UNIVERSITY OF SÃO PAULO

POLYTECHNIC SCHOOL

GABRIEL CHAVES DE MELO

ELECTROENCEPHALOGRAM SIGNAL ANALYSIS FOR BRAIN-COMPUTER INTERFACE AIMING AT MOTOR REHABILITATION APPLICATION

São Paulo

March 2023

GABRIEL CHAVES DE MELO

Electroencephalogram signal analysis for brain-computer interface aiming at motor rehabilitation application

Thesis submitted to the Polytechnic School of the University of São Paulo in fulfillment of the requirements for the degree of Doctor of Science.

São Paulo

2023

GABRIEL CHAVES DE MELO

Electroencephalogram signal analysis for brain-computer interface aiming at motor rehabilitation application

Original Version

Thesis submitted to the Polytechnic School of the University of São Paulo in fulfillment of the requirements for the degree of Doctor of Science.

Concentration area: Control and Mechanical Automation Engineering

Advisor:

Prof. Dr. Arturo Forner Cordero

São Paulo

2023

Catalogação-na-publicação

Chaves de Melo, Gabriel
Electroencephalogram signal analysis for brain-computer interface aiming at motor rehabilitation application / G. Chaves de Melo -- São Paulo, 2023.
144 p.

Tese (Doutorado) - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia Mecatrônica e de Sistemas Mecânicos.

1.Interface Cérebro-Computador 2.Eletroencefalografia 3.Eletrodo de Referência 4.Reabilitação motora I.Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia Mecatrônica e de Sistemas Mecânicos II.t.

Dedico este trabalho a todos que estão com dificuldades psicológicas ou emocionais para terminarem suas teses ou dissertações.

ACKNOWLEDGMENTS

Expresso meus sinceros agradecimentos a...

Deus, que se faz presente na minha vida de tantas formas.

Minha alma gêmea, esposa, melhor amiga, parceira, apoiadora, incentivadora e porto seguro, Ana Luísa.

Minha família, pai, mãe, Paula, Ricardo, Andrea, Guilherme, Victor, Luca, Fernanda, Arthur, Leo, Mel e Pimpim.

Professor Arturo Forner Cordero, pela orientação e oportunidade de fazer esse trabalho.

Professora Gabriela Castellano, pelo incentivo, apoio e ajuda.

Professores Maria Lúcia Machado Duarte, Leandro Soares de Oliveira e Marco Antônio Meggiolaro, por me ajudarem a trilhar o caminho que me trouxe até aqui.

Camila Taira e Lucas Suplino pela amizade, parceria e apoio.

Ricardo Saeki, por todos os ensinamentos e pelo exemplo.

Clóvis de Barros Filho e Samuel Gomes, que me permitiram encontrar a alegria, nas vezes em que ela parecia sumir nos convencionalismos do ego.

“Só levo a certeza de que muito pouco sei, ou nada sei”.

Almir Sater e Renato Teixeira

ABSTRACT

de Melo, G. C. **Electroencephalogram signal analysis for brain-computer interface aiming at motor Rehabilitation application.** 2023. 144p. Ph.D. Thesis.

The neural pathway between the brain and the rest of the body might be compromised at different levels. For instance, after a stroke a person may not be able to control his/her own arm. In such cases, a motor rehabilitation therapy might enable the patient to recover complete or partial control over the paralyzed limb. However, when there is no residual control of the affected limb, rehabilitation is extremely difficult. Brain-Computer Interface is a system that records brain signals and translates them into commands to a computer or a machine. Mainly in the past 10 years, BCIs became a promising solution for these rehabilitation cases, by using auxiliary devices, such as exoskeletons, to execute a movement with the paralyzed limb once the intention to move it is detected by processing in real-time the brain signals. Non-invasive BCIs, i.e. ones that does not require surgery to implant electrodes, are usually preferred because of safety issues. In these BCIs, the electroencephalographic (EEG) technique is the most widely adopted procedure to record brain signals. EEG-based BCIs for motor rehabilitation is the focus of this work. The major challenge to develop such systems is the variability of EEG signals. This means that recording brain activity of the same person executing the same task repeatedly, the signals from different trials present considerable differences. The head volume conduction and the impossibility of recording the signals referenced to an electrically neutral point are two of the main reasons for such variability. This work aims at developing a strategy to improve the movement intention identification in pseudo real-time. First, a study concerning motor signals is conducted. Then, a method to personalize the BCI algorithm to each individual is proposed to reduce the signal variability, thus improving the accuracy of movement detection. It was seen that the proposed method was effective, reaching an average accuracy of 95% across 15 subjects.

Keywords: Brain-Computer Interface (BCI), Electroencephalogram (EEG), Reference electrode, Motor rehabilitation.

RESUMO

de Melo, G. C. **Análise de sinais de eletroencefalograma para Interface Cérebro-Computador visando aplicação em reabilitação motora.** 2023. 144p. Ph.D. Thesis.

Há casos em que o caminho do sistema nervoso que conecta o cérebro aos membros do corpo fica comprometido. Depois de um acidente vascular cerebral (AVC), por exemplo, uma pessoa pode perder a capacidade de controlar seu braço. Nesses casos, a terapia de reabilitação motora pode devolver o controle total ou parcial sobre o membro afetado. Todavia, quando não há movimento residual no membro afetado, a reabilitação é extremamente difícil. Uma Interface Cérebro-Computador (ICC) registra sinais cerebrais e os transforma em comandos para uma máquina. As ICCs tornaram-se uma alternativa promissora para a reabilitação desses casos mais difíceis. Isso ocorre por meio de um dispositivo auxiliar, como um exoesqueleto, que movimenta o membro paralisado do paciente ao detectar essa intenção por meio de processamento em tempo real dos sinais cerebrais. ICCs não invasivas são preferíveis em função de sua segurança. Nas ICCs, a eletroencefalografia (EEG) é a técnica mais empregada para registrar os sinais cerebrais. ICC baseada em EEG é o foco deste trabalho. O maior desafio desses sistemas é a variabilidade dos sinais. Isso significa que os sinais cerebrais de uma pessoa em uma dada tarefa repetida várias vezes serão diferentes em cada repetição. O volume condutor entre o cérebro e o couro cabeludo e a impossibilidade de se registrar os sinais utilizando uma referência eletricamente neutra são duas das principais razões para a variabilidade do EEG. Este trabalho tem por objetivo desenvolver uma estratégia para melhorar a identificação do movimento em pseudo tempo real. Um estudo sobre sinais motores de membros superiores é realizado. Na sequência, é proposto um método para personalizar o algoritmo de uma ICC a cada indivíduo, visando reduzir a variabilidade do EEG e, assim, melhorar a identificação pretendida. Foi visto que o método proposto foi eficaz, proporcionando uma acurácia média de 95% entre 15 sujeitos.

Palavras-chaves: Interface Cérebro-Computador (ICC); Eletroencefalograma (EEG); Eletrodo de referência; Reabilitação Motora.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

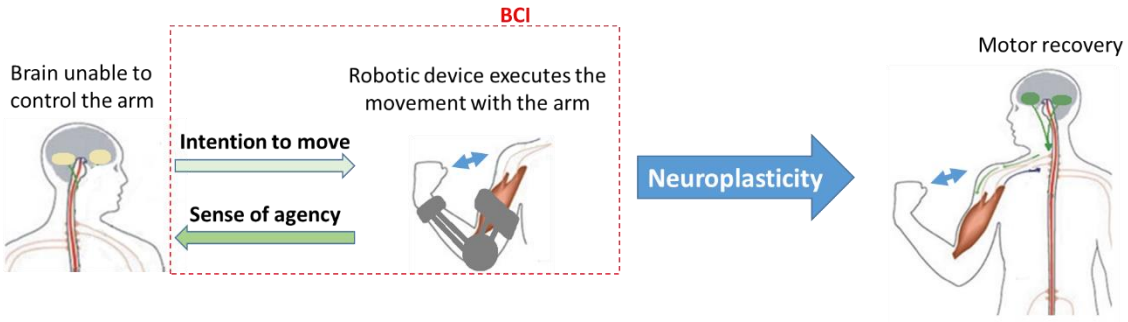
AR	Average Reference
CC	Correlation Coefficient
CNV	Contingent Negative Variation
CSP	Common Spatial Patterns
ECG	Electrocardiogram
ECoG	Electrocorticogram
EEG	Electroencephalogram
EMG	Electromyogram
ERD	Event-Related Desynchronization
ERN	Error-Related Negativity
ErrP	Error-Related Potential
ERS	Event-Related Synchronization
FLD	Fisher Linear Discriminant
fMRI	Functional Magneto Resonance Imaging
fNIRS	Functional Near Infra-Red Spectroscopy
MEG	Magnetoencephalogram
RLC	Raw Literature Channels
RM	Right Mastoid
RP	Readiness Potential
RRC	Raw Ranked Channels
SL	Surface Laplacian
TLC	Transformed Literature Channels
TRC	Transformed Raw Channels
USP	University of São Paulo

TABLE OF CONTENTS

1. INTRODUCTION

Rehabilitation is crucial to recover the quality of life for people who suffer any neuromotor impairment. There are cases in which the natural pathway between the brain and a limb becomes non-functional, making it impossible for the person to control his/her own body. In these cases, it is extremely difficult to provide assistance or rehabilitation (ORBAN et al., 2022). During the past decades, Brain-Computer Interfaces (BCIs) have become a potential solution to this problem. BCIs interpret the intention of the person and execute a certain task. It uses the neuroplasticity of the central nervous system (CNS) induced by the sense of agency to restore motor function (DALY; WOLPAW, 2008). The sense of agency is the sense that a certain action is self-generated. Figure 1 illustrates the role of a BCI in a motor rehabilitation therapy using an upper limb exoskeleton.

Figure 1 - Motor rehabilitation with a BCI.



Source: Own authorship.

Since the 1970's, when the concept of a BCI was first presented in the scientific community (VIDAL, 1973), research groups around the world have been working towards BCI's for daily living. However, this goal has not yet been accomplished, despite of the time and resources invested to reach this goal (ORBAN et al., 2022).

BCIs can be invasive (sensors implanted inside the head) or non-invasive (sensors on the scalp). The most popular type of BCI is non-invasive based on electroencephalography (EEG) signals, for practical and safety reasons (LEBEDEV; NICOLELIS, 2017). However, EEG poses an important challenge for BCI development, which is the difficulty to design algorithms that accurately

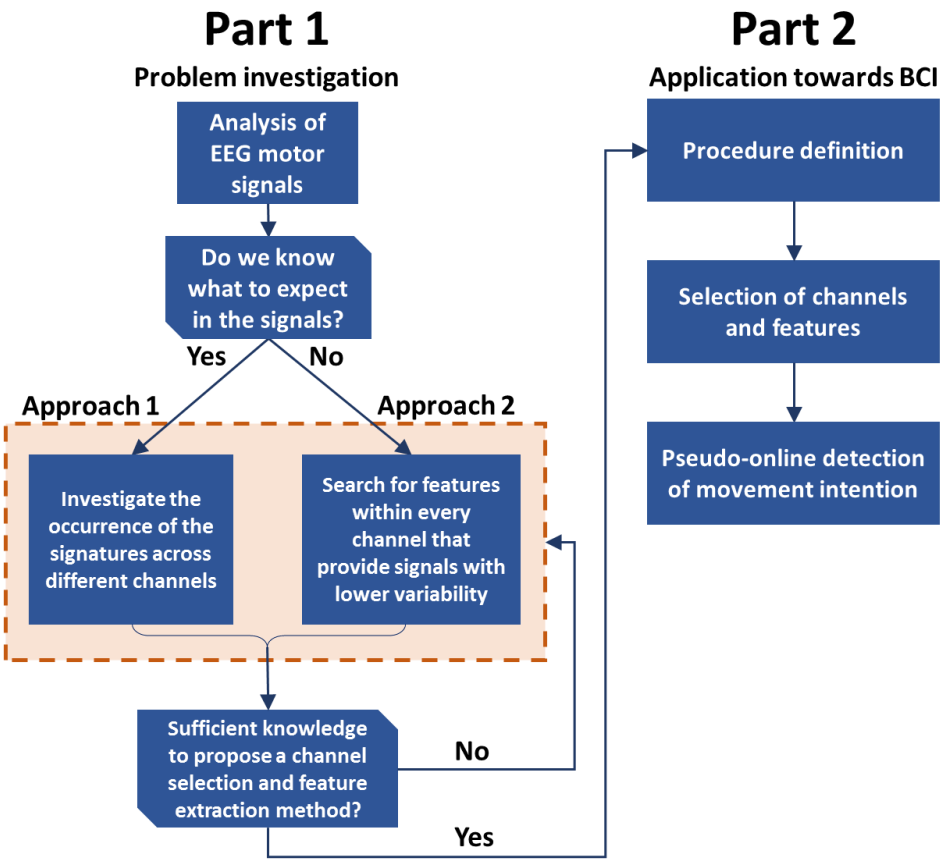
classifies mental activities. The main reason is the inter and intra-subject signal variability (LÓPEZ-LARRAZ et al., 2018; VIDAL, 1973).

Signal variability is still an unsolved issue in the BCI field and at the same time it is a critical step of a BCI system development and perhaps the most challenging obstacle in the way of functional BCIs (SAHA; BAUMERT, 2020). The present work addresses the signal variability problem aiming to improve an EEG-based BCI classification. In this work, the BCI performance is simulated offline. The major goal of this work is to build an algorithm to identify the start of the upper limb movement. The minor goal is to define a strategy to find EEG features with low intra-subject variability, which will help to achieve the major goal.

The approach is divided into two major parts, see Figure 2. The first is to perform offline analysis to find subject-specific channels with relevant information about movement intention and low intra-subject variability. The second part consists in using single trials (no averaging, so it can simulate a BCI) and proposing feature extraction, channel selection and classification techniques for the BCI algorithm. A publicly available EEG dataset recorded during the execution of motor tasks is used in this work. As a practical contribution, an experimental setup for EEG data acquisition is proposed for future work.

The rest of this thesis is organized as follows: section 2 is the literature review; section 3 presents the methodology; section 4 presents the results; section 5 presents the discussion; section 6 is the conclusion.

Figure 2 - Schematic overview of the methodology.



Source: Own authorship.

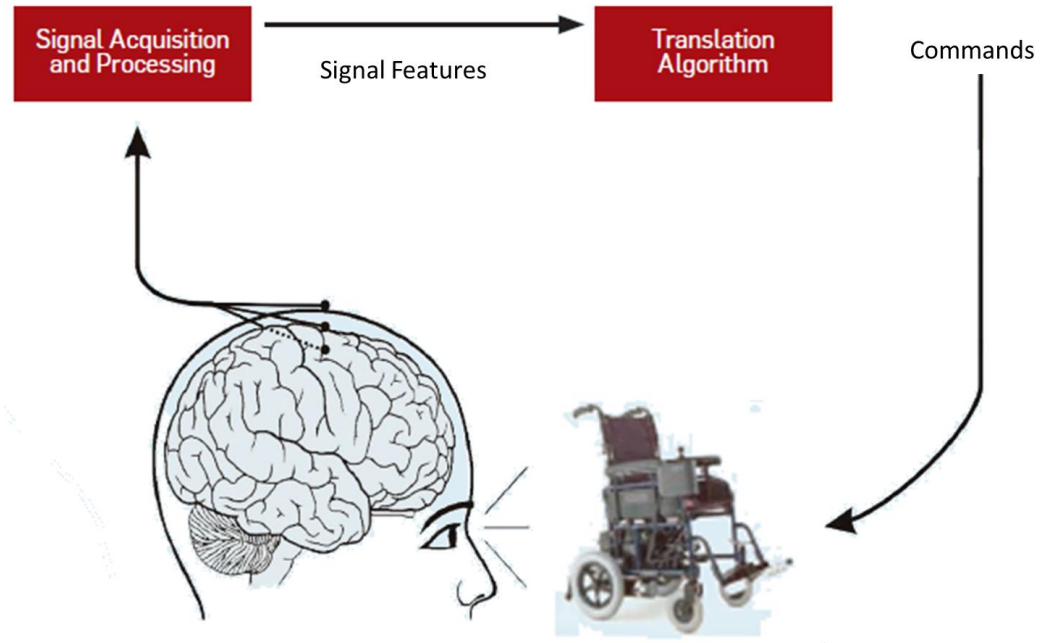
2. LITERATURE REVIEW

This section is divided into topics. They provide a literature review about the BCI field; BCIs for motor rehabilitation; EEG signals related to motor tasks; EEG technique and its challenges.

2.1. Brain-Computer Interface (BCI)

The term Brain Computer Interface (BCI) can be described as a system that acquires signals from the brain and translates them into commands to a computer or a machine in a way that the user can control part of the environment external to himself using only mental activity, as illustrated in Figure 3. Thus, a BCI system must perform a few basic tasks continuously, namely: 1) record the brain signals; 2) extract information from the signals; 3) distinguish the user's intent using the information; 4) transform the intent into commands to the machine; 5) make the machine execute the action properly. Another task that could open new scenarios is to provide adequate feedback to the user about the action, prior to the execution itself to allow for corrections.

Figure 3 - Schematic overview of a BCI system.



Source: Adapted from (MCFARLAND; WOLPAW, 2011).

There are multiple ways in which we can classify BCIs and one of them is as invasive or non-invasive BCI. The invasive type records brain signals with

implanted electrodes, which means that surgery is required to place the electrodes. The major problems are the surgery and post-surgery risks and long-term decrease in signal quality because of the possibility of electrodes changing their position with time and because of physiological reactions to the electrode. The main advantage is the signal quality once the electrodes are placed. Its high spatial and temporal resolution and signal-to-noise ratio are key for BCI performance. Such systems have already achieved great results towards the control of robotic devices only with brain activity. (NICOLAS-ALONSO; GOMEZ-GIL, 2012; RAMADAN; VASILAKOS, 2017).

The non-invasive type records brain signals with electrodes or other sensors placed on the scalp. The most used methods for non-invasive signal recording in BCI systems are electroencephalography (EEG) and functional Near Infrared Spectroscopy (fNIRS). Both of them are safe, portable and low-cost methods (WOLPAW; WOLPAW, 2012). EEG measures electric fields generated by the cortex and fNIRS measures oxygen concentration in the brain. Both phenomena reflect brain activation. EEG is the most widely used method and the main reason for that is its greater temporal resolution (in the order of milliseconds). Although the fNIRS can achieve a temporal resolution in the order of hundreds or tens of milliseconds, the phenomenon itself that is being measured (oxygen concentration) might take up to 1 or 2 s to show significant changes reflecting the cortical activation. Other techniques for non-invasive brain signal recording that are worth being mentioned, but mainly used for other applications, are magnetoencephalography (MEG), functional Magnetic Resonance Imaging (fMRI), and Positron Emission Tomography (PET) (LEBEDEV; NICOLELIS, 2017). These methods require bigger and more expensive equipment, as well as a prepared environment, thus they are not portable. PET also requires the use of tracers in the patient’s body.

A remarkable point in the history of BCIs is the investigation of “EEG biofeedback”, which consists in providing some type of feedback to the subject about their own neuronal activity (LEBEDEV; NICOLELIS, 2017). This approach made it possible for the subjects to achieve volitional control of their own EEG activity at some level. As clarified by the authors in (LEBEDEV; NICOLELIS, 2017), this type of research became very popular in the 1960’s and 1970’s. In

1973, Vidal (VIDAL, 1973) authored an article concerning what seems to be one of the first so called Brain-Computer Interface systems. In this work, Vidal says that EEG signals might be capable of being used for man-computer communication or for controlling prosthetic devices or even spaceships, and that one may suggest that such a feat could be potentially around the corner. It is interesting to note that this optimistic feeling remained present in the scientific literature over a few decades despite the very modest development throughout the years. However, Vidal remains cautious in the text towards what is to be expected in terms of the scientific development of the field, given that he exposed the fact that there was no consensus at all in many aspects of EEG signals theory and processing at the time. Perhaps the major challenge pointed out in the article is the variability of EEG signals, which represents a big obstacle for deciphering the signals and performing online analysis. Thus, Vidal advocates strongly towards the operant conditioning approach, in which the subjects are trained to provide more reliable and stable EEG patterns instead of having a complex algorithm attempting to decipher their highly non-stable patterns.

Farwell and Donchin (FARWELL; DONCHIN, 1988), 15 years after Vidal's BCI article, presented a BCI to enable communication through typing. The approach consists in using a well-known EEG evoked signal, the P300, as a label to identify which command the subject wants to select on the computer screen. This approach commonly referred to as P300 speller does not rely on the comprehension of EEG signals related to the task itself, nor on volitional control of EEG signals. Rather than that, it uses the cortical automatic response to a certain type of stimulus in order to check the subject's choice among a few presented options on the screen. Hence, although useful, this approach is limited in its applications.

Lebedev and Nicolelis (LEBEDEV; NICOLELIS, 2017) report that between the 1970's and the 1990's the BCI/BMI field did not experience much evolution mainly because of the lack of technological development in the area. However, the authors say that advances towards implantable electrodes gave new strength to the field in the 1990's, especially for invasive systems. Nevertheless, non-invasive approaches also benefit from research using invasive recordings. In 1990, in (NEAT et al., 1990) the authors described a system in which subjects

could gain some level of control over their 8-12 Hz EEG amplitude recorded in the motor cortex (Mu rhythm) by receiving a 1-dimensional cursor movement on the computer screen as a feedback. However, the system required an active operator with expertise to adjust the best time length to be analyzed for executing a command and to set the relation between Mu rhythms amplitude and cursor movement. This adaptive decoding device (ADD), as the authors named it, is used in the BCI presented by Wolpaw and colleagues (WOLPAW et al., 1991) to control a cursor with up and down movements. The authors in (PFURTSCHELLER; FLOTZINGER; KALCHER, 1993) presented a BCI system in which a classifier algorithm learned to decode whether the subject is planning to move its right or left finger in order to provide commands for a cursor control. These two BCI systems, from Wolpaw et al. (WOLPAW et al., 1991) and from Pfurtscheller et al. (1993) (PFURTSCHELLER; FLOTZINGER; KALCHER, 1993), illustrate two important approaches towards the relation between subjects and BCI systems. In the first one the subject adapts to the system's criteria for transforming EEG features into commands, while in the second one, the system adapts itself to the subject's EEG patterns for discriminating commands.

The development of robotic prosthesis and orthosis, data acquisition devices, microcontrollers, and computer software in the past three decades and these promising works in the early 1990's (among others) provided a major stimulus to the BMI field. This scenario opened a variety of possibilities for BMI applications, but till the present the transformation of EEG signals into commands by the BMI system itself has not achieved enough reliability to enable its use in the daily life of impaired individuals. As McFarland and Wolpaw (MCFARLAND; WOLPAW, 2017) state, BMIs for critical communication and control applications need to have near-perfect performance. Nevertheless, the authors highlight the reasonably novel application for these systems that does not require such perfection in order to be effective, which is neuro-rehabilitation procedures (or motor rehabilitation). It is clarified that the BMI/BCI only needs to enhance recovery of function beyond that achieved by standard rehabilitation therapies alone. Therefore, the possibility to obtain practical achievements with BCIs gain a new perspective. This might be a key factor to boost BCI development, that for

decades have seemed to be near to practical applications, but still hasn't been able to find its place outside laboratories walls.

2.2. BCI for motor rehabilitation

BCIs for motor rehabilitation purposes began to attract interest of the scientific community at the end of the first decade of this century (DALY; WOLPAW, 2008; LÓPEZ-LARRAZ et al., 2018). The most credible framework for effective motor re-learning intervention post brain injury is based on the activity-dependent central nervous system (CNS) plasticity (DALY; WOLPAW, 2008). BCI-based approaches use EEG signals to stimulate and guide CNS plasticity to improve motor function (DALY; WOLPAW, 2008). In (DALY; WOLPAW, 2008), the authors highlight two BCI-based motor learning strategies. The first is to train patients to produce more normal brain signals, which would be assessed by measured EEG features. More normal brain signals, carried out by CNS plasticity, would lead to more normal motor functions. The second strategy is to use assistive devices actuated by BCI. The improvement of motor function would produce sensory inputs that induces CNS plasticity towards normal motor control. López-Larraz and colleagues (LÓPEZ-LARRAZ et al., 2018) also suggest that the learning from both sides – human and machine – might constitute a breakthrough towards optimized motor function plasticity. This means that the BCI system should also act as an adaptive mechanism to provide an optimal and synergistic combination along with the patient.

The BCI-based rehabilitation, commonly assessed with post-stroke patients, has been shown to be safe and promising in terms of its effectiveness, in both short and long-term (BAI et al., 2020; CERVERA et al., 2018; MANSOUR et al., 2022). However, there are still a number of challenges to be addressed in the field (ORBAN et al., 2022). The neurophysiological basis to understand how BCI rehabilitation protocols contributes towards neurological improvements is still unknown and it is key to understand the correlations between BCI performances, neurophysiological changes, and functional clinical scores (GUGER et al., 2018; LÓPEZ-LARRAZ et al., 2018). There is still little consensus concerning important elements of BCI clinical trials, such as signal processing, feedback, experimental protocols and subject instructions, additional devices (e.g. exoskeletons and

others) (BAI et al., 2020; LÓPEZ-LARRAZ et al., 2018). This is a relevant problem, because many factors, such as the ones mentioned and a few others, are known to influence EEG recordings (GUGER et al., 2018). In addition, a key aspect to be mentioned is the well-known inter-subject variability concerning brain signals (specially with non-invasive recordings) and how this can be enhanced by the unique way that the stroke (or other cause) damages the brain in each subject. This reinforces the idea of BCI systems and therapies being thought as something to be personalized to each patient (LÓPEZ-LARRAZ et al., 2018).

In general, review articles highlight the inefficiency of conventional rehabilitation strategies towards motor rehabilitation for patients with severe neurological impairments, despite its cause. Also, the large number of patients in such conditions nowadays is commonly mentioned in these studies. The major motivations for investing in BCI research are of two natures: (1) practical aspects and (2) future perspectives. For the practical aspects, we can point out the growing accessibility in terms of cost, availability, portability, and easiness to use of the required hardware and fast development of new digital signal processing techniques combined with the easiness to try out different methods. In terms of future perspectives, the motivation is directly related to the promising results that has been presented in the scientific community and the lack of other promising alternatives towards the motor rehabilitation of people with severe neurological impairments. (BAI et al., 2020; CERVERA et al., 2018; DALY; WOLPAW, 2008; GUGER et al., 2018; LÓPEZ-LARRAZ et al., 2018).

2.3. EEG signals related to motor tasks

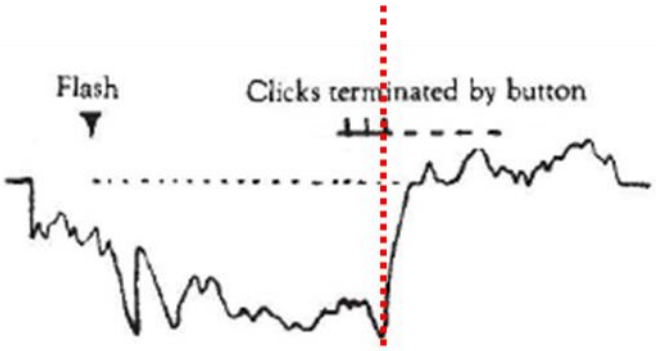
This section is divided into sub-sections addressing different signatures from EEG signals, namely: contingent negative variation (CNV), readiness potential (RP), event-related desynchronization and synchronization (ERD/ERS), event-related potentials (ERP), and error-related potentials (ErrP). These signatures are likely to be present in tasks protocol with motor activity involved.

2.3.1. Contingent Negative Variation (CNV)

This EEG signature is a slow potential (< 4 Hz) that appears as a negative shift between a warning stimulus and an imperative “go” stimulus usually no longer

than 2 s apart from each other (see Figure 4). Because of the dependency of the warning-imperative stimuli relation, the negative prolonged wave was designated as contingent negative variation (CNV) by Walter and colleagues (WALTER et al., 1964), who first reported this signal in 1964 by analyzing the vertico-mastoid EEG channel, i.e. one electrode at the vertex (Cz) and the other (reference electrode) at the mastoid. If either the warning or imperative stimulus is presented alone, or if both are presented without the need of an action by the subject, the contingent negative variation cannot be detected (WALTER et al., 1964).

Figure 4 - The CNV. The vertical red dashed line is when the movement occurred.



Source: Adapted from (WALTER et al., 1964)

Inter-stimuli intervals greater than 2 s can also elicit the CNV. In such cases, two components can be detected: initial and late CNV, designated as iCNV and lCNV, respectively (KROPP et al., 2000). The iCNV reaches its maximum amplitude between 550 ms and 750 ms after the first stimulus, while the lCNV reaches its maximum 200 ms preceding the second stimulus. A post imperative negative variation (PINV) can also be detected as a prolongation of the CNV in the interval 0-2 s after the second stimulus when analyzing CNV in intervals greater than 2 s and in the electrode Cz referenced to linked-mastoids (KROPP et al., 2000).

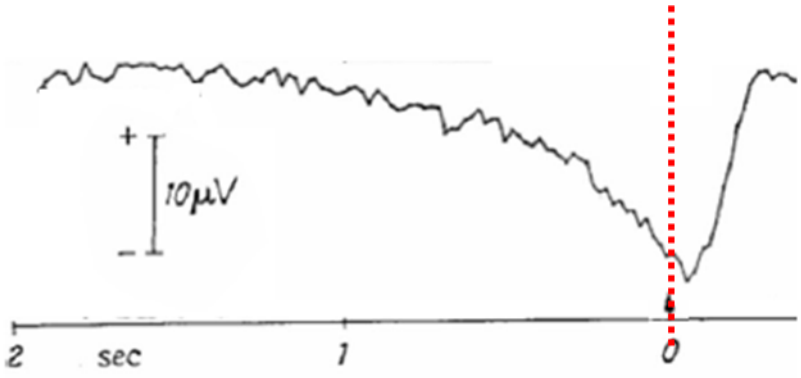
Variations of the warning-go paradigm has also been reported to elicit the CNV at electrodes all over the midline (nose to forehead direction) on the scalp, using the average reference over 128 electrodes (MASAKI; SOMMER, 2012). In (MASAKI; SOMMER, 2012), Masaki and colleagues presented an experiment in which S1 (first stimulus), presented as the appearance of a light blue circle (with a pointer at 12 o'clock position), was 2 s apart from S2. After these 2 s, the pointer initiated a clockwise movement "painting" the circle white. The onset of this

movement was considered as S2. However, subjects were instructed to press a force sensor with their finger and to achieve the peak force at the same time the pointer completed one cycle (that could take 420 or 570 ms in two experimental variations). The interval from force onset and force peak was previously defined: 100 or 200 ms, in also two experimental variations. This means that the exact movement onset by subjects were not to happen coincidently with S2, but rather be relative to it, despite the small possible time differences (220-470 ms delay). However, one can note that this paradigm maintains the subject's attention and engagement to the task, which is essential to elicit the CNV (WALTER et al., 1964).

2.3.2. Readiness Potential (RP)

The readiness potential, EEG signature originally designated as *bereitschaftspotential* (in an article written in German) in 1965 is a slow negative potential preceding self-paced movements in approximately 1.5 s, as shown in Figure 5 (KORNHUBER; DEECKE, 1965). In EEG analysis, this signal is very similar to the CNV, but has some particularities (other than its neurological meaning). For instance, when using linked-ears as reference, it is reported to rise bilaterally, widespread, and maximum over the midline at precentral and parietal region. Within 500 ms and 90 ms prior to movement onset, the slow negativity rapidly and asymmetrically increases its gradient towards the contralateral hemisphere, although it continues at the ipsilateral hemisphere until it culminates in a negative peak (SHIBASAKI et al., 1980). These two phases of BP are commonly designated as early BP and late BP, or sometimes as BP and negative slope (NS'), respectively (TARKKA; HALLETT, 1991). Shibasaki and Hallett (SHIBASAKI; HALLETT, 2006), in their review article, clarify that although BP usually starts not before 2 s preceding the movement, it may initiate *"much earlier as compared to the movement executed in more natural conditions"* depending on the experimental paradigm. The late BP, however, maintain its timing characteristics even in such cases. Authors lists many factors that are known to influence early and late BP. Among others, they include level of attention (larger early BP), preparatory state (earlier onset), learning (larger early BP), speed (later onset), and precision (larger late BP).

Figure 5 - The RP. The vertical red dashed line is when the movement occurred.

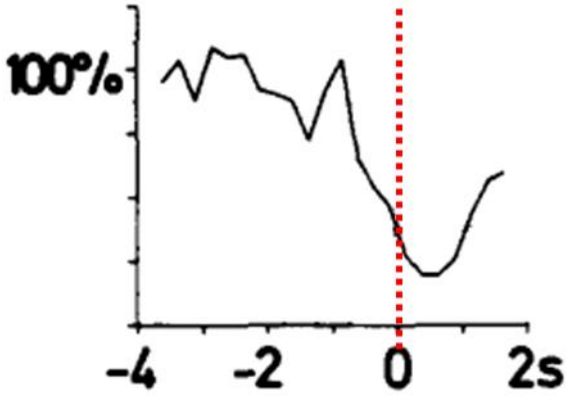


Source: Adapted from (KORNHUBER; DEECKE, 1965).

2.3.3. Event-related Desynchronization and Synchronization (ERD and ERS)

This phenomenon refers to the rhythmicity of EEG signals. As stated by Pfurtscheller and Lopes da Silva (PFURTSCHELLER; LOPES DA SILVA, 1999), it may be viewed as an event-related phase desynchronization or synchronization of neural networks oscillations within a specified frequency band. The ERD is associated with an increased activation of cortical neurons (PFURTSCHELLER, 2001). Hence, visual information processing will be characterized by ERD within the alpha band (8-12 Hz) measured over occipital region. In the same manner, motor behavior or sensorimotor activation will be characterized by ERD within the alpha band over the motor cortex, also known as Rolandic Mu rhythm, or simply Mu band. On the other hand, reduced information processing with little motor behavior will be characterized by large amplitudes of synchronized alpha band activity (PFURTSCHELLER, 2001). The ERD and ERS are seen in the EEG signals as band-specific power decrease or increase, respectively. See Figure 6 for an example of ERD.

Figure 6 - The ERD. The vertical red dashed line is when the movement occurred.



Source: Adapted from (PFURTSCHELLER; ARANIBAR, 1979).

2.3.3.1. Motor ERD

When it comes to ERD related to motor activities, Shibasaki and Hallett (SHIBASAKI; HALLETT, 2006) provide a succinct review of other works regarding lateralization evolution of ERD. With certain level of uncertainty due to the non-generalizable results throughout literature, they highlight the pre-movement contralateral occurrence of ERD within the alpha (8-12 Hz) and beta (12-25 Hz) bands, growing into a bilateral ERD over motor cortex for right hand movement by right-handed subjects. For left hand movement of right-handed subjects, ERD starts and remains bilateral. Regarding temporal aspects, motor ERD is known to be observable within the alpha and beta bands as early as 2 s prior to movement onset. This was reported by (PFURTSCHELLER; ARANIBAR, 1979) for the alpha band and in other works throughout the years. Pfurtscheller and Lopes da Silva (PFURTSCHELLER; LOPES DA SILVA, 1999) mention what is known as beta rebound within the first second after movement offset, which is characterized by a beta band ERS (i.e. increased beta band energy), while alpha band is still presenting a desynchronized behavior.

In (PFURTSCHELLER; LOPES DA SILVA, 1999), the authors draw attention to the frequency specificity concerning ERD. The alpha band, for instance, in some cases should be divided into lower (7-10 Hz) and upper (10-12 Hz) alpha bands. They explain that lower alpha ERD usually occurs in response to any type of task and is topographically widespread, whereas upper alpha ERD is usually topographically restricted. Moreover, despite commonly adopted frequency intervals for alpha and beta bands, subject-specificities must be considered in

order to define adequate frequency ranges associated with these rhythmicity activities. In this case, authors suggest the use of wavelet transform previous to ERD calculation.

2.3.3.2. ERD quantification

There are several ways to quantify ERD/ERS. Three of them are considered in the following paragraphs and they all require multiple trials of the same task, which means that they are only suitable for offline analysis. It is important to mention that ERD quantification only makes sense if there is a clear peak in the reference period (interval used as baseline) in the power spectrum (PFURTSCHELLER; LOPES DA SILVA, 1999). The ERS is only meaningful if there is also a clear peak in the power spectrum that was initially not detectable.

2.3.3.2.1. Overlapping power spectrum density (PSD)

In (PFURTSCHELLER; ARANIBAR, 1977), Pfurtscheller and Aranibar analyzed the time course of EEG power in the so called rhythmic activity within the alpha band (RAAB) with an experimental paradigm involving auditory and visual stimuli. To do so, they performed spectral analysis by means of periodograms calculated for every 1 s interval of the task with an overlap of 0.5 s. This means they first used the interval from 0 to 1 s from all trials to estimate the power spectrum density (PSD) for that particular interval of the task and obtain the energy inside it. Then they used the interval from 0.5 to 1.5 s from all trials and so on until the entire task interval was considered. The electrodes analyzed were C3 and O1 referenced to A1 (left earlobe), and C4 and O2 referenced to A2 (right earlobe).

2.3.3.2.2. Band power method

One of the most often adopted procedures was first proposed in (PFURTSCHELLER, 1977) to analyze power changes related to sensorial stimuli. It was first applied to detect power changes related to motor activities in (PFURTSCHELLER; ARANIBAR, 1979), using a bipolar reference scheme to record EEG. The scheme was a transverse electrode chain: T’4-C4-C’4-Cz-C’3-C3-T’3 (the superscript indicates a position midway between original position and the adjacent towards the vertex). However, it was slightly changed afterwards. A very clear step-by-step procedure is presented in (PFURTSCHELLER; LOPES

DA SILVA, 1999), commonly referred to as the band power method, showing how it should be applied. It consists of band-pass filtering EEG data within the frequency band of interest, followed by squaring every sample to obtain power information. The average over trials must be calculated and the resultant mean power smoothed by averaging over time. Finally, percentage values are obtained for each point in time relative to a baseline value. The baseline value is calculated as the mean power in an interval a few seconds before the event occurs. The difference between this procedure and the one proposed in (PFURTSCHELLER, 1977) is that the latter smooths the power values before averaging over trials.

2.3.3.2.3. Inter-trial variance method

The third method, presented in (KALCHER; PFURTSCHELLER, 1995), addresses a limitation of the band power method. This is the impossibility of discriminating phase-locked and non-phase-locked EEG activities when both types of activity are within the same frequency band. Thus, if a different EEG activity takes place simultaneously to motor ERD and with overlapping frequency bands, the EEG activity energy can mask EEG desynchronization or synchronization. The alternative procedure proposed by the authors consists in calculating the inter-trial variance of filtered EEG for every instant of the task, which gives the power changes over time of only non-phase-locked activities. Therefore, both methods provide power changes over time by using synchronized data from multiple trials. The difference is that the band power method calculates the mean value of power samples, and the inter-trial variance method calculates the variance of EEG samples. The former gives the total power associated with the frequency band analyzed, the latter gives the power associated only with non-phase-locked activities.

2.3.4. Event-related Potentials (ERP)

Voltage fluctuations in EEG associated with physical or mental occurrence are designated as ERPs (PICTON et al., 2000). They can be extracted from raw EEG recordings by adequately filtering and averaging. The term is usually applied to signals in time domain and characterized by negative and positive peaks related to the timing of the event. Therefore, all previously mentioned EEG signals, except for ERD/ERS can be considered to be specific types of ERPs. There are

three categories that are worth being mentioned here and do not include the previous topics: visual ERP, auditory ERP, and P300. Table 2 shows a summary regarding these three types of ERP.

Visual ERPs are expected to be larger at posterior midline regions and lateral occipital regions (LUCK, 2014). They are characterized by a positive peak with its maximum near 100 ms after the stimulus and reflect two components: C1 (with varying polarity, predominant over posterior midline, with maximum before 100 ms post stimulus) and P1 (positive peak predominant over lateral occipital areas, with maximum after 100 ms post stimulus). C1 and P1 may sum to one another if C1 is positive, or in case of C1 being negative, it will be seen as a negative peak slightly before P1 appears.

Auditory ERPs are mainly characterized by a small peak with 50 ms latency (P50), a negative peak with 100 ms latency (N100) and a larger positive peak with 160 ms latency (P160) (LUCK, 2014). Some auditory components can be recorded all over the scalp, but usually reach their maximum in specific electrode-reference montages, e.g. mastoid to mastoid or vertex to neck montage (PICTON, 2013).

The P300, or P3, was identified in 1965 as a peak at about 300 ms after visual or auditory stimuli with an electrode placed one-third of the distance along a line from the vertex to the external auditory meatus and the reference being attached to the earlobes (SUTTON et al., 1965). P3 has been related to the uncertainty of the subject with respect to the type of stimuli and its incorrect prediction. Duncan-Johnson and Donchin (DUNCAN-JOHNSON; DONCHIN, 1977) corroborated these findings with an experiment with two types of auditory stimuli. The P3 amplitude measured over the midline from frontal to occipital regions was directly related to the rarity of the stimuli and most prominent at Pz location. All EEG recordings were referenced to linked mastoids. P3 has been shown to be actually a result from two different components, P3a and P3b larger in frontal and parietal regions, respectively (SQUIRES; SQUIRES; HILLYARD, 1975). P3a was shown to be probability-dependent, but unrelated to the task itself. P3b on the other hand, although also probability-dependent, occurred only when the stimulus was important for the task. EEG recordings were made at frontal, vertex, and parietal

sites, always referenced to the right mastoid. When the designation P300 or P3 is used in literature, authors are usually referring to P3b (LUCK, 2014).

2.3.5. Error-related Potentials (ErrP)

In the BCI field many ERP components measured in response to errors are designated with the umbrella term error-related potentials (ErrP) (SPÜLER; NIETHAMMER, 2015).

Perhaps the first to be reported is the error-related negativity (ERN), in 1993 (GEHRING et al., 1993). However, what seems to be the same component had been presented two years earlier in (FALKENSTEIN et al., 1991), who referred to it as error-negativity (Ne). This component is observed when computing the difference between ERPs after incorrect and correct performances (the first minus the second). The authors in (FALKENSTEIN et al., 1991) showed, in a choice reaction task, that ERPs following correct performances present a large positive peak within -50 to 150 ms (relative to the reaction offset) in between both hemispheres. On the other hand, ERPs following incorrect performances present amplitudes severely attenuated in the same time range. A positive wave complex without a distinct peak can be detected in the range of 150 to 350 ms for ERPs following incorrect trials, meanwhile ERP after correct performance is decreasing after reaching its peak. When computing difference curves (i.e. ERP following errors – ERP following correct trials), a negative peak within 100 ms after reaction offset by the subject can be seen, followed by a positive peak at approximately 200 ms after reaction offset. In general, these effects are more clear at frontal regions.

In tasks that the subjects cannot know for themselves of their performance, thus requiring a feedback, a wave similar to the ERN, but with larger latency, can be detected (MILTNER; BRAUN; COLES, 1997). In midline electrodes referenced to linked-ears, a negative peak in the ERP after incorrect feedback at approximately 250 ms post-feedback can be detected, whereas ERP after “correct” feedback presents a positive peak. After this peak, ERP after “incorrect” feedback presents what might be the same positive peak seen previously following “correct” feedback. Therefore, the difference wave presents a large negativity (250 ms following incorrect feedback) and a positivity in its sequence

(MILTNER; BRAUN; COLES, 1997). Hoyrold and colleagues (HOLROYD et al., 2003) showed that the wrong prediction concerning the feedback has impact on the ERN. Authors showed that after negative feedback the ERN has greater amplitude when the subject expects a positive feedback when compared to when he expects a negative feedback.

The study concerning ErrPs allows the development of theories concerning error processing by the human brain. Krigolson and Hoyrold (KRIGOLSON; HOLROYD, 2006) proposed a hierarchical model for error processing. The model considers that high-level errors (goal attainment) and low-level errors (e.g. trajectory deviations during motor control) are assessed by frontal and posterior error systems, respectively. The former is the typical response ERN observed in earlier studies (after subject's response at a choice reaction task), and the latter was found to also present a negative peak in the difference curve, but with a larger latency, with its peak in the 100-150 ms interval (meanwhile the former case presented the negative peak before 100 ms). Moreover, high-level ERN was shown to be more evident at frontal region and low-level ERN at posterior region.

When it comes to BCIs, it is possible to explore ErrPs to understand the particularities of error processing in this scenario. Usually, related works analyzes error potentials associated with wrong outcomes by the BCI, i.e. the subject delivers the correct command, but the algorithm processes the wrong command (FERREZ; MILLÁN, 2008; SCHALK et al., 2000; SPÜLER; NIETHAMMER, 2015). For instance, Ferrez and Millán (FERREZ; MILLÁN, 2008) revealed different components in the difference wave calculated at FCz with EEG recordings referenced to the average of 64 electrodes. They called this new ErrP as "interaction ErrP". It was mainly characterized by a sequence of four peaks relative to the feedback: small positive (200 ms), big negative (250 ms), big positive (320 ms), and a broader negative (450 ms).

The major significance of ErrP towards BCI consists on improving the algorithms by enabling adaptive procedures to re-calibrate the algorithm or to enhance the sense of agency effect. This would be achieved by updating in real-time their EEG pattern recognition, adapting to the user and by stopping a wrong outcome once the user is informed about it. However, this approach is still not well

established in the field. Main reasons might be the difficulty to detect ErrP in real-time.

2.3.6. Summary

Tables 1, 2, and 3 show a summary of the main features concerning spatio-temporal aspects of ERD, CNV, and RP (Table 1), ERPs (Table 2), and ErrPs (Table 3). As already seen in the present literature review, it is hard to generalize these features, reason why the content in these tables should be considered cautiously. The intent is to provide a succinct and superficial comparison between the mentioned EEG signals, rather than providing robust information individually for each signal type.

Table 1 - Comparison between main spatio-temporal properties of ERD, CNV, and RP.

	ERD (alpha band)	CNV	RP
When	Prior (<2 s) and (2) during the movement	Between warning-go stimuli (<2 s)	Before self-paced movement: (1) early (<2 s) and (2) late RP (<0.5 s)
Where	Motor cortex (1) contralateral and (2) bilateral	z-line, most prominent over the central region	Central region close to z-line (1) Bilateral and (2) contralateral
How it looks	Energy decrease in frequency bands	Slow decrease in amplitude	Slow decrease in amplitude and (2) abrupt decrease

Source: Own authorship.

Table 2 - Comparison between main spatio-temporal properties of previously discussed ERPs.

	Auditory	Visual	P300 (3b)
When	Approximately (1) 50 ms, (2) 100 ms, and (3) 160 ms post-stimulus	Approximately 100 ms post stimulus	Approximately 300 ms post stimulus
Where	Mainly at mastoids	Lateral occipital regions	Most prominent at Pz location
How it looks	Small positive peak, (2) negative peak, and (3) a final positive peak	Positive peak	Positive peak

Source: Own authorship.

Table 3 - Comparison between main spatio-temporal properties between different ErrPs. The “r”, “f”, and “i” preceding “ErrP” refers to “response”, “feedback”, and “interaction”. Although the first two were previously referred as ERN instead of ErrP, the latter is an umbrella term that includes all error-related potentials.

	rErrP	fErrP	iErrP	Low-level rErrP
When	Post-reaction: (1) prior to 100 ms and (2) 200 ms	Post-feedback: (1) 250 ms and (2) prior to 500 ms	Post-feedback: (1) 200 ms, (2) 250 ms, (3) 320 ms, and (4) 450 ms	Post-error: between 100 and 150 ms

Where	Frontal regions, z-line	Frontal regions, z-line	Frontal regions, z-line	Posterior regions, z-line
How it looks	(1) Negative and (2) positive peaks	(1) Large negative peak and (2) positive peak	(1) Small positive, (2) big negative, (3) big positive, (4) broad negative peaks	Negative peaks

Source: Own authorship.

2.4. EEG technique and its challenges

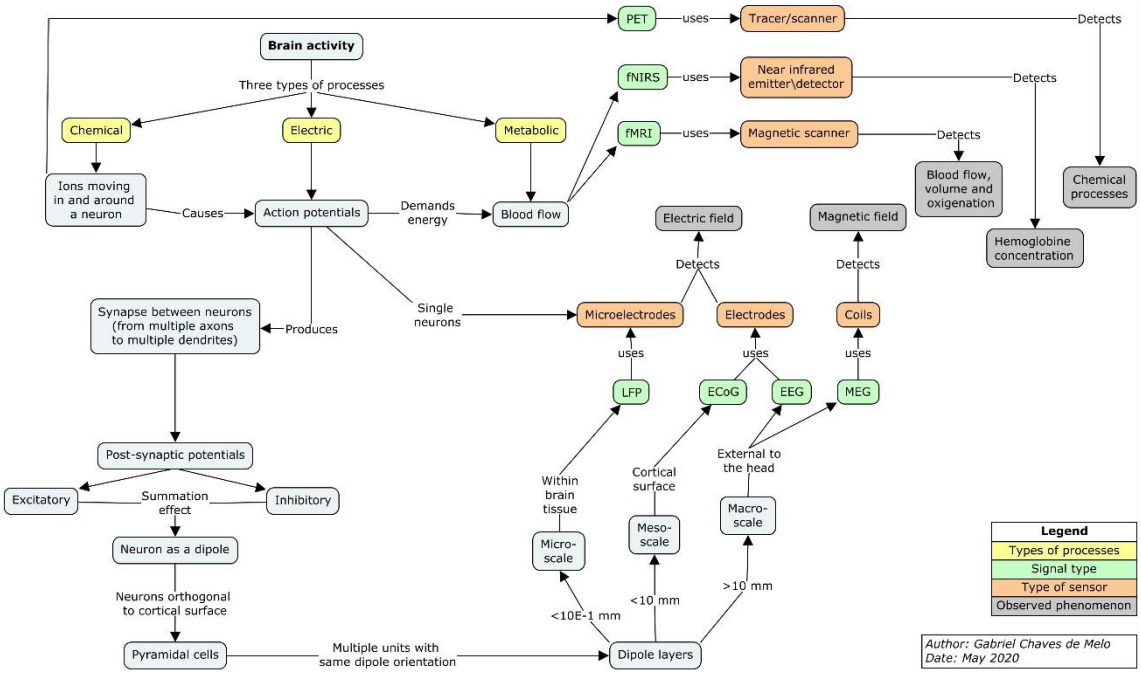
This section is divided into three sub-topics addressing, each one of them, brain signals, the reference electrode, and spatial resolution.

2.4.1. Brain signals

The electric field of a neuron is a result of the summation of multiple post-synaptic potentials in its axons, which can be excitatory or inhibitory. When a large number of neurons (up to 10^9) is displayed in layers aligned perpendicular to the cortical surface and has similar electric field orientation, they actuate as a dipole capable of having its electric field measured by electrodes on the scalp. These neurons that together generate EEG signals are called pyramidal cells because of their morphology. Between the cortex and the scalp there is a conductive medium in which the electric field generated by different dipole layers spreads and mixes into each other. Therefore, what reaches the scalp to be recorded by the EEG electrodes is a space average of the activities from different cortical areas whose electric field has superimposed each other. Moreover, because of different tissues with different resistivity properties, the conductive medium also actuates as a low-pass filter, thus EEG content is usually under 40 Hz. Important to note that in this work EEG always refers to the non-invasive EEG. Electrocorticogram (ECoG), on the other hand, which uses electrodes placed invasively on the

cortex, has a frequency spectrum up to 200 Hz. (WOLPAW; WOLPAW, 2012). Figure 7 illustrates schematically different signals that are generated by the brain and the techniques that can be used to measure them.

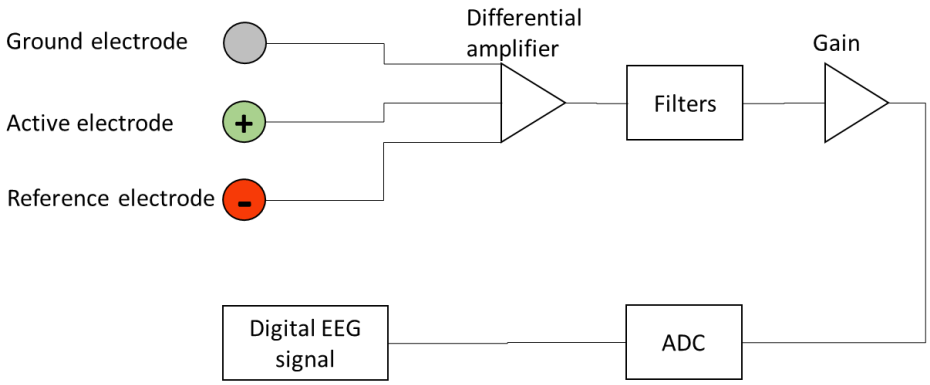
Figure 7 - Brain signals: origin and recording techniques.



Source: Own authorship.

In the case of EEG, the basic in terms of signal recording consists of at least three electrodes and an amplifier and analog-to-digital converter (ADC), see Figure 8. In this minimum configuration, one electrode is needed to actuate as the ground, so that the body surface and the amplifier are at the same potential level. The ground electrode can be placed anywhere on the body, but is usually placed on the scalp among the other electrodes. The EEG signal itself is recorded as the potential difference between the other two electrodes placed on the scalp. If there are multiple electrodes and only one of them is used as a reference for calculating the potential difference, the EEG measurement is said to be monopolar. On the other hand, if EEG is recorded from different pairs, it is said to be a bipolar measurement. There are cases in which the amplifier amplifies the signals referenced to the analog average from all the electrodes. The reference can always be changed digitally just by simple subtraction if the signals were initially referenced to the same value.

Figure 8 - Basic scheme for EEG data acquisition and digitalization.

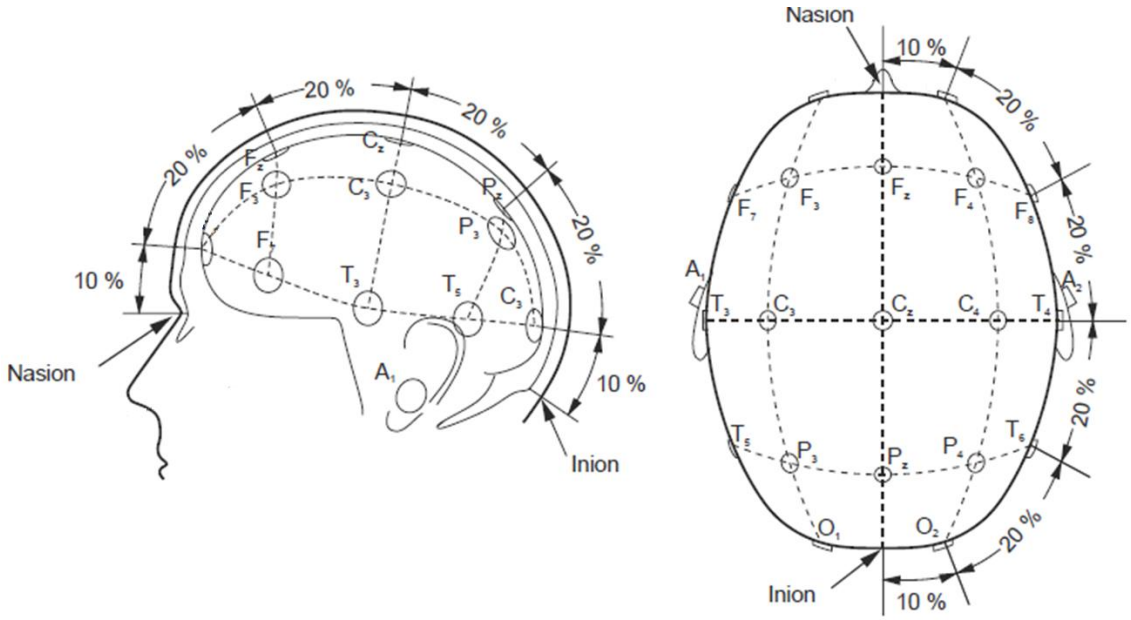


Source: Own authorship.

2.4.2. Electrodes positioning

To place the electrodes on the scalp, the international 10-20 system is commonly used to define their location (JASPER, 1958). This system was based on anatomical landmarks, allowing reasonable correspondence between electrode positions and brain structures (BÖCKER; VAN AVERMAETE; VAN DEN BERG-LENSSEN, 1994). To define the positions, two reference points are used, the nasion, located at the top of the nose, and the inion, in the bony lump at the base of the skull (NICOLAS-ALONSO; GOMEZ-GIL, 2012). Intermediate points are marked using 10% and 20% of the entire distance between the reference points, see Figure 9. The letters used to name the positions are related to the cortical region underneath (brain lobes), such as frontal, parietal, temporal, and occipital. However, the letter C is used to indicate the central region that crosses the vertex side-to-side on the scalp. The numbers increase as they move away from the middle-line between hemispheres, with odd numbers on the left and even numbers on the right. The midline is indicated by the letter z.

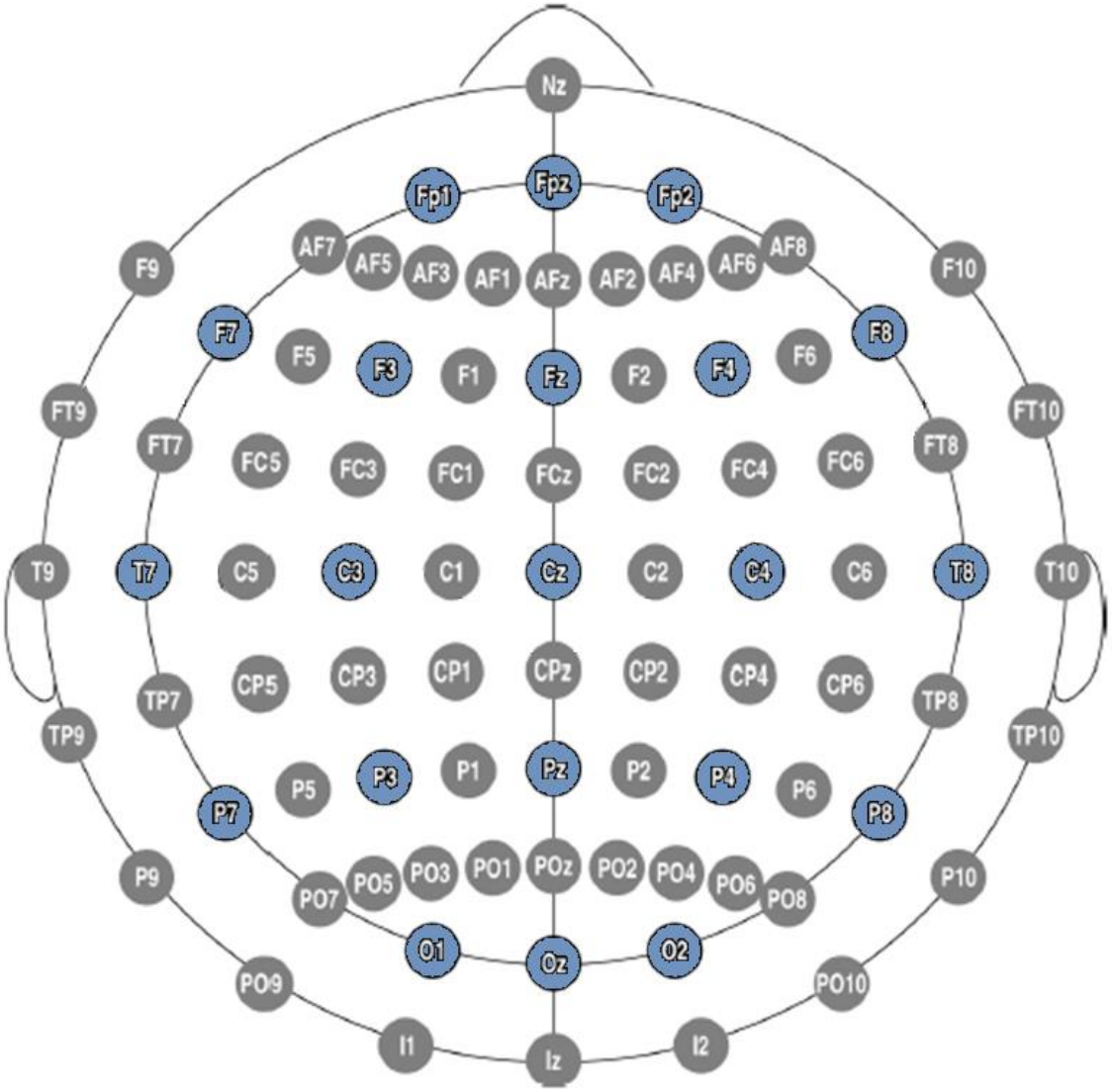
Figure 9 - The 10-20 positioning system.



Source: Adapted from (OOSTENVELD; PRAAMSTRA, 2001).

With the evolution of the field, researchers begun to use larger number of electrodes to perform EEG measurements. Hence, extended systems were proposed, such as the 10-10 and 10-5 systems, according to the percentage values used to define the positions (OOSTENVELD; PRAAMSTRA, 2001). To name the new locations in the nasion-inion direction, the letters are combined similarly to the wind rose. Intermediate locations in the left-right direction are indicated with the letter h, meaning they are half way to the next position towards the midline. See Figure 10 for the 10-10 system.

Figure 10 - The 10-10 positioning system.



Source: Adapted from (OOSTENVELD; PRAAMSTRA, 2001).

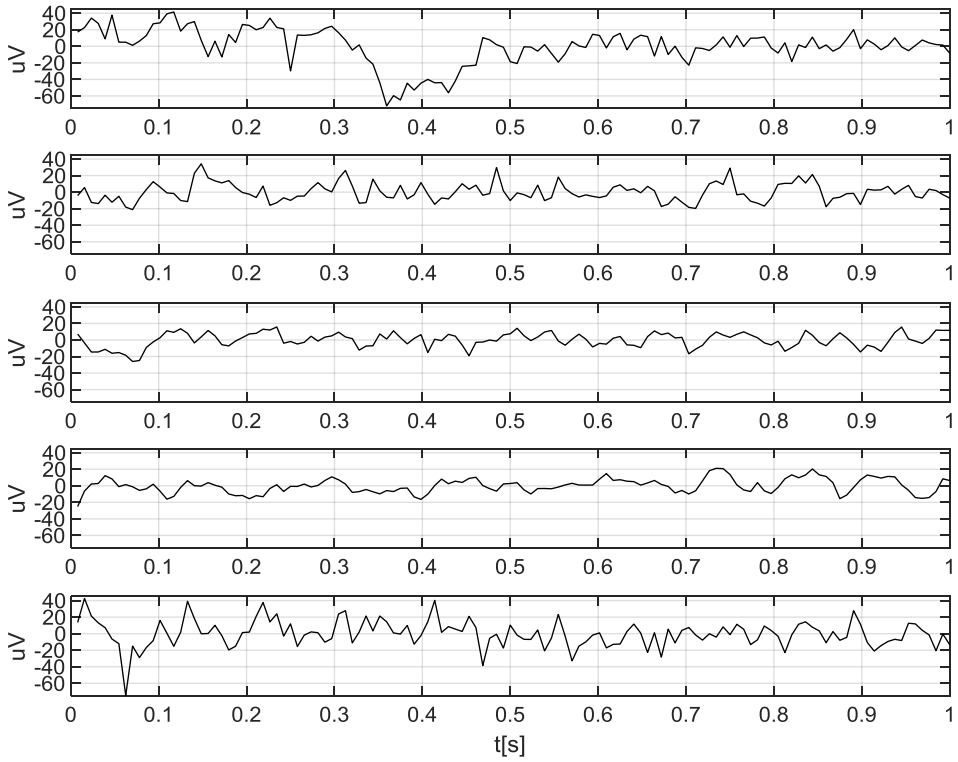
2.4.3. Signal variability – inter and intra-subject variability

Early works in the BCI field have shown that dealing with signal variability is a critical step of a BCI system development and perhaps the most challenging obstacle in the way of functional BCIs (VIDAL, 1973; WOLPAW et al., 1991). Still nowadays it is an important and unsolved issue in the BCI field (SAHA; BAUMERT, 2020). This variability implies in unstable signal's features when executing a certain task repeatedly. Thus, it becomes difficult to decode the mental activity using those features. An alternative to this problem is the operant conditioning approach, in which the subjects are trained to provide more reliable and stable EEG features (VIDAL, 1973). However, this is a complex approach

and it has been shown that it does not work for all subjects (NEAT et al., 1990; WOLPAW et al., 1991). The other approach is to develop algorithms to extract multiple features from the signals and use pattern recognition algorithms to decode the brain activity (PFURTSCHELLER; FLOTZINGER; KALCHER, 1993).

Part of the EEG signal variability comes from hardware issues such as the analog-to-digital converter (quantization noise), electrodes/skin impedance and environmental electrical fields. This can be sufficiently minimized by adequate preparation of the environment, proper technique to collect the data, and high quality equipment. Also, some level of variability is intrinsic to the natural brain functioning under certain circumstances (GARRETT et al., 2013). Other causes of variability, perhaps the most important ones, are the reference electrode and the poor spatial resolution (head volume conduction). They will include, in any EEG recording, a variety of ongoing brain activity taking place at the moment. Some of these brain activities might be directly related to the ongoing task being analyzed and others might have no relation to it at all. Hence, the activities unrelated to the task can be different over successive trials of the same task, despite being executed by the same person. Figure 11 shows EEG recordings from the same subject performing the same task repeatedly and all of them are different from each other. A more complete explanation of this problem is presented in sections 2.4.4 and 2.4.5.

Figure 11 - EEG recordings over five trials of the same task performed by the same subject.



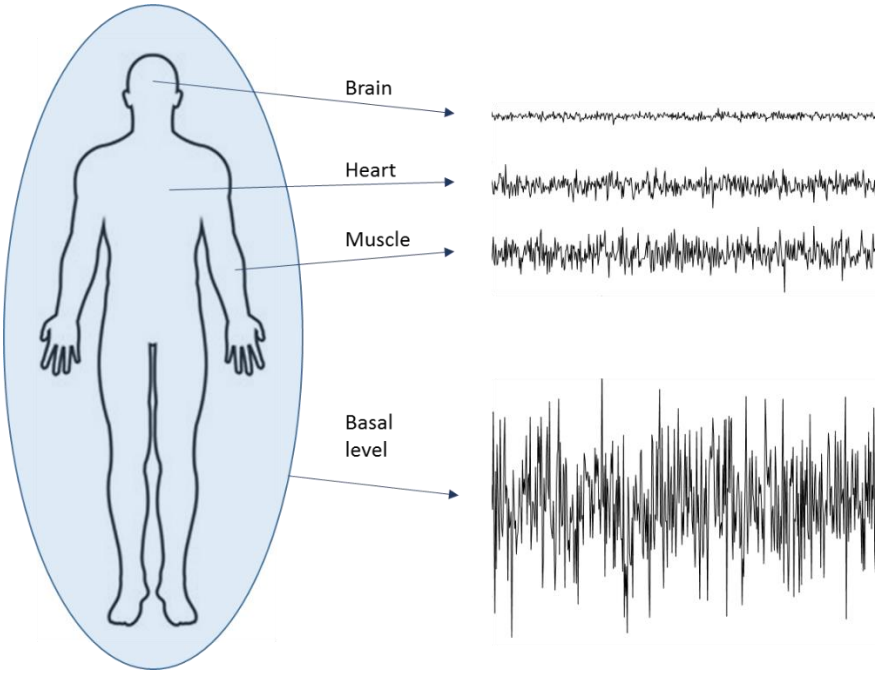
Source: Own authorship.

2.4.4. Reference electrode

In (NUNEZ; SRINIVASAN, 2006), the authors provide a didactic explanation concerning the reference electrode in EEG recordings. The main topics addressed by the authors are clarified in the following. The EEG equipment measures the electrical potential difference between two sites by means of a current passing through a parallel circuit which starts at one electrode and finishes at the other. The goal in terms of brain signal recording would be to measure the potential of certain spot of the brain with respect to an electrically neutral point that works as a perfect reference, often designated as quiet reference. However, given what has been exposed about EEG equipment, the reference electrode must actuate as an extremity of a parallel circuit that uses a very small current to infer the potential difference that originates this current. Hence, some arbitrary electrically neutral point not located on the body surface would not work if it does not allow a current to pass through, because in such case the potential measured would be null. Even if it was possible to measure

scalp potential referenced to a neutral site, it would also not be useful, in this case, because the human body is always under influence of environmental electrical fields from multiple sources (electric equipment, power line, etc.), and the voltages are often considerably higher than brain potentials, see Figure 12. This means that measuring the potential of a certain site on the scalp with respect to a neutral point would provide voltages that include all electrical fields influencing the body.

Figure 12 - Illustration of different electrophysiological signals' amplitudes and interference from external environment on the human body.

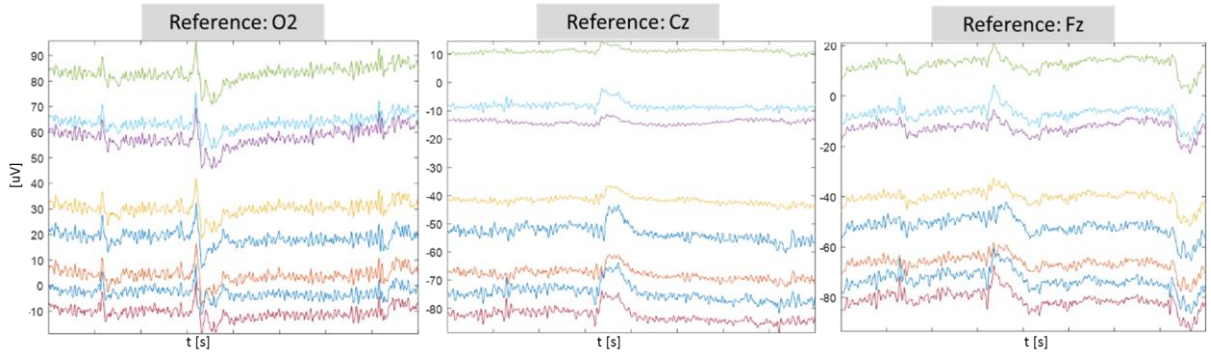


Source: Own authorship.

Finally, it is important to mention that even without external electric fields influencing the human body, there is no neutral point on its surface (GESELOWITZ, 1998), and the brain potentials acquired at the scalp are usually much smaller than other electrophysiological potentials recorded with electrodes placed on the skin, such as electromyography (EMG) and electrocardiography (ECG) (URIGÜEN; GARCIA-ZAPIRAIN, 2015). Thus, to properly acquire EEG signals the so called reference electrode must be placed on the scalp so as to avoid masking brain potentials with non-brain potentials. Moreover, given that it is not possible to know in advance where exactly every active EEG source is located on the cortex during an activity, every EEG recording must be seen as

bipolar, which means it measures the potential difference between two active transient sites. Figure 13 shows the EEG recordings from eight active electrodes referenced to three different electrodes (the dataset is from (BLANKERTZ et al., 2007)). The differences are notable, although they all represent the same interval of cortical activation. The head volume conduction, later discussed in the present work, also contributes to this problem.

Figure 13 - EEG recordings from the same eight active electrodes (CCP8 to CCP7) referenced to three different electrodes (O2, Cz, and Fz).



Source: Own authorship.

The main alternatives found in literature to deal with the impossibility of obtaining real monopolar EEG recordings are the average reference (AR) and reference electrode standardization technique (REST). Historically, linked-mastoids were used in many early EEG studies as a non-official standard.

The AR has a physical background that consists in the nullity of the surface integral of the potential created by dipoles in a sphere (BERTRAND; PERRIN; PERNIER, 1985). However, it also assumes a spherical head geometry and requires a large number of electrodes to cover equally the entire head's surface to be consistent, which is impossible because of the neck.

The REST is an algorithm proposed in (YAO, 2001) with the intention of transforming a scalp point to a reference point at infinity (infinity in this context means electrically neutral). The technique is sensitive to the electrode density, electrode coverage (both limitations shared with AR and CSD technique), and to the head model uncertainties, given that the method requires a head model (NUNEZ, 2010).

The evaluation of reference choice in EEG recordings has been a research topic for a long time and remains an important topic nowadays. Five articles addressing this issue are briefly commented on the following.

In (JUNGHÖFER et al., 1999), the authors demonstrated by means of simulation the polar average reference effect (PARE), as they called it. It was seen that electrodes in the center of the electrode array had smaller potential amplitudes than those located at the edge. This effect was non-negligible even when the head was covered by an electrode net extending across 270^{o}. In (QIN; XU; YAO, 2010), the authors compared the AR, REST, linked-mastoids, and left mastoid as reference options for real and simulated data. In their analyses, they found that REST provided less relative error than the others with simulated data, and that the three other reference alternatives introduced significant changes to spectral analysis. In (LIU et al., 2015), the authors also assessed the effect of AR, REST, and linked mastoids. They report that both AR and REST outperforms linked-mastoids. They also show that AR is more sensitive to the number of electrodes used than REST, but REST is more sensitive to uncertainties relative to electrode's location. In (CHELLA; PIZZELLA; ZAPPASODI, 2016) it was compared the AR, REST, linked-mastoids, and Cz as references in view of EEG functional connectivity. Their results showed descendent distortions to the connectivity patterns for references Cz, linked-mastoids, AR, and REST, respectively. They also report a high degree of dependency of REST relative to realistic head models and high density electrode arrays. In the work of (HU et al., 2018), the authors included four other monopolar references to their comparison with simulated data relative to (CHELLA; PIZZELLA; ZAPPASODI, 2016): left ear lobe, Fz, Pz, and Oz. Neither of the monopolar references are recommended by the authors due to the large distortions it puts on the EEG. In agreement with results from the other articles, they report the best results for REST and the worst for linked-mastoids, with AR in between. AR in this case showed significant dependency relative to head coverage, but was not sensitive to the number of electrodes.

Besides these attempts to re-reference EEG data in order to approximate recorded signals to ideal potentials relative to infinity there is another alternative commonly adopted to avoid arbitrary reference choices: surface Laplacian (SL)

for current source density (CSD) estimation. This technique is taken as a spatial filter and eliminates the reference from the signals, reason why it is not considered a referencing or re-referencing technique such as the others, although it constitutes an alternative for the reference electrode issue.

The CSD via Laplace, originally proposed by (HJORTH, 1975), in fact eliminates the reference electrode effect, but assumes that the scalp surface is flat for nearby electrodes. However, for the addition of an electrode to be significant, it may sometimes require more than 5 cm separation from its neighbors, which may cause the approximation of real head surface into a flat surface somewhat arbitrary (NUNEZ et al., 1994). Perrin and colleagues (PERRIN et al., 1989) used spherical splines interpolation to estimate a continuous potential surface and calculate the SL. In this case the head geometry is assumed to be spherical and it requires a large number of electrodes (e.g. 64 or more) to provide a theoretically meaningful result (SRINIVASAN et al., 1996). Mathematical approximation inherent to the interpolation process demonstrated in (PERRIN et al., 1989) and inherent to the discretization applied by (HJORTH, 1975) also represents sources of uncertainty to the CSD estimation via SL. This approach is also discussed in the spatial resolution section.

One can see that in fact there is no definite solution to the problem of the active reference in EEG recordings. The supposedly best options (AR, REST, and SL) require large numbers of electrodes (more than 64), and they also include uncertainties due to head geometry assumptions and electrodes placement. Table 4 summarizes sources of uncertainties in the most common solutions throughout literature.

Table 4 - Sources of uncertainties in the solutions to the reference electrode problem. The * indicates that the assumption involved in the SL solution concerning the head electrical properties is not usually considered to be an issue.

	Head geometry	Head electrical properties	Source distribution	Number of electrodes	Electrodes distribution	Mathematical uncertainties
Surface Laplacian	Yes	**No***	**No**	Yes	Yes	Yes
Average reference	Yes	**No**	**No**	Yes	Yes	Yes
REST	Yes	Yes	Yes	Yes	Yes	**No**

Source: Own authorship.

The large number of electrodes implies significant practical difficulties for executing experiments and even more for continuous use of a BCI. Moreover, the incorrect assumption of head geometry and the non-precise electrodes placement introduces quantitative distortions in the solutions. On the other hand, choosing a specific monopolar reference, although commonly adopted in EEG analysis, has poor or none technical/physiological justification. Hence, this is an unsolved issue in EEG analysis, especially towards BCI applications, given that the system is highly affected by EEG variability.

2.4.5. Spatial resolution

EEG is a technique known to provide poor spatial resolution. The poor spatial resolution can be understood by the following statement: the minimal distance that separates two notably different EEG recorded signals is greater than the minimal distance that separates two EEG sources in the cortex. EEG has poor spatial resolution because of three factors, namely (1) limited spatial sampling, (2) head volume conductor, and (3) reference electrode (NUNEZ et al., 1994). The limited spatial sampling is explained by the 10-20 positioning system, in which electrodes spacing is 6 cm in average. Moreover, the addition of extra electrodes doesn't add information content to raw EEG, unless further pre-processing takes place (NUNEZ et al., 1994). However, the head volume conductor and reference electrode are the main problem. The reference

electrode issue was already exposed, thus in the following it is discussed with more details the head volume conductor and the most common alternatives for overcoming its undesired effects.

The volume conduction is the name given to the manner in which the electric current spreads from brain sources through brain, CSF, skull, and scalp tissue (WOLPAW; WOLPAW, 2012). The geometric and electric characteristics of the different medium are said to determine the volume conduction. Authors explain that the skull tissue has high resistivity, which causes the local cortical current to spread widely, and they also draw attention to the spacing between the cortex and the scalp, which increases this spreading effect, causing a space-averaging in EEG signals. An active dipole, by means of postsynaptic potentials in a set of neurons, creates an essentially instantaneous voltage field throughout the head with no meaningful delay (LUCK, 2014). Also, the electricity does not just run from one pole to another of a dipole in a conductive medium, but spreads across the conductor. This is the reason why a signal generated by a certain dipole is not recorded only by the electrode placed directly above it, but also by other electrodes distributed over the scalp (LUCK, 2014).

Thus, one can see that there is a problem to be addressed, which is characterized by multiple cortical sources generating different signals that will spread through the volume conductor, mix into each other, and only then will be recorded by EEG electrodes on the scalp. The problem is well stated: how can one obtain the original signals by analyzing their mixture? Many approaches have been presented in literature and the most remarkable ones are introduced here: principal component analysis (PCA) and independent component analysis (ICA), common spatial patterns (CSP), current source density (CSD) estimation via surface Laplacian, and inverse problem solutions. PCA and CSP does not actually attempt to find supposedly original cortical signals, but rather provides statistical approaches to help towards improved spatial analysis of EEG signals. Each one of these approaches is introduced in the following paragraphs.

2.4.5.1. Common Spatial Patterns

CSP was presented by (FUKUNAGA, 1972) in his book and first applied to EEG by (KOLES; LAZAR; ZHOU, 1990) to differentiate between two classes of EEG

signals. The method is based on the diagonalization of a combined covariance matrix obtained from EEG signals from each class. It is a statistical method that creates a projection matrix that linearly transforms raw EEG data so it can be further processed towards better results. The raw EEG is projected onto a space in which the variance of EEG from one class is maximum when the variance from the other class is minimal, hence providing class-discriminable signals. It is worth noting that this approach does not intend to properly obtain original cortical signals, nor lays on any neurophysiological assumption. Moreover, it is only applied towards two-classes problems. It has been widely adopted in EEG analysis with great success in the BCI field. For instance, a search in Scopus database (www.scopus.com, accessed in September/2020) with the words “EEG” and “common spatial patterns”, limited to articles and limited to the articles that contain these words in their keywords list, found 401 items. It is worth to mention that maybe even a greater amount of papers uses the CSP technique with EEG as a tool to address something else, and in such cases the words referring to CSP would not appear in their title or in their keywords list.

2.4.5.2. Surface Laplacian

The Hjorth’s method, as it is sometimes referred to, estimates the current source density with surface Laplacian (HJORTH, 1975). It is based on the physics of electrical current being radially dispersed along the surface of a conductive volume. This current will generate convex and concave curvatures in the potential field that will perfectly balance each other in perpendicular directions (x and y) in areas where no source is active. This balance is described by Laplace’s equation. Hence, when there is no balance it means that there is an orthogonal current entering the corresponding area (along z-axis), and the density of this current can be estimated by using the Laplacian operator. Although it is originally applied to continuous surfaces, the central difference operator (finite difference) can be used to estimate the current density in a discrete potential surface provided by the electrodes (HJORTH, 1975). The method has a few important assumptions, namely: conductive properties as a function of depth are the same for any area of the scalp; the distances between electrodes are constant and the grid is homogeneous; and the individual group formed by neighbor electrodes used to estimate CSD at one point is geometrically small enough to be considered plane.

An important drawback is the estimation at the edges, since some neighbor electrodes will be missing. The major advantages of this method are the physical basis and the independence regarding the reference electrode, which is a consequence of the second derivative of Laplace's equation. The transformation can be done in real-time by linearly transforming the raw-EEG signal matrix, all channels at once, for each time instant, just like the CSP method.

As an alternative for the finite difference approach, it was proposed the application of the spherical spline interpolation, presented in (WAHBA, 1981), to EEG signals in order to obtain a continuous potential distribution over the scalp, then the surface Laplacian can be used to estimate the CSD (PERRIN et al., 1989). The current density can be estimated on any point of the sphere that represents the scalp. Although this method requires more complex mathematics when compared to Hjorth's method, it may also generate a transformation matrix to linearly transform data in real-time for the places where there are electrodes. One new challenge that comes along with this approach is the uncertainty of a few parameters during the calculation process, e.g. a couple of series truncation, uncertainty towards the spline flexibility value and the regularization parameter. It is also worth to mention that the interpolation is sensitive to the number of electrodes used to record the data. Both methods, finite difference and spherical interpolation are widely adopted techniques in the problem of estimating the current density of EEG sources. For the mathematical details of the SL, see Appendix C.

2.4.5.4. Principal Components Analysis and Independent Component Analysis

Some techniques rely on statistical properties of EEG signals. PCA, for instance, is a multivariate statistic technique commonly applied to EEG signals for dimensionality reduction. One of the first works (if not the first) to present a scientific article using PCA to study EEG (specifically average evoked potentials) was (DONCHIN, 1966). PCA is an orthogonal transformation that decomposes the original *n*-dimensional data into a *m*-dimensional set of uncorrelated components, where $m \leq n$ (ZENG; LI, 2016). The uncorrelated components are called principal components. The first component contains the signal's largest

variance, which is understood as the largest amount of information. The second component contains the second largest variance, and so on. In this way it is possible to obtain only a few uncorrelated components that represent, for instance, 90% of total variance of the raw multi-channel EEG. ICA extracts statistical sources, called independent components, from the raw multivariate recordings (ZENG; LI, 2016). It is a blind source separation (BSS) technique. It consists in creating an unmixing matrix that will linearly transform raw data, just like the other methods already described. There are many different algorithms to create the unmixing matrix. Although PCA and ICA are both quite similar approaches, there are substantial differences in what they provide as results. A simple analogy may help to understand. If a set of images with human faces is provided for the application of PCA and ICA separately, the first few principal components will provide something related to an average face. On the other hand, one independent component might be related to the eyes, another one to the nose, to the mouth, and so on. These two approaches and their many variations are extensively found in the scientific literature for EEG signal analysis to overcome issues inherent to the volume conduction.

2.4.5.5. Inverse Problem

A different approach towards the problem consists in developing EEG source models. This is designated as the forward problem. Electric and geometric properties of the head and volume conductor are estimated to produce these source models and generate signals that would be recorded on the scalp due to these sources. This makes it feasible, for instance, to generate artificial EEG signals. On the other hand, with actual EEG recordings and with these sources already modeled, one can estimate multiple dipoles position, orientation, and magnitude, i.e. perform a source localization procedure. This approach is designated as the inverse problem.

In the 1950’s there was already a published paper concerning source modeling for electrophysiological signals (FRANK, 1952). The author assessed electrical potential produced by a dipole in a homogeneous conducting sphere for electrocardiographic applications. Later, many papers were published assessing the inverse problem in EEG analysis. For instance, (HENDERSON; BUTLER; GLASS, 1975) and (ARY; KLEIN; FENDER, 1981) presented strategies to

identify a single dipole (position, magnitude and orientation) by means of EEG signals and a multilayer conductive sphere model, already considering in some level the anisotropy of the human head volume conductor. In (STOK, 1987), the author analyzed in which way some model parameters would influence the inverse problem solution, also using a multilayer sphere model and a single dipole. (MUNCK; DIJK; SPEKREIJSE, 1988) investigated if a mathematical dipole description would properly suit the source modeling in the analyses of visual evoked potentials in EEG, rather than other approaches that better represents cortical sources in terms of extension. In their scenario, the authors found that even when the source is large and the electrode grid does not cover entirely the source area this would not result in large errors. Pascual-Marqui and colleagues (PASCUAL-MARQUI; MICHEL; LEHMANN, 1994) proposed a new approach towards the source localization problem. Their method does not assume a specific number of dipoles nor a distribution on a known surface. The neurophysiological assumption in which the procedure relies is that neighboring neurons are simultaneously and synchronously activated. The method is called Low Resolution Electromagnetic Tomography, referred to as LORETA. This strategy for solving the inverse problem, in which the dipoles are distributed in the whole brain volume or cortical surface, with no previous restrictions in terms of number of sources and their distribution is called the non-parametric approach or Distributed Source Models (GRECH et al., 2008). There are a number of other non-parametric methods that are explained by (GRECH et al., 2008). The authors also mention that the former approach, in which the number of dipoles is assumed and a head model (that could be spherical or even realistic, for instance) is used is called parametric approach or Equivalent Current Dipole methods.

2.5. State-of-the-art and work justification

The BCI field has existed in the scientific literature for about 50 years. In the first four decades, very often the works claim to be on the edge of some breakthrough that could lead to BCIs capable of being used in daily life activities. However, BCIs are still mainly restricted to laboratories and other very simple applications. On the other hand, the increasing number of people with severe neurological impairments and the lack of successful neurological rehabilitation therapies for such cases opened a new perspective to the BCI field. BCIs for

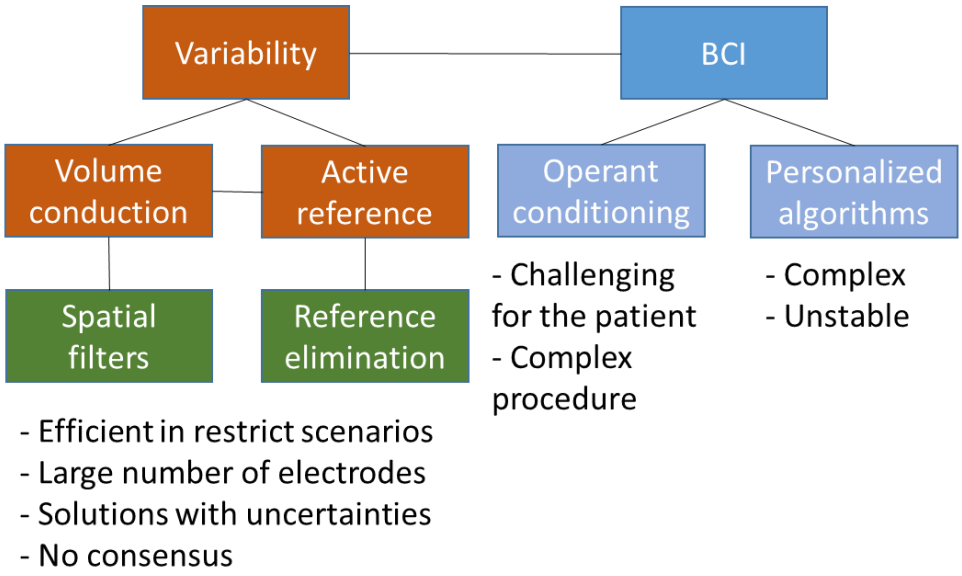
motor/neurological rehabilitation do not require the same level of accuracy when compared to BCIs for other purposes (e.g. for controlling a robotic prosthesis or a wheelchair). Therefore, motor rehabilitation became a strong motivation for BCI research, especially in the past 10 years. Hence, this work focuses on this type of BCI.

EEG is the most commonly used brain signal recording technique for BCIs. It has major practical strengths, such as: low-cost, portability, safeness, easiness to use, and high temporal resolution. EEG attends to almost every requisite for an ideal BCI system, but with the current state-of-the-art on EEG signal processing and recording, it still lacks a key feature: reliability. The low reliability from EEG towards BCI systems comes from the inter and intra-subject variability of the signals.

One of the reasons for this variability is because EEG signals have very low magnitudes compared to other electrophysiological signals (or other electrical signals in general), so a high signal-to-noise ratio (SNR) is a natural challenge. Besides hardware issues, two major causes of such variability are the head volume conduction and the reference electrode that must be placed on the scalp among the other electrodes. These problems cannot be avoided, because they are natural characteristics of the human body and of the EEG measurement. For neither of these problems the scientific community has achieved a minimum consensus towards alternatives to minimize their impact. Among the most commonly adopted pre-processing techniques to address these issues one can cite the CSP, SL, PCA, ICA, inverse problem, AR and REST. All of them have important limitations, which includes the necessity of a large number of electrodes for EEG recordings.

The BCI systems may be of two categories, one in which the system adapts to the user, and another in which the user adapts to the system. The first is highly affected by the signal variability, usually requiring complex algorithms to pre-process, extract features and perform patter recognition. The latter is highly dependent on the person’s capacity to learn to provide the correct signals and on the strategy for teaching this ability to the user. Figure 14 summarizes schematically this scenario.

Figure 14 - Context of this work within the EEG-based BCI field.



Source: Own authorship.

The methodology of this work is of the first type, in which the system adapts to the user. However, to minimize the necessity of complex algorithms, often highly unstable over time, in this work a strategy to reduce signal variability with offline processing prior to the development of the BCI is proposed. The goal is to find subject-specific sub-set of channels and signal features that are stable over different trials of the same task. This knowledge is then used to configure the online signal processing algorithms to achieve the major goal of this work, which is to operate a BCI in real-time aiming at a motor rehabilitation activity. The methodology is developed to overcome the signal variability relying on the following hypothesis: for a particular task and person there must be specific regions of the scalp with low inter-trial variability relative to EEG signals.

These regions will likely be reflecting specific regions of the cortex involved in the task in different ways. Therefore, by an adequate strategy to select pairs of electrodes (i.e. EEG channels), it is possible to find these regions and reduce the intra-subject variability. As for the inter-subject variability, it is assumed to be a natural phenomenon, thus requiring the strategy to be applied individually for each person. This becomes highly significant if one considers the application towards people with brain damage.

It is known that many EEG features are subject-specific, and even when a feature is common to different subjects (e.g. the ERD), they have their own specificities for each individual (e.g. the timing relative to movement onset and the frequency range). The brain morphology, head shape and, likely, the volume conduction properties are highly individualized as well. Given this scenario, using an arbitrary reference electrode for processing multiple subject's signals (which is a widely adopted strategy) might increase the challenges for the EEG analysis.

Besides the goals accomplishment, two intended contributions of this work are:

1) To develop a methodology that avoids relying on assumptions that could be inaccurate, such as: the lack of activity in the reference electrode, the shape of the head, the number of dipoles generating the signals and the electrical properties of the volume conductor modelling the head.
2) To enable the use of a BCI with a low number of electrodes, which is crucial for practical applications.

3. METHODOLOGY

The methodology is divided into two major parts that are executed sequentially (see Figure 2).

Part 1 of the methodology evaluates EEG motor signals by two approaches to gather important knowledge to guide the second part (Part 2). Because Part 1 does not directly aim at real-time performances signal processing, the use of trial averaging is considered. The first approach (Approach 1) of part 1 considers well-known EEG motor signatures and assumes that the scalp location of the electrodes to best detect these signatures are subject-specific and cannot be precisely known a priori. This holds for the main electrode and for the reference electrode in a channel. The second approach (Approach 2) of Part 1 looks for transformations in the EEG signals that provides lower intra-subject variability. In this case, not only the best channels are considered to be subject-specific and previously unknown, but also the features that most likely provide movement related information.

In Part 2, the aim is at real-time movement intention detection. Therefore, it considers only single trial signals. The method considers a pseudo-online processing. The implementation of the methods of Part 1 and its result analysis guided the proposal of the methods for Part 2.

Prior to presenting Part 1 and Part 2, the procedures that are common to both of them are presented, such as the public EEG dataset that is used in the analysis.

3.2. Publicly available dataset

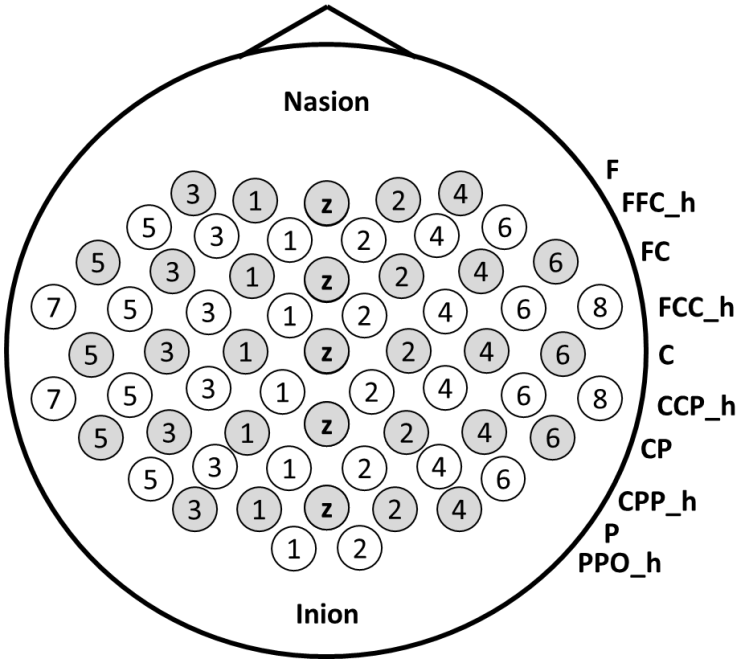
The publicly available dataset (OFNER et al., 2017), from the Graz BCI group, contains EEG signals from 15 subjects related to elbow flexion and extension, forearm supination and pronation, wrist closing and opening, and resting, which totalizes seven classes. The movements were executed with a passive exoskeleton. In all movement cases, there are imagined and real executions. Each subject executed 10 runs of the experiment (different experiments for real and imagined movements). Each run consisted of 42 trials, with six trials for each class, randomly distributed. Thus, there are 60 trials of each class per subject. Each trial consists of 2 s with a blank screen with a cross in the middle, 3 s with

the movement cue indicating what should be executed, and finally 2 to 3 s of a completely blank screen before a new trial starts. The beginning of the new trial is indicated by a short beep and also by the cross in the middle of the blank screen, see Figure 15. The dataset contains movement data, such as the elbow angle of the exoskeleton and movement data of the fingers. It also provides cue information, so it is possible to know exactly when each activity was supposed to take place. The EEG was recorded in 61 positions according to the 10-20 extended system (see Figure 16). The ground electrode was placed at the AFz, and the reference electrode at the right mastoid. Signals were acquired at a 512 Hz rate and filtered with an 8th order Chebyshev band-pass filter from 0.01 to 200 Hz. A notch-filter at 50 Hz was used to suppress power line interference.

Figure 15 - Sequence of one trial of the public dataset.

Source: Adapted from (OFNER et al., 2017).

Figure 16 - The 61 electrodes positions used in the EEG dataset from the Graz BCI group (OFNER et al., 2017).



Source: Own authorship.

3.3. Part 1

Two approaches (Approach 1 and Approach 2) are considered for executing part 1. Both of them are executed offline. The pre-processing the is common to both approaches is presented in the following.

3.3.1. Data pre-processing for part 1

All the signal processing was done using the MATLAB software (R2015a The Mathworks, Inc., Natick, Massachusetts, United States). The EEG data was first separated into trials. For the public dataset from Graz BCI group, the trials of interest consisted in the ones with elbow movement (flexion and extension). The trials last for 8 s, starting with 1 s of blank screen followed by a short beep and the appearance of a cross in the screen. Two seconds later ($t = 3\,s$) the movement cue appears. They are synchronized by the time instant that the cross appeared on the screen. This means that after 3 s of each trial the cue indicating the movement was presented.

To eliminate trials contaminated with EOG artifacts, the procedure described by (SANNELLI et al., 2019) was applied. It consists of eliminating trials with standard

deviation greater than two times the average standard deviation in an iterative procedure, i.e. after eliminating the first trials, the procedure is repeated with the new set of trials, until no more trials are eliminated.

Afterwards, a re-referencing procedure was conducted. The original EEG signal was transformed into all possible channel combination. So, if the dataset is acquired with N electrodes, the number of possible channels results from the combination 2-by-2 of N.

An important part of the analysis is to detect the movement onset during the tasks. For the current analysis, only elbow flexion and extension were considered. The dataset informs the cue's onset instants and type. However, there is an unavoidable delay from the cue appearance to the movement initiation. To detect the movement onset, the signal with the elbow joint angle is used. In general, the neutral position corresponds to approximately 45º, although there are significant variations of this value over time. During the flexion movement the angle value of the sensor reduced and during the extension it increased. In both cases the total variation regarding the entire movement showed to be very similar after a visual inspection.

The procedure to detect the onset for each trial followed eight steps: (a) The signal corresponding to the elbow joint angle is segmented in the same way the EEG signals are for each trial; (b) the trials with elbow extension had their signals inverted, i.e. multiplied by -1, so in both, flexion and extension, the angle should decrease to indicate the movement execution; (c) the sensor signal is set to have its lowest value in the trial equal to zero, so all trials are approximately in the same range of values; (d) Because of the variations in the neutral angle during the experiment, a reference value for the neutral position is determined by the mean value in the interval corresponding to the 0.5 s prior to the cue appearance; (e) the total range in the angle values (difference between maximum and minimum values) in the interval corresponding to 0.5 s prior to the cue and 2 s after the cue is determined and 10% of this angle variation is used as tolerance; (f) the trial interval is inspected point-by-point starting at the cue appearance, and once the angle value is lower than the reference value minus four times the tolerance, it is considered that the current point is already past the onset; (g)

starting from this point, the interval is inspected moving backwards point-by-point until the angle value is closer to the reference by a distance smaller than the tolerance, this is then defined as the movement onset point of the trial; (h) the average over all onset (time) points is calculated and used as the onset for the average EEG signal over all trials. For the analysis of the grand averages, the onset instant is calculated as the average over all subjects as well.

Further steps of the signal processing depend on the features to be extracted. Therefore, they are presented in the following subsections.

3.3.2. Approach 1 – searching for channels

Given well-known EEG features that are likely present during the motor task, a search to detect the channels that contain these features is executed.

Assumptions:

a) Task-related EEG activity is known;
b) The best pair of electrodes to detect these activities is not known.

Procedures:

- Define criteria for detecting known signals;
- Find best channels to detect these features;

Major challenge:

- Specify the timing of the EEG activities related to the movement.

3.3.3. Approach 2 – searching for channels and features

A search to find features and channels that present high stability over trials was performed. In other words, the search intent is to find features among every channel with the lowest inter-trial variability for each subject.

Assumptions:

a) Task-related EEG activity is not known;
b) The best pair of electrodes to detect these activities is not known;
c) There are features that characterizes task-related EEG activity.

Procedures:

- Use a sliding window in time domain;
- Extract multiple features for every window;

Major challenges:

- Choose a large initial set of parameters/features;
- Quantify inter-trial variability.

3.3.4. Signal processing

The signal processing is specific to each of the two approaches; thus, they are presented separately.

3.3.4.1. Part 1 – Approach 1

Prior to presenting the processing procedures for the first approach, it is important to define the hypothesis regarding what is expected to be found in the data to be analyzed.

For the Graz dataset, the beep/cross associated with the movement cue to be performed 2 s later can be seen as a warning-go paradigm, which is known to elicit a CNV. Nevertheless, there is no need for the subject to respond fast to the cue, which may attenuate the CNV by requiring less engagement within the inter-stimuli interval. Moreover, the subject's response lacks any feedback, which differs from usual CNV paradigms. This might also attenuate the subject's engagement, which increases the possibility of an unidentifiable CNV. On the other hand, although the task is not self-paced (which would elicit the RP), the lack of these important aspects related to the CNV paradigm approximates the activity to a self-paced type. Therefore, in some level the RP might be present as well. Before the movement cue appears on the screen, it would usually be possible to identify ERD related to motor planning. However, the subject does not know which motor action will be performed within a small repertoire, thus it is reasonable to think that the subject is unable to plan specifically the movement. Hence, ERD might not happen prior to the cue appearance.

Signal processing procedures

The processing consists of highlighting the main features of the delta and alpha bands oscillations, and then quantifying the presence of these features in the

signals. For each EEG channel, two separate procedures were performed, one for the delta band (CNV and RP) and another for the alpha band (ERD). All procedures are executed separately for each subject, unless explicitly informed otherwise.

Delta band signals (CNV/RP)

For the delta band, the signals are first band-pass filtered between 0.1 and 4 Hz. Then, the average of all trials is calculated. For each channel, the maximum EEG value within the 2 s interval with the cross on the screen is detected, then the first negative peak after the movement onset is detected. The difference between them is calculated and is used as a score to quantify the negative shift of the delta band, as illustrated in Figure 17. The channel with the largest negative shift is thus selected.

Alpha band signal (ERD)

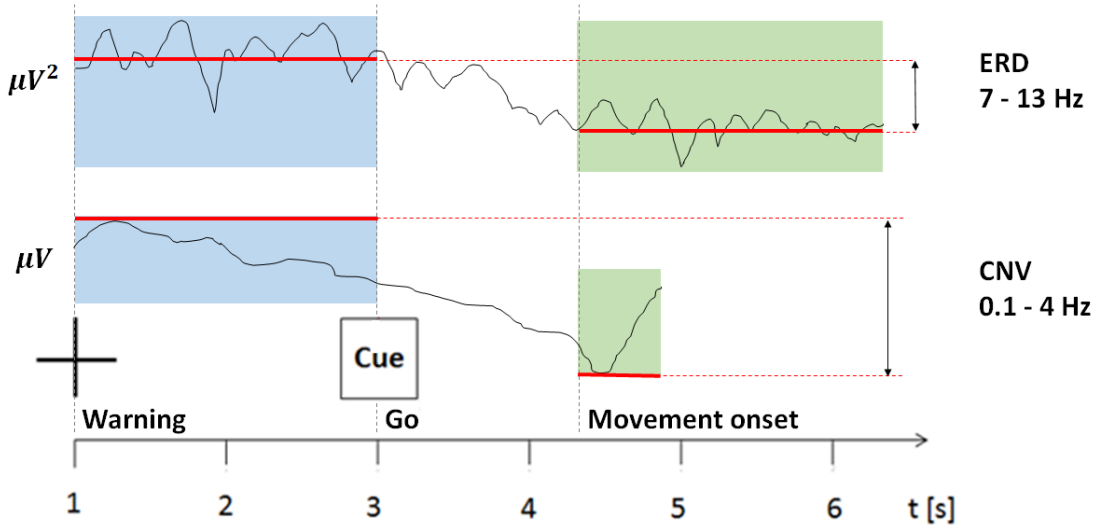
For the ERD, also for each channel, signals are band-pass filtered in the 7 – 13 Hz band, then the samples are squared, and the average of all trials is calculated, providing the average power. The average power is smoothed with a sliding window of 12 samples (approximately 0.1 s). The mean energy in the 2 s interval with the cross on the screen is calculated, the mean energy in the first 2 s after the movement onset is calculated, then the difference between them is used as the score to quantify the ERD, see Figure 17. In this case, the channel with the largest energy difference is selected.

Once the channels are selected for each frequency band and subject (one channel for each case), the signals for each channel of a given subject are averaged among all trials. Then, the average signals are individually normalized by setting the highest value within the 8 s interval corresponding to one trial as 1 and the lowest value as 0. This is done by subtracting from the entire signal its lowest value, then dividing the entire signal by its highest value. For the alpha band energy average signals, the instant in which the energy reaches a 75% decay compared to any past instant is defined, and for the delta band signals the first valley (negative peak) after the movement onset is defined.

The average (over subjects) among all individual averages for each frequency band is calculated and referred to as the grand average. The grand averages are calculated for the selected channels (even if these are different for every subject) and also for other channels, depending on the frequency band, for adequate comparison. For the delta band signals, the other channels are Cz referenced to the average reference (AR) and Cz referenced to the right mastoid (RM). For the alpha band energy signals the channels are C3 (contra-lateral to the moving hand) referenced to AR and RM. These channels are frequently used to assess signals such as CNV, RP and motor ERD. The mean time instant of the movement onset among all trials is determined, as well as the instant in which the subjects reach the maximum amplitude of the movement and start to move back to the initial position. The movement offset is not explicitly determined, but in general it is approximately at t=8 s, which coincides with the end of the trial interval considered in the analysis.

To compare the subjects' signals with each other for the two frequency bands separately, the Pearson correlation coefficient was calculated between every pair of signals of each band.

Figure 17 - Time intervals of the task that were used to quantify the ERD and CNV in the signals. The red lines indicate the value that is considered. For the ERD case, the value is the average of the colored intervals, for the CNV it is the maximum in the blue interval and the first negative peak in the green interval.

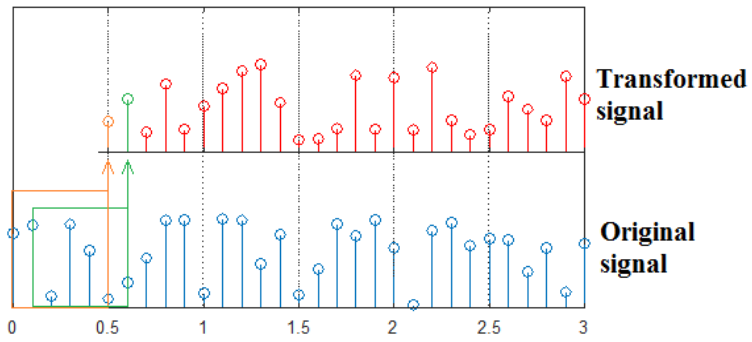


Source: Own authorship.

3.3.4.2. Part 1 – Approach 2

Given what is presented in the literature review about EEG signals in motor tasks (section 2.3), several features were chosen to be calculated from a sliding window, as explained in section 3.3.3. The features are intended mainly to represent a few important things about the signals, such as band-specific energy and its change across time, negative or positive shifts, and slope. The original signal is then substituted by a new set of transformed signals, according to the features calculated at every time interval defined by the sliding window. This window is always used entirely, i.e. if its 0.5 s long, the first 0.5 s of the original signal will be lost (substituted by zeros) when transformed. The sliding window moves one sample at a time till the final 0.5 s interval of the trial. It is important to mention that for different features there are different down-sample processes, so moving the sliding window one sample at a time does not necessarily means moving the same time interval for all features. See Figure 18 for an illustration of the transformation with the sliding window process. However, in one specific case the transformation did not require a sliding window, as it is explained in the following paragraphs.

Figure 18 - Sliding window process to transform the signals.



Source: Own authorship.

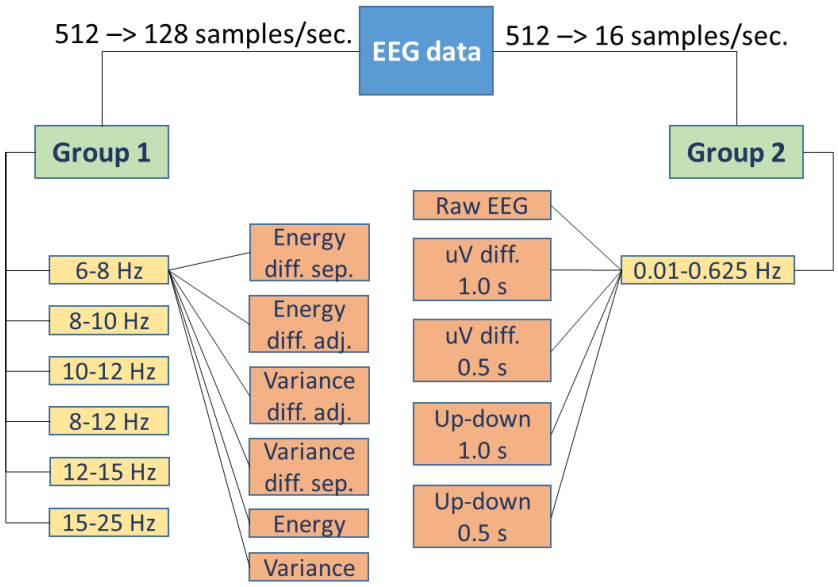
The features calculated with the samples from each sliding window contain different information about energy, variance, and slope. They are divided into two groups considering their frequency content. Group 1 seeks information on frequencies related to motor cortex activation (6 – 25 Hz) and group 2 extracts data from frequencies lower than 1 Hz, usually related to cortical activities prior to movement execution or movement information (KOBLER; SBURLEA; MÜLLER-PUTZ, 2018).

In both groups, when the terms short window and long window are mentioned, they refer to 0.5 s and 1.0 s length windows. When the relation between two windows is described as adjacent windows it means that the windows share one sample (the last sample from one of them is the first sample from the other one); when it is described as separated windows it means that there is exactly one short window (0.5 s length) in between them and both windows are necessarily short windows.

Group 1 down-samples the signal to 128 Hz (1 sample every 7.8 ms). It uses only short windows and is divided into six sub-bands: 6 – 8 Hz, 8 – 10 Hz, 10 – 12 Hz, 8 – 12 Hz, 12 – 15 Hz, and 15 – 25 Hz. For each sub-band, also six features are calculated: 1) total energy; 2) variance; 3) energy difference between adjacent windows; 4) variance difference between adjacent windows; 5) energy difference between separated windows; and 6) variance difference between adjacent windows. Thus, group 1 transforms the original signal into 36 new signals. See Figure 19 for a visual representation.

Group 2 always uses the 0.01 – 0.625 Hz frequency band. To enable proper filtering and avoid numerical errors due to the low cut-off frequencies, the signal is downsampled to 16 samples per second. Five transformations are executed: 1) raw EEG signal; 2) summation of the derivatives in short windows; 3) summation of the derivatives in long windows; 4) number of samples with positive derivative minus the quantity with negative derivatives in short windows; and 5) number samples with positive derivative minus the quantity with negative derivatives in long windows. Thus, group 2 transforms the original signal into five new signals. Both groups together generate 41 new transformed signals. See Figure 19 for a visual representation.

Figure 19 - Visual representation of the 41 transformations separated in two groups. Group 1 concentrates on motor-related frequency sub-bands and consists in six transformations for each sub-band. Group 2 concentrates in very low frequencies and consists in five transformations in the corresponding frequency band.

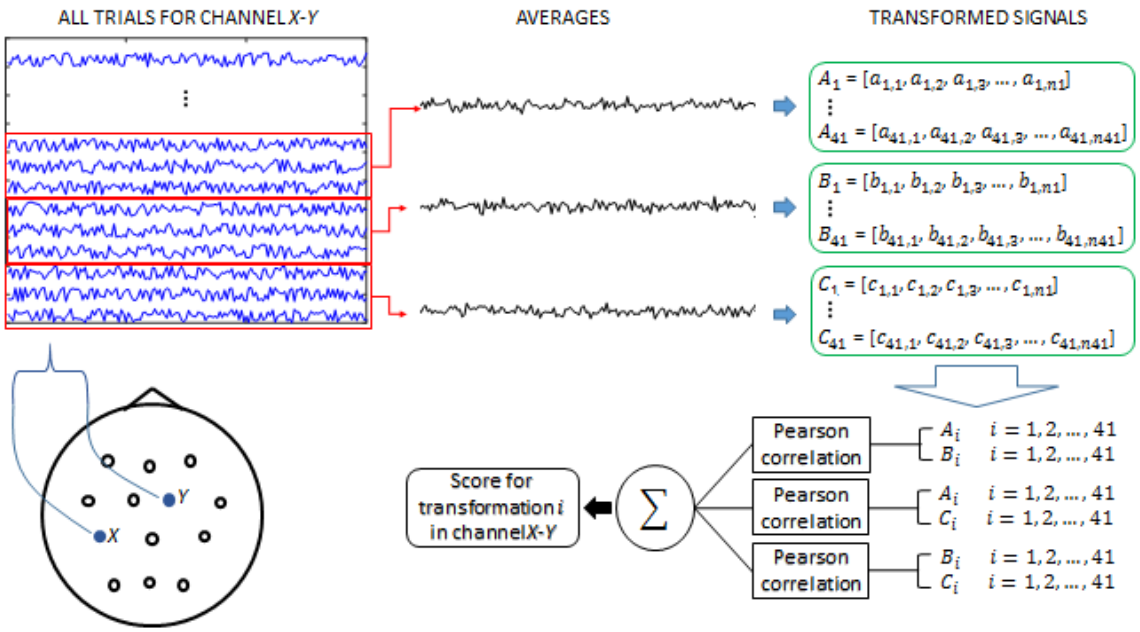


Source: Own authorship.

All these transformations are applied to average signals of every possible channel. For each channel, three sets of trials randomly picked (and without sharing any trial) are used to calculate three average signals. These three average signals are transformed each one into 41 new signals. For a specific signal transformation, the Pearson correlation coefficient ρ (see Equation 1, where x and y are the two signals for which the coefficient is being calculated and n is the number of samples in the signals) is calculated in pairs between the three transformed signals for one specific channel. The coefficients are summed and provides the score for that one transformation in that one specific channel. By doing so, it is expected to find features (which includes the channels) with lower inter-trial variability. See Figure 20 for a graphical explanation of these steps.

$$\rho = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} \quad (1)$$

Figure 20 - Steps executed for offline analysis in part 1 - approach 2, using EEG averages.



Source: Own authorship.

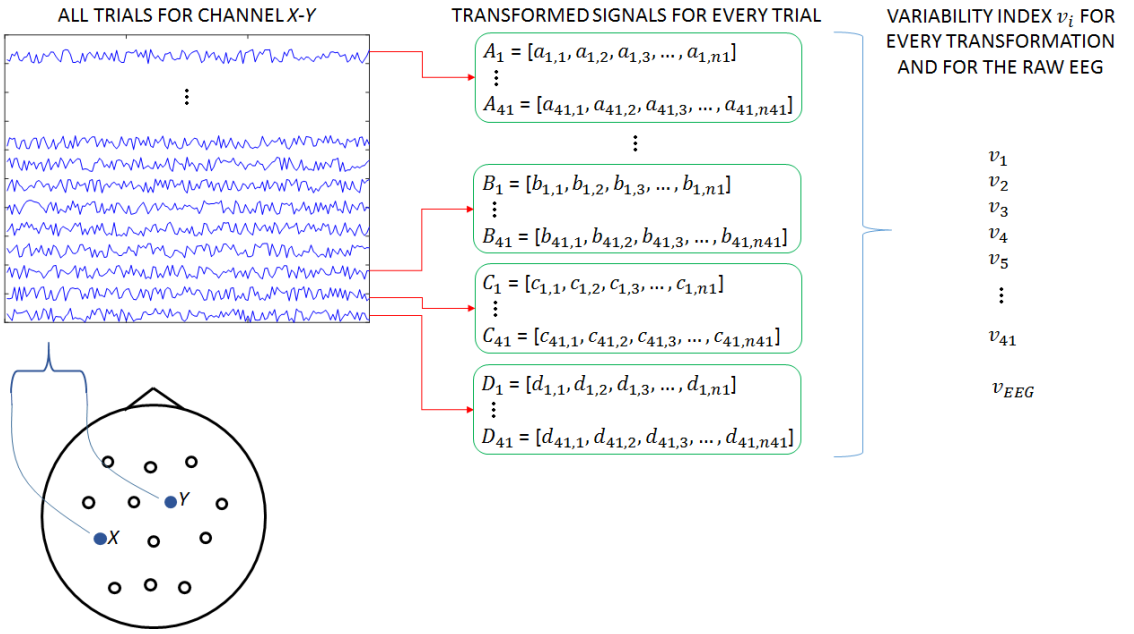
The procedure described in the previous paragraph is executed three times, which we refer to as rounds. The intention is to check the similarity between round's results as another strategy to guarantee that results are in fact overcoming intra-subject variability. The entire procedure is executed for different numbers of trials used to calculate the average signals. To define the number of trials to be considered, a preliminary test was performed. An algorithm was designed to calculate average signals starting with two trials and increasing the number of trials until the Pearson correlation coefficient between two consecutive averages (i.e. using only one trial more than the other) was higher than 0.9. This led to the number of 15 trials. Therefore, the first execution of the procedures uses 15 trials to calculate an average signal, then in the following execution it uses 7 trials (because it is an intermediate number between 15 and a single trial).

Finally, an execution is performed without using average signals. For each electrode pair, all transformations are applied to every single trial separately and an inter-trial variability index is determined for each transformation and for the original EEG signal (see Figure 21). The inter-trial variability index v (Equation 2) is calculated as described: the inter-trial standard deviation is calculated for each time instant and it is divided by the maximum amplitude difference in the entire

trial interval of the average signal. Then, the mean value of this ratio across the entire interval is calculated. This procedure is used instead of using the standard deviation alone because the transformations applied to the signals generate new signals with different units and scales. A natural alternative would be the coefficient of variation (ratio between standard deviation and mean), but it could be problematic because in some cases the mean would be very close to zero. Moreover, the mean value in some cases does not represent cortical activities, but some basal potential influencing the EEG oscillations or even might be scalp-electrode impedance artifacts, therefore it is not relevant at all.

$$v = \frac{1}{n}\sum_{i=1}^{n}\frac{\sigma_i}{\max(\bar{x}) - \min(\bar{x})} \tag{2}$$

Figure 21 - Steps executed for offline analysis of part 1 - approach 2, using single trials.



Source: Own authorship.

3.4. Part 2

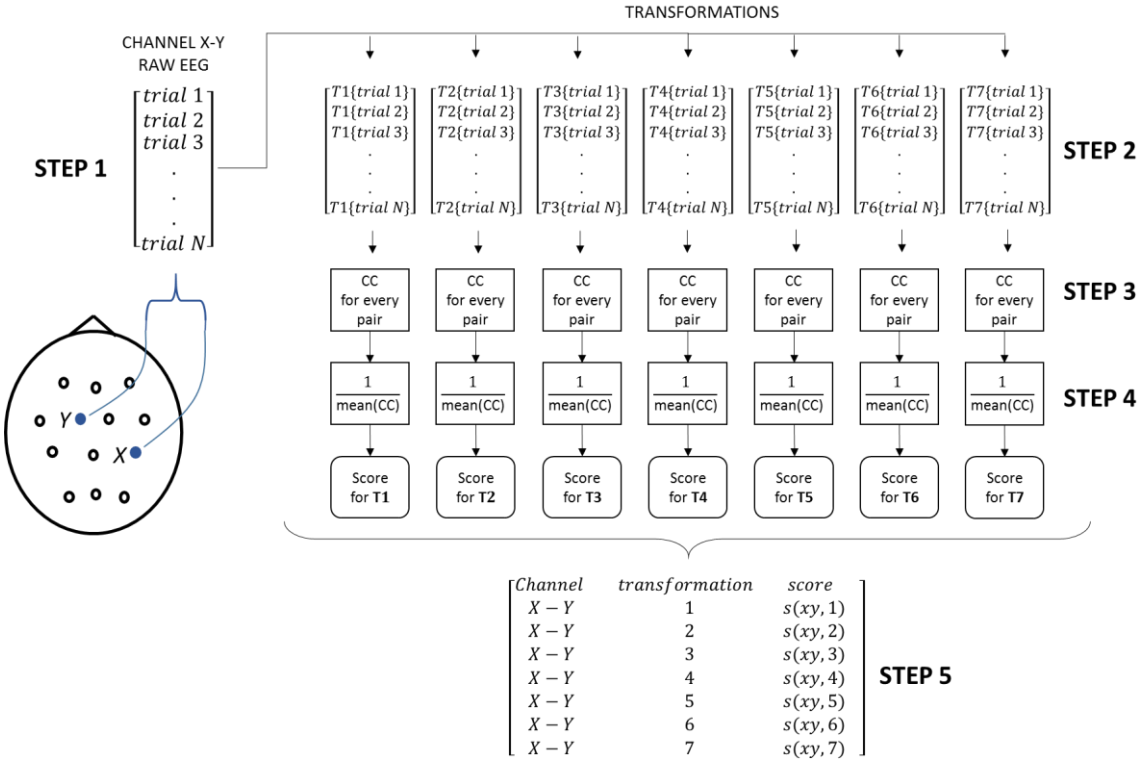
The methodology allows to find a subject-specific sub-set of channels and signal features that are stable over different trials of the same task, i.e. that have low inter-trial variability. The channels and their respective features are later tested in a pseudo-online classification algorithm.

The methods for Part 2 were proposed after analyzing Part 1 results. Also the implementation of the algorithms was considered for defining the strategies for Part 2, given that real-time (or pseudo-real time) signal processing cannot benefit from offline strategies in some cases.

3.4.1. Finding channels with low variability

All channels are considered, which means that every pair of electrodes is analyzed. All the procedures are executed separately for each subject. The variability (or its inverse, reproducibility) is evaluated by means of the Pearson correlation coefficient between two signals corresponding to different trials of the same task. Not only the raw EEG signal is analyzed, but also band-specific energy signals and a low-frequency subsampled EEG as well. These different signals are referred to as transformed EEG. Each transformed EEG signal and raw EEG signal of a single channel is analyzed and associated to a corresponding variability score to enter a ranking. The procedures executed are: 1) all the trials of a given channel are stored in a matrix where each row is a different trial; 2) all the transformations are applied to the EEG matrix, generating new matrices, one for each transformed signal and one for the original signal; 3) the Pearson correlation coefficients (CC) are calculated for all combinations two-by-two of the trials, separately for each matrix; 4) the inverse of the mean of all coefficients gives the variability score of the corresponding matrix (i.e. lower scores indicate low variability); 5) all the variability scores from the different matrices, with the corresponding transformations, are organized in a new matrix that will generate the final ranking. This procedure is repeated until all channels are analyzed. When step 5 is repeated, the new information is always concatenated in the same matrix, generating new lines. Figure 22 shows schematically the procedures for one channel from one subject.

Figure 22 - Procedures to rank the channels according to their inter-trial variability. Step 1 is segmenting the signal into task trials and storing them as lines in a matrix. Step 2 is transforming this matrix into 7 new matrices, each one representing a signal transformation. Step 3 is calculating the CC for every pair of trials separately for each matrix. Step 4 is calculating the variability score for each matrix (the inverse of the mean CC value). Step 5 is storing the information in a matrix that will later become the variability ranking after being ordered according to the scores. When a new channel is considered, step 5 allocates the information as new lines in the same matrix.



Source: Own authorship.

After all channels have been analyzed, the matrix is complete and then sorted to put the variability scores in ascending order, originating the final ranking. Each line of the ranking for each subject is associated to one channel, one transformation, and the corresponding variability score.

The transformations are:

1) energy signal in the 0.1-4 Hz band;

2) energy signal in the 4-8 Hz band;

3) energy signal in the 8-10 Hz band;

4) energy signal in the 10-12 Hz band;

5) energy signal in the 8-12 Hz band;

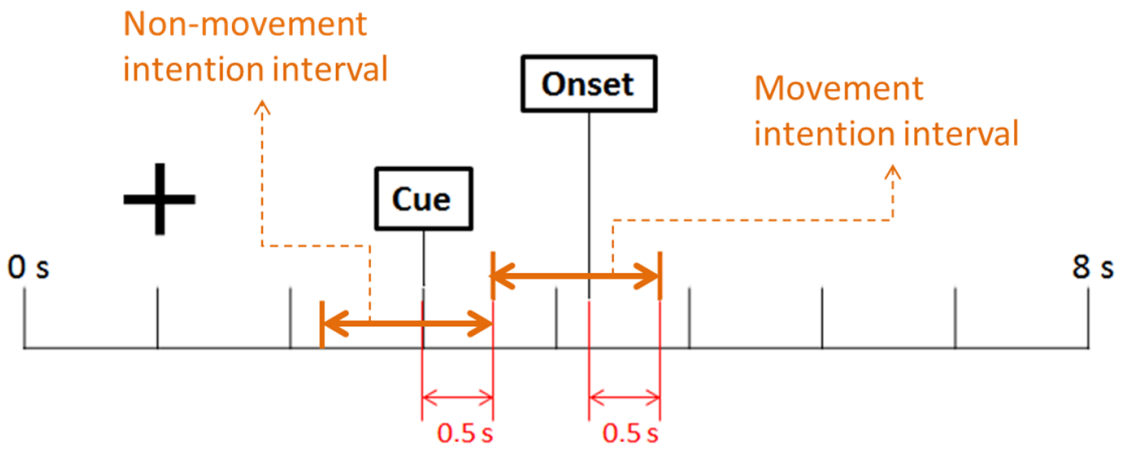
6) energy signal in the 12-25 Hz band;
7) EEG signal in 0-0.625 Hz band and subsampled to 16 samples per second.

Prior to the transformations being applied, the original raw EEG signal was subsampled to 128 Hz. Oscillating signals, even after being transformed into energy signals by squaring every sample, might lead to low CC values due to phase differences even if they have the same amplitude and frequency. Because the focus of this analysis is on the energy, the amplitude is more important than the phase. To attenuate misleading CC values due to phase differences, the energy signals are smoothed with a 0.5 s window prior to the CC calculation.

3.4.2. Classification

After building the ranking for each subject, the pseudo-online classification takes place to simulate an experiment with the intent to identify when the movement intention occurs during the trial. All the procedures are executed separately for each subject. Data from the elbow movement (flexion and extension) is used for calibration, and data from the hand (opening and closing) is used for testing the classifier. The reasons for using different movements for calibrating and for testing are: 1) both are executed with the same limb and same side (right); 2) it increases the number of points to be tested; 3) supposedly, this scenario should be more difficult than using data from the same movement. The interval within a trial that is considered to have movement intention starts 0.5 s after the movement instruction appears and finishes 0.5 s after the movement initiates. The reasons for considering a 0.5 s delay after the cue is that it was seem in a preliminary analysis that the subject usually does not start the movement before 1.0 s after the cue and the subject does not know in advance which movement will be indicated. The non-movement interval is immediately before the movement interval and selected so both classes have the same size (Figure 23).

Figure 23 - Task intervals used for each class: non-movement and movement intention. The interval that is considered to have movement intention starts 0.5 s after the movement instruction appears and finishes 0.5 s after the movement actually initiates. The non-movement interval is immediately before the movement interval and it is selected so both classes have the same size.



Source: Own authorship.

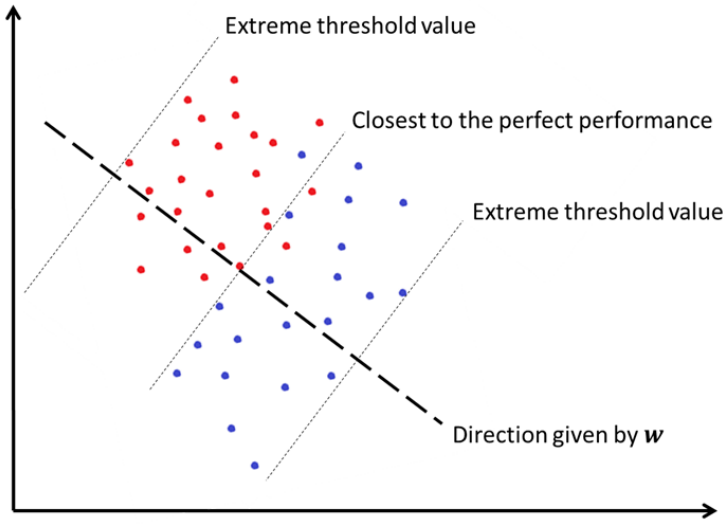
The feature vectors are simply the transformed signals in a 1 s length window. This window slides forward in time with a step equal to 0.0625 s (which is equivalent to 8 samples in the original EEG or only 1 sample in the subsampled EEG) to generate a new feature vector. This will lead to more than 2000 feature vector per class in most cases.

To execute the classification procedure there are some possibilities to be considered, such as neural networks, linear discriminants, linear regression, etc. Nevertheless, when only two categories are to be distinguished, a linear discriminant is the most suitable option. First, because its efficacy has been largely demonstrated in EEG classification. Second, because it is very simple to implement. Moreover, the intended contribution of this work is not in the classification procedure itself, but in the procedures that anticipate this step. Therefore, demonstrating the efficacy of the strategies proposed in this work using a simple classifier becomes more relevant. It is expected that a better performance can be achieved with other classification algorithms, such as artificial neural networks, for instance. In this work there are only two categories of EEG signals: movement intention or not. The Fisher's Linear Discriminant (FLD) is the technique selected for classifying the signals. The FLD and threshold calculation are explained in the following topic.

3.4.3. Fisher's Linear Discriminant

The FLD is a linear classifier that separates the input data into two classes. It consists in finding a projection vector $\boldsymbol{w}$ that maximizes the distance between the class average of the projected points and minimizes the intra-class variance. A threshold value must be defined to separate the two classes of the projected points. See Figure 24.

Figure 24 - Representation of the FLD calculation. Blue and red dots are two-dimensional inputs. Each color corresponds to a different class and each dot corresponds to an input vector. The FLD calculation consists of finding the vector $\boldsymbol{w}$ that gives the direction (dashed line) in which the inputs projections are best separated in their classes. The thin lines orthogonal to the thicker dashed line are examples of threshold values. The lateral lines are thresholds that puts all inputs in the same class, and the central line is the optimized threshold for this case.



Source: Own authorship.

In practical terms, we define $\boldsymbol{w}$, then we multiply $\boldsymbol{w}$ with the input vector (same dimensionality) to obtain a single scalar value. This value is compared to the threshold to see in which class it belongs.

The vector $\boldsymbol{w}$ can be calculated as described by (DUDA, RICHARD O. AND HART, PETER E. AND STORK, 2000). Only the main procedures to obtain the projection vector are described in this topic. For more details concerning the FLD calculation, see Appendix A. The calculation of the threshold of the discriminant is described in this topic as well.

First, the mean feature vector ($\boldsymbol{m}_i$) of each class (i) is calculated by (eq. 3)

$$\boldsymbol{m}_i = \frac{1}{N_i}\sum_{\eta=1}^{N_i}\boldsymbol{\delta}_\eta \tag{3}$$

where $i = 1$ or $i = 2$, $\boldsymbol{\delta}_\eta$ is the feature vector and N_i is the total number of feature vectors from class i.

The, the dispersion matrices $\boldsymbol{S}_i$ from each class and the matrix $\boldsymbol{S}_W$ should be defined as:

$$\boldsymbol{S}_i = \sum_{\eta=1}^{N_i}(\boldsymbol{\delta}_\eta - \boldsymbol{m}_i)(\boldsymbol{\delta}_\eta - \boldsymbol{m}_i)^T \tag{4}$$

and

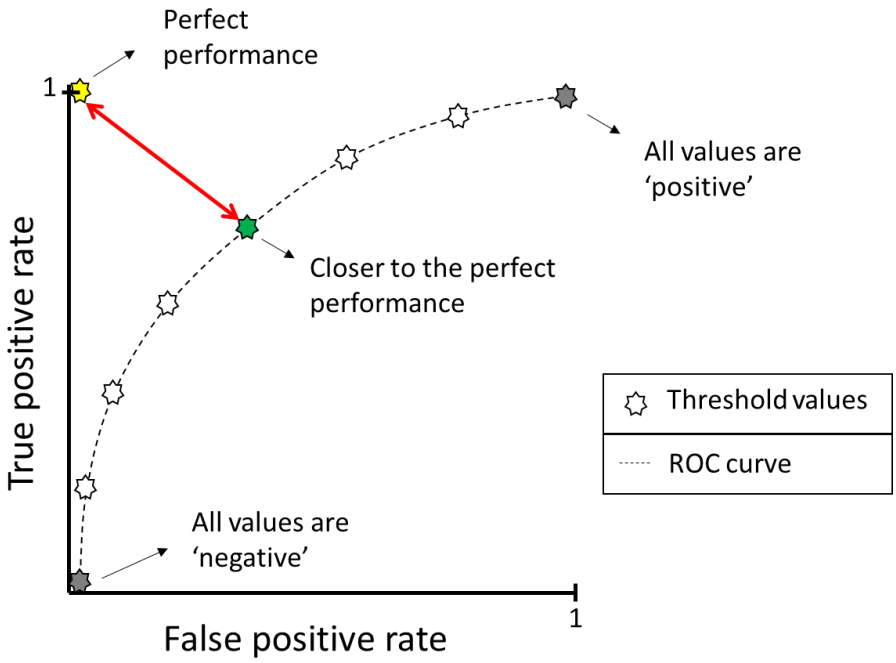
$$\boldsymbol{S}_W = \boldsymbol{S}_1 + \boldsymbol{S}_2 \tag{5}$$

respectively. The projection vector $\boldsymbol{w}$ is calculated by (eq. 6)

$$\boldsymbol{w} = \boldsymbol{S}_W^{-1}(\boldsymbol{m}_1 - \boldsymbol{m}_2) \tag{6}$$

Finally, we have to define the threshold. To do so, the Receiver Operator Characteristic (ROC) curve is used with the calibration data. The ROC curve is obtained by defining extreme threshold values that puts all projected points in the same class (one value for each class), as shown in Figure 24. Then, many different thresholds within these two values are tested. The ROC curve corresponds to all these threshold values (see Figure 25). The curve is plotted with the x-axis representing the false positive rate and the y-axis representing the true positive rate. ‘Positive’ in this case refers to a specific class. In this work, ‘positive’ would be when the movement intention is occurring. In this work, 50 different threshold values between the two extreme values were used. The threshold that provides the minor distance to the perfect result (100% true positive rate and 0% false positive rate) is chosen.

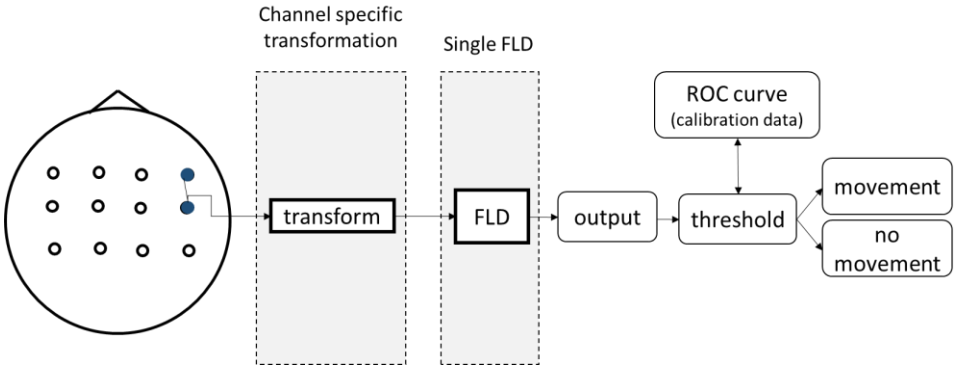
Figure 25 - ROC curve with eight threshold values (indicated by the stars). Vertical and horizontal axes are true positive rates and false positive rates, respectively. The gray stars are the extreme threshold values (represented by the lateral thin lines in Figure 24). The green star indicates the threshold value that would be chosen (represented by the central thin line in Figure 24), because it is the closest to the yellow star representing the perfect ideal threshold (this ideal threshold does not exist in Figure 24).



Source: Own authorship.

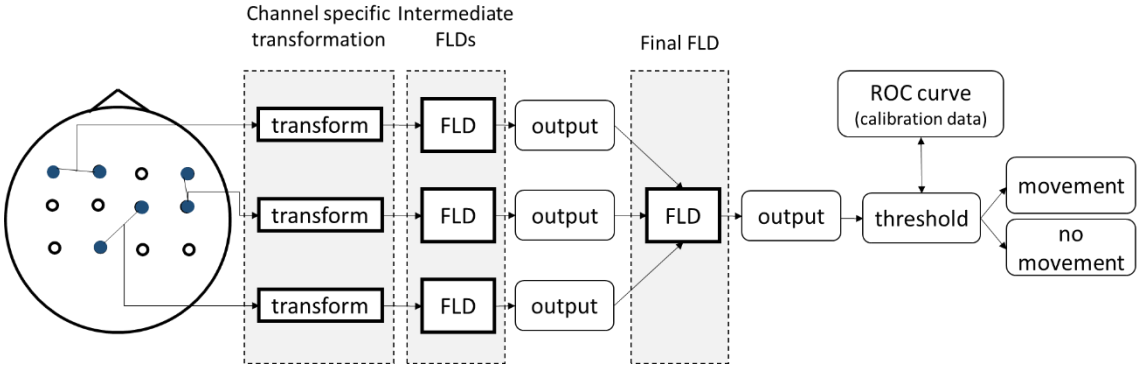
The classification is first performed using only the first channel of the ranking. Then, the number of channels is increased one-by-one until the first 10 channels of the ranking are used. Hence, there are 10 cases to be analyzed. When only one channel is used, there is only one FLD in the process (see Figure 26). When multiple channels are used, there is one FLD specific to each channel (in this case they are referred to as intermediate FLDs) and their outputs at a certain instant are used as the feature vector for one final FLD (see Figure 27). For the final FLD, the receiver operator characteristic (ROC) curve is built to define the threshold.

Figure 26 - Classification when only one channel is used.



Source: Own authorship.

Figure 27 - Classification when multiple channels are used. In this illustration, three channels are considered.



Source: Own authorship.

3.4.4. Scenarios for comparison

To enable a more complete analysis, three other scenarios are considered for comparison purposes. They are described in topics a) through c). All scenarios are summarized in Table 5.

a) Identical to the proposed method, except by the fact that only raw EEG is considered (no transformations are used). In Figure 22 it would be represented by skipping step 2, by having only one line for channel X-Y in step 5.

b) Twelve electrodes are selected to provide the signals for the classifier (FC1, FC3, FC2, FC4, C1, C3, C2, C4, CP1, CP3, CP2, CP4). These 12 electrodes are also used in three separated cases: 1) using the right mastoid (RM) as a monopolar reference; 2) with the SL filter (with all 61

electrodes); and 3) with the AR (with all 61 electrodes). Thus, there are 3 cases within this scenario. Each channel has their signals transformed to whatever transformation that provides the smallest variability. In Figure 22, it would correspond to executing all 5 steps separately for each of the 12 channels and selecting the transformation with the smallest variability for each channel.

c) The same as scenario b), but only raw EEG is considered (no transformations).

Overall, there are 26 scenarios for each subject, with 10 being for the proposed method, other 10 for scenario a), three for scenario b) and other three for scenario c). They are referred to as: transformed ranked channels (TRC), raw ranked channels (RRC), transformed literature channels (TLC), raw literature channels (RLT), respectively. See Table 5.

Table 5 - The four different scenarios analyzed in the present work.

Scenario	Abreviation	Channels considered	Reference scheme	Number of transformations considered	Channels used for classification	Results per subject
Transformed Ranked Channels	TRC	All channel pairs combinations	Bipolar	7	All quantities from 1 to 10	10
Raw Ranked Channels	RRC	All channel pairs combinations	Bipolar	None	All quantities from 1 to 10	10
Transformed Literature Channels	TLC	61 (original dataset)	SL AR RM	7	12	3
Raw Literature Channels	RLC	61 (original dataset)	SL AR RM	None	12	3

Source: Own authorship.

3.4.5. Surface Laplacian and Average reference

The calculation of AR and SL are briefly presented in this topic. For more details, see Appendix B and C. The AR is executed by simply subtracting from the raw EEG, at every instant of the recorded signal, the average potential across all channels. See (eq. 7).

$$V'_{t,i} = V_{t,i} - \frac{1}{N}\sum_{i=1}^{N} V_{t,i} \tag{7}$$

where N is the number of channels, $V_{t,i}$ is the electrical potential at instant t, channel i, and $V'_{t,i}$ is the new potential with the AR scheme. For the complete mathematical explanation about the AR, see Appendix B.

The SL, on the other hand, involves more mathematical procedures. In this work, the SL is applied as proposed by (PERRIN et al., 1989). In this case, splines are used for estimating a continuous spherical surface from discrete electric potentials measured with EEG electrodes, then the Laplacian operator is applied to the surface.

The full mathematical explanation is shown in Appendix C. In this section, only the essential steps for implementing the algorithm is shown.

To smooth the signals with splines, it is necessary to find a solution to minimize a function. After a few steps, we come to a linear system (WAHBA, 1981):

$$\begin{pmatrix} \boldsymbol{K} + N\lambda\boldsymbol{I} & \boldsymbol{T} \\ \boldsymbol{T}' & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{c} \\ \boldsymbol{d} \end{pmatrix} = \begin{pmatrix} \boldsymbol{V} \\ \boldsymbol{0} \end{pmatrix} \tag{8}$$

Where N is the number of EEG channels, $\boldsymbol{V}$ is a vector Nx1 with the EEG potentials at a certain instant in all N channels, and λ is the spline regularization parameter. The other variables are later explained. The first equation of the system gives the smoothed potentials $\boldsymbol{V}_\lambda$:

$$\boldsymbol{Kc} + N\lambda\boldsymbol{Ic} + \boldsymbol{Td} = \boldsymbol{V} \tag{9}$$

$$\boldsymbol{V}_\lambda = \boldsymbol{V} - N\lambda\boldsymbol{Ic} = \boldsymbol{Kc} + \boldsymbol{Td} \tag{10}$$

We can now apply the Laplacian to the smoothed signals. The transformed variables are indicated with the $\sim$:

$$Lap_S(\boldsymbol{V}_\lambda) = \widetilde{\boldsymbol{K}}\boldsymbol{c} + \widetilde{\boldsymbol{T}}\boldsymbol{d} \tag{11}$$

For geodesic distances and spherical coordinates in spherical splines (WAHBA, 1990), we have:

$$(\mathbf{T})_{\mathrm{i}} = 1 \tag{12}$$

$$\left(\widetilde{\boldsymbol{T}}\right)_{\mathrm{i}} = 0 \tag{13}$$

Then (eq. 11) becomes

$$Lap_S(\boldsymbol{V}_\lambda) = \widetilde{\boldsymbol{K}}\boldsymbol{c} \tag{14}$$

The vector $\boldsymbol{c}$ is defined as $\boldsymbol{c} = \boldsymbol{C}\boldsymbol{V}$. Re-writing Equation 14, we have

$$Lap_S(\boldsymbol{V}_\lambda) = \widetilde{\boldsymbol{K}}\boldsymbol{C}\boldsymbol{V} \tag{15}$$

By defining $\boldsymbol{L} = \widetilde{\boldsymbol{K}}\boldsymbol{C}$, the transformation of raw EEG signals $\boldsymbol{V}$ into the new signals $Lap_S(\boldsymbol{V}_\lambda)$ can be done with a linear transformation:

$$Lap_S(\boldsymbol{V}_\lambda) = \boldsymbol{L}\boldsymbol{V} \tag{16}$$

Therefore, it is necessary to find $\boldsymbol{C}$ and $\widetilde{\boldsymbol{K}}$ to obtain $\boldsymbol{L}$.

According to (WAHBA, 1990) (for spherical splines and spherical coordinates),

$$\boldsymbol{C} = \boldsymbol{Q_2}[\boldsymbol{Q_2}^T(\boldsymbol{K} + N\lambda\,\boldsymbol{I})\boldsymbol{Q_2}]^{-1}\boldsymbol{Q_2}^T \tag{17}$$

And

$$\left(\widetilde{\boldsymbol{K}}\right)_{ij} = -\frac{g_m(\boldsymbol{r_j}, \boldsymbol{r_i})}{r^2} \tag{18}$$

where $\boldsymbol{Q_1}_{N\mathrm{xM}}$, $\boldsymbol{Q_2}_{N\mathrm{x}(N-M)}$ and $\boldsymbol{R}_{M\mathrm{x}M}$ are matrices obtained with the QR decomposition of $\boldsymbol{T}_{NxM}$, which also includes the null matrix $\boldsymbol{O}_{(N-M)xM}$. The r is the head's radius, $\boldsymbol{r_i}$ and $\boldsymbol{r_i}$ are channel location vectors, and m is the spline flexibility. Also,

$$g_m(\boldsymbol{r}, \boldsymbol{r_i}) = \frac{1}{4\pi}\sum_{l=1}^{\infty}\frac{2l+1}{l^m(l+1)^m}P_l(\hat{\boldsymbol{r}}.\hat{\boldsymbol{r}}_{\boldsymbol{i}}) \quad m > 1 \tag{19}$$

where $P_l(.)$ are Legendre Polynomials with one variable and degree l. It is obtained with

$$P_l(a) = \frac{1}{2^l l!}\frac{d^l}{da^l}[(a^2-1)^l] \tag{20}$$

Due to a singularity, it is defined that $P_l(1) = 1$. To calculate the other values of the polynomial, Eq. (21) can be used (MEZIANI, 2016):

$$P_l(a) = \frac{1}{2^l}\sum_{l_0=0}^{l/2}\frac{-1^{l_0}(2l-l_0)!}{l_0!\,(l-l_0)!\,(l-2l_0)!}a^{l-2l_0} \tag{21}$$

The parameters m, r, l, and λ must be defined. As recommended by (KAYSER; TENKE, 2015), it is defined $\lambda = 10^{-5}$ and $m = 4$. The radius is $r = 0.15$ and, after preliminary tests, it was set $l = 15$.

4. RESULTS

This chapter is divided into two major parts corresponding to Part 1 and Part 2 of the methodology.

4.1. Part 1 – Approach 1

4.1.1. Grand averages

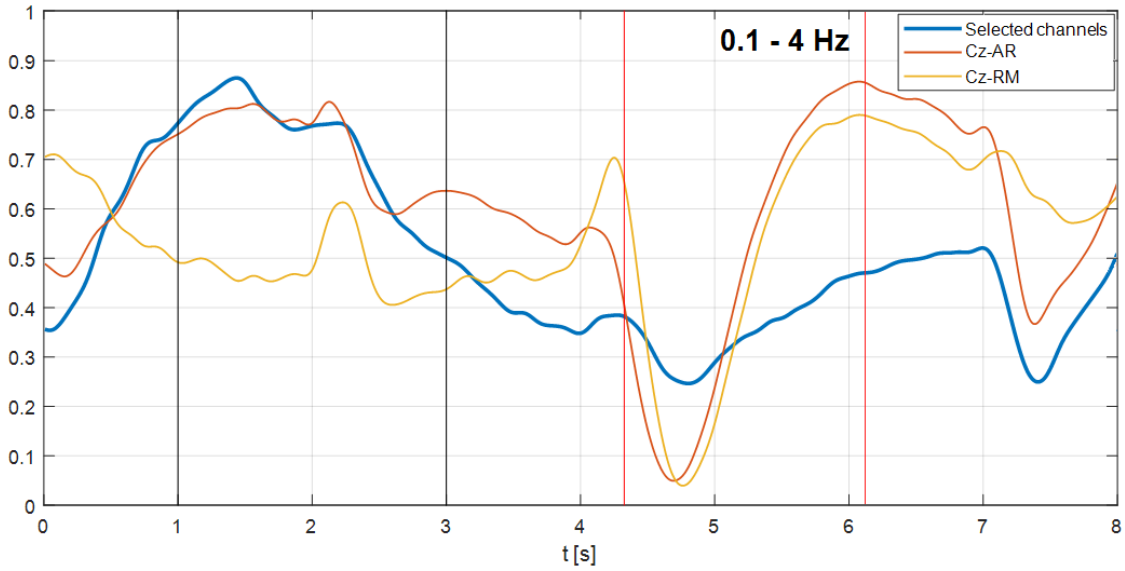
In Figures 28 and 29**,** the grand average signals corresponding to the task interval are presented. In these figures, the two solid vertical black lines indicate the instant in which the cross appears on the screen (t=1 s) and the instant in which the movement cue appears (t=3 s). The two solid vertical red lines indicate the average-over-subjects instants when the movement starts (first line) and when the movement reaches its maximum amplitude, immediately before initiating the returning movement. In general, the end of the movement was approximately at t=8 s, that is why it is not indicated in the figures. To generate the presented signals, the individual average signal of each subject was normalized between zero and 1. Then, the grand average was calculated across all subjects.

Figure 28 shows the grand average of the delta band signals obtained with the channels selected individually for each subject, and with the signals at Cz-AR, and Cz-RM channels, respectively (see Table 1). Prior to the grand average calculation, the individual signals were normalized between 0 and 1.

A few similarities among the three grand averages were noted. In every signal there is a negative peak after the movement onset (varying approximately between 0.40 s to 0.50 s post-onset), which marks what would be the end of the negative shift that characterizes the CNV or the RP. The signals then begin to increase at least until the second red line, which indicates the instant when the maximum movement amplitude occurs. Between this instant and the end of the trial window, a significant decrease in the signal takes place. With the grand average of the selected channels, the negativity occurs in one stage, starting approximately at t=1.50 s (0.50 s after the cross) and finishing approximately at t=4.80 s (0.50 s after the onset). The normalized potential decay is 0.60. In the grand average with Cz-AR, the negativity occurs in three stages, each of them starts and finishes at instants t=2.12 s and t=2.62 s (with 0.22 normalized

potential decay), t=3.00 s and t=3.87 s (with 0.10 normalized potential decay), t=4.05 s and t=4.70 s (with 0.51 normalized potential decay). In the grand average with Cz-RM, the negativity starts at t=0 s, but at approximately t=1.50 s it ends (with a 0.24 normalized potential decay), then the signal presents a very notable peak, followed by a slight increase until t=4.00 s, when another notable peak occurs immediately followed by a steep valley. This peak is at t=4.25 s (0.05 s prior to the movement onset), and the valley that follows is at t=4.75 s (0.45 s after the onset). The decay between the peak and the valley is approximately 0.67.

Figure 28 - The average signals are calculated with subject-specific selected channel, with channel Cz-AR and Cz-RM. Black lines indicate the instant in which the cross appears on the screen (t=1 s) and the instant in which the movement cue appears (t=3 s). Red lines indicate when the movement actually starts (first line) and when the movement reaches its maximum amplitude, immediately before initiating the return movement.



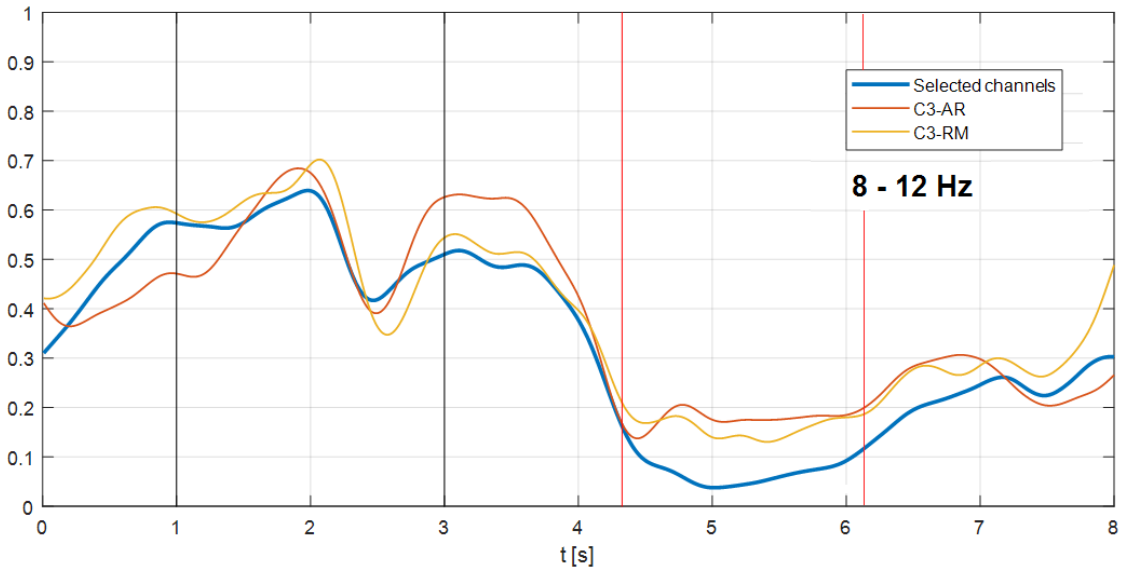
Source: Own authorship.

Figure 29 shows the grand averages of the alpha band energy signals obtained with the channels selected individually for each subject and the signals at C3-AR, and C3-RM channels, respectively (see Table 1). Prior to the grand average calculation, the individual signals were normalized between 0 and 1.

In all cases, within the 2 s interval with the cross on the screen, the energy rises, falls, and begins to increase again prior to the appearance of the movement cue. It then starts to decay significantly a little before half the interval between the cue

appearance and the movement onset. By the time it reaches the movement onset, the ERD is very clear. The energy remains considerably low until the subject's movement reaches its maximum amplitude (indicated by the second red line, see Fig. 29), then an increase in the energy level becomes notable.

Figure 29 - The average signals are calculated with subject-specific selected channel, with channel C3-AR and C3-RM. Black and red lines are as for Figure 28.



Source: Own authorship.

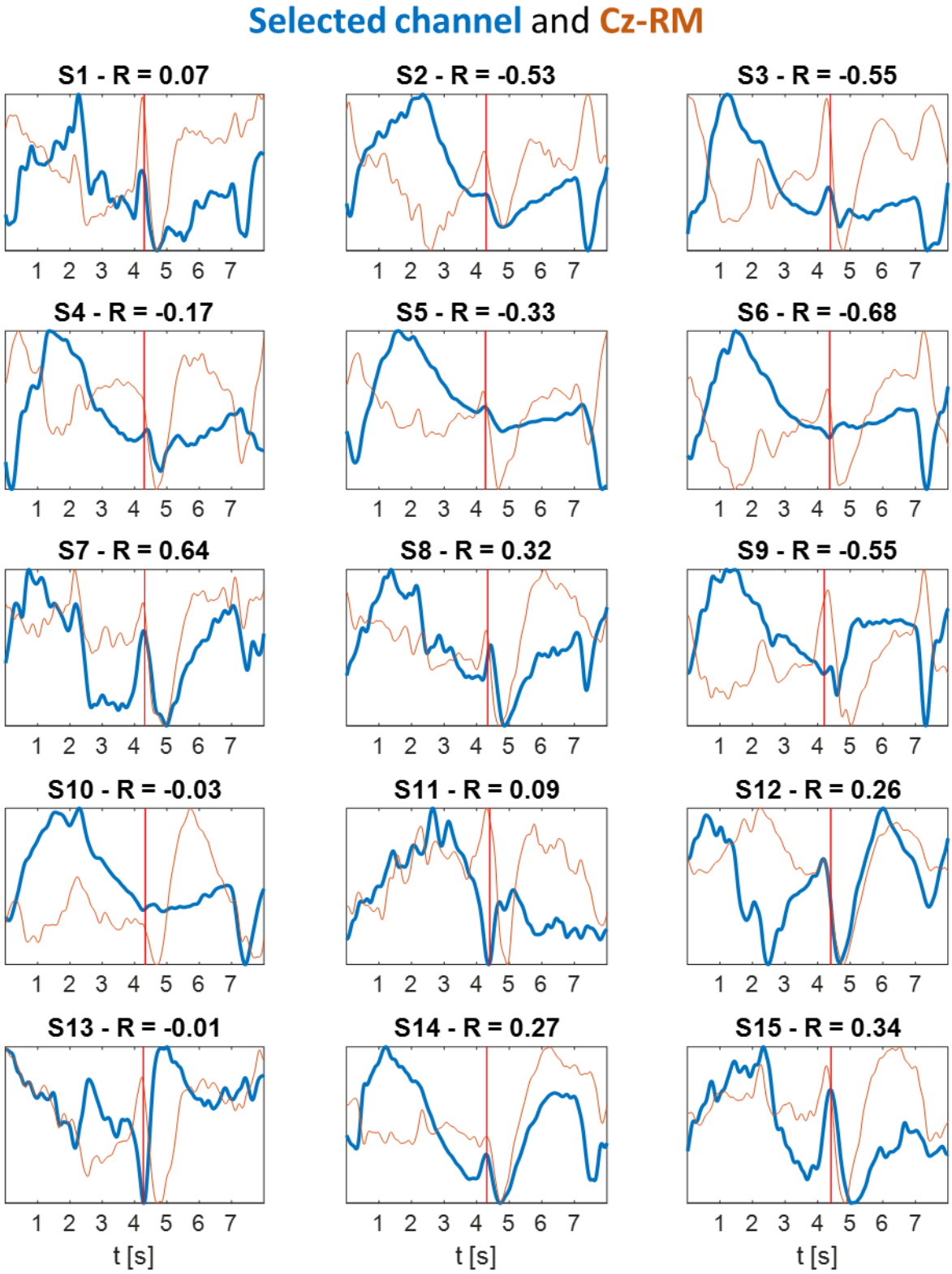
4.1.2. Individual averages

Figures 30-33 show the individual trial average for each subject from the individually selected channels using the proposed method and from the channels commonly used in the literature (Cz-RM and Cz-AR for the delta band – Figures 30 and 31; C3-RM and C3-AR for the alpha band – Figures 32 and 33). Before averaging the trials in each case, the signals were normalized between 0 and 1. The solid vertical red line indicates the movement onset.

In Figures 32 and 33, the largest anticipation of the ERD (considering the energy reduction of 75%) relative to the movement onset is 2.89 s (subject 6); the shortest anticipation is actually a delay, since it happened 0.07 s after the onset (subject 5). Analyzing the mean value, the average anticipation is approximately 1 s, in agreement with results reported in other works (e.g. (PFURTSCHELLER; ARANIBAR, 1979)). For the delta band averages (Figures 30 and 31), the valley after the movement onset has a mean delay of 0.38 s. The smallest delay is

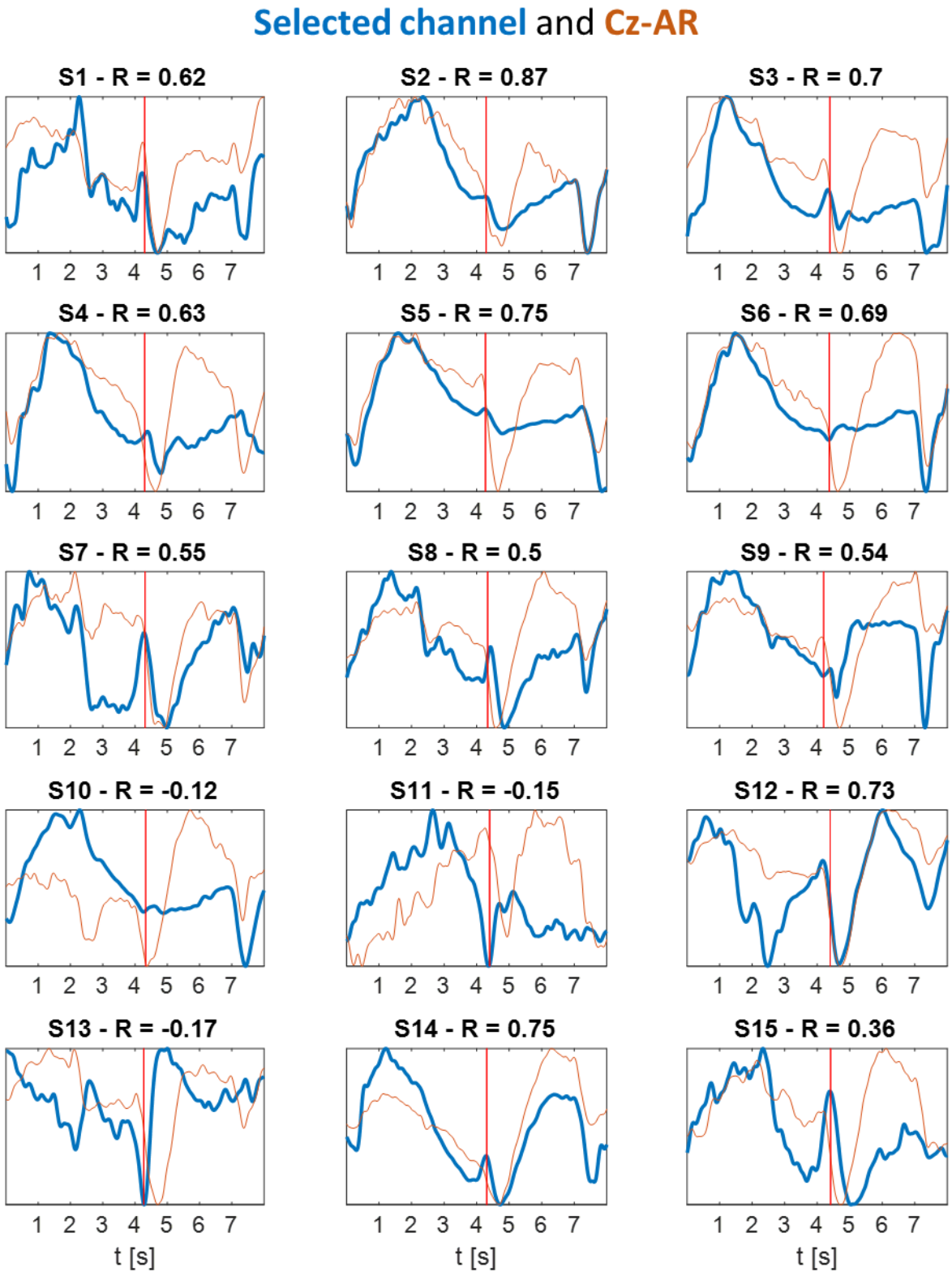
approximately coincident with the onset (0.01 s, subjects 6, 11, and 13), while the largest is 0.67 s (subject 7). When analyzing the Pearson coefficient (R) in the delta band signals, there is no high correlation (>0.90) between the signals in the selected channels and in Cz-RM and five for Cz-AR. In the alpha band signals, there are eight for the channel C3-RM and nine for C3-AR.

Figure 30 - Individual averages in the delta band with the selected channels (in blue) and Cz-RM (in orange). Letter S followed by 1 to 15 indicates the subjects and R is the Pearson correlation between the presented signals.



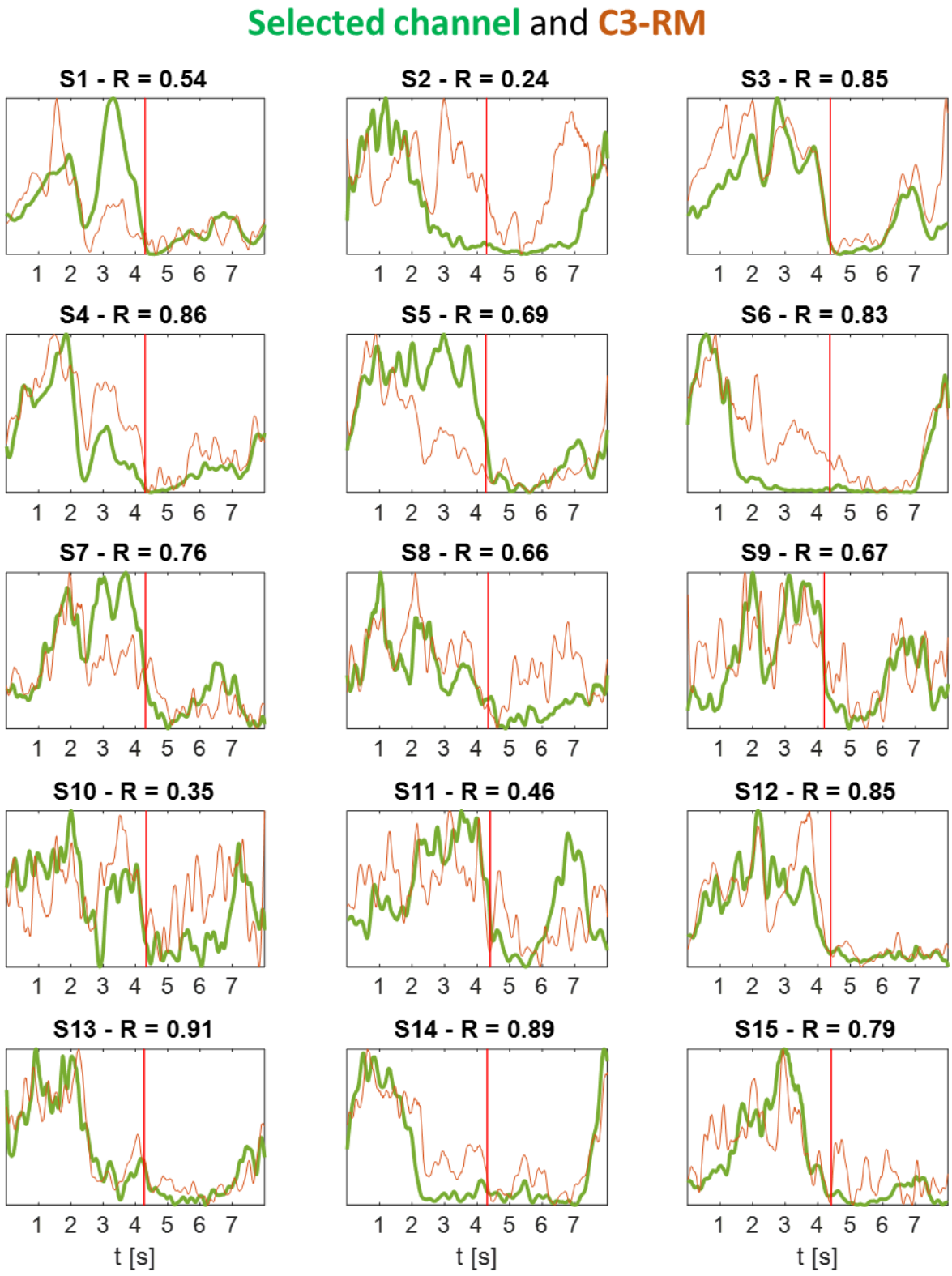
Source: Own authorship.

Figure 31 - Individual averages in the delta band with the selected channels (in blue) and Cz-AR (in orange). Letter S followed by 1 to 15 indicates the subjects and R is the Pearson correlation between the presented signals.



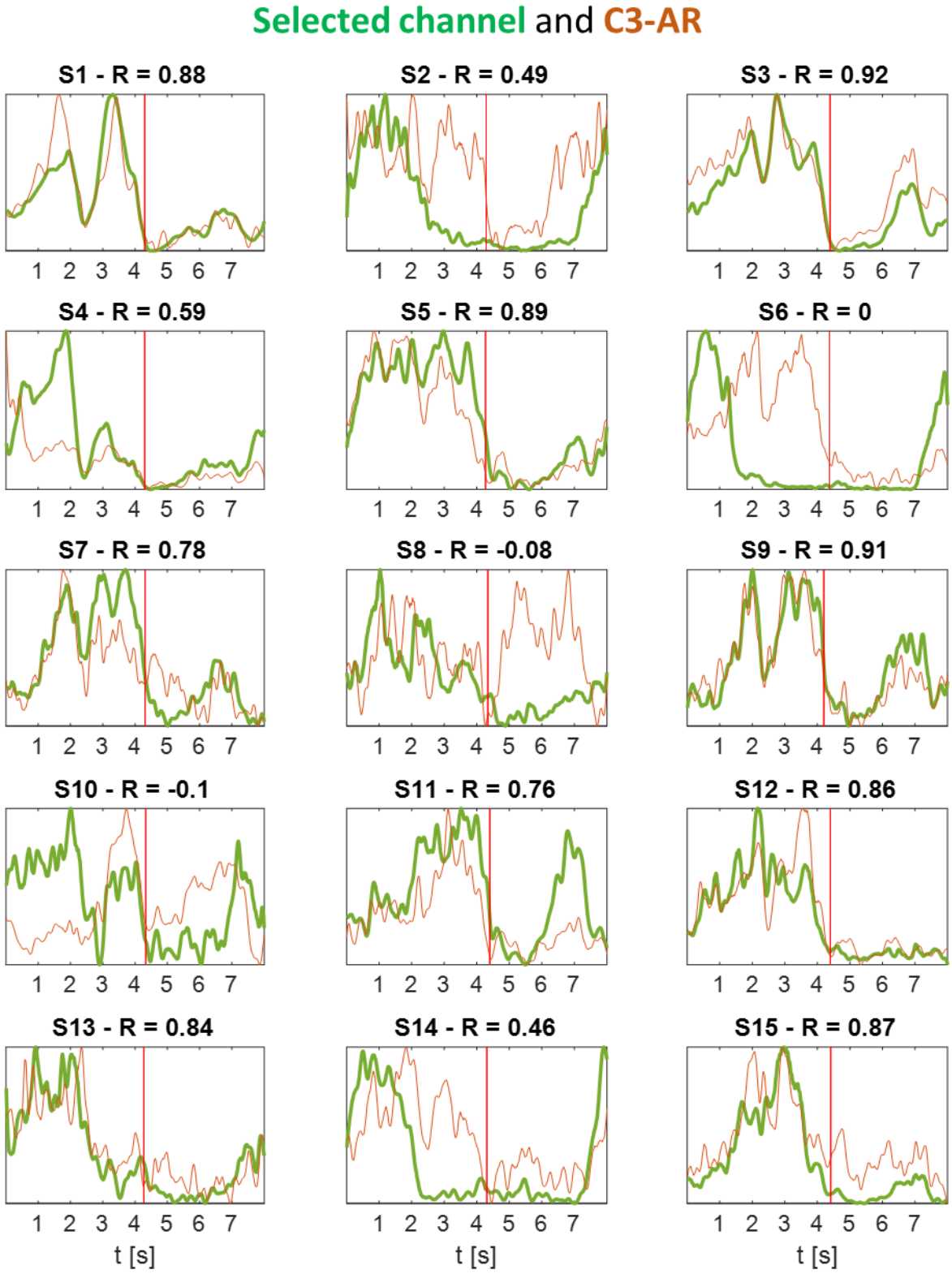
Source: Own authorship.

Figure 32 - Individual averages in the alpha band with the selected channels (in green) and C3-RM (in orange). Letter S followed by 1 to 15 indicates the subjects and R is the Pearson correlation between the presented signals.



Source: Own authorship.

Figure 33 - Individual averages in the alpha band with the selected channels (in green) and C3-AR (in orange). Letter S followed by 1 to 15 indicates the subjects and R is the Pearson correlation between the presented signals.



Source: Own authorship.

4.1.3. Correlation between subjects

The Pearson correlation coefficients for every pair of the individual trial average signals were calculated separately for the energy signal in the alpha band and for the signal in the delta band of the selected channels. In Table 6, the bottom triangle values are the correlation results for the energy on the alpha band; in the upper triangle of the table are the values for the delta band signals. Numbers from 1 to 15 refer to the subjects. The orange cells correspond to very high positive correlation values (from 0.90 to 1.00), yellow cells correspond to high positive correlation values (from 0.70 to 0.90), and not colored correspond to moderate, weak or negligible values (lower than 0.70). Numbers in red represent negative values. The coefficients indicate how similar are two signals compared to each other. If they are identical, the correlation coefficient is 1 (maximum value), if they are the exact opposite, the coefficient is -1 (minimum value), and if they share no similarities at all, the coefficient is 0. In Table 6, there isn’t any negative value out of the moderate, weak or negligible category.

Table 6 - All correlation values calculated between the signals of delta and alpha bands separately. Each value refers to a pair of subjects in a given frequency band. The bottom triangle values are the correlation results for the energy on the alpha band; in the upper triangle of the table are the values for the delta band signals. Numbers from 1 to 15 refer to the subjects. The orange cells correspond to very high positive correlation values (from 0.90 to 1.00), yellow cells correspond to high positive correlation values (from 0.70 to 0.90), and not colored corresponds to moderate, weak or negligible values (lower than 0.70). Numbers in red indicate negative values.

Alpha waves / Sub.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1		0.87	0.73	0.77	0.60	0.76	0.56	0.82	0.62	0.80	0.43	0.01	0.43	0.67	0.78
2	0.18		0.84	0.84	0.79	0.86	0.41	0.79	0.66	0.96	0.70	0.12	0.30	0.72	0.81
3	0.77	0.22		0.85	0.79	0.92	0.58	0.81	0.78	0.90	0.48	0.06	0.23	0.79	0.79
4	0.51	0.83	0.50		0.90	0.84	0.49	0.80	0.63	0.84	0.55	0.13	0.33	0.75	0.78
5	0.77	0.48	0.89	0.66		0.75	0.25	0.58	0.46	0.83	0.69	0.30	0.40	0.60	0.68
6	0.06	0.79	0.08	0.49	0.23		0.45	0.79	0.84	0.93	0.54	0.02	0.18	0.74	0.70
7	0.84	0.02	0.89	0.39	0.78	0.30		0.74	0.56	0.36	0.25	0.56	0.12	0.80	0.68
8	0.41	0.68	0.56	0.58	0.77	0.48	0.37		0.71	0.74	0.30	0.24	0.27	0.82	0.81
9	0.79	0.10	0.78	0.27	0.60	0.38	0.88	0.19		0.73	0.29	0.39	0.05	0.79	0.48
10	0.51	0.65	0.47	0.66	0.64	0.42	0.40	0.66	0.41		0.71	0.13	0.27	0.70	0.75
11	0.61	0.16	0.76	0.00	0.60	0.27	0.74	0.36	0.78	0.32		0.46	0.15	0.18	0.31
12	0.67	0.41	0.79	0.59	0.89	0.08	0.75	0.74	0.55	0.60	0.52		0.02	0.41	0.07
13	0.32	0.85	0.44	0.82	0.66	0.55	0.27	0.80	0.14	0.77	0.08	0.66		0.10	0.32
14	0.08	0.92	0.05	0.74	0.35	0.89	0.13	0.53	0.24	0.52	0.31	0.22	0.71		0.71
15	0.78	0.09	0.87	0.39	0.79	0.22	0.83	0.44	0.69	0.30	0.63	0.76	0.33	0.10	

Delta waves

Source: Own authorship.

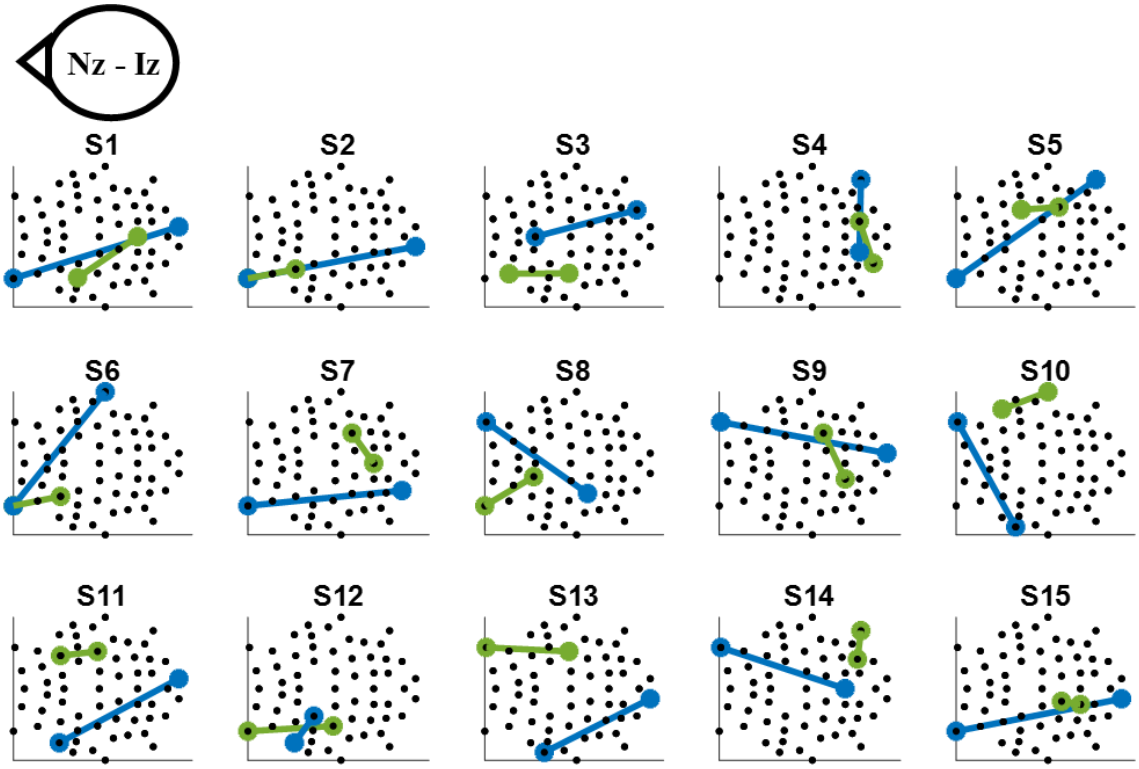
Table 7 presents the selected channels for each subject separated by the two frequency bands (delta and alpha), and Figure 34 shows the spatial distribution (with a view from the top) of the selected channels named in Table 6. The circles represent the available electrodes and the filled circles connected with a solid line (blue for delta band and green for alpha band) represent the channels.

Table 7 - Selected channels for each subject and each frequency band (delta and alpha).

Sub.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Delta	PPO2	PPO1	P2	P4	P4	TTP8h	P1	CPP3h	PPO2	FTT7h	FC5	FCC1h	FTT7h	CPz	PPO1
	F3	F3	FCz	CPP1h	F3	F3	F3	F4	F4	F4	PPO2	FC5	PPO1	F4	F3
Alpha	CPz	FC3	C3	P1	CCP4h	FC3	CPz	FC1	CP1	TTP8h	C4	C3	C4	P4	CP1
	FCC3h	F3	FFC5h	CPP2h	FCC2h	F3	CCP4h	F3	CCP4h	FC6	FC4	F3	F4	CCP4h	CCP1h

Source: Own authorship.

Figure 34 - Spatial representation of the selected channels for each subject and each frequency band. The blue circles are the electrodes that form the channel selected for the delta band, the green circles are the electrodes that form the channel selected for the alpha band. The lines connecting the filled circles represent the channels.



Source: Own authorship.

In the case of delta waves, there are five correlation values with very high positive coefficients (i.e. greater than or equal to 0.90). In general, they occurred between channels with notable spatial differences between them. Subjects 4 and 5 (P4-CPP1h and P4-F3): both share the first electrode, but the references are very far from each other. Subjects 6 and 3 (TTP8h-F3 and P2-FCz): the similarities are that both connect the posterior to the anterior parts (relative to the coronal plane),

the first electrode is in the right hemisphere and the reference is dislocated to the left. However, for subject 3 the channel is very close to the midline and closer to the inion, while for subject 6 it is in the extremities left-right and closer to the nasion. Subjects 10 and 2 (FTT7h-F4 and PPO1-F3): the references are in the same cortical line, but opposite to each other. First electrodes are in anterior and posterior parts, both in the left hemisphere, with one in the extremity and the other close to the midline. Subjects 10 and 3 (FTT7h-F4 and P2-FCz): one channel is in the anterior part with electrodes separated by two cortical lines in the left-right extremities, the other connects the posterior to the anterior part and is very close to the midline. Subjects 10 and 6 (FTT7h-F4 and TTP8h-F3): the references are in the same cortical line, but opposite to each other; first electrodes are close to the central cortical line (deviated one line towards the nasion and one line towards the inion) in the extremity of opposite hemispheres. They are almost perfectly inverted from each other in the left-right direction. Channels with notable spatial similarities, such as the ones for subjects 1, 2, and 15 (PPO2-F3, PPO1-F3, and PPO1-F3 again), presented lower correlation coefficients, but still in the 'high' range (0.87 for S1 and S2, 0.78 for S1 and S15, and 0.81 for S2 and S15). This also happened for S8 (CPP3h-F4) and S14 (CPz-F4) with 0.82 correlation, and for S5 (P4-F3) and S6 (TTP8h-F3) with 0.75 correlation.

When it comes to the alpha waves, only subjects 2 and 14 generated a very high correlation value. The channels are FC3-F3 and P4-CCP4h. Although both channels have electrodes very close to each other (which is the case for most channels), they have almost every other spatial characteristic different from one another: one is oriented in the inion-nasion direction and is located at front-left; the other is oriented in the right-left direction and is located at back-right.

Subjects 2 and 6 have the same channel (FC3-F3) and a high correlation value between their signals, but still lower than 0.80. Subject 8 relative to subjects 2 and 6 has correlation coefficients lower than 0.70, although subject 8's channel is very close (FC1-F3). Subjects 12 and 13 have channels with mirrored location in opposite hemispheres (C3-F3 and C4-F4), which could indicate the possibility of similar signals, since the ERD is very often seen bilaterally. Nevertheless, the correlation value is lower than 0.70.

4.2. Part 1 – Approach 2

A first attempt for part 1 – approach 2 is executed with the publicly available dataset from the Graz BCI group (Ofner et al., 2017) with EEG signals in an elbow flexion/extension activity performed by 15 subjects. Tables 4 and 5 shows the occurrence of each feature in the top 100 correlation results, all together for all subjects in the three rounds for the case with 15 trials generating averages and 7 trials generating averages respectively. Results from tab. 4 shows the first two transformations corresponding to motor-frequency bands (#3 and #6) with relatively close values of occurrence (21% and 18%). The following two have also relatively close values with each other and with the first two (18% and 15%), but in this case they are both low-frequency transformations (#39 and #41). They are followed by five other transformations related to motor-frequencies and one related to low-frequency. In tab. 5, the raw down-sampled EEG (#37) has 52% of occurrence, while the other nine greatest occurrences are all with 10% or less and correspond to seven motor-frequencies transformations and two low-frequency. In tab. 6 the results using single trials instead of averages are presented. Only two results among the highest 10 values are related to low-frequency transformations, one of them using 0.5 s windows (#38) and the other using 1.0 s windows (#39). Every other result uses motor-frequencies transformations and two 0.5 s windows with a gap in between, i.e. 1.5 s intervals to execute the transformation.

Table 8 - Most recurrent transformation among all rounds and subjects using 15 trials to calculate the average signal.

Transformation	Code	Occurrence
(0.5 s) 10-12 Hz energy	#3	21%
(0.5 s) 15-25 Hz energy	#6	18%
(1.0 s) ups & downs	#39	18%
(1.0 s) slope	#41	15%
(0.5 s) 6-8 Hz energy	#1	7%
(0.5 s) 6-8 Hz energy difference with 0.5 s gap	#25	5%
(0.5 s) 8-12 Hz energy	#4	4%
(0.5 s) down-sampled EEG	#37	3%
(0.5 s) 8-10 Hz energy	#2	3%
(0.5 s) 10-12 Hz energy difference with 0.5 s gap	#27	2%

Source: Own authorship.

Table 9 - Most recurrent transformation among all rounds and subjects using 7 trials to calculate the average signal.

Transformation	Code	Occurrence
(0.5 s) down-sampled EEG	#37	52%
(0.5 s) 10-12 Hz energy	#3	10%
(1.0 s) slope	#41	7%
(0.5 s) 6-8 Hz energy difference with 0.5 s gap	#25	6%
(1.0 s) ups & downs	#39	6%
(0.5 s) 15-25 Hz energy	#6	4%
(0.5 s) 6-8 Hz energy	#1	4%
(0.5 s) 8-10 Hz energy	#2	2%
(0.5 s) 10-12 Hz energy difference with 0.5 s gap	#27	2%
(0.5 s) 8-12 Hz energy	#4	1%

Source: Own authorship.

Table 10 - Most recurrent transformation using single trials.

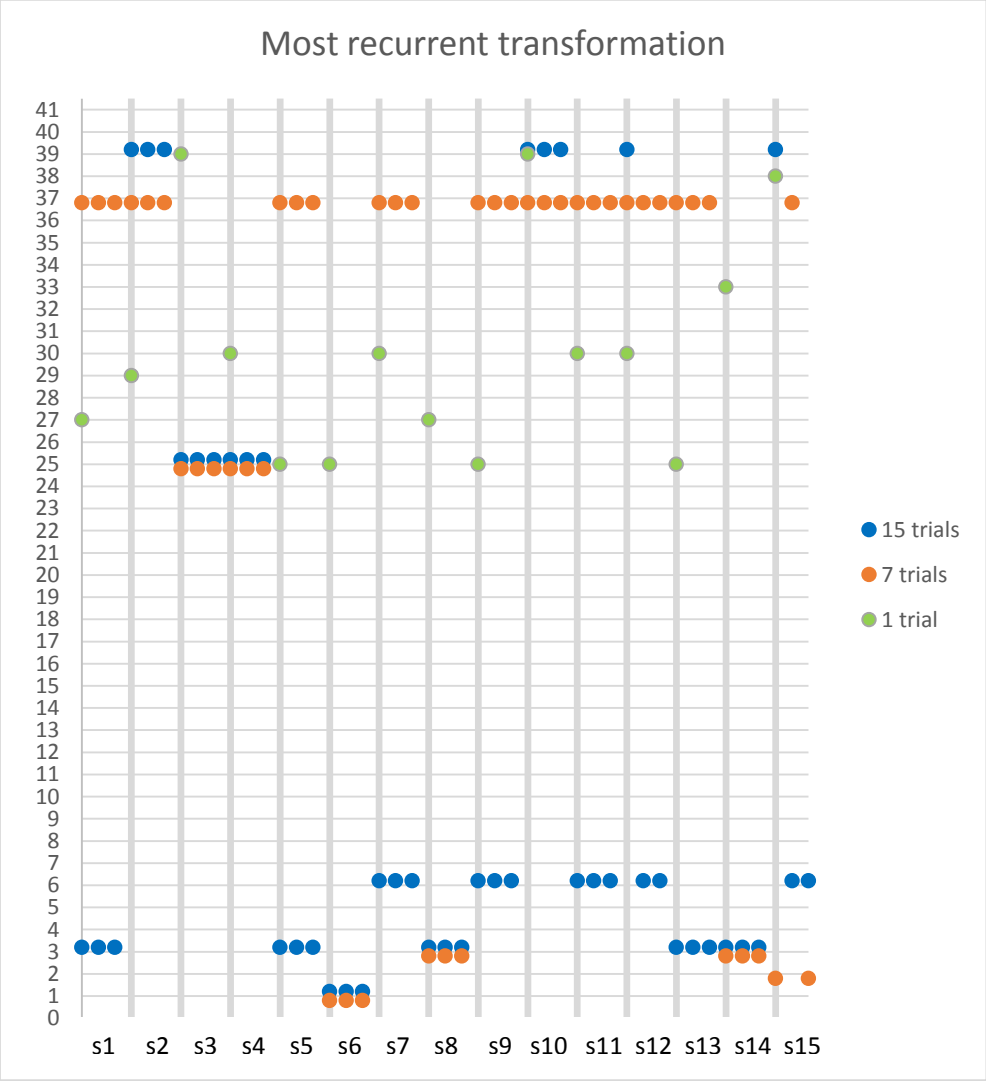
Transformation	Code	Occurrence
(0.5 s) 6-8 Hz energy difference with 0.5 s gap	#25	27%
(0.5 s) 15-25 Hz energy difference with 0.5 s gap	#30	21%
(0.5 s) 10-12 Hz energy difference with 0.5 s gap	#27	15%
(1.0 s) ups & downs	#39	10%
(0.5 s) ups & downs	#38	6%
(0.5 s) 10-12 Hz variance difference with 0.5 s gap	#33	6%
(0.5 s) 12-15 Hz energy difference with 0.5 s gap	#29	5%
(0.5 s) 8-12 Hz energy difference with 0.5 s gap	#28	4%
(0.5 s) 6-8 Hz variance difference with 0.5 s gap	#31	3%
(0.5 s) 12-15 Hz variance difference with 0.5 s gap	#35	1%

Source: Own authorship.

In Figure 35 the single most recurrent transformation in the top 100 results separated by subjects and rounds are shown for the cases with 15 and 7 trials being used to calculate the average signals. The cases in which the result for all three rounds is not the same are: subjects 12 (15 trials) and 15 (7 and 15 trials). They represent 10% (3 in 30). The case in which single trials are used and only one round is performed is also shown in Figure 35. The most recurrent transformation varies from #25 to #39 across all subjects. For subjects 3, 10, and

15 the transformations are related to low-frequency EEG (#39, #39, and #38, respectively). All the others use two 0.5 s windows with a gap in between.

Figure 35 - Most recurrent transformation separated by subjects and rounds. Blue dots refer to results using 15 trials to calculate the average, orange dots to 7 trials, and green dots to single trials.



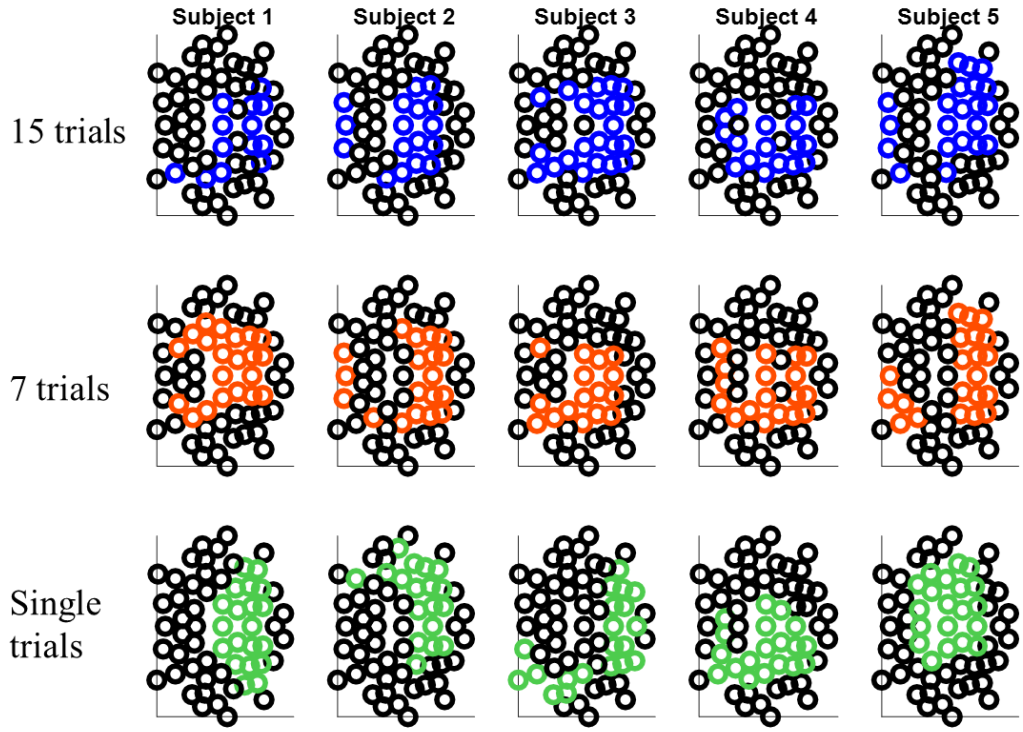
Source: Own authorship.

In Figure 36, Figure 37, and Figure 38, the colored channels are within the 3-dimensional space defined by the average of the electrodes positions summed with its standard deviation in all three directions. The electrodes are the ones in the top 100 results for round 1 when using 15 trials (blue circles) and using 7 trials (orange circles) to calculate the average signal, and the case where single trials are used (green circles). Two regions were calculated for each subject in each

figure, although they are not distinguished, one for the first electrode in the pair and the other for the second electrode.

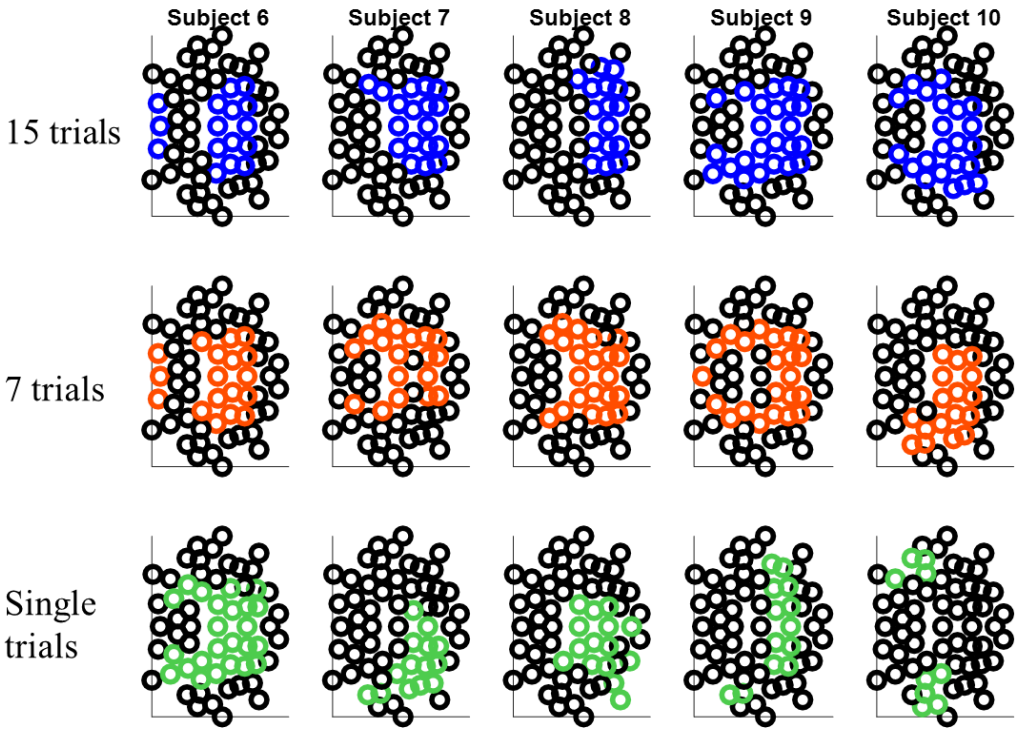
For almost every subject there is a lot of similarity between the electrodes obtained with 15 trials and with 7 trials to calculate the averages. In general, it seems that electrodes corresponding to the central, parietal and occipital regions are more present. Frontal electrodes appear also, but in a minor number. A qualitative analysis seems to show that frontal-central regions, especially near the midline, are absent in most subjects. The single trials case usually differs from the others results. However, some similarity can be seen in S4, S6, S13, and S14.

Figure 36 - Scalp regions (10-20 system) in the 100 best results of round 1 separated by subjects. Subjects 1 to 5. The nasion is on the left and the inion is on the right.



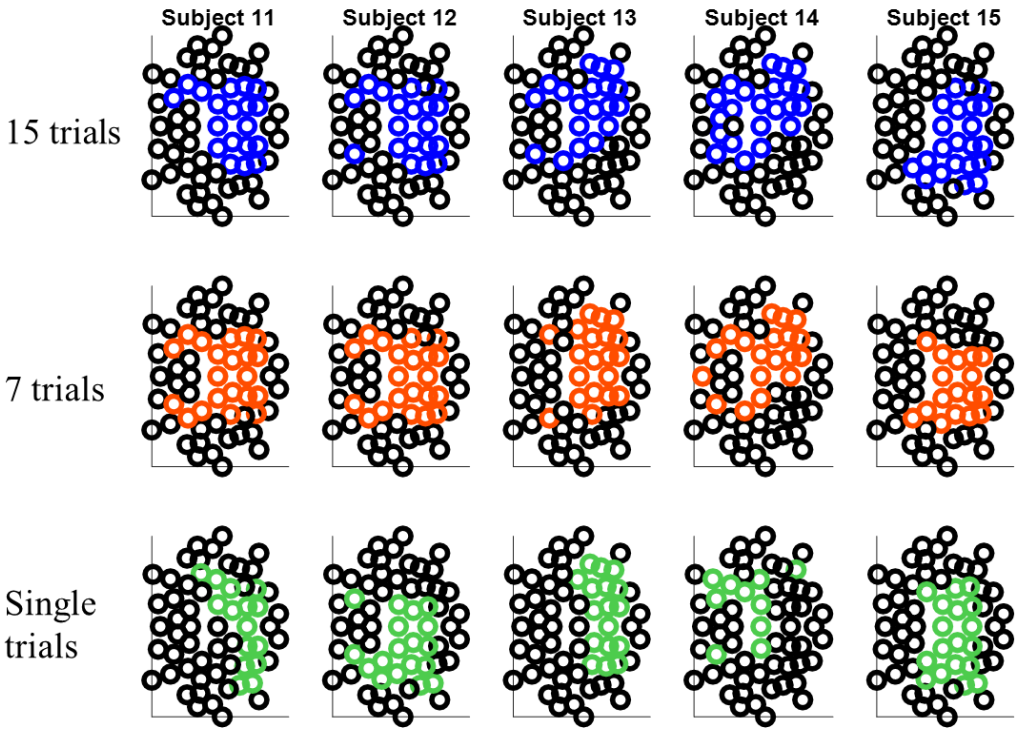
Source: Own authorship.

Figure 37 - Scalp regions (10-20 system) in the 100 best results of round 1 separated by subjects. Subjects 6 to 10.



Source: Own authorship.

Figure 38 - Scalp regions (10-20 system) in the 100 best results of round 1 separated by subjects. Subjects 11 to 15.

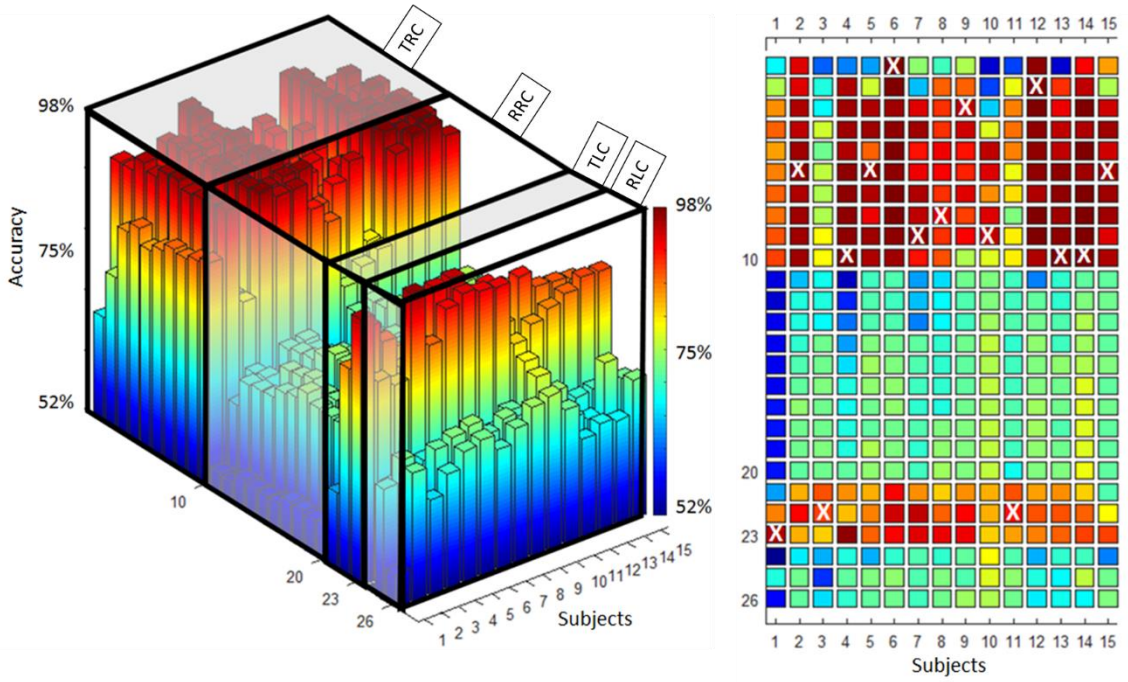


Source: Own authorship.

4.3. Part 2

Figure 39 shows all the accuracy results for the four scenarios analyzed with every variation within them (see Table 5) and for every subject. The left side of Figure 39 shows a 3-dimensional view of the results, and the right side shows the top view of the same object. The TRC and TLC scenarios contain the highest accuracy for all subjects (marked with an X in the figure to the right). The individual best result for each subject is with the TRC for 12 subjects, while the remaining three subjects (subjects 1, 3, and 11) have their best result with TLC. The minimum overall accuracy is 52%, for subject 1, RLC scenario, using SL; the highest accuracy is 98%, it occurred for subjects 4, 6, 12, 13, and 14, TRC scenario, using 10, 1, 2, 10, and 10 channels, respectively.

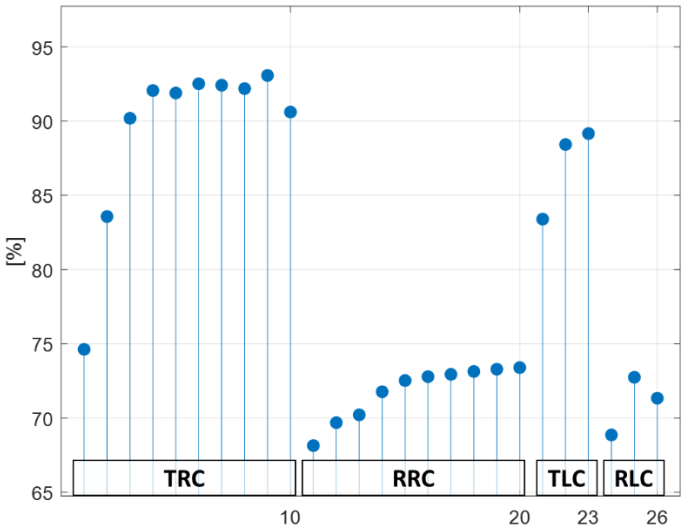
Figure 39 - Individual performances for each subject in all 26 scenarios. The box in the left shows the results 3-dimensionally, on the right it is the top view of the same figure. The best accuracy result of each subject is marked with a capital X in white. TRC: transformed ranked channels; RRC: raw ranked channels; TLC: transformed literature channels; RLT: raw literature channels.



Source: Own authorship.

Figure 40 shows the average accuracy across the subjects for each case. The lowest mean value is for RRC using 1 channel, and the best is for TRC using 9 channels. For the TRC, only when using 1 and 2 channels the mean value was not higher than 90%. Moreover, in no other scenario (i.e. RRC, TLC, RLC) the average achieved 90%. In the scenarios without transforming the signals (RRC and RLC), the average accuracies are always lower than 75%.

Figure 40 - Mean accuracy values over subjects for each case. TRC: transformed ranked channels; RRC: raw ranked channels; TLC: transformed literature channels; RLC: raw literature channels.



Source: Own authorship.

In Table 10 the best accuracy result with the proposed method (TRC) is presented for each subject, along with the corresponding number of electrodes and channels. Also, the best overall result (i.e. regardless of the scenario) is also presented for the three subjects that had their best accuracy in other scenarios (subjects 1, 3, and 11, all with TLC). The corresponding sensibility, specificity and precision are presented as well. The mean values and standard deviations (S.D.) of every parameter is presented at the bottom, always calculated considering the results with the proposed method, TRC. The best individual accuracy was 98% and occurred for five subjects (4, 6, 12, 13, 14). The lowest individual accuracy was for subject 3, 80%. The average number of electrodes and channels with TRC were 10 and 7, respectively, while the mean accuracy was 95%. For the subjects with the best result in TLC scenario, the reference choices were the RM (subject 1) and the AR (subjects 3 and 11). Thus, the total number of electrodes were 13 for subject 1, and 61 for subjects 3 and 11. When considering only the TRC scenario, the minimum number of electrodes and channels occurred for subject 6, with the lowest possible number: 2 electrodes and 1 channel. The maximum number of electrodes was 16, for subject 10, and the maximum number of channels was 10 (the highest possible), for subjects 1, 3, 4, 13.

Considering only the best results in TRC scenario, the accuracy was higher than or equal to 94% whenever the quantity of electrodes was higher than 11. With lower number of electrodes, there were three cases in which the accuracy was lower than 94%. These were using 11, 6, and 8 electrodes, for subjects 1, 3, and 11, respectively. These three subjects had accuracies lower than 90% with TRC and their best individual result occurred when using TLC. The number of channels, on the other hand, does not seem to be related in any means to the best accuracy for each subject. For instance, the best and worst individual accuracies using TRC happened both with 10 channels (subjects 3 and 4). Also, the lowest numbers of channels (1 and 2, for subjects 6 and 12, respectively) yielded the highest overall accuracy, 98%.

Table 11 - Best result for each subject. The * indicates that the average considers only the results in the TRC. The letters and numbers in italic refer to results in any scenario other than TRC.

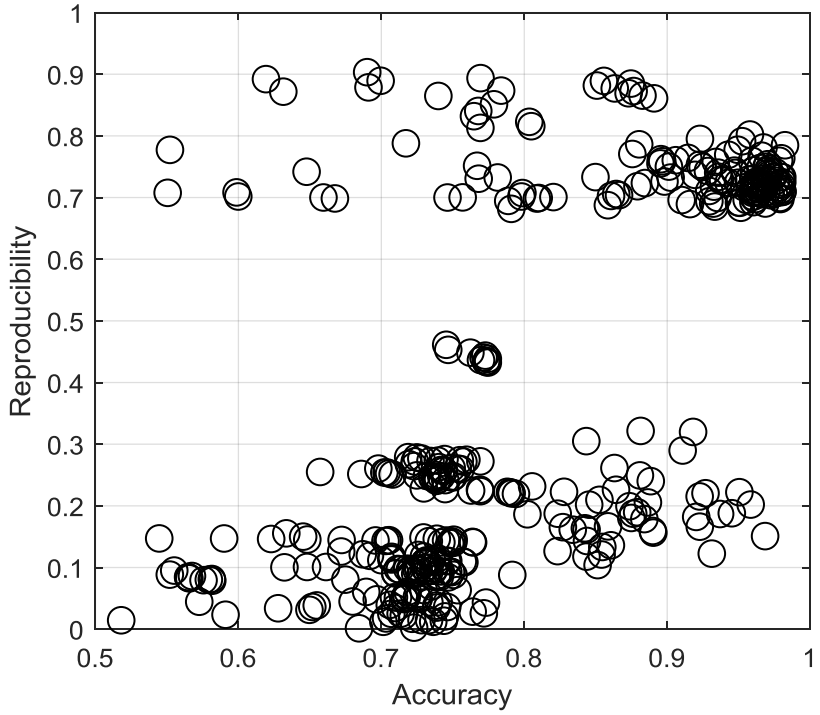
Subject	Electrodes	Channels	Accuracy	Sensibility	Specificity	Precision
1	11	10	89%	92%	86%	87%
	13 (RM)	*12*	*96%*	*96%*	*95%*	*95%*
2	9	6	97%	98%	95%	96%
3	6	10	80%	93%	68%	74%
	61 (AR)	*12*	*89%*	*89%*	*89%*	*89%*
4	18	10	98%	98%	97%	97%
5	7	6	97%	98%	96%	96%
6	2	1	98%	98%	99%	99%
7	13	9	97%	97%	96%	96%
8	13	8	94%	93%	94%	94%
9	5	3	95%	94%	96%	96%
10	16	9	95%	97%	93%	94%
11	8	4	87%	90%	84%	85%
	61 (AR)	*12*	*91%*	*88%*	*94%*	*94%*
12	4	2	98%	98%	98%	98%
13	14	10	98%	99%	97%	97%
14	13	10	98%	99%	97%	97%
15	12	6	97%	97%	96%	96%
Average*	10	7	95%	96%	93%	93%
S.D.*	4,3	3,2	5%	3%	8%	7%

Source: Own authorship.

Figure 41 shows the accuracy and its respective reproducibility value (the inverse of the variability) obtained in each scenario and in its internal variation for every subject. Therefore, every circle in Figure 41 corresponds to a bar shown in Figure 40. Because the reproducibility is channel-specific, in cases with more than one

channel, the reproducibility considered here is the mean value across the channels. Most of the circles are concentrated in two reproducibility intervals, 0-0.25 and 0.60-1.00. Most of them are also concentrated mainly in two accuracy intervals, 0.70-0.80 and 0.90-1.00. The lower accuracy region corresponds to the lower reproducibility region, and the higher accuracy region corresponds to the higher reproducibility region as well.

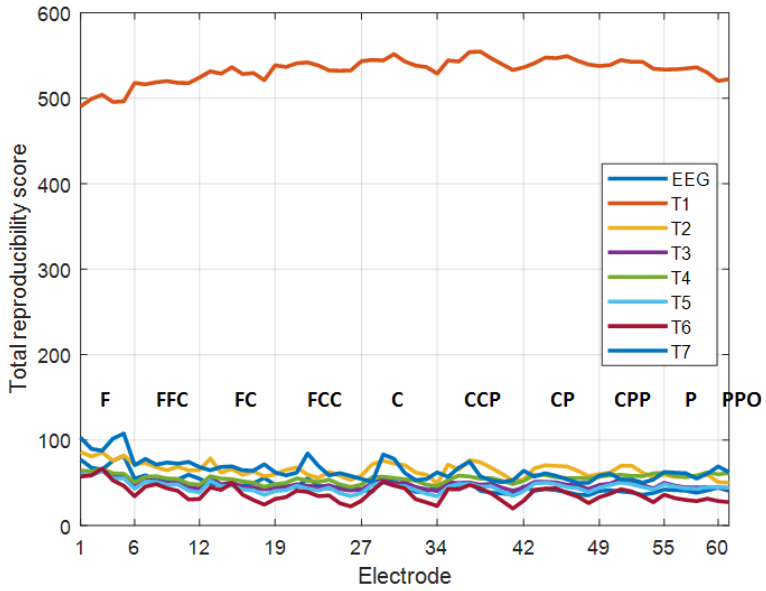
Figure 41 - The mean reproducibility value for each case and the corresponding accuracy. All subjects are included. Each circle represents a specific case and a specific subject (26 cases per subject).



Source: Own authorship.

Figure 42 compares the effect of the transformations on the reproducibility, separated by the electrode's position. The rankings of all subjects were analyzed together for every electrode. For a given transformation, whenever a certain electrode appeared in a channel, its reproducibility score was summed. The reproducibility for the raw EEG was also included in the figure. The transformation 1, which is the energy of the delta band, outstands all the others, including the raw EEG, for every scalp location. Also, it is possible to note an oscillating behavior of almost every graphic, with higher reproducibility near the midline positions in each cortical line, instead of positions close to the hemispheres' extremity.

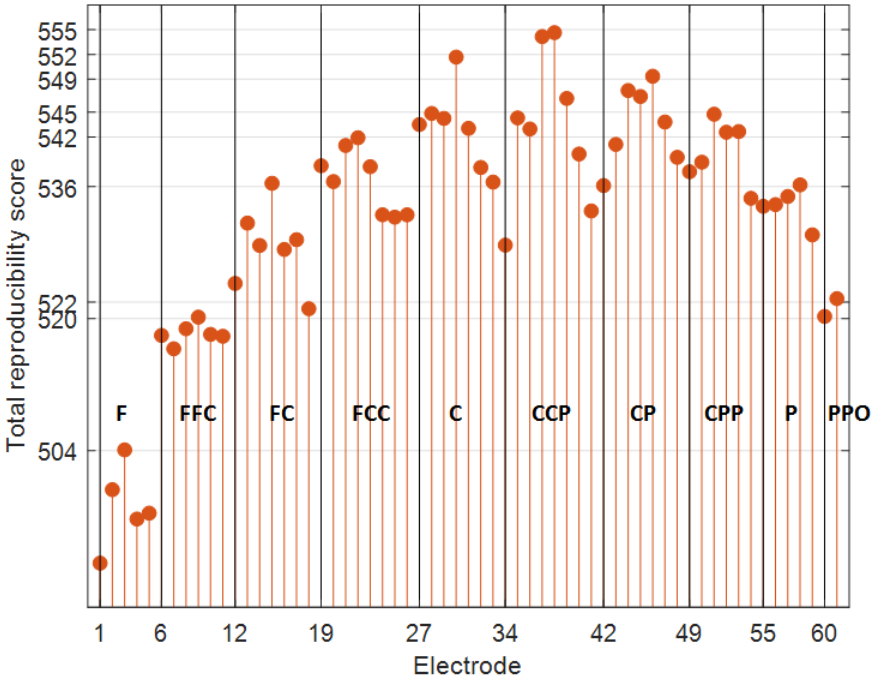
Figure 42 - Total reproducibility scores for each electrode position and each transformation, including raw EEG. T1 to T7 are the different transformations. The letters in the figure indicate the cortical line associated to the corresponding electrodes. For each line, the number of the x-axis represents the electrode position moving from the extreme left to the extreme right.



Source: Own authorship.

Figure 43 shows the delta band energy (T1) reproducibility alone with more details. It is the same result presented in Figure 42. In this figure, the oscillating behavior is more evident. The reproducibility scores are always greater when the electrode position is far from the hemispheres' extremity in every cortical line. The general behavior of the graphic shows that higher reproducibility scores are associated to the CCP line. The other cortical lines seem to have reproducibility scores gradually lower as they become further away from CCP. In fact, a visual inspection suggests that the values are symmetrical relative to the CCP line. Thus, it does not matter if the region is closer to the nasion or to the inion, but rather how distant they are from CCP.

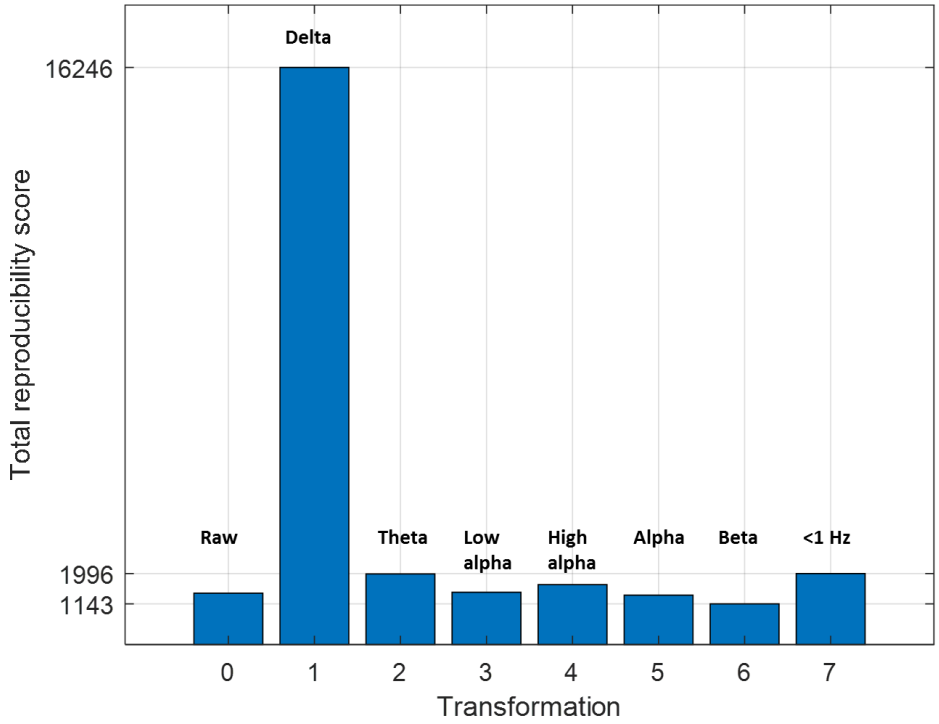
Figure 43 - Same as Fig. 42, but zoomed in to the delta band energy information. Instead of a continuous line to represent the values, in this figure stems are used for a better view of the different electrode positions. The values in the y-axis correspond to the maximum of each cortical line.



Source: Own authorship.

The total reproducibility of each transformation disregarding the scalp positions is shown in Figure 44. The delta band total reproducibility score (16246) is more than eight times the second-best score (1996 - very low frequencies of sub-sampled EEG). The lowest score was for beta band energy, with 1143. When considering only the energy signals in the traditional EEG frequency bands, an inverse relation is observed between the frequency and the reproducibility. Ascendant frequency bands, delta (>4 Hz), theta (4-8 Hz), alpha (8-12 Hz), and beta (12-25 Hz), have descendent reproducibility scores (16246, 1986, 1386, and 1144, respectively). However, the alpha sub-bands did not follow the same rule. Low alpha (8-10 Hz) has less reproducibility than high alpha (10-12 Hz), 1470 and 1686, respectively.

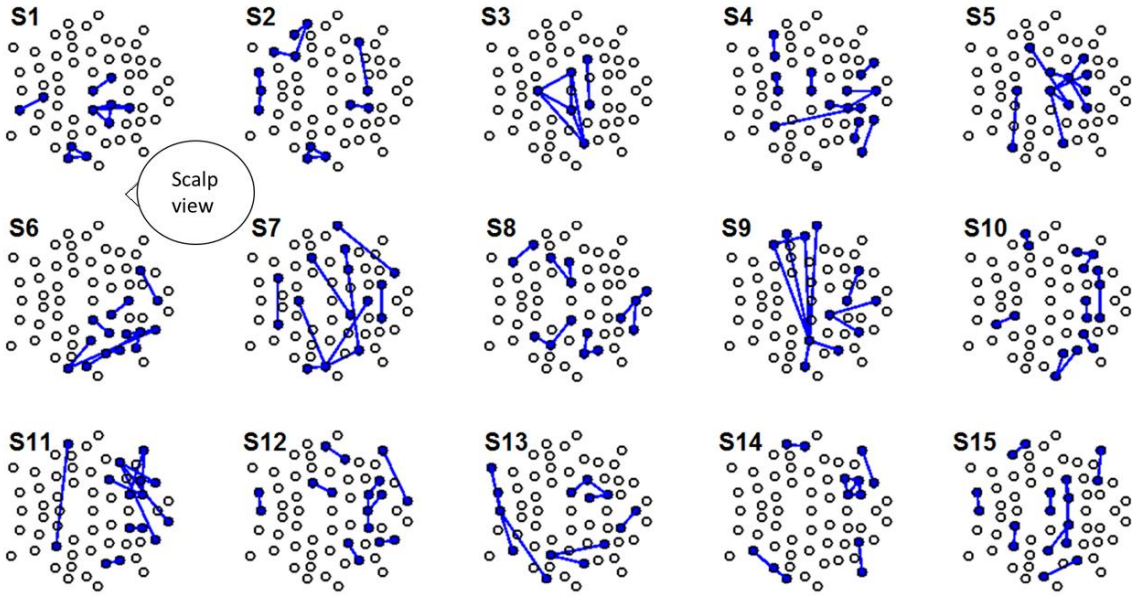
Figure 44 - Total reproducibility score separated by transformation and raw EEG.



Source: Own authorship.

Figure 45 shows the first 10 channels ranked for each subject (TRC). By visual inspection, one can see that the spatial distribution differs considerably across different subjects. However, channels with electrodes relatively close to each other seem to be the general pattern, with a few exceptions. For instance, only 29 of the 150 channels (19.3%) connect cortical lines that are not adjacent to each other (according to the extended 10-20 system). Also, only 6 channels connect a left or right extremity to the opposite hemisphere, with 4 of them being in subject 9's ranking, 1 in subject 7's ranking and 1 in subject's 11 ranking.

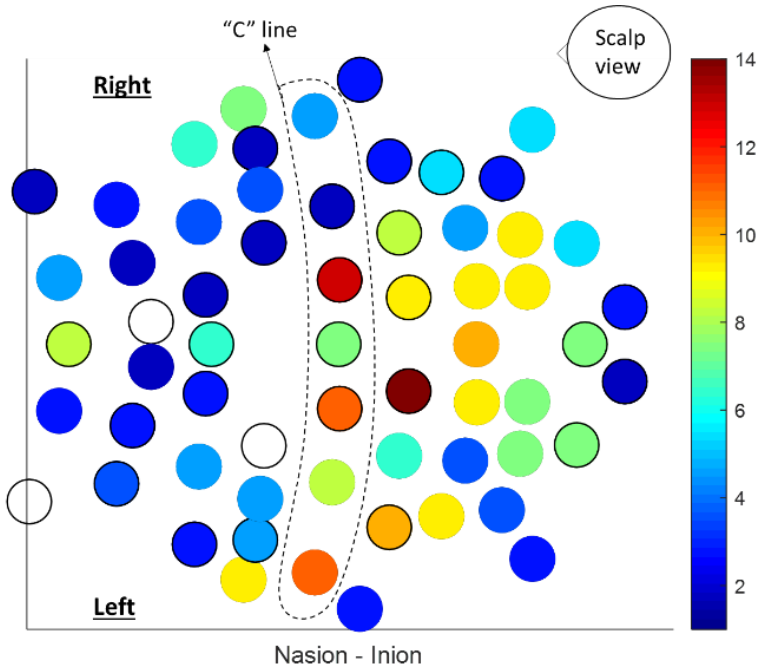
Figure 45 - The top 10 channels using transformed EEG (TRC) and separated by subject. The channels are represented by the connection between two electrodes. In all scalp maps, the nasion is to the left and the inion is to the right.



Source: Own authorship.

A visual representation of the incidence of each electrode in the top 10 channels from all subjects together is shown in Figure 46. Each electrode position could appear up to 150 times (10 times for each of the 15 subjects). The electrode that appeared most is CCP1h, with 14 appearances. Only three electrode positions did not appear at all, all of them in frontal regions (F3, FFC2h, and FCC1h). Overall, there is a predominance of the colors representing more than 8 appearances in the central and central-parietal regions (lines C, CCP, CP, and CPP).

Figure 46 - Visual representation of the incidence of each electrode position in the top 10 channels of all subjects together. The number of times the electrode appeared is indicated by its color in the scalp map, according to the colorbar in the figure.



Source: Own authorship.

5. DISCUSSION

Part 1 aims to obtain robust information (features and channels) about EEG signals related to specific motor tasks and specific for each subject. So far, the EEG data analyzed is the publicly available dataset provided by the Graz BCI group. Part 1 is divided into two approaches. Approach 1 attempts to detect the CNV and the ERD phenomena by searching every channel with re-referencing procedure. Approach 2 also applies the re-referencing procedure and in each case transforms raw re-referenced EEG into 41 new signals and attempts to identify the transformation that leads to a more stable signal across different trials of the same task, and to which channels it should be applied. Part 2 aims to improve classification accuracy in a simulated BCI by reducing EEG intra-subject variability.

5.1. Part 1 – approach 1

In this work we tested the identification of EEG motor signatures across different electrodes and references. We found the motor signatures of CNV/RP and ERD with some electrode location patterns across subjects. However, it became apparent that subject-specific electrode choice could improve the motor signatures detection.

5.1.1. Delta band grand averages, the RP and the CNV

The grand average using the channels selected individually for each subject for the delta band shows a negative shift that begins 2.75 s before the movement onset and reaches its peak approximately 0.50 s after the movement onset. This timing, although it has some differences compared to the RP reported by Kornhuber and Deecke (KORNHUBER; DEECKE, 1965), it agrees with other RP works, as shown in Table 11. When it comes to the CNV, our results also show agreement with other related works, as shown in Table 12. Therefore, the temporal marks observed using the channel selection method we presented seems to agree with important timing characteristics of both EEG signatures, the CNV and the RP.

Table 12 - Timing features of the RP. The minus sign indicates that it refers to the interval before the movement onset.

Negative shift with respect to movement onset	Our results	(KORNHUBER; DEECKE, 1965)	(SHIBASAKI; HALLETT, 2006)	(LEW et al., 2012)
Start	-2.75 s	-1.50 s	< -2.00 s	
Negative peak	0.50 s	0.09 s		> 1.00 s

Source: Own authorship.

Table 13 - Timing features of the CNV. Where it says, “prior to movement onset”, the movement had a delay of less than 0.50 s relative to the second stimulus.

	Our results	(WALTER et al., 1964)	(KROPP et al., 2000)	(MASAKI et al., 2012)
Peak after movement onset	1.75 s after the second stimulus	Prior to movement onset	2.00 s after the second stimulus	Prior to movement onset

Source: Own authorship.

In terms of the spatial distribution, the channels selected in the present work show a tendency to connect posterior and anterior scalp positions. This agrees with the RP presented in (KORNHUBER; DEECKE, 1965), in which the negativity is more evident when using channels with electrodes parallel to the sagittal plane.

The other grand averages presented here using other reference schemes (Cz-AR and Cz-RM) show similar timing for the negative peak compared to the grand average using the selected channels.

It is important to note that the experimental protocol that generated the current dataset is a combination of a self-paced movement type (typical for RP) and a warning-go type (typical for CNV). As there was a cue to start a new trial (beep sound and blank screen) along with the freedom to start the movement, we considered that the task involved some of the same cortical processes. The neurophysiological origin of the CNV and RP patterns are still not well established and there is discussion even about the possibility of one being part of the other (KONONOWICZ; PENNEY, 2016; SHIBASAKI; HALLETT, 2006). Nevertheless, despite of the difficulties to categorize the EEG delta band motor signatures found

in this work (Figure 28), the evidence indicates that they are indeed the CNV and/or RP. Thus, guaranteeing that the procedure presented here was useful to identify these EEG movement-related parameters.

5.1.2. Alpha band grand averages and the ERD

The ERD in the grand average signals (Figure 29) are very similar to the ERD presented Pfurtscheller and Aranibar (PFURTSCHELLER; ARANIBAR, 1979) in the C4-Cz and C3-Cz channels. They reported the ERD occurs in two stages (first the energy decays, then rises, and finally decays again) prior to the movement onset. In the grand average using the selected channels, the first and second waves start approximately 2.30 s and 0.80 s prior to the movement onset, respectively. The ERD has also been reported to occur significantly only during movement execution (RAMOS-MURGUIALDAY; BIRBAUMER, 2015). However, the differences with our results could be explained by experimental protocol differences as they used a BCI system with a hand orthosis and analyzed the band between 6 and 14 Hz.

The grand average using the channels selected individually for each subject is very similar to the other grand averages we found when using “traditional” ERD channels for motor tasks. Therefore, it can be concluded that the selected channels for each subject do, in fact, account for the motor ERD.

5.1.3. Spatial differences with timing similarities across subjects

With respect to the delta waves found for each subject it can be concluded that similar channels in different subjects do not imply similar signals. The reverse also holds, similar signals between different subjects (correlation greater than 0.90) does not imply similar channels.

These findings provide strong evidence of subject-specificities concerning the head volume conduction and the cortical electrical activity for a certain task.

In the alpha waves found for each subject with the channel selection procedure, the signals are converted to energy by squaring all samples. Therefore, two electrodes forming a channel provides the same signal regardless which one of them is the reference electrode.

In Figure 32 (or Figure 33), three ERD patterns can be observed: 1) two stages; 2) early; 3) late. The first category is when there are two energy peaks prior to the movement onset (similar to what was reported in (PFURTSCHELLER; ARANIBAR, 1979)), approximately at t=2 s and in between t=3 s and t=3.5 s. This pattern was found in subjects 1, 4, 7, 9, 10. The second is when the 75% energy decay occurs more than 1 s prior to the movement onset and it was found in subjects 2, 6, 8, 13, 14. The third was found in subjects 3, 5, 11, 12, 15. The subject's individual energy signals are equally divided among the three categories.

5.1.4. Spatial characteristics of the selected channels

For 80% of the subjects, the channels selected for the delta band have electrodes far from each other and usually in opposite sides of the coronal plane. The reference electrode is usually in the frontal region and the main electrode a little behind the central line. Only three subjects (4, 10 and 12) did not show this spatial characteristic, but they showed different patterns.

For 67% of the subjects (10 out of 15) the reference electrodes were either F3 (six cases) or F4 (four cases). Thus, it is reasonable to hypothesize that this region has an important role in slow potentials related to the task. It can be responsible for a positive shift as a consequence of some cognitive/behavioral process, or it can have an approximately null inter-trial average signal, acting as a good reference option for EEG analysis.

The channels selected for the alpha band signals are well distributed over the scalp. In most cases both electrodes within a channel are very close to each other. The subjects with electrodes that are most separated are subjects 1, 12, and 13, but even in these cases they are not too far (compared, e.g., to the delta band channels). Perhaps, the major finding here is that the channels must be chosen with electrodes very close to each other to properly analyze ERD in the alpha band. This might be a consequence of the small amplitudes of these oscillations compared to other EEG signatures that could be detected along the scalp. It also allows to hypothesize that the ERD is highly localized.

5.1.5. Temporal marks on the grand averages of the selected channels

In the delta band signals there are important similarities between the grand averages of the selected channels and the channels Cz-AR and Cz-RM (Figure 28). In all cases there are two notable bumps, the first is close to t=2.12 s and the second is in the final 0.20 s interval prior to the movement onset. In the channel selection case, however, these bumps are not statistically significant. Therefore, one can attribute these marks to an activity inherent to the Cz site. For the second bump, this agrees with the so called P-50 reported in (SHIBASAKI et al., 1980), which is a positive peak 0.05 s prior to the onset on Cz site referenced to linked-ears. There is a delay of approximately 0.15 s, but it is important to note that the mentioned work (SHIBASAKI et al., 1980) uses the electromyogram (EMG) signal to set temporal marks related to the movement. In the present work an exoskeleton with an angle sensor to detect the movement is used, which might lead to timing differences.

The negative shift of the delta band grand averages of the selected channels and Cz-AR starts near 0.50 s after the cross appearance. For Cz-RM it starts earlier, prior to the appearance of the cross on the screen. All three cases share the negative peak near t=4.75 s, which is approximately 0.50 s after the movement onset. In Cz-AR and Cz-RM grand averages, the signal increases after the negative peak until the movement reaches its maximum amplitude. In the grand average of the selected channels, the signal increases a little further, approximately until half of the interval of the final phase of the movement (when the arm is returning to its initial position). At this final phase of the movement, the grand averages of the selected channels and Cz-AR show a very clear negative peak almost coincident with each other. The same is not seen with statistical significance in the Cz-RM grand average. This final negative peak draws particular attention because the signal forms a ‘w’ shape connecting the beginning and the end of the movement.

In the alpha energy grand average signals, the energy level is higher before the movement and starts to decay significantly as the movement onset gets closer. The lowest energy value occurs little after the movement onset. However, at the final phase of the movement, there is a notable increase in the energy. In the selected channels case and C3-RM, the energy keeps increasing till the end of

the movement. For the other cases, the energy increases, maintains a higher level for an interval between 0.50 s and 1.00 s and then decays again. Because the movement is still occurring, although it has already reached its maximum amplitude and is returning to the final/start position, one would expect that the energy level would remain low. In the literature it is reported that the alpha ERD may even last longer than the movement itself (PFURTSCHELLER; LOPES, 1999). Given that the ERD expresses a local cortical activation, this may suggest that, at the final part of the movement, the corresponding cortical region is not fully involved with the task. During the experiment, the task was repeated multiple times by each subject and there was no performance requirement or feedback. Thus, it is possible that the movement became automatic after the first few repetitions, requiring less activity from the cortex.

5.1.6. Application to BCI

The goal of a BCI is to detect temporal events in brain signals that can be used to identify the users' intention or to control a device. Therefore, with the results presented here, we can propose three recommendations to improve BCIs. First, the channels should be chosen individually for each person, even when the BCI is designed to detect well-known EEG signatures such as the ones explored in this paper. Second, although many BCI-based protocols usually focus on the movement initiation intention, it was found that the end of the movement is preceded by some specific features as well. In the delta band the 'w' wave anticipates the end of the movement by the detection of the second valley. The small energy increase in the alpha band can also do the same. Third, it is necessary to characterize the pattern of the ERD for a specific subject before defining the feature extraction procedure.

5.1.7. Summary

This part of the project (Part 1 – Approach 1) aimed at improving the robustness of the identification of two known EEG motor-related signatures in a widely adopted experimental paradigm. The key was to search the electrode combinations that provided the more robust CNV or RP and alpha ERD detection. To accomplish this goal, two criteria were proposed to quantify the presence of the mentioned signals in the EEG recordings. With these criteria all possible

channels (pairs of electrodes) were analyzed computationally to detect these signals.

The results showed that the channels selected with the proposed method presented the intended motor signatures. It was shown that the alpha band ERD can occur in different ways (two stages, early and late ERD). The channels that best presents the motor signatures are highly subject-specific, although there are some spatial patterns specific to the frequency band. In the delta band, the channels usually connect electrodes distant from each other and placed in opposite sides of the coronal plane. The reference is usually in the anterior part, most likely in F3 or F4 locations. In the alpha band the channels can be widely distributed over the scalp with electrodes very close to each other. The end of the movement can be anticipated by specific events in the signals, such as the ‘w’ shape in the delta band and the energy increase in the alpha band. Evidence suggesting a reduced participation of the cortex at the final phase of a cyclic movement was presented with the alpha band signals. These findings can help to understand EEG signals during motor tasks, and they also provide spatial-temporal information of the signals for BCI development.

To improve the analysis, it would be interesting to analyze not only the best channel for each case, but a number of channels to see if the findings are consistent among these other locations. Moreover, conclusions regarding the meaning of the signatures must be considered very carefully. For instance, in many cases the location of the signature is a factor considered together with its temporal features for its proper identification.

5.2. Part 1 – approach 2

When analyzing results for all subjects together, it is seen that a few signal transformations dominate the best results. When using 15 trials to calculate averages, the energy in short windows in the 10 – 12 Hz and 15 – 25 Hz are responsible for 39% of the 4500 best results (100 results per round and per subject). The transformations called ups and downs, and slope comes next, summing 33%. When using seven trials to calculate averages, the down-sampled EEG is 52% of the best results. With single trials, energy difference between separated windows in the 6 – 8 Hz, 15 – 25 Hz, and 10 – 12 Hz frequency bands

are responsible for 63% of the results. Ups and downs come next with 10%. Interesting to note that results were considerably different between the three cases when considering the highest occurrences in each case (e.g. the first values to sum up to 50%). However, only 17 of the 41 transformations appeared in Tables 7-9, and all ten transformations of Table 7 are repeated in Table 8. In Table 9 there are seven new transformations. Some similarity between the first two cases were expected, because both involved averaging to enhance the SNR and used the same procedure.

Figure 35 shows the single most frequent transformation in the top 100 of each subject and separated by rounds, i.e. if one is to select a single transformation to apply to the signals, this one could be it. All rounds follow the same procedure, but the trials used to calculate the averages are probably different, as they are selected randomly out of almost 120 trials. Except for S12 and S15, all other subjects have the same result in all three rounds, which shows consistency within each case. S12 when using 15 trials has a different result in one round, and S15 has a different result in one round using 15 trials and one using 7 trials. This is evidence of the robustness of the method. It can be seen that when using 15 trials, the single most frequent transformation uses total motor frequency energy in short windows for 11 out of 15 subjects (73%). For two subjects the most frequent transformation uses motor frequency energy difference between separated windows, and for the other two it was low frequency transformation. Although in this case the results are dominated by motor frequencies transformation, the fact that two subjects have a low frequency transformation as their most frequent result highlights the importance of working with different possibilities to account properly for the subject-specificities. Analyzing the case with seven trials being used to calculate the averages, the down-sampled EEG dominates the results (9 out of 15 subjects), as seen previously. In five cases (S3, S4, S6, S8, and S14) the results completely match the ones using 15 trials, which shows consistency in the sense that averaging 15 trials and seven trials should lead to similar results. On the other hand, results using single trials are somehow similar to each other, but considerably different from the other cases. S3, S10, and S15 have low frequency transformations as the most frequent, and all the other subjects (12 out of 15: 80%) have transformations involving two

separated windows. These transformations use 1.5 s of the original signal, which is a longer interval compared to the other transformations. It could be that using a longer window helps to compensate the variability of single trials in some level.

Figures 36-38 show, in the electrode space, the main regions involved in the top 100 results of each subject in each case (using 15 trials, seven trials, and single trials). The results using 15 and seven trials shows some strong resemblance for the same subjects, and once again the subject specificity can be seen. Nevertheless, still analyzing the cases with averages and analyzing all subjects together it is possible to see that peripheral electrodes and frontal-central electrodes close to the midline are almost never included in the marked regions. In the case using single trials the results are considerably different, but still shows a tendency to avoid peripheral and frontal-central regions close to the midline. Subject 4 and subject 15, in the single trial case, have similarities with results from the other two cases. In terms of differences and similarities between the three cases these results are in accordance with the previous results concerning the transformations.

A few more analyses that can be done includes checking if there is a significant occurrence of particular electrodes in the reference position for each subject and check the spatial relation in the electrodes pairs (channels), e.g. if the electrodes are usually close to each other, or in different hemispheres, etc. Also, regarding the transformations it is interesting to analyze the supposedly best signals to see if there is any remarkable characteristic that allows to identify the timing of the movement onset. The inter-subject variability concerning the transformed signals can also be assessed. Moreover, because many of the proposed transformations seems not to be relevant, they can be ignored in a further attempt and substituted by other transformations relative to the features that seems to be more meaningful.

5.3. Part 2

In Part 2 it was proposed a procedure to reduce the EEG variability that selects the channels individually for each subject and transforms the EEG signals to improve a pseudo-online movement intention detection. The goal was to reduce the inter-trial variability of the signals. For a comparison, the classification

procedure (movement intention detection) was also tested without applying transformations to the signals and with channels selected by means of literature information. This led to the four scenarios shown in Table 5: TRC, RRC, TLC, RLC.

In Figure 39, the level difference in the accuracy results across all subjects and scenarios shows the effectiveness of the channel selection and transformation procedures. TRC and TLC are both dominated by accuracies above 80%, opposed to the other scenarios without transformations. In Figure 40, with the averages across subjects, the advantage of the channel selection procedure over the literature channels becomes clearer. The latter did not achieve 90% of average accuracy in any case, while the TRC always had above 90% average accuracy when using three or more channels. The single best result of each subject (Table 6) confirms that transforming the channels and analyzing their variability is key for a better performance. Even when the subject had their best result with literature channels instead of with the proposed method (20% of the cases), the best result was still using the transformations. For all the other subjects, the best result was using the proposed method for selecting channels and transformations.

The proposed method not only showed to be effective for improving the accuracy, but also for enabling a small number of electrodes and channels in signal acquisition and classification. The mean number of electrodes (10), channels (7) and accuracy (95%) sets a satisfying global performance by the proposed method (TRC scenario). The minimum and maximum over the best individual results were 80% and 98%, respectively. In recent works from the literature, satisfying accuracy results have been reported when classifying between rest and movement, but usually with mean accuracy lower than what was obtained with the method proposed in the present work. For instance, the work of Duan and colleagues (DUAN et al., 2021) used the same public dataset that was used in the present work. The authors classified between trials with movement execution and trials with rest, obtaining the best accuracy for elbow flexion vs rest classification, with an average across subjects of 90%. They also present an interesting search for more stable signals by using a technique designated as Task-Related Component Analysis, with either the SL or AR as reference

schemes. In (MASCOLINI; KHAN; MESIN, 2022), Mascolini and colleagues detected movement intention using 9 and 6 electrodes referenced to the right ear lobe. Both motor execution and motor imagery were analyzed for 16 subjects, with the best individual accuracy result being 100%, but with an average across all subjects of 85%. In (ALIAKBARYHOSSEINABADI et al., 2021), the authors performed a classification procedure to distinguish between trials with hand movement and trials without movement. The EEG data was recorded from 30 amyotrophic lateral sclerosis patients in 9 electrodes referenced to the right ear lobe. The authors performed a subject-specific channel, feature and classifier selection. Only one channel was selected per subject. The best accuracy result was 91% and the average across subjects was 81%. The authors reported that the individual configuration for each subject varied significantly. Thus, the comparison of the present work's results to other recent work in the literature reveals that the proposed method is promising.

The method was developed on the hypothesis that it is possible to find specific scalp locations that provide signals with low variability. Such signals would enable a better classification accuracy. Figure 41 validates this hypothesis by showing that lower variability (indicated by its inverse, the reproducibility) corresponds to higher accuracy and vice-versa. Therefore, transforming the signals and selecting the adequate channels was an effective strategy for providing more stable signals and for highlighting movement intention information.

It is important to note that the proposed method and the calibration of the classifier were executed with data corresponding to the elbow movement, while the classification performance is evaluated with hand movement. This suggests a certain level of robustness. On the other hand, the method should be further validated with an online classification experiment. It would also be interesting to apply the method to data collected in different days, which might provide even more robustness to the results.

Analyzing the effect of the transformations in the reproducibility of the signals, Figure 44 shows that the energy in the delta band is alone the most effective. All the other cases seem very close to each other, compared to the delta band total score. The correlation coefficient is sensitive to phase differences between signals with equal oscillating power and frequency. To minimize this effect, the

energy signals were always smoothed prior to the correlation coefficient calculation. In this case, slow frequencies could lead to higher correlation coefficient compared to high frequencies with the same phase difference. However, due to the major difference in the total reproducibility scores of the delta band energy and all the others, it is reasonable to consider that a possible bias would not be significant. Figure 42 shows the same effect separated by electrode positions. The discrepancy between delta band energy and the others occurs regardless of the electrode position. Motor related EEG signatures in the delta band, such as CNV and RP, can justify its effectiveness. The same was expected for the alpha band, due to the ERD, but the results showed otherwise. This becomes even more interesting because the CNV and RP are elicited in specific task scenarios that do not match the one used to generate the signals analyzed in the present work. The alpha band ERD, on the other hand, is expected in any task involving conscious movement. A hypothesis that can explain this result is the existence of other cortical activities unrelated to the movement in the analyzed frequency bands. In this case, there would be more of these activities in the alpha band than in the delta band, causing this difference in the reproducibility scores. These activities are usually properly eliminated with trial averaging, which is not possible in realistic BCI scenarios. Thus, for BCI purposes, delta band features are better suited for movement intention detection.

In Figure 43, a closer look to the delta band results shows that the CCP line is the one with higher reproducibility. Also, the reproducibility in different cortical lines (according to the positioning system) seems to be higher as they approximate CCP. In Figure 46, with only the first 10 channels of each subject, the prevalence of the CCP line is also evident. The electrode position with most appearances is CCP1h. It appeared 14 times, which shows that the distribution of the electrode across the top 10 of all subjects was relatively dispersed. This is confirmed by Figure 45, because the top 10 channels across different subjects do not seem to share any pattern other than the prevalence of central-parietal regions. This region is above the motor cortex, so its importance is expected. Given that the delta band energy was the most used transformation, it suggests that slow potentials in this scalp region provide important task-related information. Further investigation towards this hypothesis considering neurophysiological

aspects is encouraged. The frontal region is more affected by eye artifacts (URIGÜEN; GARCIA-ZAPIRAIN, 2015), which can explain the lower reproducibility of its signals. In Figure 43, within each cortical line, the reproducibility is higher in positions close to the midline and lower near the left or right extremities. This is an interesting result that also encourages further investigation. Nevertheless, the proximity to the CCP line is more significant than the proximity to the midline between hemispheres. In practical terms for BCIs, these results suggest that opting for delta band signals in electrodes placed at the central-parietal region might be a good choice in order to have more stable signals across different trials.

5.3.1. Summary

In this work we proposed a method to reduce the EEG variability by selecting subject-specific channels and channel-specific transformations that provide more stable signals across different trials of a motor task. The goal was to improve the movement intention detection in a pseudo-online classification procedure. The stability of the signals was measured with the Pearson correlation coefficient. The transformations were all band-specific. The results showed that the proposed method succeeded in providing high accuracy values (95% average across subjects). Also, the results showed that lower inter-trial variability did provide better classification performances. Signals in the delta band provided the lowest inter-trial variability for all subjects and scalp positions. The electrodes in the CCP line associated to the delta band energy were, in general, the best choices according to the proposed method. However, the best channels were highly subject-specific. Also, the reproducibility of the delta band signals shows that the scalp positions close to the midline outstand the others close to the extremities within each cortical line in the 10-20 system.

The fact that the data used for each subject was acquired all in one day is a limitation of the presented analysis. This is due to the intra-subject variability that occurs between different days and might differ from the inter-trial variability in one single experiment. Moreover, the classification using a real system that interacts with the subject is expected to be different, especially because it introduces new cortical activities originated from this interaction. An also important point to

consider is the use of signals from real patients, given that the signals might present different spatio-temporal features.

6. CONCLUSION

In this work, the main challenge for BCI development that is addressed is the intra-subject variability of EEG signals. Inappropriate hypothesis concerning the reference electrode's best location and the possibility of a neutral reference are considered to be important reasons for the difficulty when dealing with the signal variability. In section 2, literature review, these issues were consistently presented along with other related topics. To define the approach, this work relied on the following statements:

- All EEG recordings are bipolar;
- The correct choice of pairs of electrodes and signal transformations can reduce intra-subject variability;
- Lower variability can lead to higher BCI accuracy;
- Inter-subject variability cannot be avoided, thus solutions are always subject-specific.

The work was divided into two parts. Part 1 consisted in studying EEG signals related to motor activities and signal variability, both cases using trial average signals. Two different approaches were adopted in Part 1. The goal was to gather information and improve the knowledge towards the dataset that was selected for this work. Part 2 consisted in a strategy to improve a pseudo-online movement intention detection by finding channels and transformations that provided signals with lower intra-subject variability. The dataset used in this work is publicly available and contains data from motor tasks with the upper limb from 15 healthy subjects.

In Part 1, Approach 1, it was shown that well-known motor signals, namely the ERD and CNV or RP, were better detected in different channels for each subject. Some similarities were found in the channels for each EEG signature. For instance, in the ERD case, channels were usually with electrodes placed very close to each other, although the pair location varied significantly among subjects. In the delta band signatures (CNV and RP), channels were usually with electrodes distant from each other and connecting the two sides of the coronal plane. Approach 2 of Part 1 showed that, in general, band-specific energy features have less variability and the bands varies across subjects. It also showed

that the signals varied significantly when averaged across 15 and 7 trials, and even more when using single trials.

In Part 2, it was proposed a method for selecting channels (pairs of electrodes) and channel-specific transformation based on band-specific energy to find signals with low intra-subject variability. The aim was to achieve a high-level accuracy for movement intention detection in a pseudo-online classification task. Results showed that the method was effective compared to other state-of-the-art work. Delta band features outperformed all the others. Also, the variability-based channel selection procedure showed that solutions must be subject-specific. Moreover, it showed that it is important to consider channel options beyond the ones commonly related to the motor cortex, such as the C and FC lines in the 10-20 system. Overall, the channels were highly subject-specific, but mainly placed in the central-parietal line.

This work contributes to the BCI field in different ways. In Part 1, the identification of motor signatures across different pairs of electrodes and the identification of novel features suggesting the anticipation of the end of a cyclic movement contributes to the study of EEG-related signatures. In Part 2, the good classification results per itself is a valuable contribution. Also, drawing attention to the unsolvable problem (so far) of the active reference electrode and the relevance of not neglecting it allowed an efficient strategy to be developed to overcome the limitations that it causes. Moreover, the method proposed in this work can incorporate other transformation, more channels and different variability measures.

For future works, especially for Part 2, it is encouraged to boost the methodology by inserting more channels and other variability criteria. More signal transformation is also suggested. The use of more sophisticated classification algorithms can lead to even better results as well. Finally, the use of data acquired in different days to find more robust subject-specific configuration and testing the methodology with a real BCI system are the natural next steps to be taken. After these steps, the applicability of the system to assist a motor rehabilitation therapy should be investigated.

APPENDIX A (Fisher's Linear Discriminant)

Fisher's Linear Discriminant – (DUDA, RICHARD O. AND HART, PETER E. AND STORK, 2000)

Let us consider a set of η d-dimensional samples $\boldsymbol{\delta}_1, \boldsymbol{\delta}_2, \dots, \boldsymbol{\delta}_\eta$ (i.e. each $\boldsymbol{\delta}$ is a vector column with d elements), where N_1 is the number of samples of the sub-set D_1 corresponding to class $\boldsymbol{\tau}_1$ and N_2 is the same for sub-set D_2 from class $\boldsymbol{\tau}_2$. A linear combination of the elements of $\boldsymbol{\delta}$ gives us

$$y = \boldsymbol{w}^T \boldsymbol{\delta} \qquad (1)$$

Doing this for all samples $\boldsymbol{\delta}_1, \boldsymbol{\delta}_2, \dots, \boldsymbol{\delta}_\eta$, we will have a corresponding set with η samples $y_1, y_2, \dots, y_\eta$ in which each y is a scalar. This set can be divided into two sub-sets Y_1 e Y_2, each of them corresponding to a specific class. Geometrically, if $\|\boldsymbol{w}\| = 1$, each y_i is the projection of $\boldsymbol{\delta}_i$ in a line with the same direction as $\boldsymbol{w}$. The goal of the FLD is to make the projections on this line separable according to their classes $\boldsymbol{\tau}_1$ e $\boldsymbol{\tau}_2$.

The separability of the projected points can be measured by the distance between their class-average position. If $\boldsymbol{m}_i$ is the mean sample vector of class i given by

$$\boldsymbol{m}_i = \frac{1}{N_i} \sum_{p=1}^{N_i} \boldsymbol{\delta}_p \qquad (2)$$

then the mean projected position $\tilde{m}_i$ can be written as

$$\tilde{m}_i = \frac{1}{N_i} \sum_{y \in Y_i} y = \frac{1}{N_i} \sum_{y \in Y_i} \boldsymbol{w}^T \boldsymbol{\delta} = \boldsymbol{w}^T \boldsymbol{m}_i \qquad (3)$$

Therefore, the distance between the average position of the projected points from each class is

$$|\tilde{m}_1 - \tilde{m}_2| = |\boldsymbol{w}^T (\boldsymbol{m}_1 - \boldsymbol{m}_2)| \qquad (4)$$

To obtain a good separability of the points, the distance between their class-average must be big compared to the standard deviation of the classes. The dispersion of each class is the measure used to quantify this. It is given by

$$\tilde{s}_i^2 = \sum_{y \in Y_i} (y - \tilde{m}_i)^2 \quad (5)$$

Hence, $\frac{1}{\eta}(\tilde{s}_1^2 + \tilde{s}_2^2)$ is an estimate of the combined variance and $\tilde{s}_1^2 + \tilde{s}_2^2$ is called intra-class dispersion of the projected points. The FLD technique uses the linear function $\boldsymbol{w}^T\boldsymbol{\delta}$ for which

$$J(\boldsymbol{w}) = \frac{|\tilde{m}_1 - \tilde{m}_2|^2}{\tilde{s}_1^2 + \tilde{s}_2^2} \quad (6)$$

Is maximum and independent from $\|\boldsymbol{w}\|$. While $\boldsymbol{w}$ that maximizes $J(.)$ provides the best class separability, it is necessary to also define a threshold to delimit each class space.

To obtain $J(.)$ As na explicit function of $\boldsymbol{w}$, the dispersion matrices $\boldsymbol{S}_i$ e $\boldsymbol{S}_W$ must be defined as

$$\boldsymbol{S}_i = \sum_{\boldsymbol{\delta} \in D_i} (\boldsymbol{\delta} - \boldsymbol{m}_i)(\boldsymbol{\delta} - \boldsymbol{m}_i)^T \quad (7)$$

$$\boldsymbol{S}_W = \boldsymbol{S}_1 + \boldsymbol{S}_2 \quad (8)$$

Thus,

$$\tilde{s}_i^2 = \sum_{\boldsymbol{\delta} \in D_i} (\boldsymbol{w}^T\boldsymbol{\delta} - \boldsymbol{w}^T\boldsymbol{m}_i)^2$$

$$= \sum_{\boldsymbol{\delta} \in D_i} \boldsymbol{w}^T (\boldsymbol{\delta} - \boldsymbol{m}_i)(\boldsymbol{\delta} - \boldsymbol{m}_i)^T\boldsymbol{w} \quad (9)$$

$$= \boldsymbol{w}^T\boldsymbol{S}_i\boldsymbol{w}$$

Hence, the summation of these dispersion can be written as

$$\tilde{s}_1^2 + \tilde{s}_2^2 = \boldsymbol{w}^T\boldsymbol{S}_W\boldsymbol{w} \quad (10)$$

In a similar way, for the separation of the average projected points

$$(\tilde{m}_1 - \tilde{m}_2)^2 = (\boldsymbol{w}^T\boldsymbol{m}_1 - \boldsymbol{w}^T\boldsymbol{m}_2)^2$$

$$= \boldsymbol{w}^T(\boldsymbol{m}_1 - \boldsymbol{m}_2)(\boldsymbol{m}_1 - \boldsymbol{m}_2)^T\boldsymbol{w} \quad (11)$$

$$= \boldsymbol{w}^T \boldsymbol{S}_B \boldsymbol{w}$$

where

$$\boldsymbol{S}_B = (\boldsymbol{m}_1 - \boldsymbol{m}_2)(\boldsymbol{m}_1 - \boldsymbol{m}_2)^T \qquad (12)$$

Matrix $\boldsymbol{S}_W$ is named as intra-class dispersion matrix. It is proportional to the covariance of the grouped d-dimensional samples. Matrix $\boldsymbol{S}_B$ is called inter-class dispersion matrix. We can write $J(.)$ as a function of these matrices:

$$J(\boldsymbol{w}) = \frac{\boldsymbol{w}^T \boldsymbol{S}_B \boldsymbol{w}}{\boldsymbol{w}^T \boldsymbol{S}_W \boldsymbol{w}} \qquad (13)$$

This expression is known as Rayleigh ratio. By using Lagrange multipliers, it can be demonstrated that the vector that maximizes $J(.)$ Must satisfy

$$\boldsymbol{S}_B \boldsymbol{w} = \lambda \boldsymbol{S}_W \boldsymbol{w} \qquad (14)$$

For a constant λ, which is an eigenvalue problem. If $\boldsymbol{S}_W$ is not singular, the conventional eigenvalue problem can be obtained by

$$\boldsymbol{S}_W^{-1} \boldsymbol{S}_B \boldsymbol{w} = \lambda \boldsymbol{w} \qquad (15)$$

Because $\boldsymbol{S}_B \boldsymbol{w}$ will always be in the direction of $\boldsymbol{m}_1 - \boldsymbol{m}_2$, it is not necessary to solve the eigenvalue problem. Because $\boldsymbol{w}$ does not depend on its localization in space, the solution for $\boldsymbol{w}$ that optimizes $J(.)$ Can be written as

$$\boldsymbol{w} = \boldsymbol{S}_W^{-1}(\boldsymbol{m}_1 - \boldsymbol{m}_2) \qquad (16)$$

APPENDIX B (Average Reference)

This appendix is referenced to (BERTRAND; PERRIN; PERNIER, 1985)

For dipoles inside a sphere:

The surface integral of electric potentials $\emptyset(\boldsymbol{r})$ on the sphere with surface S is null:

$$\oint_S \emptyset(\boldsymbol{r})dS = 0 \tag{17}$$

where $\emptyset(\boldsymbol{r})$ is the "true" potential at position $\boldsymbol{r}$ referenced to "infinity". The integral can be approximated by a summation for discrete measurements in positions indicated by $n = (1, \dots, N)$:

$$\sum_{n=1}^{N} \emptyset(\boldsymbol{r}_n) = 0 \tag{18}$$

In the case of EEG measurements, the potential V_n is always recorded as the difference between the true electrical potential $\emptyset(\boldsymbol{r}_n)$ and the true potential at a reference location $\emptyset(\boldsymbol{r}_R)$:

$$V_n = \emptyset(\boldsymbol{r}_n) - \emptyset(\boldsymbol{r}_R) \tag{19}$$

Calculating the average of both sides of eq. (19), we have

$$\frac{1}{N}\sum_{n=1}^{N} V_n = \frac{1}{N}\sum_{n=1}^{N}[\emptyset(\boldsymbol{r}_n) - \emptyset(\boldsymbol{r}_R)] \tag{20}$$

The left side is the average potential V_{AVG} across all EEG channels. The right side can be separated into two summations. The potential at the reference position does not depend on the location of the other electrode, so eq. (20) can be re-written as

$$V_{AVG} = \left(\frac{1}{N}\sum_{n=1}^{N} \emptyset(\boldsymbol{r}_n)\right) - \left(\frac{1}{N} \times N \times \emptyset(\boldsymbol{r}_R)\right) \tag{21}$$

than as

$$V_{AVG} = \left(\frac{1}{N}\sum_{n=1}^{N} \emptyset(\boldsymbol{r}_n)\right) - \emptyset(\boldsymbol{r_R}) \quad (22)$$

The first term of the right side vanishes, according to eq. (18). Thus,

$$V_{AVG} = -\emptyset(\boldsymbol{r}_R) \quad (23)$$

Substituting eq. (23) in eq. (19) and isolating the true electrical potential referenced to infinity, we obtain

$$V_n - V_{AVG} = \emptyset(\boldsymbol{r}_n) \quad (24)$$

Therefore, the true potential $\emptyset(\boldsymbol{r}_n)$ at position $\boldsymbol{r}_n$ can be estimated by subtracting the average over all measurements from the potential V_n recorded at $\boldsymbol{r}_n$ referenced to any arbitrary electrode.

APPENDIX C (Surface Laplacian)

a. Physical meaning of the Surface Laplacian (CARVALHAES; DE BARROS, 2015)

All the content of this appendix is referenced to (CARVALHAES; DE BARROS, 2015).

The electric field can be quantified as the negative of the gradient of the electric potential:

$$\boldsymbol{E}(\boldsymbol{r}) = -\nabla V(\boldsymbol{r}) = -\frac{\partial V}{\partial x}i - \frac{\partial V}{\partial y}j - \frac{\partial V}{\partial z}k \tag{25}$$

The divergent of the electric field is defined as the negative of the Laplacian surface of the electric field. Thus,

$$Div(\boldsymbol{E}) \equiv -Lap(V) \tag{26}$$

$$Div(\boldsymbol{E}) = Div(-\nabla V) = -\frac{\partial^2 V}{\partial x^2} - \frac{\partial^2 V}{\partial y^2} - \frac{\partial^2 V}{\partial z^2} \tag{27}$$

$$Div(-\boldsymbol{E}) \equiv Lap(V) = \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} \tag{28}$$

For EEG measurements, the scalp is the surface to be analyzed and the cortex is the place where the sources are at. Hence, there are no sources at the surface, which gives us

$$Lap(V) = 0 \quad \Rightarrow \quad \frac{\partial^2 V}{\partial^{\ 2}} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = 0 \tag{29}$$

From (1), we know that

$$E_z = -\frac{\partial V}{\partial z} \quad \Rightarrow \quad \frac{\partial E_z}{\partial z} = -\frac{\partial^2 V}{\partial z^2} \quad \Rightarrow \quad -\frac{\partial E_z}{\partial z} = \frac{\partial^2 V}{\partial z^2} \tag{30}$$

Substituting (5) in (4), we have

$$\frac{\partial^2 V}{\partial x^2}+\frac{\partial^2 V}{\partial y^2}-\frac{\partial E_z}{\partial z}=0 \quad \Rightarrow \quad \frac{\partial^2 V}{\partial x^2}+\frac{\partial^2 V}{\partial y^2}=\frac{\partial E_z}{\partial z} \tag{31}$$

The electric field in a volume conductor can also be defined as $\boldsymbol{E}=\rho \boldsymbol{J}$, where ρ is the resistivity (constant) and $\boldsymbol{J}$ is the current density. Thus, $\boldsymbol{E_z}=\rho \boldsymbol{J_z}$. This gives us

$$\frac{\partial^2 V}{\partial x^2}+\frac{\partial^2 V}{\partial y^2}=\rho \frac{\partial \boldsymbol{J_z}}{\partial z} \tag{32}$$

The left side of (7) is defined as the Laplacian surface of V:

$$Lap_S(V)=\frac{\partial^2 V}{\partial x^2}+\frac{\partial^2 V}{\partial y^2} \tag{33}$$

Then

$$Lap_S(V)=\rho \frac{\partial \boldsymbol{J_z}}{\partial z} \tag{34}$$

This means that $Lap_S(V)$ is proportional to the variation of the current density perpendicular to the surface (radial direction in a spherical geometry). Therefore, if $Lap_S(V) \neq 0$, then there is current flowing from underneath the scalp, which is associated to the presence of an electrical source inside the head (in this case, on the cortex).

b. Calculation of the Surface Laplacian – (CARVALHAES; DE BARROS, 2015; CARVALHAES; SUPPES, 2011)

The spline function is the unique solution for the problem of finding a function f that minimizes the equation

$$SS(f,\lambda)=\frac{1}{N}\sum\left(V_i-f(\boldsymbol{r_i})\right)^2+\lambda\, J_m[f(\boldsymbol{r})] \tag{35}$$

where N is the number of channels, V_i is the EEG potential recorded at electrode i, $\boldsymbol{r_i}$ is the position vector of electrode i, λ is the spline regularization parameter, $J_m[f]$ is a measure of the roughness of f in terms of its partial derivative of m-th order (WAHBA, 1990) and m is the flexibility of the adjustment. If $\lambda=0$, then it occurs an interpolation ($f(\boldsymbol{r_i})=V_i$), if $\lambda>0$ the data is smoothed instead.

The function that minimizes $SS(f,\lambda)$ is

$$f(\boldsymbol{r}) = \sum_{i=1}^{N} c_i \left\|\boldsymbol{r} - \boldsymbol{r_i}\right\|^{2m-3} + \sum_{l=1}^{M} d_l \phi_l(\boldsymbol{r}) \tag{36}$$

where m e M are integers that satisfies the conditions

$$\begin{gathered} 2m > 3 \\ M = \binom{m+2}{3} \\ M < N \end{gathered} \tag{37}$$

and $\phi_l(\boldsymbol{r})$ is given by

$$\phi_l(\boldsymbol{r}) = x^{i-j} y^{j-k} z^k \tag{38}$$

where

$$\begin{gathered} l = i + j + k + 1 \\ 0 \leq i \leq m - 1 \\ 0 \leq j \leq i \\ 0 \leq k \leq j \end{gathered} \tag{39}$$

To use a matrix notation, we can define

$$\boldsymbol{V} = (f(\boldsymbol{r}_1), \ldots, f(\boldsymbol{r}_N))^T \tag{40}$$

$$\boldsymbol{c} = (c_1, \ldots, c_N)^T \tag{41}$$

$$\boldsymbol{d} = (d_1, \ldots, d_M)^T \tag{42}$$

$$K_{ij} = \left\|\boldsymbol{r_i} - \boldsymbol{r_j}\right\|^{2m-3} \tag{43}$$

$$T_{ij} = \phi_j(\boldsymbol{r_i}) \tag{44}$$

We can write the system:

$$\begin{pmatrix} \boldsymbol{K} + N\lambda \boldsymbol{I} & \boldsymbol{T} \\ \boldsymbol{T}^T & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{c} \\ \boldsymbol{d} \end{pmatrix} = \begin{pmatrix} \boldsymbol{V} \\ 0 \end{pmatrix} \tag{45}$$

Column vector $\boldsymbol{c}$ with N elements and $\boldsymbol{d}$ with M elements are the coefficients that expands the function $f(\boldsymbol{r})$. $\boldsymbol{V}$ is the column vector with the N instantaneous EEG potentials in positions $\boldsymbol{r_1}, \boldsymbol{r_2}, \ldots, \boldsymbol{r_N}$. Matrices $\boldsymbol{K}$ and $\boldsymbol{T}$ have dimensions of NxN and NxM, respectively.

The formal solution of the system is given by (WAHBA, 1990):

$$\boldsymbol{c} = \boldsymbol{Q_2}[\boldsymbol{Q_2^T}(\boldsymbol{K} + N\lambda\, \boldsymbol{I})\boldsymbol{Q_2}]^{-1}\boldsymbol{Q_2^T}\,\boldsymbol{V} \tag{46}$$

$$\boldsymbol{Rd} = \boldsymbol{Q_1^T}(\boldsymbol{V} - \boldsymbol{Kc} - N\lambda \boldsymbol{c}) \tag{47}$$

where $\boldsymbol{Q_1}_{N\text{x}M}$, $\boldsymbol{Q_2}_{N\text{x}(N-M)}$ and $\boldsymbol{R}_{M\text{x}M}$ are matrices obtained by the QR decomposition of $\boldsymbol{T}_{NxM}$, eq. (48), which also includes the null matrix $\boldsymbol{O}_{(N-M)xM}$.

$$\boldsymbol{T} = (\boldsymbol{Q_1} \quad \boldsymbol{Q_2}) \begin{pmatrix} \boldsymbol{R} \\ \boldsymbol{O} \end{pmatrix} \tag{48}$$

By defining the smoothed potentials as $\boldsymbol{V_\lambda} = \boldsymbol{V} - N\lambda \boldsymbol{Ic}$ and rearranging the terms of the first line of the system in eq. (45), we have

$$\boldsymbol{V_\lambda} = \boldsymbol{Kc} + \boldsymbol{Td} \tag{49}$$

Defining $\widetilde{K}_{ij} = Lap_s(K_{ij})$ and $\widetilde{T}_{ij} = Lap_s(T_{ij})$, we have

$$Lap_s(\boldsymbol{V_\lambda}) = \widetilde{\boldsymbol{K}}\boldsymbol{c} + \widetilde{\boldsymbol{T}} \tag{50}$$

However, when applied to a quadratic surface, such as the sphere, this approach faces some problems. When $m > 2$, matrix $\boldsymbol{T}$ will have columns expanded by the terms x^2, y^2 e z^2 ($\phi_l(\boldsymbol{r}) = x^{i-j}y^{j-k}z^k$), that are linearly dependent due to the restriction $x^2 + y^2 + z^2 = 1$, inherent to the surface geometry.

To avoid this problem, (WAHBA, 1990) propose the use of geodesic distances for spherical surfaces instead of Euclidean distances. This gives us:

$$f(\boldsymbol{r}) = \sum_{i=1}^{N} c_i \, g_m(\boldsymbol{r}, \boldsymbol{r_i}) + \boldsymbol{d} \tag{51}$$

where $\boldsymbol{c}$ e $\boldsymbol{d}$ are still obtained with eq. (46) e (47). In this new scenario, we have

$$K_{ij} = g_m(\boldsymbol{r_i}, \boldsymbol{r_j}) \tag{52}$$

$$T_{\mathrm{i}} = 1 \tag{53}$$

where

$$g_m(\boldsymbol{r}, \boldsymbol{r_i}) = \frac{1}{4\pi} \sum_{l=1}^{\infty} \frac{2l+1}{l^m (l+1)^m} P_l(\hat{\boldsymbol{r}}.\hat{\boldsymbol{r}}_{\boldsymbol{i}}) \quad m > 1 \tag{54}$$

and P_l are Legendre polynomial functions with one variable and degree l (the symbol ^ indicates unit vectors), given by

$$P_l(a) = \frac{1}{2^l l!} \frac{d^l}{da^l} [(a^2 - 1)^l] \tag{55}$$

Due to a singularity, it is defined $P_l(1) = 1$.

Using spherical coordinates with the convention illustrated in Figure 47 (where $\theta \in [0, \pi]$ and $\varphi \in [0, 2\pi]$), and applying the Laplacian surface to $f(\boldsymbol{r})$, we have

$$Lap_s(f) = \frac{1}{r^2 \sin\theta} \left[\frac{\partial}{\partial \theta} \left(\sin\theta \frac{\partial f}{\partial \theta} \right) + \frac{1}{\sin\theta} \frac{\partial^2 f}{\partial \varphi^2} \right] \tag{56}$$

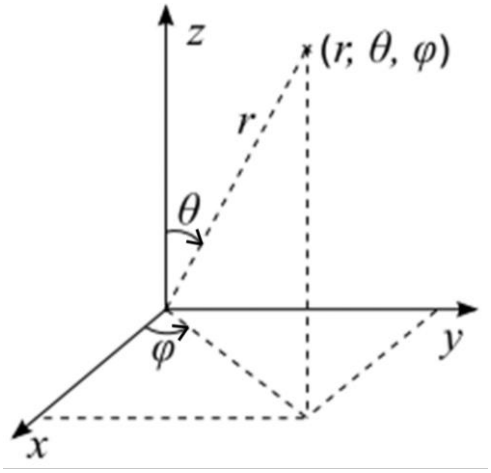


Figure 47 - Spherical coordinate system.

According to Jackson (1999), p.110 *apud* (CARVALHAES; DE BARROS, 2015), Legendre polynomials satisfies the relation

$$Lap_s\left(P_l(\hat{\boldsymbol{r}}\,.\,\hat{\boldsymbol{r}}_{\boldsymbol{i}})\right) = -\frac{l(l+1)}{r^2}P_l(\hat{\boldsymbol{r}}\,.\,\hat{\boldsymbol{r}}_{\boldsymbol{i}}) \tag{57}$$

Thus, the solution for the Laplacian surface of $f(\boldsymbol{r})$ is given by

$$Lap_s\left(f(\boldsymbol{r})\right) = -\frac{1}{r^2}\sum_{i=1}^{N} c_i g_{m-1}(\boldsymbol{r}, \boldsymbol{r}_{\boldsymbol{i}}), \qquad m > 1 \tag{58}$$

The eq. (58) is known as the spherical Laplacian surface of $f(\boldsymbol{r})$. For the geodesic distances and spherical coordinates, it follows that $\widetilde{\boldsymbol{T}}_{Nx1}$, e $\widetilde{\boldsymbol{K}}_{NxN}$ are:

$$\tilde{T}_{\mathrm{i}} = 0 \tag{59}$$

$$\widetilde{K}_{ij} = -\frac{g_m(\boldsymbol{r_j}, \boldsymbol{r_i})}{r^2} \tag{60}$$

Now, the transformed smoothed potential $Lap_s\left(f(\boldsymbol{r})\right)$ at an arbitrary point $\boldsymbol{r}$ can be calculated.

To be able to linearly transform the EEG data in the electrode locations, (CARVALHAES; SUPPES, 2011) define matrices $\boldsymbol{C}$ and $\boldsymbol{D}$ as

$$\boldsymbol{c} = \boldsymbol{CV} \tag{62}$$

$$\boldsymbol{d} = \boldsymbol{DV} \tag{63}$$

From eq. 46 and 47, we obtain

$$\boldsymbol{C} = \boldsymbol{Q_2}[\boldsymbol{Q_2^T}(\boldsymbol{K} + N\lambda\, \boldsymbol{I})\boldsymbol{Q_2}]^{-\mathbf{1}}\boldsymbol{Q_2^T} \tag{64}$$

$$\boldsymbol{D} = \boldsymbol{R}^{+}\boldsymbol{Q_1^T}(\mathbf{1} - \boldsymbol{KC} - N\lambda\boldsymbol{C}) \tag{65}$$

where $\boldsymbol{R}^{+}$ the pseudo-inverse of $\boldsymbol{R}$. Thus, from eq. (58), eq. (60) e eq. (62), we have

$$Lap_s(\boldsymbol{V}_\lambda) = \widetilde{\boldsymbol{K}}\boldsymbol{CV} \tag{66}$$

$\widetilde{\boldsymbol{K}}\boldsymbol{C}$ does not depend on the measured potentials and linearly transforms $\boldsymbol{V}$. Hence, we can define the transformation matrix $\boldsymbol{L}$ that will act as the Laplacian filter:

$$\boldsymbol{L} = \widetilde{\boldsymbol{K}}\boldsymbol{C} \tag{67}$$

Finally, we can represent the smoothed transformed potentials by the linear transformation

$$Lap_S(\boldsymbol{V}_\lambda) = \boldsymbol{LV} \tag{68}$$

Therefore, $\boldsymbol{L}$ has to be defined once, then it can be continuously used to transform the signals using eq. (68).

REFERENCES

ALIAKBARYHOSSEINABADI, Susan; DOSEN, Strahinja; SAVIC, Andrej M.; BLICHER, Jakob; FARINA, Dario; MRACHACZ-KERSTING, Natalie. Participant-specific classifier tuning increases the performance of hand movement detection from EEG in patients with amyotrophic lateral sclerosis. **Journal of Neural Engineering**, *[S. l.]*, v. 18, 2021.

ARY, James P.; KLEIN, Stanley A.; FENDER, Derek H. Location of Sources of Evoked Scalp Potentials: Corrections for Skull and Scalp Thicknesses. **Biomedical Engineering Online**, *[S. l.]*, v. 28, n. 6, p. 447–452, 1981.

BAI, Zhongfei; FONG, Kenneth N. K.; ZHANG, Jack Jiaqi; CHAN, Josephine; TING, K. H. Immediate and long-term effects of BCI-based rehabilitation of the upper extremity after stroke: a systematic review and meta-analysis. **Journal of NeuroEngineering and Rehabilitation**, *[S. l.]*, v. 17, n. 1, p. 1–20, 2020. DOI: 10.1186/s12984-020-00686-2.

BERTRAND, O.; PERRIN, F.; PERNIER, J. A Theoretical Justification of the Average Reference in Topographic Evoked Potential Studies. **Electroencephalography and clinical neurophysiology**, *[S. l.]*, n. 62, p. 462–464, 1985.

BLANKERTZ, Benjamin; DORNHEGE, Guido; KRAULEDAT, Matthias; MÜLLER, Klaus Robert; CURIO, Gabriel. The non-invasive Berlin Brain-Computer Interface: Fast acquisition of effective performance in untrained subjects. **NeuroImage**, *[S. l.]*, v. 37, n. 2, p. 539–550, 2007. DOI: 10.1016/j.neuroimage.2007.01.051.

BÖCKER, Koen B. E.; VAN AVERMAETE, Jurgen A. G.; VAN DEN BERG-LENSSEN, Margaretha M. C. The international 10-20 system revisited: Cartesian and spherical co-ordinates. **Brain Topography**, *[S. l.]*, v. 6, n. 3, p. 231–235, 1994. DOI: 10.1007/BF01187714.

CARLSEN, Anthony N.; CHUA, Romeo; TIMOTHY INGLIS, J.; SANDERSON, David J.; FRANKS, Ian M. Motor preparation in an anticipation-timing task. **Experimental Brain Research**, *[S. l.]*, v. 190, n. 4, p. 453–461, 2008. DOI: 10.1007/s00221-008-1487-5.

CARVALHAES, Claudio; DE BARROS, J. Acacio. The surface Laplacian technique in EEG: Theory and methods. **International Journal of Psychophysiology**, *[S. l.]*, v. 97, n. 3, p. 174–188, 2015. DOI: 10.1016/j.ijpsycho.2015.04.023.

CARVALHAES, Claudio G.; SUPPES, Patrick. **A spline framework for estimating the EEG surface Laplacian using the Euclidean metric**. **Neural Computation**, 2011. DOI: 10.1162/NECO_a_00192.

CERVERA, María A.; SOEKADAR, Surjo R.; USHIBA, Junichi; MILLÁN, José del R.; LIU, Meigen. Brain-computer interfaces for post-stroke motor rehabilitation: a meta-analysis. **Annals of Clinical and Translational Neurology**, *[S. l.]*, p. 1–13, 2018. DOI: 10.1002/acn3.544.

CHELLA, Federico; PIZZELLA, Vittorio; ZAPPASODI, Filippo. Impact of the reference choice on scalp EEG connectivity estimation. **Journal of Neural Engineering**, *[S. l.]*, v. 13, n. 3, p. 1–21, 2016. DOI: 10.1088/1741-2560/13/3/036016. Disponível em: http://dx.doi.org/10.1088/1741-2560/13/3/036016.

DALY, Janis J.; WOLPAW, Jonathan R. Brain – computer interfaces in neurological rehabilitation. **The Lancet Neurology**, *[S. l.]*, v. 7, n. 11, p. 1032–1043, 2008. DOI: 10.1016/S1474-4422(08)70223-0. Disponível em: http://dx.doi.org/10.1016/S1474-4422(08)70223-0.

DE MELO, Gabriel Chaves; MARTES STERNLICHT, Vitor; FORNER-CORDERO, Arturo. EEG Analysis in Coincident Timing Task Towards Motor Rehabilitation. **Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS**, *[S. l.]*, v. 2020-July, p. 3027–3030, 2020. DOI: 10.1109/EMBC44109.2020.9175851.

DONCHIN, Emanuel. Multivariate Approach to the Analysis Average Evoked Potentials. **IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING**, *[S. l.]*, v. 13, n. 3, p. 131–139, 1966.

DUAN, Feng; JIA, Hao; SUN, Zhe; ZHANG, Kai; DAI, Yangyang; ZHANG, Yu. Decoding Premovement Patterns with Task - Related Component Analysis. **Cognitive Computation**, *[S. l.]*, p. 1389–1405, 2021. DOI: 10.1007/s12559-021-09941-7. Disponível em: https://doi.org/10.1007/s12559-021-09941-7.

DUDA, RICHARD O. AND HART, PETER E. AND STORK, David G. **Pattern Classification**. 2nd. ed. New York: Wiley-Interscience, 2000.

DUNCAN-JOHNSON, Connie C.; DONCHIN, Emanuel. **On Quantifying Surprise: The Variation of Event-Related Potentials With Subjective Probability**. **Psychophysiology**, 1977. DOI: 10.1111/j.1469-8986.1977.tb01312.x.

FALKENSTEIN, M.; HOHNSBEIN, J.; HOORMANN, J.; BLANKE, L. Effects of crossmodal divided attention on late E R P components . II . Error processing in choice reaction tasks. **Electroencephalography and clinical neurophysiology**, *[S. l.]*, n. 78, p. 447–455, 1991.

FARWELL, L. A.; DONCHIN, E. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. **Electroencephalography and Clinical Neurophysiology**, *[S. l.]*, v. 70, p. 510–523, 1988.

FERREZ, Pierre W.; MILLÁN, José R. Error-Related EEG Potentials Generated During Simulated Brain – Computer Interaction. **IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING**, *[S. l.]*, v. 55, n. 3, p. 923–929, 2008.

FORNER-CORDERO, Arturo; QUADRADO, Virgínia H.; TSAGBEY, Sitsofe A.; SMITS-ENGELSMAN, Bouwien C. M. Improved Learning a Coincident Timing Task With a Predictable Resisting Force. **Motor Control**, *[S. l.]*, 2017.

FRANK, Ernest. Electric Potential Produced by Two Point Current Sources in a Homogeneous Conducting Sphere. **Journal of Applied Physics**, *[S. l.]*, v. 23, n. 1225, 1952. DOI: 10.1063/1.1702037.

FUKUNAGA, K. **Introduction to Statistical Pattern Recognition**. 2. ed. New York: Academic Press, 1972.

GARRETT, Douglas D.; SAMANEZ-LARKIN, Gregory R.; MACDONALD, Stuart W. S.; LINDENBERGER, Ulman; MCINTOSH, Anthony R.; GRADY, Cheryl L. Moment-to-moment brain signal variability: A next frontier in human brain mapping? **Neuroscience and Biobehavioral Reviews**, *[S. l.]*, v. 37, n. 4, p. 610–624, 2013. DOI: 10.1016/j.neubiorev.2013.02.015. Disponível em: http://dx.doi.org/10.1016/j.neubiorev.2013.02.015.

GEHRING, William J.; GOSS, Brian; COLES, Michael G. H.; MEYER, David E.; DONCHIN, Emanuel. A Neural System for Error Detection and Compensation. **Psychological Science**, *[S. l.]*, v. 4, n. 6, p. 385–390, 1993. DOI: 10.1111/j.1467-9280.1993.tb00586.x.

GESELOWITZ, David B. The zero of potential. **IEEE Engineering in Medicine and Biology Magazine**, *[S. l.]*, v. 17, n. 1, 1998. DOI: 10.1109/51.646230.

GRECH, Roberta; CASSAR, Tracey; MUSCAT, Joseph; CAMILLERI, Kenneth P.; FABRI, Simon G.; ZERVAKIS, Michalis; XANTHOPOULOS, Petros; SAKKALIS, Vangelis; VANRUMSTE, Bart. Review on solving the inverse problem in EEG source analysis. **Journal of NeuroEngineering and Rehabilitation**, *[S. l.]*, v. 5, p. 1–33, 2008. DOI: 10.1186/1743-0003-5-25.

GUGER, Christoph et al. Brain-computer interfaces for stroke rehabilitation : summary of the 2016 BCI Meeting in Asilomar Brain-computer interfaces for stroke rehabilitation : summary of the 2016 BCI Meeting in Asilomar. **Brain-Computer Interfaces**, *[S. l.]*, v. 00, n. 00, p. 1–17, 2018. DOI: 10.1080/2326263X.2018.1493073. Disponível em: https://doi.org/10.1080/2326263X.2018.1493073.

HENDERSON, C. J.; BUTLER, S. R.; GLASS, A. The Localization of Equivalent Dipoles of EEG Sources by the Application of Electrical Field Theory. **Electroencephalography and Clinical Neurophysiology**, *[S. l.]*, v. 39, p. 117–130, 1975.

HJORTH, Bo. An on-line transformation of EEG scalp potentials into orthogonal source derivations. **Electroencephalography and Clinical Neurophysiology**, *[S. l.]*, v. 39, n. 5, p. 526–530, 1975. DOI: 10.1016/0013-4694(75)90056-5.

HOLROYD, Clay B.; NIEUWENHUIS, C. A. Sander; NICK, Y.; COHEN, Jonathan D. Errors in reward prediction are reflected in the event-related brain potential. **COGNITIVE NEUROSCIENCE AND NEUROPSYCHOLOGY**, *[S. l.]*, v. 14, n. 18, 2003. DOI: 10.1097/01.wnr.0000099601.41403.a5.

HU, Shiang; LAI, Yongxiu; VALDES-SOSA, Pedro A.; BRINGAS-VEGA, Maria L.; YAO, Dezhong. How do reference montage and electrodes setup affect the measured scalp EEG potentials? **Journal of Neural Engineering**, *[S. l.]*, n. 15, 2018. DOI: 10.1088/1741-2552/aaa13f. Disponível em: https://doi.org/10.1088/1741-2552/aaa13f.

JASPER, H. H. The Ten-Twenty Electrode System of the International Federation. **Electroencephalography and Clinical Neurophysiology**, *[S. l.]*, v. 10, p. 371–375, 1958.

JUNGHÖFER, M.; ELBERT, T.; TUCKER, D. M.; BRAUN, C. The polar average reference effect: A bias in estimating the head surface integral in EEG recording. **Clinical Neurophysiology**, *[S. l.]*, v. 110, n. 6, p. 1149–1155, 1999. DOI: 10.1016/S1388-2457(99)00044-9.

KALCHER, J.; PFURTSCHELLER, G. Discrimination between phase-locked and non-phase-locked event-related EEG activity. **Electroencephalography and clinical neurophysiology**, *[S. l.]*, v. 94, p. 381–384, 1995.

KAYSER, Jürgen; TENKE, Craig E. **Issues and considerations for using the scalp surface Laplacian in EEG/ERP research: A tutorial review**. **International Journal of Psychophysiology**, 2015. DOI: 10.1016/j.ijpsycho.2015.04.012.

KOBLER, Reinmar J.; SBURLEA, Andreea I.; MÜLLER-PUTZ, Gernot R. Tuning characteristics of low-frequency EEG to positions and velocities in visuomotor and oculomotor tracking tasks. **Scientific Reports**, *[S. l.]*, v. 8, n. 1, p. 1–14, 2018. DOI: 10.1038/s41598-018-36326-y.

KOLES, Zoltan J.; LAZAR, Michael S.; ZHOU, Steven Z. Spatial Patterns Underlying Population Differences in the Background EEG Rp-NP. **Brain Topography**, *[S. l.]*, v. 2, n. 4, p. 275–284, 1990.

KONONOWICZ, Tadeusz W.; PENNEY, Trevor B. The contingent negative variation (CNV): timing isn't everything. **Current Opinion in Behavioral Sciences**, *[S. l.]*, n. 8, p. 231–237, 2016.

KORNHUBER, Hans H.; DEECKE, Lüder. Brain potential changes in voluntary and passive movements in humans:readiness potential and reafferent potentials. **Pflügers Archiv - European Journal of Physiology**, *[S. l.]*, p. 1115–1124, 1965. DOI: 10.1007/s00424-016-1852-3. Disponível em: http://dx.doi.org/10.1007/s00424-016-1852-3.

KRIGOLSON, O. E.; HOLROYD, C. B. Evidence for Hierarchical Error Processing in the Human Brain. **Neuroscience**, *[S. l.]*, v. 137, p. 13–17, 2006. DOI: 10.1016/j.neuroscience.2005.10.064.

KROPP, Peter; KIEWITT, Angela; HARTMUT, G.; VETTER, Peter; GERBER, Wolf-Dieter. Reliability and Stability of Contingent Negative Variation. **Applied Psychophysiology and Biofeedback**, *[S. l.]*, v. 25, n. 1, p. 33–41, 2000.

LEBEDEV, Mikhail A.; NICOLELIS, Miguel A. L. Brain-machine Interfaces: from basic science to neuroprosthesis and neurorehabilitation. **Physiological Reviews**, *[S. l.]*, n. 37, p. 767–837, 2017. DOI: 10.1152/physrev.00027.2016.

LEW, Eileen; CHAVARRIAGA, Ricardo; SILVONI, Stefano; MILLÁN, José R.; IFFT, Peter J. Detection of self-paced reaching movement intention from EEG signals. **Frontiers in Neuroengineering**, *[S. l.]*, v. 5, n. July, p. 1–17, 2012. DOI: 10.3389/fneng.2012.00013.

LIU, Quanying; BALSTERS, Joshua H.; BAECHINGER, Marc; VAN DER GROEN, Onno; WENDEROTH, Nicole; MANTINI, Dante; GROEN, Onno Van Der; WENDEROTH, Nicole; MANTINI, Dante. Estimating a neutral reference for electroencephalographic recordings: The importance of using a high-density montage and a realistic head model. **Journal of Neural Engineering**, *[S. l.]*, v.

12, n. 5, 2015. DOI: 10.1088/1741-2560/12/5/056012. Disponível em: http://dx.doi.org/10.1088/1741-2560/12/5/056012.

LÓPEZ-LARRAZ, E.; SARASOLA-SANZ, A.; IRASTORZA-LANDA, N.; BIRBAUMER, N.; RAMOS-MURGUIALDAY, A. Brain-machine interfaces for rehabilitation in stroke: A review. **NeuroRehabilitation**, *[S. l.]*, v. 43, n. 1, p. 77–97, 2018. DOI: 10.3233/NRE-172394.

LUCK, Steven J. **An Introduction to the Event-Related Potential Technique**. Second ed. [s.l.] : The MIT Press, 2014.

MANSOUR, Salem; ANG, Kai Keng; NAIR, Krishnan P. S.; PHUA, Kok Soon; ARVANEH, Mahnaz. Efficacy of Brain–Computer Interface and the Impact of Its Design Characteristics on Poststroke Upper-limb Rehabilitation: A Systematic Review and Meta-analysis of Randomized Controlled Trials. **Clinical EEG and Neuroscience**, *[S. l.]*, v. 53, n. 1, p. 79–90, 2022. DOI: 10.1177/15500594211009065. Disponível em: https://doi.org/10.1177/15500594211009065.

MASAKI, Hiroaki; SOMMER, Werner. Cognitive neuroscience of motor learning and motor control. **The Journal of Physical Fitness and Sports Medicine**, *[S. l.]*, v. 1, n. 3, p. 369–380, 2012. DOI: 10.7600/jpfsm.1.369.

MASAKI, Hiroaki; SOMMER, Werner; TAKASAWA, Noriyoshi; YAMAZAKI, Katuo. Neural mechanisms of timing control in a coincident timing task. **Experimental Brain Research**, *[S. l.]*, v. 218, n. 2, p. 215–226, 2012. DOI: 10.1007/s00221-012-3052-5.

MASCOLINI, Alessio; KHAN, Imran; MESIN, Luca. Non-linear optimized spatial filter for single-trial identification of movement related cortical potential. **Biocybernetics and Biomedical Engineering**, *[S. l.]*, v. 42, n. 1, p. 426–436, 2022. DOI: 10.1016/j.bbe.2022.02.013. Disponível em: https://doi.org/10.1016/j.bbe.2022.02.013.

MCFARLAND, D. J.; WOLPAW, J. R. EEG-based Brain-computer Interfaces. **Current Opinion in Biomedical Engineering**, *[S. l.]*, v. 4, p. 194–200, 2017. DOI: 10.1016/j.cobme.2017.11.004. Disponível em: https://doi.org/10.1016/j.cobme.2017.11.004.

MEZIANI, A. **Legendre polynomials and applications**. 2016. DOI: 10.1090/amsip/006/37. Disponível em: https://faculty.fiu.edu/~meziani/Note13.pdf. Acesso em: 25 jul. 2021.

MILTNER, Wolfgang H. R.; BRAUN, Christoph H.; COLES, Michael G. H. Event-Related Brain Potentials Following Incorrect Feedback in a Time Estimation Task: Evidence for a “Generic” Neural System for Error Detection. **Journal of Cognitive Neuroscience**, *[S. l.]*, v. 9, n. 6, p. 788–798, 1997.

MOLTENI, Erika; CIMOLIN, Veronica; PREATONI, Ezio; RODANO, Renato; GALLI, Manuela; BIANCHI, Anna M. Towards a biomarker of motor adaptation: Integration of kinematic and neural factors. **IEEE Transactions on Neural Systems and Rehabilitation Engineering**, *[S. l.]*, v. 20, n. 3, p. 258–267, 2012. DOI: 10.1109/TNSRE.2012.2189585.

MUNCK, Jan C.; DIJK, Bobo W. Van; SPEKREIJSE, Henk. Mathematical Dipoles

are Adequate to Describe Realistic Generators of Human Brain Activity. **IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING**, *[S. l.]*, v. 35, n. 11, 1988.

NEAT, Gregory W.; MCFARLAND, Dennis J.; WOLPAW, Jonathan R.; FORNERIS, Catherine A. EEG-Based Brain-to-Computer Communication: System Description. *In*: ANNUAL INTERNATIONAL CONFERENCE OF THE IEEE ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY 1990, **Anais** [...]. [s.l: s.n.] p. 2298–2300.

NICOLAS-ALONSO, Luis Fernando; GOMEZ-GIL, Jaime. Brain computer interfaces, a review. **Sensors**, *[S. l.]*, v. 12, n. 2, p. 1211–1279, 2012. DOI: 10.3390/s120201211.

NUNEZ, P. L.; SILBERSTEIN, R. B.; CADUSCH, P. J.; WIJESINGHE, R. S.; WESTDORP, A. F.; SRINIVASAN, R. A theoretical and experimental study of high resolution EEG based on surface Laplacians and cortical imaging. **Electroencephalography and Clinical Neurophysiology**, *[S. l.]*, v. 90, n. 1, p. 40–57, 1994. DOI: 10.1016/0013-4694(94)90112-0.

NUNEZ, Paul L. REST: A good idea but not the gold standard. **Clinical Neurophysiology**, *[S. l.]*, v. 121, n. 12, p. 2177–2180, 2010. DOI: 10.1016/j.clinph.2010.04.029.

NUNEZ, Paul L.; SRINIVASAN, Ramesh. **Electric Fields of the Brain: The Neurophysics of EEG**. 2nd. ed. Oxford: Oxford University Press, 2006.

OFNER, Patrick; SCHWARZ, Andreas; PEREIRA, Joana; MÜLLER-PUTZ, Gernot R. Upper limb movements can be decoded from the time-domain of low-frequency EEG. **PLoS ONE**, *[S. l.]*, v. 12, n. 8, p. 1–24, 2017. DOI: 10.1371/journal.pone.0182578.

OOSTENVELD, Robert; PRAAMSTRA, Peter. The five percent electrode system for high-resolution EEG and ERP measurements. **Clinical Neurophysiology**, *[S. l.]*, v. 112, n. 4, p. 713–719, 2001. DOI: 10.1016/S1388-2457(00)00527-7.

ORBAN, Mostafa; ELSAMANTY, Mahmoud; GUO, Kai; ZHANG, Senhao; YANG, Hongbo. A Review of Brain Activity and EEG-Based Brain–Computer Interfaces for Rehabilitation Application. **Bioengineering**, *[S. l.]*, v. 9, n. 12, 2022. DOI: 10.3390/bioengineering9120768.

PASCUAL-MARQUI, R. D.; MICHEL, C. M.; LEHMANN, D. Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. **International Journal of Psychophysiology**, *[S. l.]*, v. 18, n. 1, p. 49–65, 1994. DOI: 10.1016/0167-8760(84)90014-X.

PERRIN, F.; PERNIER, J.; BERTRAND, O.; ECHALLIER, J. F. Spherical splines for scalp potential and current density mapping. **Electroencephalography and Clinical Neurophysiology**, *[S. l.]*, v. 72, n. 2, p. 184–187, 1989. DOI: 10.1016/0013-4694(89)90180-6.

PFURTSCHELLER, G. Graphical Display and Statistical Evaluation of Event-Related Desynchronization (ERD). **Electroencephalography and clinical neurophysiology**, *[S. l.]*, n. 43, p. 757–760, 1977.

PFURTSCHELLER, G. Functional brain imaging based on ERD/ERS. **Vision**

Research, *[S. l.]*, v. 41, p. 1257–1260, 2001.

PFURTSCHELLER, G.; ARANIBAR, A. Event-Related Cortical Desynchronization Detected by Power Measurements of Scalp EEG. **Electroencephalography and clinical neurophysiology**, *[S. l.]*, n. 42, p. 817–826, 1977.

PFURTSCHELLER, G.; ARANIBAR, A. Evaluation of Event-Related Desynchronization (ERD) Preceding and Following Voluntary Self-paced Movement. **Electroencephalography and clinical neurophysiology**, *[S. l.]*, v. 46, n. 3535, p. 138–146, 1979.

PFURTSCHELLER, G.; LOPES DA SILVA, F. H. Event-related EEG/MEG synchronization and desynchronization: basic principles. **Clinical Neurophysiology**, *[S. l.]*, v. 110, p. 1842–1857, 1999.

PFURTSCHELLER, G.; LOPES, F. H. Event-related EEG/MEG Synchronization and Desynchronization: Basic Principles. **Clinical Neurophysiology**, *[S. l.]*, v. 110, p. 1842–1857, 1999.

PFURTSCHELLER, Gert; FLOTZINGER, Doris; KALCHER, Joachim. Brain-Computer Interface - a new communication device for handicapped persons. **Journal of Microcomputer Applications**, *[S. l.]*, n. 16, p. 293–299, 1993.

PICTON, T. W. et al. Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. **Psychophysiology**, *[S. l.]*, p. 127–152, 2000.

PICTON, Terence. Hearing in Time: Evoked Potential Studies of Temporal Processing. **Ear and Hearing**, *[S. l.]*, p. 385–401, 2013.

PORT, Nicholas Lindman; LEE, Daeyeol; DASSONVILLE, Paul; GEORGOPOULOS, Apostolos P. Manual interception of moving targets. I. Performance and movement initiation. **Experimental Brain Research**, *[S. l.]*, v. 116, p. 406–420, 1997. Disponível em: papers3://publication/uuid/384A5F21-CB72-4AA9-B496-E8E6B4A1BD03.

POULTON, E. E. **On prediction in skilled movements**. **Psychological Bulletin**, 1957. DOI: 10.1037/h0045515.

QIN, Yun; XU, Peng; YAO, Dezhong. A comparative study of different references for EEG default mode network : The use of the infinity reference. **Clinical Neurophysiology**, *[S. l.]*, v. 121, n. 12, p. 1981–1991, 2010. DOI: 10.1016/j.clinph.2010.03.056. Disponível em: http://dx.doi.org/10.1016/j.clinph.2010.03.056.

RAMADAN, Rabie A.; VASILAKOS, Athanasios V. Brain computer interface: control signals review. **Neurocomputing**, *[S. l.]*, v. 223, 2017. DOI: 10.1016/j.neucom.2016.10.024.

RAMOS-MURGUIALDAY, Ander; BIRBAUMER, Niels. Brain oscillatory signatures of motor tasks. **Journal of neurophysiology**, *[S. l.]*, v. 113, n. 10, p. 3663–82, 2015. DOI: 10.1152/jn.00467.2013. Disponível em: http://www.ncbi.nlm.nih.gov/pubmed/25810484%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4468978.

SAHA, Simanto; BAUMERT, Mathias. Intra- and Inter-subject Variability in EEG-Based Sensorimotor Brain Computer Interface: A Review. **Frontiers in Computational Neuroscience**, *[S. l.]*, v. 13, n. January, p. 1–8, 2020. DOI: 10.3389/fncom.2019.00087.

SANNELLI, Claudia; VIDAURRE, Carmen; MÜLLER, Klaus Robert; BLANKERTZ, Benjamin. **A large scale screening study with a SMR-based BCI: Categorization of BCI users and differences in their SMR activity**. [s.l: s.n.]. v. 14 DOI: 10.1371/journal.pone.0207351.

SCHALK, Gerwin; WOLPAW, Jonathan R.; MCFARLAND, Dennis J.; PFURTSCHELLER, Gert. EEG-based communication: presence of an error potential. **Clinical Neurophysiology**, *[S. l.]*, v. 111, p. 2138–2144, 2000.

SCHMIDT, Richard A.; WHITE, Janice L. Evidence for an error detection mechanism in motor skills: A test of adams' closed-loop theory. **Journal of Motor Behavior**, *[S. l.]*, v. 4, n. 3, p. 143–153, 1972. DOI: 10.1080/00222895.1972.10734930.

SHIBASAKI, H.; BARRETT, G.; HALLIDAY, Elise; HALLIDAY, A. .. Components of the movement-related cortical potential and their scalp topography. **Electroencephalography and Clinical Neurophysiology**, *[S. l.]*, v. 49, n. 3–4, p. 213–226, 1980. DOI: 10.1016/0013-4694(80)90216-3.

SHIBASAKI, Hiroshi; HALLETT, Mark. What is the Bereitschaftspotential? **Clinical Neurophysiology**, *[S. l.]*, v. 117, n. 11, p. 2341–2356, 2006. DOI: 10.1016/j.clinph.2006.04.025.

SPÜLER, Martin; NIETHAMMER, Christian. Error-related potentials during continuous feedback : using EEG to detect errors of different type and severity. **Frontiers in Human Neuroscience**, *[S. l.]*, v. 9, n. March, p. 1–10, 2015. DOI: 10.3389/fnhum.2015.00155.

SQUIRES, Nancy K.; SQUIRES, Kenneth C.; HILLYARD, Steven A. Two Varieties of Long-Latency Positive Waves Evoked by Unpredictable Auditory Stimuli in Man. **Electroencephalography and clinical neurophysiology**, *[S. l.]*, n. 38, p. 387–401, 1975.

SRINIVASAN, Ramesh; NUNEZ, Paul L.; TUCKER, Don M.; SILBERSTEIN, Richard B.; CADUSCH, Peter J. Spatial sampling and filtering of EEG with spline laplacians to estimate cortical potentials. **Brain Topography**, *[S. l.]*, v. 8, n. 4, p. 355–366, 1996. DOI: 10.1007/BF01186911.

STOK, Cees J. Single Dipole Source Estimation. **IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING**, *[S. l.]*, v. 34, n. 4, p. 289–296, 1987.

SUTTON, Samuel; BRAREN, Margery; ZUBIN, Joseph; R., John E. Evoked-Potential Correlates of Stimulus Uncertainty. **Science**, *[S. l.]*, p. 1187–1188, 1965.

TARKKA, I. M.; HALLETT, M. Topography of Scalp-Recorded Motor Potentials in Human Finger Movements. **Journal of Clinical Neurophysiology**, *[S. l.]*, v. 8, n. 3, p. 331–341, 1991. DOI: 10.1097/00004691-199107010-00009.

TRESILIAN, James R. The accuracy of interceptive action in time and space.

Exercise and Sport Sciences Reviews, *[S. l.]*, v. 32, n. 4, p. 167–173, 2004. DOI: 10.1097/00003677-200410000-00008.

URIGÜEN, Jose Antonio; GARCIA-ZAPIRAIN, Begoña. EEG artifact removal - State-of-the-art and guidelines. **Journal of Neural Engineering**, *[S. l.]*, v. 12, n. 3, 2015. DOI: 10.1088/1741-2560/12/3/031001.

VIDAL, Jacques Jean Jacques. Toward Direct Brain-Computer Communication. **Annu. Rev. Biophys. Bioeng.**, *[S. l.]*, v. 2, p. 157–180, 1973.

WAHBA, G. **Spline Interpolation and Smoothing on the Sphere**. **Siam Journal on Scientific and Statistical Computing**, 1981. DOI: 10.1137/0902002. Disponível em: http://gateway.webofknowledge.com/gateway/Gateway.cgi?GWVersion=2&SrcAuth=mekentosj&SrcApp=Papers&DestLinkType=FullRecord&DestApp=WOS&KeyUT=A1981LU00400002%5Cnpapers2://publication/uuid/1D6F027F-1B11-4DBC-82DE-01187E5B45C0.

WAHBA, Grace. Spline models for observational data. **CBMS-NSF Regional Conference Series in Applied Mathematics Regional Conference Series in Applied Mathematics**, *[S. l.]*, v. 59, p. 13, 1990. DOI: 10.1016/0021-9045(91)90041-8.

WALTER, W. Grey; COOPER, R.; ALDRIDGE, V. J.; MCCALLUM, W. C. Contingent Negative Variation: an Eletric Sign of Sensori- Motor Association and Expectancy in the Human Brain. **Nature**, *[S. l.]*, v. 203, p. 380–384, 1964.

WOLPAW, Jonathan R.; MCFARLAND, Dennis J.; NEAT, Gregory W.; FORNERIS, Catherine A. An EEG-based brain-computer interface for cursor control. **Electroencephalography and Clinical Neurophysiology**, *[S. l.]*, v. 78, n. 3, p. 252–259, 1991. DOI: 10.1016/0013-4694(91)90040-B.

WOLPAW, Jonathan R.; WOLPAW, Elizabeth Winter. **Brain-computer Interfaces: principles and practice**. [s.l: s.n.].

YAO, D. A method to standardize a reference of scalp EEG recordings to a point at infinity. **Physiological Measurement**, *[S. l.]*, v. 22, n. 4, p. 693–711, 2001. DOI: 10.1088/0967-3334/22/4/305.

ZENG, Ke; LI, Xiaoli. Artifact Removal in EEG Recordings. *In*: **Signal Processing in Neuroscience**. [s.l: s.n.]. p. 288.