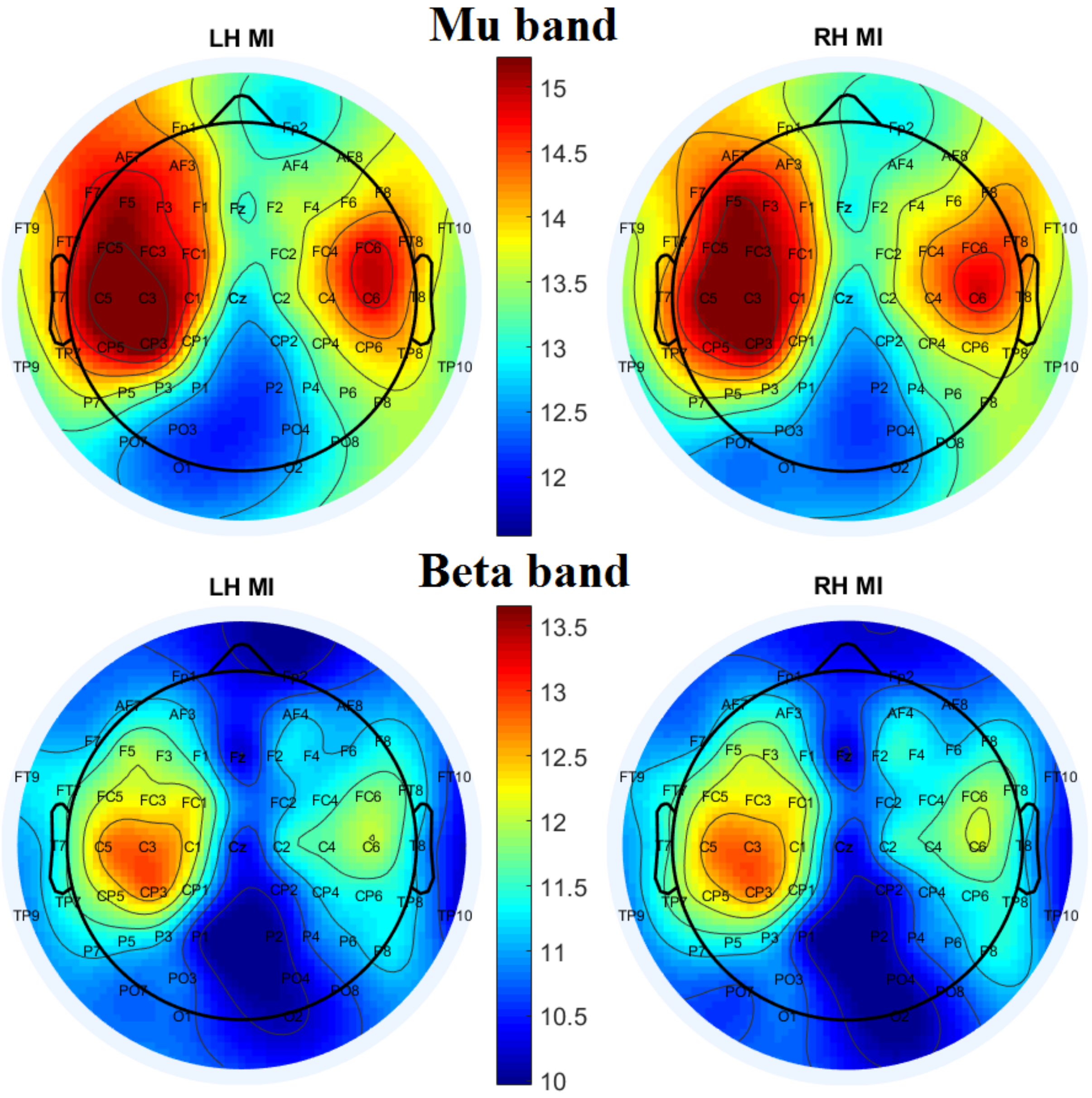


Figure B.4. Strength values distribution over the scalp (subject 4).

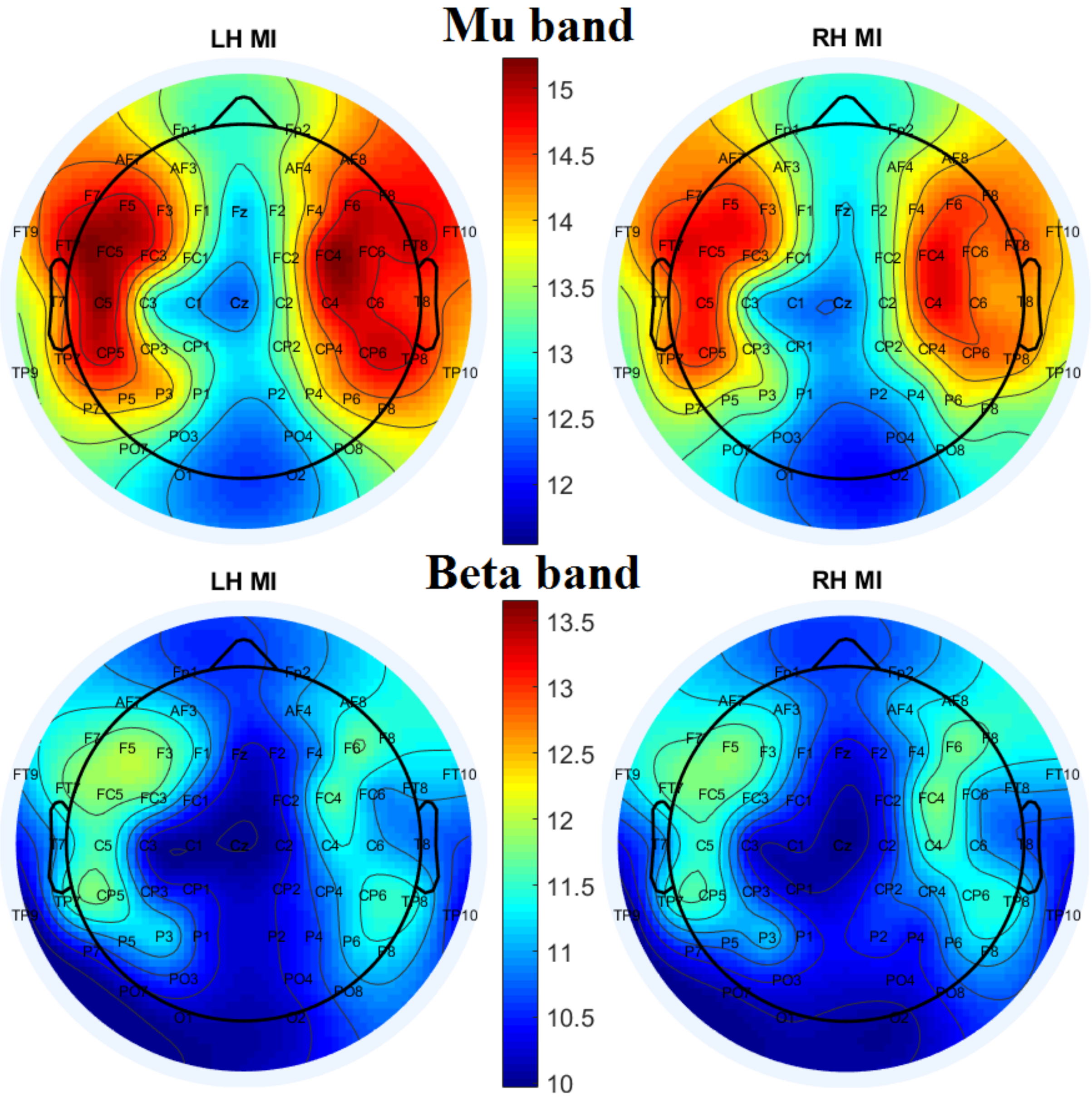


Figure B.5. Strength values distribution over the scalp (subject 5).

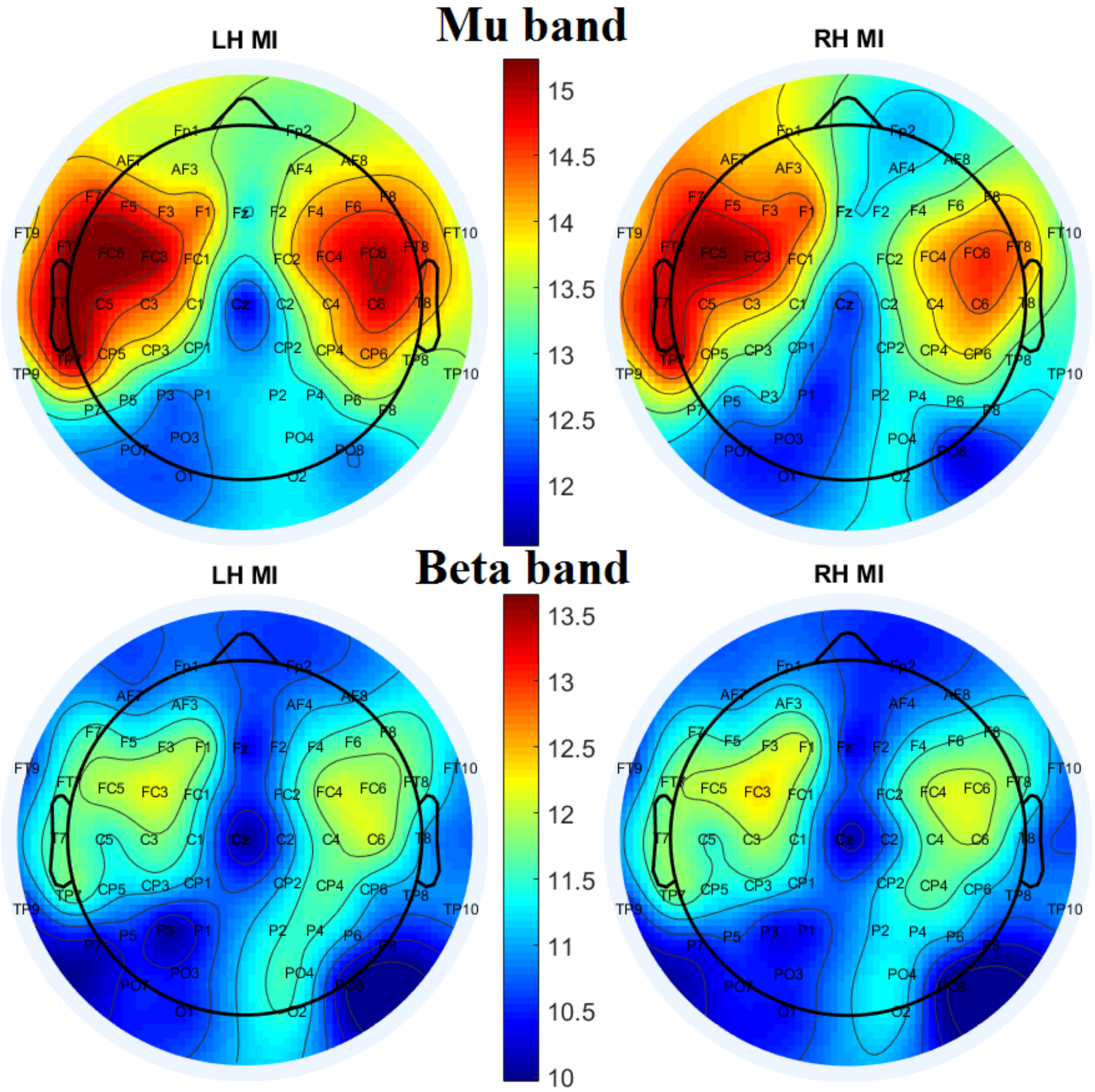


Figure B.6. Strength values distribution over the scalp (subject 6).

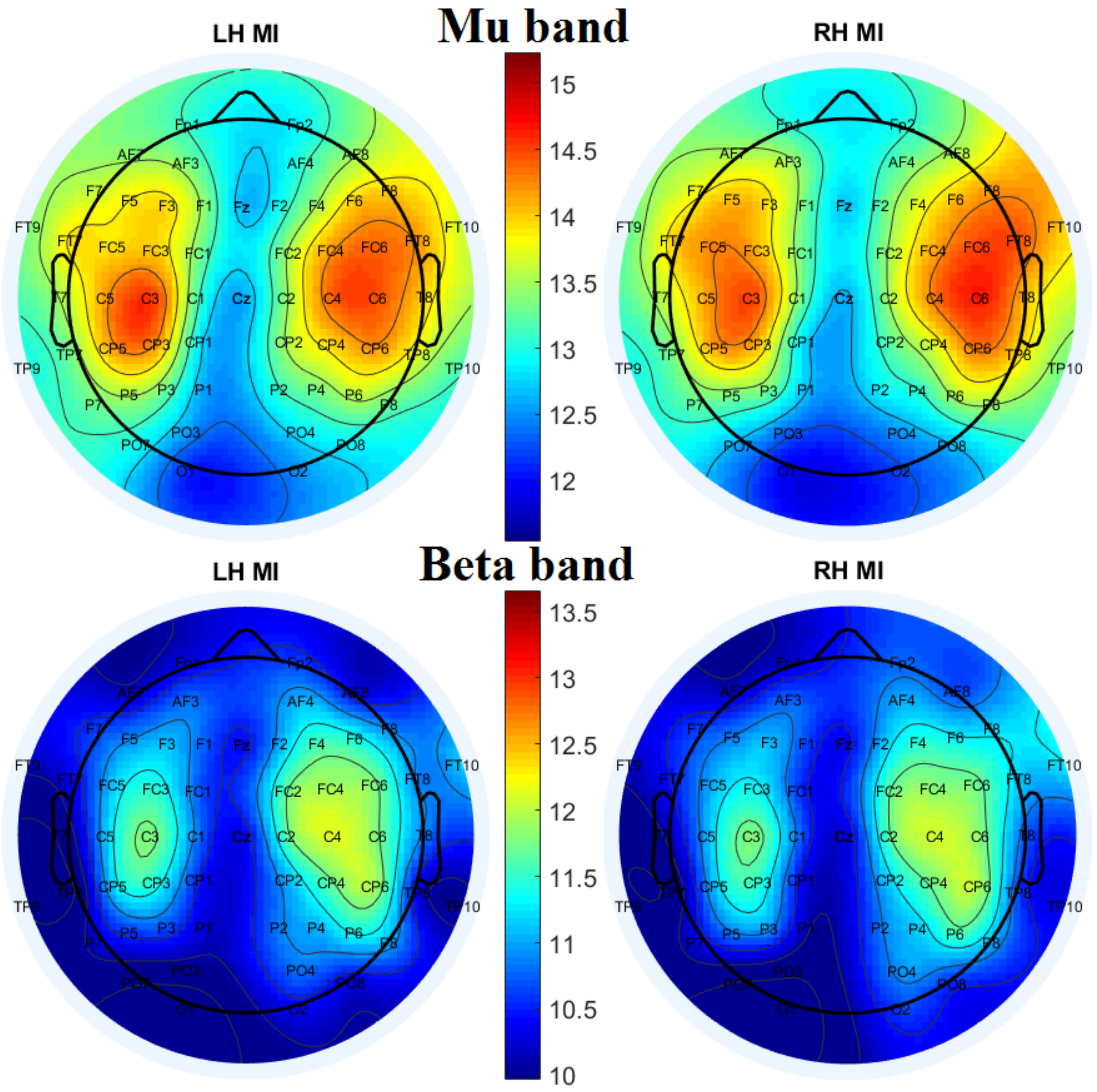


Figure B.7. Strength values distribution over the scalp (subject 7).

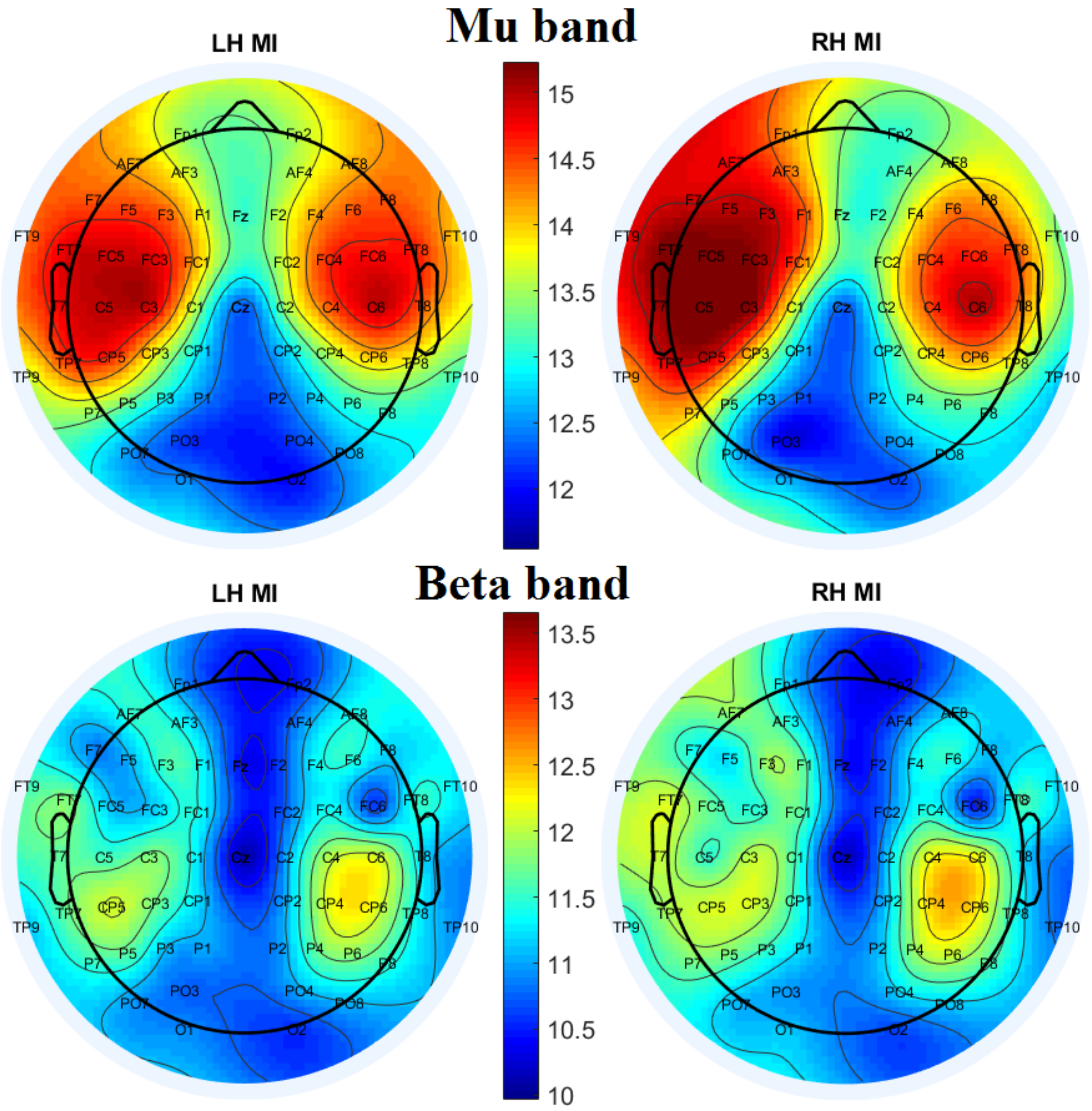


Figure B.8. Strength values distribution over the scalp (subject 8).

Appendix C

Individual classification results

This appendix presents the complete set of classification results for each subject individually.

C.1. Using all graphs nodes

In this section, tables presented below contain results when all graphs nodes are considered for classification. Tables C.1, C.2 and C.3 present classification rates for the strength, clustering coefficient and characteristic path length, respectively. Note that the three classification methods are used: LSLDA, LSVM and PSVM. Also, the effect of using just one hemisphere for classification or the use of both in the same fashion as descripted in Chapter 4 is explored. These results are averaged over both acquisitions for each subject. “RH” and “LH” stand for “right hemisphere” and “left hemisphere”, respectively, meaning that only this specific hemisphere had been used for data classification. “Diff” refers to using the difference of a metric’s value on a node on the RH minus its value on its contralateral part over the LH. Note that results are also shown for each band within the same table.

Table C.1. Individual classification results (strength, all nodes). The three classification approaches and the effect of using each hemisphere separately are shown.

	Strength – individual classification results											
	LSLDA				**LSVM**				**PSVM**			
Band	**Subject**	Classification rate (%)			**Subject**	Classification rate (%)			**Subject**	Classification rate (%)		
		RH	LH	Diff		RH	LH	Diff		RH	LH	Diff
	1	70	72	72	**1**	66	69	71	**1**	50	48	66
	2	65	63	59	**2**	64	63	61	**2**	48	51	59
	3	62	58	57	**3**	63	53	64	**3**	52	48	63
μ	**4**	66	68	70	**4**	68	77	66	**4**	48	47	61
	5	70	58	59	**5**	71	61	70	**5**	56	48	63
	6	62	61	64	**6**	68	68	66	**6**	50	53	61
	7	57	64	70	**7**	50	65	74	**7**	53	48	63
	8	59	66	63	**8**	59	63	61	**8**	50	52	56
	Subject	Classification rate (%)			**Subject**	Classification rate (%)			**Subject**	Classification rate (%)		
		RH	LH	Diff		RH	LH	Diff		RH	LH	Diff
	1	77	66	81	**1**	82	73	80	**1**	49	46	79
	2	73	67	67	**2**	74	70	66	**2**	49	49	71
	3	65	67	65	**3**	58	64	66	**3**	51	48	66
β	**4**	62	49	66	**4**	63	55	65	**4**	48	49	63
	5	70	76	76	**5**	70	72	73	**5**	57	49	70
	6	55	59	57	**6**	59	54	58	**6**	50	50	58
	7	63	66	69	**7**	65	70	68	**7**	49	48	73
	8	48	72	66	**8**	55	71	62	**8**	61	58	65

Table C.2. Individual classification results (clustering coefficient, all nodes). The three classification approaches and the effect of using each hemisphere separately are shown.

	Clustering coefficients – individual classification results											
	LSLDA				**LSVM**				**PSVM**			
Band	**Subject**	Classification rate (%)			**Subject**	Classification rate (%)			**Subject**	Classification rate (%)		
		RH	LH	Diff		RH	LH	Diff		RH	LH	Diff
μ	**1**	58	58	66	**1**	39	55	46	**1**	41	58	45
	2	59	55	54	**2**	53	56	55	**2**	55	57	55
	3	55	50	52	**3**	55	51	48	**3**	55	50	50
	4	55	63	59	**4**	52	55	51	**4**	54	52	51
	5	55	52	52	**5**	54	59	50	**5**	57	55	49
	6	53	55	63	**6**	61	52	64	**6**	63	49	63
	7	62	59	71	**7**	54	45	57	**7**	53	44	59
	8	65	57	70	**8**	54	59	57	**8**	52	59	59
	Subject	Classification rate (%)			**Subject**	Classification rate (%)			**Subject**	Classification rate (%)		
		RH	LH	Diff		RH	LH	Diff		RH	LH	Diff
β	**1**	73	70	73	**1**	68	58	62	**1**	65	56	62
	2	64	63	59	**2**	54	53	53	**2**	53	53	52
	3	63	51	59	**3**	59	57	52	**3**	58	57	51
	4	56	59	61	**4**	41	48	53	**4**	40	49	53
	5	63	57	56	**5**	57	43	50	**5**	55	42	51
	6	54	55	61	**6**	50	49	53	**6**	47	46	53
	7	52	59	56	**7**	52	50	38	**7**	49	52	38
	8	63	66	58	**8**	58	51	60	**8**	57	50	59

Table C.3. Individual classification results (characteristic path length, all nodes). The three classification approaches and the effect of using each hemisphere separately are shown.

	Characteristic path length – individual classification results											
	LSLDA				**LSVM**				**PSVM**			
Band	**Subject**	Classification rate (%)			**Subject**	Classification rate (%)			**Subject**	Classification rate (%)		
		RH	LH	Diff		RH	LH	Diff		RH	LH	Diff
	1	54	55	50	**1**	54	56	45	**1**	54	55	45
	2	54	53	52	**2**	53	54	55	**2**	47	53	52
	3	55	39	46	**3**	54	40	50	**3**	52	41	50
μ	**4**	52	60	55	**4**	53	59	53	**4**	52	63	54
	5	54	53	43	**5**	57	51	42	**5**	56	52	50
	6	57	50	58	**6**	57	48	56	**6**	55	48	57
	7	58	54	62	**7**	49	49	58	**7**	55	52	59
	8	52	57	58	**8**	50	62	55	**8**	52	62	56
	Subject	Classification rate (%)			**Subject**	Classification rate (%)			**Subject**	Classification rate (%)		
		RH	LH	Diff		RH	LH	Diff		RH	LH	Diff
	1	54	46	55	**1**	57	45	56	**1**	54	45	53
	2	56	56	40	**2**	58	60	40	**2**	59	55	45
	3	61	60	40	**3**	61	59	38	**3**	62	60	37
β	**4**	52	54	53	**4**	50	51	65	**4**	53	51	50
	5	42	58	65	**5**	45	60	67	**5**	43	57	70
	6	51	55	59	**6**	46	56	58	**6**	49	54	58
	7	56	38	58	**7**	59	38	61	**7**	59	36	55
	8	52	61	60	**8**	52	62	60	**8**	53	59	60

Tables C.4 and C.5 show individual classification results for betweenness and eigenvector centralities. Degree centrality is not shown, since it is the same as the strength, already shown in Table C.1.

Table C.4. Individual classification results (betweenness centrality, all nodes). The three classification approaches and the effect of using each hemisphere separately are shown

	Betweenness centrality – individual classification results											
	LSLDA				**LSVM**				**PSVM**			
Band	**Subject**	Classification rate (%)			**Subject**	Classification rate (%)			**Subject**	Classification rate (%)		
		RH	LH	Diff		RH	LH	Diff		RH	LH	Diff
	1	56	50	63	**1**	57	48	55	**1**	56	52	44
	2	58	48	48	**2**	52	52	48	**2**	48	59	48
	3	41	53	49	**3**	45	52	40	**3**	51	56	52
μ	**4**	59	49	54	**4**	62	52	53	**4**	55	48	49
	5	51	46	53	**5**	52	49	47	**5**	56	48	50
	6	57	57	63	**6**	58	53	60	**6**	52	59	51
	7	62	51	51	**7**	58	45	54	**7**	57	49	62
	8	58	58	59	**8**	55	59	55	**8**	54	52	59
	Subject	Classification rate (%)			**Subject**	Classification rate (%)			**Subject**	Classification rate (%)		
		RH	LH	Diff		RH	LH	Diff		RH	LH	Diff
	1	59	56	53	**1**	57	57	55	**1**	51	54	55
	2	44	55	51	**2**	45	65	55	**2**	52	56	51
	3	66	49	52	**3**	61	49	52	**3**	53	52	51
β	**4**	52	64	64	**4**	52	53	60	**4**	59	49	51
	5	49	61	56	**5**	52	58	57	**5**	50	59	55
	6	52	56	55	**6**	48	49	60	**6**	55	55	55
	7	58	63	49	**7**	60	61	52	**7**	49	52	52
	8	43	57	51	**8**	45	56	59	**8**	45	50	58

.

Table C.5. Individual classification results (eigenvector centrality, all nodes). The three classification approaches and the effect of using each hemisphere separately are shown.

	Eigenvector centrality – individual classification results											
	LSLDA				**LSVM**				**PSVM**			
Band	**Subject**	Classification rate (%)			**Subject**	Classification rate (%)			**Subject**	Classification rate (%)		
		RH	LH	Diff		RH	LH	Diff		RH	LH	Diff
μ	**1**	70	69	70	**1**	57	48	55	**1**	63	66	63
	2	66	67	63	**2**	52	52	48	**2**	54	59	59
	3	64	55	58	**3**	63	66	63	**3**	64	55	56
	4	69	72	72	**4**	55	57	59	**4**	52	59	48
	5	75	58	57	**5**	64	55	57	**5**	59	53	52
	6	62	60	63	**6**	52	59	48	**6**	60	59	60
	7	48	65	68	**7**	59	53	52	**7**	57	59	62
	8	59	60	51	**8**	57	62	63	**8**	59	60	60
	Subject	Classification rate (%)			**Subject**	Classification rate (%)			**Subject**	Classification rate (%)		
		RH	LH	Diff		RH	LH	Diff		RH	LH	Diff
β	**1**	78	71	77	**1**	74	68	73	**1**	74	68	73
	2	70	70	71	**2**	56	63	58	**2**	56	63	59
	3	64	66	63	**3**	50	63	66	**3**	50	63	66
	4	64	45	63	**4**	59	47	55	**4**	59	47	55
	5	67	78	74	**5**	63	66	61	**5**	63	66	63
	6	55	62	57	**6**	53	54	54	**6**	53	54	53
	7	59	70	72	**7**	59	64	56	**7**	59	64	56
	8	45	73	63	**8**	56	59	58	**8**	56	59	57

C.2. Using specific node pairs

By a “node pair”, it is meant that the metric’s value on a node on the right hemisphere minus its value on its contralateral part on the left hemisphere is used as input to the classifier. At this point, only the LSLDA method was tested, for reasons already stated within this dissertation (lower computational cost and better performance for most cases).

Table C.6. Individual classification results (strength, specific nodes). Classification method: difference as input to the LSLDA classifier.

Strength – individual classification results with specific nodes										
Band	**Pairs number**	**Subject**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
	1	Classification	63	65	61	61	57	65	62	63
µ	**2**	rate	74	71	74	70	73	72	76	70
	3	(%)	82	71	75	78	75	77	77	77
Band	**Pairs number**	**Subject**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
	1	Classification	74	59	64	63	64	60	63	65
β	**2**	rate	84	74	71	72	74	71	72	73
	3	(%)	85	78	77	73	79	73	78	77

Table C.7. Individual classification results (clustering coefficient, specific nodes). Classification method: difference as input to the LSLDA classifier.

Clustering coefficient – individual classification results with specific nodes										
Band	**Pairs number**	**Subject**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
	1	Classification	60	63	60	52	56	70	53	64
µ	**2**	rate	72	70	69	68	67	75	76	70
	3	(%)	76	70	70	70	68	78	78	77
Band	**Pairs number**	**Subject**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
	1	Classification	68	59	57	59	60	62	56	63
β	**2**	rate	82	69	71	67	74	70	66	70
	3	(%)	89	72	76	70	77	73	72	73

Table C.8. Individual classification results (characteristic path length, specific nodes). Classification method: difference as input to the LSLDA classifier.

Characteristic path length – individual classification results with specific nodes										
Band	**Pairs number**	**Subject**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
	1	Classification	59	59	59	63	57	63	59	61
µ	**2**	rate	76	73	72	72	70	73	74	75
	3	(%)	80	74	74	77	75	77	77	79
Band	**Pairs number**	**Subject**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
	1	Classification	70	59	59	66	63	56	61	66
β	**2**	rate	80	73	70	73	76	69	73	76

	3	(%)	83	77	77	75	81	72	77	80

Table C.9. Individual classification results (betweenness centrality, specific nodes). Classification method: difference as input to the LSLDA classifier

Betweenness centrality – individual classification results with specific nodes										
Band	**Pairs number**	**Subject**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
	1	Classification	57	53	55	55	55	55	54	53
μ	**2**	rate	70	66	67	70	69	72	70	71
	3	(%)	74	72	70	73	73	76	72	73
Band	**Pairs number**	**Subject**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
	1	Classification	55	52	55	57	54	56	55	59
β	**2**	rate	72	70	70	72	68	70	69	73
	3	(%)	74	76	73	76	73	75	73	77

Table C.10. Individual classification results (eigenvector centrality, specific nodes). Classification method: difference as input to the LSLDA classifier.

Eigenvector centrality – individual classification results with specific nodes										
Band	**Pairs number**	**Subject**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
	1	Classification	63	63	62	65	60	63	61	63
μ	**2**	rate	77	73	73	71	69	72	73	73
	3	(%)	82	77	76	77	77	76	78	77
Band	**Pairs number**	**Subject**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
	1	Classification	67	59	65	66	66	59	63	65
β	**2**	rate	79	76	72	70	74	68	72	71
	3	(%)	81	80	79	73	77	71	75	73

Appendix D

Optimum electrodes for classification

Optimum electrode pairs for data classification are shown for each subject and individually across both acquisitions. Sometimes, the same subject in the same acquisition presented multiple pairs (or combinations of) that yielded the same optimum classification rate. In these cases, pairs are separated by commas. Combinations are indicated by the sign "+": for instance, if the pairs C3/C4 and PO7/PO8 were used, then in the tables they are shown as "C3 + PO7". Only electrodes from one side (left) are shown, since the pair is always composed of the electrode and its contralateral part. Also, results for both bands are shown as separated sections – D.1 (mu band) and D.2 (beta band).

D.1. Mu band

D.1.1. Strength

Table D.1. Optimum electrodes for classification per subject and acquisition (strength, mu band).

Subject (acquisition)	Number of pairs to use for classification		
	One Pair	**Two Pairs**	**Three Pairs**
1 (1)	PO7	FT7 + F7, F7+ P7, P1 + P5	T7 + P1 + FT9
1 (2)	P3	FC1+ P1	F1 + P1 + F5
2 (1)	FP1, F7	C1 + FT9/	CP5 + C1 + FT9
2 (2)	FP1, F1, F3, Cz	P5/+ C1	CP5 + C1 + F2
3 (1)	P5	P7 + TP7	FP1 + T7 + TP7, P7 + CP5 + TP7
3 (2)	CP3, F5	C3 + FC1	C3 + T7 + FC1, C3 + FC1 + FC5
4 (1)	C3, TP6	FC1 + PO7, F1 + PO7	F7 + FC5 + PO7, T7 + FC5 + P5
4 (2)	O1, T7	FC1 + PO7	CP3 + FT7 + PO7
5 (1)	P3	T7 + TP9	T7 + P7 + TP9
5 (2)	P3, C1	F3 + AF7	F3 + P3 + AF7, P3 + F7 + Fz, F5 + AF7 + FT9, F5 + AF7 + Fz
6 (1)	T7	P3 + TP9	P3 + TP9 + P5, P3 + FT7 + TP7
6 (2)	FC1, FC5, TP9, C1, C5, FT7, Fz	CP5 + AF3, F7 + AF, FC1 + AF7, C1 + AF7, AF7 + Fz	FC5 + CP5 + AF7, P5 + AF7 + FT7
7 (1)	Fz	F1 + F5	CP1 + F1 + F5
7 (2)	P7	P3 + FT9	F3 + P3 + FT9, P3 + O1 + FT9, P3 + AF3 + FT9, P3 + CP3 + FT9, P3 + C5 + FT9, P3 + C5 + Cz
8 (1)	C1	T7 + Cz	O1 + PO3 + Cz
8 (2)	TP9	P7 + FT	P7 + AF7 + Cz

D.1.2. Clustering coefficient

Table D.2. Optimum electrodes for classification per subject and acquisition (clustering coefficient, mu band).

Subject (acquisition)	Number of pairs to use for classification		
	One Pair	**Two Pairs**	**Three Pairs**
1 (1)	PO7	AF3 + FP1	CP1 + C1 + FC3
1 (2)	F7, T7, FT7, TP7	F7 + FC3	F7 + C1 + PO7
2 (1)	F3	O1 + P7	O1 + P7 + CP5
2 (2)	P5	C1 + C3, C1 + T7	C1 + C3 + C5
3 (1)	FC3	F5 + FC5	F3 + F5 + TP9, AF3 + F7 + C1, P3 + TP9 + F5, FC5 + TP9 + CP3, TP9 + F1 + F5
3 (2)	P7, CP1, CP3, CP5, P5	P7 + CP1	FP1 + P7 + CP1, P7 + CP1 + FT9, FC1 + FC5 + CP1
4 (1)	CP1, CP3, P1, Cz	F7 + FT7	C3 + F7 + FT7, P3 + F7 + FT7, P1 + P3 + C1, O1 + P1 + CP1, O1 + F7 + FT9, O1 + P1 + F1, O1 + C1 + F1, O1 + P1 + C1, O1 + P1 + Cz, F7 + CP1 + AF7, F7 + CP5 + FT7, T7 + CP3 + FT7, F7 + FT7 + PO3
4 (2)	FC3, F5	P3 + O1, FP1 + TP9	C3 + F5 + FT9, C3 + F5 + CP1
5 (1)	P3, T7, CP5	O1 + TP7	F5 + FC5 + AF7, F5 + TP9 + AF7
5 (2)	FP1, C5, AF7	FC5 + Cz	O1 + C1 + F5, FC5 + C1 + PO7, CP3 + PO4 + FT9
6 (1)	FC3, AF7	P3 + C1	P3 + C1 + Cz
6 (2)	P1, P5	F3 + F7, F3 + FC5, FP1 + F1, FC3 + FC5, F7 + FT9, TP9 + F1, C1 + F1, C5 + F1, AF7 + F1, FT9 + F1, Fz + F1	FP1 + FC1 + AF3, F3 + FC5 + T7, F3 + FC1 + AF3, C3 + FC5 + P1, O1 + FC5 + TP7, O1 + C1 + AF3, F7 + T7 + C5, T7 + FC5 + PO3, FC1 + FC5 + C1, FC5 + PO3 + TP7, C1 + AF3 + PO3
7 (1)	FC1, FC3, C1, AF3, F5, FT7	P3 + P5	FP1 + P3 + PO7, P3 + P5 + TP7
7 (2)	P1, P5, PO3, PO7, FT7, FT9	FC1 + AF3	F7 + FC1 + AF3, FC1 + AF3 + F1, FC1 + AF3 + FT7
8 (1)	T7, FC5, TP9, FT7, FT9	P3 + F5	CP1 + F5 + Fz
8 (2)	F3, CP5, AF3, FC3, CP3	FC1 + Cz	C1 + C3 + CP5, C3 + C5 + Cz

D.1.3. Characteristic path length

Table D.3. Optimum electrodes for classification per subject and acquisition (characteristic path length, mu band).

Subject (acquisition)	**Number of pairs to use for classification**		
	One Pair	**Two Pairs**	**Three Pairs**
1 (1)	Fz	TP9 + P1	T7 + P1 + FT9
1 (2)	C3, CP1, FC5	FC1 + P1	FP1 + P1 + FT7
2 (1)	FP1, F7, T7, F5, AF7, FT9	T7 + FT9	FP1 + FC1 + C1
2 (2)	F1, Fz	CP1 + FC5	C3 + F7 + TP9, F7 + CP1 + FC5, F7 + FC1 + FC5, F1 + P5 + AF
3 (1)	P5, P7	P7 + TP7	FP1 + FC5 + TP7, FP1 + PO3 + TP7, FP1 + P5 + TP7, P7 + FC5 + TP7, P1 + P7 + TP7, P7 + PO3 + TP7
3 (2)	AF7	C3 + FC1	C3 + FC1 + CP1
4 (1)	F7	PO7 + Fz	F7 + FC5 + PO7, F1 + F7 + PO7
4 (2)	O1	PO7 + CP3, FT7 + TP7	CP3 + FT7 + PO7
5 (1)	F3, FC1, FC3, FC5, CP3, FT9	T7 + TP9	FC3 + T7 + TP9
5 (2)	FP1, T7, AF3, AF7, TP7, FT9	F5 + AF7	AF7 + C1 + F5
6 (1)	C5, TP7	P7 + AF3	P3 + P5 + TP9
6 (2)	F1, F5	FC5 + AF7, AF3 + C5	FC5 + CP5 + AF7, AF3 + CP3 + P5, FC3 + AF7 + FT7
7 (1)	P5, P7, PO7, C1, FC3	F1 + FC5, P5 + Cz	AF7 + F1 + P5
7 (2)	P5, P7, FT9, TP9, C3	O1 + Cz	FP1 + FT9 + P3, P3 + C5 + FT9. O1 + CP5 + Cz, O1 + F5 + Cz, O1 + FT7 + Cz
8 (1)	C1, TP7, FT9	CP5 + Fz, AF7 + Fz, AF3 + Cz, PO7 + Fz	P3 + PO7 + Fz, T7 + CP5 + Fz, AF3 + CP5 + Cz
8 (2)	P7	O1 + P7	P7 + TP7 + FT9

D.1.4. Betweenness centrality

Table D.4. Optimum electrodes for classification per subject and acquisition (betweenness centrality, mu band).

Subject (acquisition)	Number of pairs to use for classification		
	One Pair	Two Pairs	Three Pairs
1 (1)	C5	T7 + CP5	T7 + CP5 + FT7
1 (2)	FP1, CP5, TP9, AF3, P5, PO7, Fz	F1 + P1, P1 + P5	F3 + FC3 + P1
2 (1)	FP1, T7, F1, P1, AF3, F5, Cz	C3 + FT7, P1 + TP9, F5 + FT7	FT7 + FP1 + F5
2 (2)	C5, PO7	C3 + FT7	C3 + P1 + FT7
3 (1)	F3, F7, Fz, TP7	FC5 + PO7	F1 + FC5 + PO7, CP5 + PO7 + Cz
3 (2)	F5	CP5 + Fz	F3 + TP9 + Cz, CP3 + CP5 + Fz, CP3 + FT9 + Fz
4 (1)	F7	F7 + FP1, F7 + P1, F7 + AF3	F7 + FC3 + O1
4 (2)	FC3, CP3	FC1 + FT7, CP1 + PO7, FC5 + Cz	FC1 + FT7 + Cz
5 (1)	FT7	P7 + TP7	P5 + P7 + TP7
5 (2)	F7, C3, P3, AF7, FT9	F7 + C1, CP3 + C1, Fz + CP3	C1 + CP3 + Fz
6 (1)	F1	C5 + O1	C3 + O1 + Fz
6 (2)	PO3	CP1 + FC3, FC3 + FC5	FC3 + C1 + CP1
7 (1)	F5, C3, O1, P5, AF7, PO7	F7 + C1	F5 + FT7 + O1, AF3 + F5 + CP3
7 (2)	PO3	FC1 + C3, P7 + CP5, CP5 + Cz	F3 + FC3 + CP5, FT9 + FC1 + C3, P7 + CP5 + Cz, CP5 + P5 + Cz
8 (1)	F3, CP1, P5, PO7, Fz	F5 + F7, F7 + PO7, F5 + FC3	F7 + TP9 + P1, F5 + FC3 + Cz
8 (2)	P3, O1, TP9, PO3, P5, PO7	FC1 + TP7	FC1 + TP7 + Fz

D.1.5. Eigenvector centrality

Table D.5. Optimum electrodes for classification per subject and acquisition (eigenvector centrality, one pair, mu band).

Subject (acquisition)	Number of pairs to use for classification		
	One Pair	**Two Pairs**	**Three Pairs**
1 (1)	FC5, TP9, PO7, FT9	FP1 + P1, P1 + AF3	T7 + TP9 + F1, T7 + P1 + FT9
1 (2)	FT9	FC1 + P1	F1 + F5 + P1
2 (1)	F5	FC1 + PO7	FC1 + F1 + PO7, FC1 + C1 + PO7, FC1 + CP3 + PO7
2 (2)	TP9	T7 + TP7, CP5 + FT9	CP1 + TP9 + F1
3 (1)	P5	P7 + TP7	F7 + P7 + TP7
3 (2)	T7	C3 + CP1	C3 + T7 + FC1
4 (1)	FT9	FC5 + PO3	FP1 + FC5 + T7, T7 + FC5 + PO3,
4 (2)	P5	PO3 + PO7	FC1 + TP9 + PO7, FC1 + AF7 + PO7, FC1 + FT7 + PO7
5 (1)	P3, O1	F3 + P7, T7 + TP9	T7 + TP9 + FC3
5 (2)	O1	F5 + AF7	P3 + F5 + AF7
6 (1)	FP1	AF3 + Fz	P3 + CP1 + FT7
6 (2)	PO3	AF7 + Fz	P3 + AF7 + FT7, T7 + AF7 + P1
7 (1)	P3, PO7	F1 + PO3, F1 + P5	C3 + F1 + F5, CP1 + F1 + F5
7 (2)	P5	P3 + FT9	C3 + FT7 + FT9, C5 + Cz + P3
8 (1)	Cz	P3 + FT7	P3 + CP5 + FT7, CP5 + FC3 + Fz
8 (2)	TP9	P7 + PO3	O1 + PO3 + P7

D.2. Beta band

D.2.1. Strength

Table D.6. Optimum electrodes for classification per subject and acquisition (strength, beta band).

Subject (acquisition)	Number of pairs to use for classification		
	One Pair	**Two Pairs**	**Three Pairs**
1 (1)	FT9	F3 + Fz	F3 + F7 + Fz
1 (2)	CP5	CP5 + AF7	C3 + P5 + AF7, FC3 + P5 + AF7
2 (1)	CP5	C1 + FT9, C5 + FT9	C1 + FT9 + Cz
2 (2)	C5	C3 + TP9	F1 + F7 + TP9
3 (1)	F3, C5, P5, TP7	O1 + AF7, F7 + AF7, CP5 + AF7	O1 + F1 + AF7
3 (2)	T7	C3 + AF3, AF3 + P5	C3 + AF3 + P5
4 (1)	PO7	F1 + FC3	FC3 + FC5 + F1, F1 + FC3 + PO3, F1 + FC3 + F5, FC3 + AF7 + FT9
4 (2)	F3, TP9	FC1 + CP1, CP1 + CP3	O1 + F5 + Fz, P7 + FC1 + CP1, FC1 + CP1 + C5, FC1 + TP7 + Cz, P1 + P5 + TP7
5 (1)	CP3	C3 + Fz, F5 + CP5, Fz + Cz	C3 + CP3 + Fz
5 (2)	FP1, O1, CP3	CP1 + CP3	FP1 + FC3 + PO3
6 (1)	F1, T7	AF3 + CP3	P3 + AF3 + C5
6 (2)	C3	P3 + FC1	P3 + FT7 + Cz
7 (1)	FT9	F1 + PO3	FP1 + P3 + PO3
7 (2)	AF3	P5 + Cz	P3 + O1 + CP1
8 (1)	P7	FP1 + C3, FP1 + O1	FP1 + P3 + PO7, FC5 + PO7 + Cz
8 (2)	FP1, P7, F1, FC5, FT7, TP7	F7 + TP9, F7 + FT9	FC5 + CP5 + C5

D.2.2. Clustering coefficient

Table D.7. Optimum electrodes for classification per subject and acquisition (clustering coefficient, beta band).

Subject (acquisition)	Number of pairs to use for classification		
	One Pair	**Two Pairs**	**Three Pairs**
1 (1)	P7 + F1 + AF3, P5, PO7, FT9, Fz	F3 + F5, FC1 + FT7	FC3 + F5 + P5
1 (2)	O1, PO3, P5, PO7	FP1 + F1	FP1 + P3 + FC1, FP1 + C5 + FC1, FP1 + FC1 + PO3
2 (1)	O1, F5, F7, T7, FC5, PO7, FT9	C1 + AF7	O1 + C1 + AF7
2 (2)	P3, O1, P5, P7	P7 + PO3	P7 + AF7 + PO7
3 (1)	FP1, C1	TP7 + Cz	F3 + F5 + C5, P5 + TP7 + Cz
3 (2)	AF7	PO7 + FT9	P3 + CP1 + TP9, F7 + PO7 + FT9
4 (1)	F3, C3, P3, CP3, CP5, C1, C5, Fz	AF3 + AF7	FC1 + AF7 + Fz
4 (2)	FT9	T7 + PO7	O1 + C1 + C5, C1 + PO3 + C5, C1 + P5 + C5
5 (1)	FP1, F3, C3, P3, O1, P7, CP1, CP5, FC5, TP9, P1, AF3, FC3, CP3, PO3, C5, P5, AF7, TP7, PO7	CP1 + PO7	F3 + CP1 + PO7, F1 + F7 + C5, CP1 + FT7 + PO7, CP1 + PO7 + Fz
5 (2)	AF7	FP1 + FC1	FP1 + FC1 + F5
6 (1)	TP9	TP7 + FT9	C3 + F5 + Fz, F7 + C5 + FT9, CP3 + F5 + Fz
6 (2)	P3, FC1, FC5, AF3, F5, PO7, FT9, Fz	CP1 + CP3	F7 + CP1 + FT7, CP1 + CP3 + AF3
7 (1)	C3, TP9, F1, C5, TP7, FT9, Fz	P5 + PO7	F7 + FC1 + FC3, F7 + FC3 + TP7, CP1 + P5 + PO7, P1 + C5 + PO7, P1 + AF7 + PO7
7 (2)	O1, AF3, PO3	F3 + AF3, FC3 + FC5	C3 + F1 + AF3
8 (1)	F1, Fz	FP1 + AF7, T7 + FT9	FC1 + C5 + F5
8 (2)	FP1, F5, F7, AF7, FT7, FT9	F3 + F5, P3 + F5, FC3 + F5	P1 + P7 + FC1

D.2.3. Characteristic path length

Table D.8. Optimum electrodes for classification per subject and acquisition (characteristic path length, beta band).

Subject (acquisition)	Number of pairs to use for classification		
	One Pair	Two Pairs	Three Pairs
1 (1)	FT9	FC3 + Fz	P1 + FC3 + Fz, FC3 + P5 + Fz, FC3 + AF7 + Fz
1 (2)	F3, C3, P3, CP5, C5, FC3, F5, P5	FC3 + AF7	C3 + AF3 + AF7, FC3 + AF3 + AF7
2 (1)	C5	CP1 + FT9	FC3 + FC5 + FT9
2 (2)	FP1	F1 + TP9	TP7 + TP9 + Cz
3 (1)	AF3	CP1 + PO3, CP3 + PO3, F5 + PO3	CP5 + PO3 + AF7, CP3 + PO3 + Cz
3 (2)	CP1	CP3 + P5	C3 + CP3 + P5
4 (1)	F1, PO3, P5	FC3 + C5	CP5 + FC3 + C5, FC3 + TP7 + FT9
4 (2)	C5	P7 + TP7, F5 + TP7	F5 + F7 + FT7, P7 + F7 + FT7, FC1 + F7 + FT7
5 (1)	C3	F1 + C3, F7 + Fz, F3 + CP3, FT7 + Fz	F3 + CP3 + FT7, F3 + CP3 + Cz
5 (2)	FP1, CP3, TP7, FT9	O1 + CP1, P7 + CP1, FC3 + C5	P3 + PO3 + CP1
6 (1)	FC5, TP9, AF3, FT7	F5 + AF3	O1 + AF3 + F5, AF3 + PO3 + F5
6 (2)	P5	C3 + PO3, C3 + O1, P3 + P5, T7 + P7	P3 + FC5 + FT7, P3 + TP7 + FT7
7 (1)	P5	FP1 + PO3	F1 + PO3 + AF7
7 (2)	P3, AF3, PO3	O1 + PO3	P3 + AF3 + C5
8 (1)	F7	FC5 + PO7	T7 + FC5 + PO7
8 (2)	C1, C3, C5, P1, P3, CP1, FC3, CP3, Fz	FP1 + FT9, F7 + FT9, F7 + CP3, C5 + FT9	TP9 + P1 + AF7

D.2.4. Betweenness centrality

Table D.9. Optimum electrodes for classification per subject and acquisition (betweenness centrality, beta band).

Subject (acquisition)	Number of pairs to use for classification		
	One Pair	**Two Pairs**	**Three Pairs**
1 (1)	C3	F3 + TP7, O1 + FC3	O1 + FC3 + Fz
1 (2)	C3, F7, TP9, P1, F5, P5, AF7, FT7, TP7, PO7	AF3 + CP5	F7 + CP5 + AF3, C1 + CP5 + AF3, Fz + CP5 + AF3, AF3 + AF7 + FC3
2 (1)	FC1, CP5, P1, F5	F3 + P7, F7 + P7, T7 + P7	T7 + P7 + Cz
2 (2)	F3, F5, F7, TP7, TP9, Cz	C3 + Fz, CP1 + PO7	C3 + PO7 + Fz
3 (1)	P5	FP1 + FT9	FC1 + FT7 + FT9
3 (2)	P5	AF3 + CP5, F5 + FT7	AF3 + F1 + C1
4 (1)	O1	AF3 + F3, F3 + Cz, C1 + CP1	F3 + P5 + TP7, F3 + AF3 + Cz, CP1 + C1 + TP7
4 (2)	C3	CP1 + CP3	F3 + CP1 + CP3
5 (1)	CP5, P5	P7 + Cz, C1 + Cz	P7 + C1 + Cz
5 (2)	F5, FC1, FC5, C5, PO7, Cz	FC5 + CP3, FT9 + Cz	F5 + FC5 + PO3
6 (1)	F5, F7, C1, C3, P1, P3, P7, O1, CP5, TP9, AF3	P3 + P5, F5 + P5	F5 + P3 + P5, F5 + CP5 + P5
6 (2)	P7, Fz	F5 + FC1	F5 + FC1 + C3
7 (1)	F5	C5 + P3	P3 + C5 + FT7
7 (2)	PO3	F1 + P1, F1 + CP3, F1 + PO7	AF3 + F1 + P7
8 (1)	O1, FC5, CP5, F1, P1, C5, P5, AF7	P5 + FT9	AF7 + F7 + P7, FT9 + F7 + P7, AF7 + P5 + P7
8 (2)	FT7	FC1 + FC5	AF3 + FC1 + FC5

D.2.5. Eigenvector centrality

Table D.10. Optimum electrodes for classification per subject and acquisition (eigenvector centrality, beta band).

Subject (acquisition)	Number of pairs to use for classification		
	One Pair	**Two Pairs**	**Three Pairs**
1 (1)	F3	F3 + Fz	C3 + P1 + O1
1 (2)	P3, F5, P5	AF7 + TP9	F1 + F3 + FC5, FC5 + AF7 + TP7
2 (1)	F7, FC3, C3, FT7, FT9	C1 + FT9	FT9 + C1 + Cz
2 (2)	P5	C3 + TP9	C3 + T7 + TP9, C3 + FC5 + TP9, C3 + FC5 + C5, T7 + TP9 + PO7
3 (1)	F3, C5	FC1 + PO3, AF7 + C5	AF7 + F5 + CP1
3 (2)	FC5	AF3 + C3, CP3 + Cz	AF3 + C3 + P5
4 (1)	PO7	FC3 + P3	FC3 + C3 + O1
4 (2)	PO7	FC1 + Cz, P5 + TP7	C1 + TP7 + Cz, P5 + TP7 + Fz
5 (1)	P1	FC1 + Fz	FT9 + P7 + Fz
5 (2)	P5	CP1 + P1	FP1 + FC3 + CP1
6 (1)	T7, Fz	AF3 + C3, AF7 + F5	AF3 + C3 + P3
6 (2)	C1, TP7	FC5 + Cz	FT7 + P3 + Fz
7 (1)	TP7	P3 + PO3, T7 + TP9	F5 + CP3 + T7
7 (2)	P3	C3 + T7	F3 + CP1 + AF7, F3 + PO3 + AF7, C5 + CP1 + P3, P3 + O1 + Cz, AF7 + CP1 + O1
8 (1)	AF7	FP1 + C1, FP1 + Cz, F3 + CP5	P3 + CP5 + TP7
8 (2)	CP1	FC5 + C5	F5 + FC5 + C5

Attachment A

Informed Consent Form

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Voluntários saudáveis

DesTINe: Desenvolvimento de Tecnologia de Informação para Neurologia

Pesquisadores responsáveis: Prof. Dr. Li Li Min e Profa. Dra. Gabriela Castellano

Você está sendo convidado a participar da pesquisa “Desenvolvimento de Tecnologia de Informação para Neurologia”, que tem como objetivo geral avaliar o uso de sistemas de computadores em pacientes com lesão do cérebro para a reabilitação. Espera-se com esta pesquisa desenvolver novos sistemas de computadores e equipamentos que auxiliem pessoas com algum tipo de lesão cerebral ou paralisia para que tenham a possibilidade de se mover ou se comunicar novamente com esses equipamentos. Para isso, serão feitos experimentos também com voluntários saudáveis, como é o seu caso.

Você poderá ser submetido a três tipos de exames: ressonância magnética (RM), eletroencefalografia (EEG) e tomografia de luz próximo da faixa de infravermelho (NIRS), para avaliar as estruturas cerebrais e seu funcionamento. Você poderá ser convidado mais de uma vez para realizar esses exames para aumentar a precisão dos resultados. Além disso, você poderá ser solicitado para testar alguns equipamentos em desenvolvimento, neste caso seriam realizados os exames de EEG e/ou NIRS com o intuito de captar sinais cerebrais para comandar os equipamentos, por exemplo, comandar uma cadeira de rodas à distância, sem estar sentado nela e sem fazer esforço físico. Todos esses exames são não invasivos, e nenhum deles utiliza radiação ionizante.

No exame de RM, você entrará na máquina de RM e lá permanecerá, deitado e imóvel, por volta de 30 a 60 minutos. Durante esse tempo a máquina medirá sinais provenientes do seu cérebro e fará imagens do mesmo. O exame não causa nenhuma dor e também não possui nenhum efeito nocivo para o corpo humano – em particular, nesta pesquisa não será feito nenhum exame de RM usando contraste. No entanto, você poderá sentir desconforto devido ao grande barulho que a máquina faz (para isso me serão fornecidos tampões de ouvido), e ao fato de que você deverá permanecer o mais imóvel possível dentro da máquina, para que os sinais possam ser medidos de forma correta. Além disso, devido a que o campo magnético usado na máquina de RM é alto (como um forte ímã), caso você possua algum metal dentro do seu corpo (como pinos, marcapassos, próteses metálicas, balas de revólver), não poderá, de nenhuma maneira, realizar o exame, e NÃO DEVE DEIXAR DE AVISAR o pesquisador responsável (que estiver realizando os exames). Por esse motivo também, antes de entrar na sala da RM, você deverá remover do corpo qualquer objeto metálico, como anéis, brincos, pulseiras e óculos, etc. Também é importante que você remova objetos magnetizados como cartões de crédito, pois eles podem ser apagados ou danificados pelo campo magnético.

No caso dos exames de EEG e NIRS, uma touca com vários sensores será colocada sobre sua cabeça, para a medida dos seus sinais cerebrais. Esses exames também não possuem nenhum efeito nocivo para o corpo humano, mas você poderá sentir dor e desconforto devido ao contato apertado e prolongado dos sensores sobre o couro cabeludo, devido a que a touca deve ser colocada bem apertada para poder medir bem o sinal. Para evitar ou minimizar esse desconforto, é importante que você avise a pessoa que estiver colocando a touca em você, sobre qualquer desconforto que sentir no momento da colocação da mesma. Você também poderá, se quiser, interromper a aquisição em qualquer momento que deseje e os sensores serão retirados, sem que isso acarrete nenhum prejuízo para sua pessoa.

A sua participação nesta pesquisa não implicará em nenhum benefício pessoal, não é obrigatória e não trará riscos previsíveis. Os riscos possíveis são mínimos, relacionados aos procedimentos clínicos não invasivos, descritos acima. Você não receberá nenhum pagamento por sua participação nesta pesquisa, mas caso venha a ter despesas de transporte ou alimentação para poder realizar estes exames, elas serão ressarcidas.

Caso queira, você poderá desistir da sua participação a qualquer momento, sem que isso lhe cause prejuízo. Você será acompanhado e assistido pelo pesquisador responsável e a sua equipe durante esses procedimentos, podendo fazer perguntas sobre qualquer dúvida que apareça durante todo o estudo. Os dados coletados estarão sob o resguardo científico e o sigilo profissional, e contribuirão para o alcance dos objetivos deste trabalho e para posteriores publicações dos dados.

Para quaisquer dúvidas, você pode contactar os pesquisadores responsáveis deste projeto: o Dr. Li Li Min, no ambulatório de Neurologia (tel: 19 3521 7754, email: limin@fcm.unicamp.br, endereço: Departamento de Neurologia, Faculdade de Ciências Médicas, UNICAMP, Rua Tessália Vieira de Camargo, 126, CEP 13083-887, Campinas, SP), ou a Dra. Gabriela Castellano (tel: 19 3521 5519, email: gabriela@ifi.unicamp.br, endereço: Departamento de Raios Cósmicos e Cronologia, Instituto de Física Gleb Wataghin, UNICAMP, Rua Sérgio Buarque de Holanda, 777, CEP 13083-859, Campinas, SP).

Para denúncias ou reclamações referentes aos aspectos éticos você pode contactar o Comitê de Ética em Pesquisa da UNICAMP (tel: 3521-8936, email: cep@fcm.unicamp.br, endereço: Rua Tessália Vieira de Camargo, 126, CEP 13083-887, Campinas, SP).

Você receberá uma cópia deste termo de consentimento esclarecido.

Eu li, entendi, e aceito participar voluntariamente desta pesquisa.

____ _______

Nome legível do participante Idade

____

Assinatura do participante

____

Assinatura do responsável

____

Assinatura do pesquisador

Campinas, ______/______/20____