NTNU

Norwegian University of
Science and Technology

A Unified Real-Time Feature Extraction and Classification Process for a BCI Based on Empirical Mode Decomposition and Support Vector Machine

Joar Molvær
Fredrik Worren

Master of Science in Cybernetics and Robotics
Submission date: June 2016
Supervisor: Marta Maria Cabrera Molinas, ITK

Norwegian University of Science and Technology
Department of Engineering Cybernetics

ABSTRACT

This thesis has studied and implemented a data-driven process combining EEG feature extraction based on empirical mode decomposition (EMD) and classification using a support vector machine (SVM). The experimental paradigm followed a persistent approach using professional EEG acquisition equipment. Eight healthy, right-handed subjects were recruited, all signing consent forms prior to their respective recording session. Each session consisted of 70 trials for three different conditions, namely left fist clenching, right fist clenching, and a neutral condition in which no movement occurred. The EEG data were recorded using a 256-electrode Geodesic Sensor Net (GSN).

The data analysis were performed using the adaptive EMD method, in which the oscillatory modes of the EEG signals were extracted and expressed by intrinsic mode functions (IMFs). The IMFs were normalized and their corresponding Hilbert transforms were computed. This process is referred to as the normalized Hilbert transform, from which a meaningful instantaneous frequency and amplitude can be obtained. The time-frequency solutions were visualized in Hilbert spectra, and a decrease in the signal power was recognized during left and right hand movements, known as event-related desynchronization (ERD). The mean ERD during the three conditions were calculated and utilized as features for the classification scheme. A given number of features were employed as support vectors, and the features of the remaining trials were predicted using an SVM classifier. Classification accuracies of up to 95% were acquired from multiple subjects, along with an average accuracy of 87.86% when differentiating left and right fist clenching. The computational effort of this unified process is found to be within the real-time demands of a brain-computer interface (BCI).

SAMMENDRAG

Denne masteroppgaven omhandler forskning og utvikling av en data-drevet prosess. Prosessen kombinerer ekstrahering av relevante egenskaper i EEG-data ved hjelp av empirical mode decomposition (EMD) og klassifisering av håndbevegelser ved bruk av en support vector machine (SVM). Det eksperimentelle paradigmet fulgte en konsistent prosedyre, og data ble målt med profesjonellt EEG-utstyr. Åtte friske, høyrehendte deltakere ble rekruttert, hvorav alle signerte samtykkeskjema før eksperimentet ble gjennomført. Eksperimentet bestod av tre forskjellige oppgaver, med 70 forsøk for hver oppgave. De tre oppgavene var enten å åpne og lukke høyre eller venstre knyttneve, eller å ikke utføre noen håndbevegelse. Data ble målt og lagret ved bruk av et Geodesic Sensor Net (GSN) med 256 elektroder.

Dataanalysen ble utført ved å bruke den adaptive EMD-metoden, hvor forskjellige oscillasjonsmoduser ble ekstrahert og uttrykt som intrinsic mode functions (IMFs). Alle IMFer ble normalisert, og den korresponderende Hilbert-transformasjonen ble kalkulert. Denne prosessen er kjent som normalisert Hilbert-transformasjon, hvor en meningsfull momentan frekvens og amplitude kan bli anskaffet. Tid-frekvens-responsen til signalene ble representert i et Hilbert spektrum, og en demping i signalets energi ble identifisert under både venstre og høyre håndbevegelser. Denne hendelsen er kjent som event-related desynchronization (ERD). Den gjennomsnittlige ERD ble brukt som egenskap i klassifiseringsprosessen, og en nøyaktighet på opptil 95% ble oppnådd hos flere av deltakerne. Videre ble en gjennomsnittlig nøyaktighet på 87.86% oppnådd ved differensiering mellom venstre og høyre håndbevegelse. Prosesseringstiden for den komplette metoden som er implementert i denne masteroppgaven er innenfor sanntidsrammene relatert til et brain-computer interface (BCI).

ACKNOWLEDGEMENTS

This thesis ends an enriching five-year Master's degree program in Cybernetics and Robotics.

We would first like to thank our supervisor Professor Marta Molinas of The Department of Engineering Cybernetics at The Norwegian University of Science and Technology (NTNU). Marta is the most dedicated and insightful Professor we have had the privilege to meet. She has, from day one, been seeking to expand our academic network and without her there would be no cooperation between NTNU and the Research Center for Adaptive Data Analysis at the National Central University in Taiwan. Because of this cooperation, we got the opportunity to spend two weeks in Taiwan, learning from one of the pioneers in adaptive data analysis, Professor Norden Huang. We would like to express our profound gratitude to Prof. Huang and his colleagues for all the hospitality and the unique lectures during our stay.

We would also like to acknowledge the scientists at the Developmental Neuroscience Laboratory at NTNU for providing access and support during our EEG experiments.

Also, thankfulness should be expressed to the guys in our office for providing a social and engaging work environment. Finally, we would like to thank our girlfriends for their support and continuous encouragement during our research.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

BCI Brain-computer interface

EEG Electroencephalogram

EKG Electrocardiogram

EOG Electrooculogram

EMG Electromyogram

EMD Empirical mode decomposition

ERP Event-related potential

ERD Event-related desynchronization

ERS Event-related synchronization

HHT Hilbert-Huang transform

HS Hilbert spectrum

IMF Intrinsic mode function

ICA Independent component analysis

LSF Laplacian spatial filter

MEG Magnetoencephalogram

NHT Normalized Hilbert transform

SD Standard deviation

SVM Support vector machine

1

INTRODUCTION

The first human electroencephalogram (EEG) was, in 1929, recorded by a German psychiatrist named Hans Berger. Since then, the EEG has been involved with numerous studies including head injuries, brain tumors, memory problems, sleep disorders, strokes, dementia, and seizure disorders such as epilepsy. Previous research has examined the possibilities of using the recorded EEG activities for actuation purposes, for instance controlling a wheelchair or a prosthetic arm. If the electrical activity of the brain during a motor-related event can be decoded, people with impaired motor skills may be assisted by neurofeedback. A form of neurofeedback can provide a real-time display of the brain activity to study and train different brain functions. This interaction between the cortical activity and computerized technology is referred to as a brain-computer interface (BCI). A BCI has several application fields, for instance, neurorehabilitation in stroke patients using EEG-based motor imagery [3]. Furthermore, no risks are linked to the noninvasive EEG, in which the electrodes are placed on the scalp.

When analyzing the cortical activity in different parts of the brain, both temporal and spectral investigations have to be conducted, by calculating the time-frequency solutions of the EEG signals. Different techniques have been developed to transform a signal from the time domain to the frequency domain, revealing its spectral information. An approach stated in [44], combines the Fourier transform and machine learning to detect epileptic seizures in EEG signals. The wavelet transform, an extension of the Fourier transform, has also been utilized in an EEG-based seizure detection scheme [13]. Both of these transforms have certain restrictions when considering the nonlinearity and nonstationarity of the EEG signals. Consequently, this thesis will focus on the application of the normalized Hilbert transform, in which the empirical mode decomposition (EMD) extracts meaningful oscillatory modes from the EEG signals. The method is adaptive, meaning that no prior knowledge about the signal is required. Moreover, such a data-driven approach may be more suitable for nonlinear and nonstationary signals. With this in mind, the following research questions will be investigated throughout this thesis:

- **RQ1:** *Can empirical mode decomposition (EMD) be used to extract features suitable to detect and differentiate motor-related events in EEG data?*
- **RQ2:** *Can the normalized Hilbert transform be applied to fully represent the time-frequency solutions of EEG data? If so, are the event-related (de)synchronizations observable in the corresponding Hilbert spectra?*
- **RQ3:** *Will the unified feature extraction and classification process be able to operate under the real-time demands of a brain-computer interface (BCI)?*

In a study conducted by [46], the detection and classification of both executed and imagined fist clenching, both in healthy subjects and stroke patients, have been performed using only one electrode. The reported classification accuracy was approximately 75%. When considering such a classification problem, it is important to extract proper features to separate the conditions. One study has proposed a methodology for detecting seizures such as epilepsy based on the extraction of temporal and spectral features by applying EMD to EEG signals, demonstrating the strength of the data-driven analysis [47].

The present study explores the neural activities of motor-related tasks, trying to differentiate between left and right fist clenching. The study provides information that contributes to the understanding of event-related desynchronizations (ERDS) during muscular movements, in addition to the application of adaptive signal analysis on EEG data. Moreover, the proposed procedure will investigate the possibility of being suitable for real-time BCI applications. It was hypothesized that induced ERDs would be observed during hand movements, and that the corresponding frequency spectra would resemble those of the brain rhythms recognized in [43].

The unified feature extraction and classification process implemented and studied throughout this research is divided into several modules. These modules will be explained in detail both in the theory chapters, *Chapter* 3, 4, 5, and 6, in addition to the method chapter, that is, *Chapter* 7. The different modules, showing the entire process from the raw EEG data to the classification of hand movements, are illustrated using the flow chart in Figure 1.

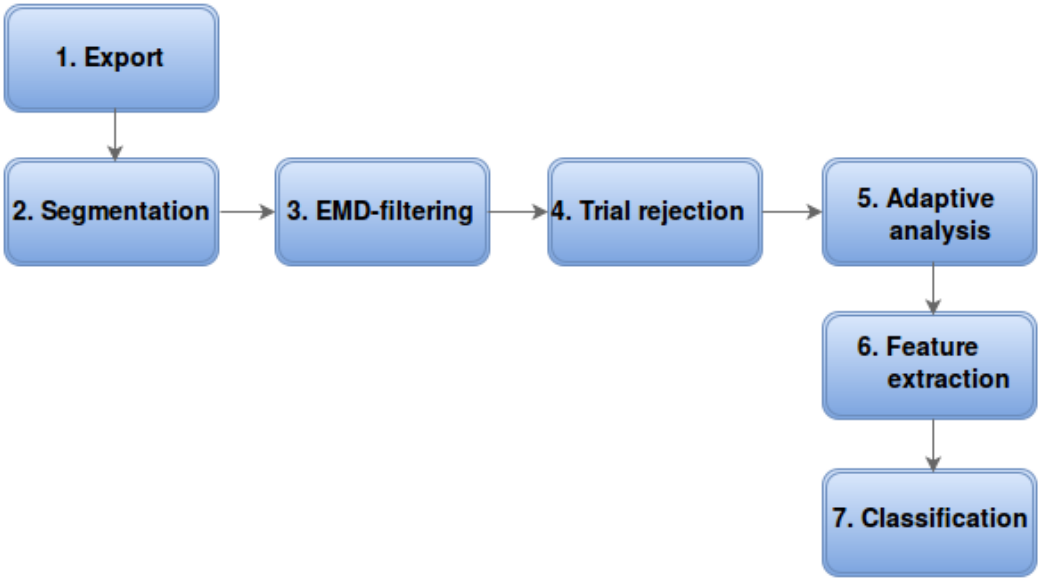

Figure 1.: The sequential steps in the unified feature extraction and classification process, from raw EEG data to the classification of hand movements.

Part I

BACKGROUND LITERATURE AND THEORETICAL FOCUS

2

ELECTROENCEPHALOGRAM

The following chapter explains briefly how numerous nerve cells in the brain can produce electric fields measurable from electrodes placed on the scalp, known as scalp EEG. Three of the internationally recognized standards for the naming and placing of the electrodes are outlined, namely the 10/20, 10/10, and 10/5 systems. In addition to discussing the smearing effect related to the scalp EEG, the different brain rhythms and their respective roles compared to the human consciousness are presented. The interpretations of evoked event-related potential (ERP) and induced event-related desynchronization (ERD) are clarified, and how the two differ from each other.

2.1 FROM ACTION POTENTIAL TO EEG

Nerve cells in the brain are called neurons, and they produce time-dependent electrical currents when activated [14]. The two primary forms of activation are fast depolarization of the neuronal membranes, referred to as *action potentials*, and slower changes in the membrane due to synaptic activation. These slow changes are called *postsynaptic potentials*, and they are mainly divided into inhibitory postsynaptic potentials (IPSPs) and excitatory postsynaptic potentials (EPSPs), which purposes are to either repress or initiate action potentials, respectively.

Regarding EEG measurements, the foremost contributors are the neurons that form *open fields*. These fields, as described in [42], are aligned perpendicularly to the cortical surface. When such neurons are activated synchronously, they produce electric and magnetic fields similar to those of the *dipole* illustrated in Figure 2. The potential voltage difference between the dipoles can be measured by electrodes placed on the scalp, whereas the magnetic activity can be detected using magnetometers. These two procedures are known as an *electroencephalogram* (EEG) and a *magnetoencephalogram* (MEG), in that order. As stated in [41], a single scalp electrode measures the sum of synaptic potential from between roughly 100 million and 1 billion neurons, which renders a large-scale measure of cortical activity. Potentials that

cancel each other out, like asynchronous activity and dipoles of opposite orientation, are not contributing to the EEG measurements.

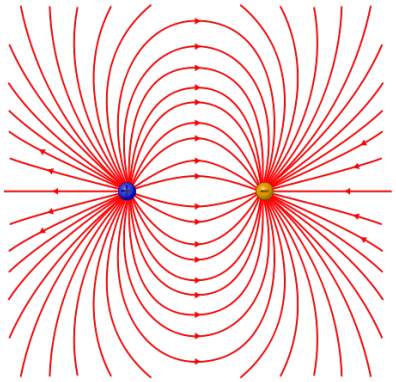

Figure 2.: Electric dipole (Photo: [19])

2.2 Scalp EEG

Different types of EEG recording electrodes may vary in configuration, size, or sensitivity. The number of electrodes can range from a single electrode to a couple of hundreds, and there exist both dry electrodes and electrodes that have to be soaked in some form of saline solution before being placed on the scalp. Moreover, the measured signal is the potential voltage difference between a given electrode and a specific *reference electrode*, and the method is therefore called *referential EEG recording*. Most EEG acquisition equipment can record using a high sampling frequency, resulting in a high temporal resolution. In contrast to this, the natural conductivity of the brain makes the process of identifying the source of the EEG signal difficult. This is known as the *smearing effect*, and results in low spatial resolution [56]. The Laplacian spatial filter (LSF) can be utilized to enhance local activity, hence increasing the spatial resolution. The LSF will be further discussed in Section 5.3.

Given the various recording methods and number of electrodes, some internationally recognized standards for the naming and placing of the electrodes have been developed. The most prevalent standard is called the *10/20* system, which signifies that the distances between adjacent electrodes are either 10% or 20% of the total distance from the front to the back of the skull or the distance from right to left. The 10/20 system is illustrated in Figure 3. The identification letters refer to the lobes, namely, **F**rontal, **T**emporal, **P**arietal, and **O**ccipital. Moreover, the letter **C** stands for central, and the **z** (zero) refers to an electrode placed on the midline. The odd and even numbers represent electrodes placed on the left and right hemisphere, in that order. The *10/20* system provides electrode naming and placement for 21 electrodes. If in need for an increased number of electrodes, one can employ the *10/10* system

(81 electrodes) or the *10/5* system, which can represent the placement and naming of over 300 electrodes [34]. Figure 4 and Figure 5 illustrate these two additional systems.

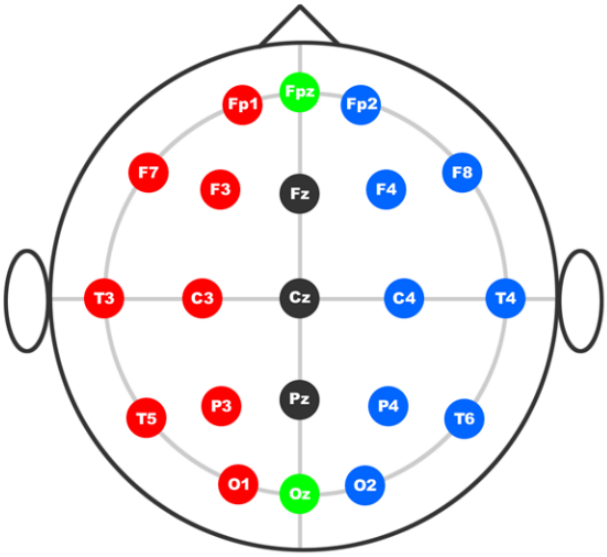

Figure 3.: The standard *10/20* system (Photo: [1])

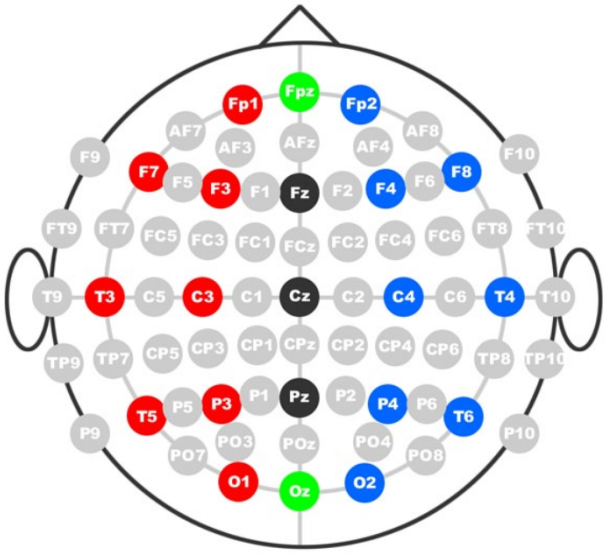

Figure 4.: The *10/10* system (Photo: [1])

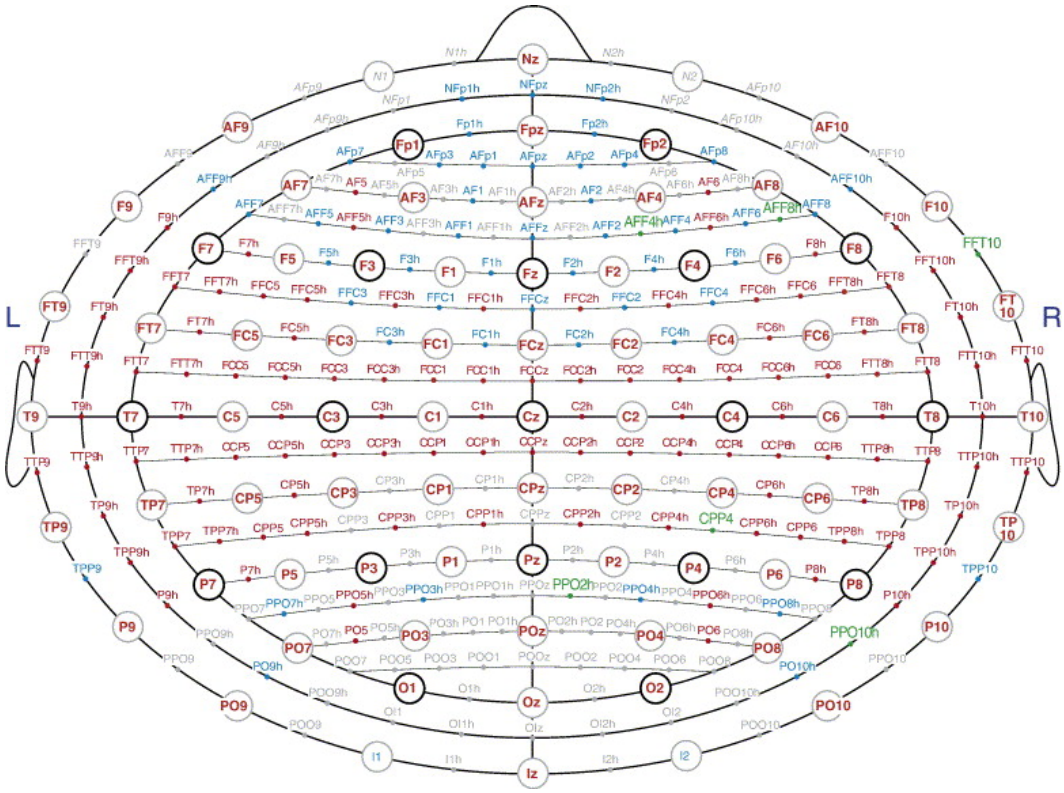

Figure 5.: The *10/5* system (Photo: [34])

2.3 BRAIN RHYTHMS

The oscillatory activities in the brain are also known as brain rhythms. Hans Berger, the inventor of EEG, discovered in 1929 the most notable rhythm. This rhythm was centered around 10 Hz and recorded from the occipital regions [7]. It became known as the *alpha* rhythm or *Berger's* wave and is associated with relaxation and closed eyes. Since then, the remaining brain rhythms have been divided into several frequency-specific bands. These bands are labeled with Greek letters and encapsulate the frequency ranges specified in Table 1. Figure 6 illustrates the frequencies of these brain rhythms using sine waves with a specific frequency and a time-varying amplitude.

Brain rhythm	**Frequency range [Hz]**
delta (δ)	0.5-4
theta (θ)	4-8
alpha (α)	8-12
beta (β)	12-30
gamma (γ)	>30

Table 1.: The different brain rhythms and their respective frequency range.

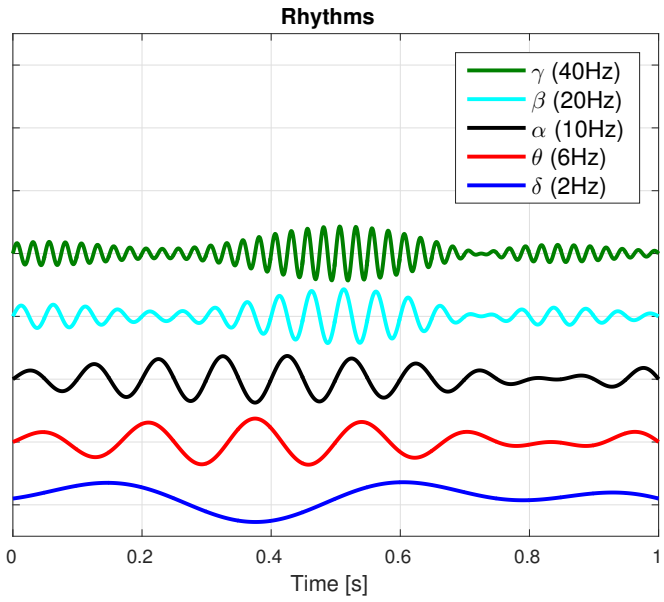

Figure 6.: The frequencies of the brain rhythms exemplified using sine waves with a time-varying amplitude.

The different brain rhythms play distinctive roles, especially compared to the level of consciousness. The *delta* rhythm, for instance, is linked with deep sleep and is prominent over the frontal regions, whereas the frontal midline *theta* rhythm is, in humans, found to be related to the working memory load [55] [24]. Moreover, occurring at both the occipital lobe and the primary motor cortex are the high-frequency *beta* and *gamma* activities, which are assumed to be related to attention. Increased energy in the beta/gamma frequency range has been detected when a dog was looking at a visual stimulus [50], and increased energy around 30 Hz was found in the visual cortex of a monkey when responding to a visual stimulus [22]. The different brain regions and areas mentioned above are presented in Figure 7.

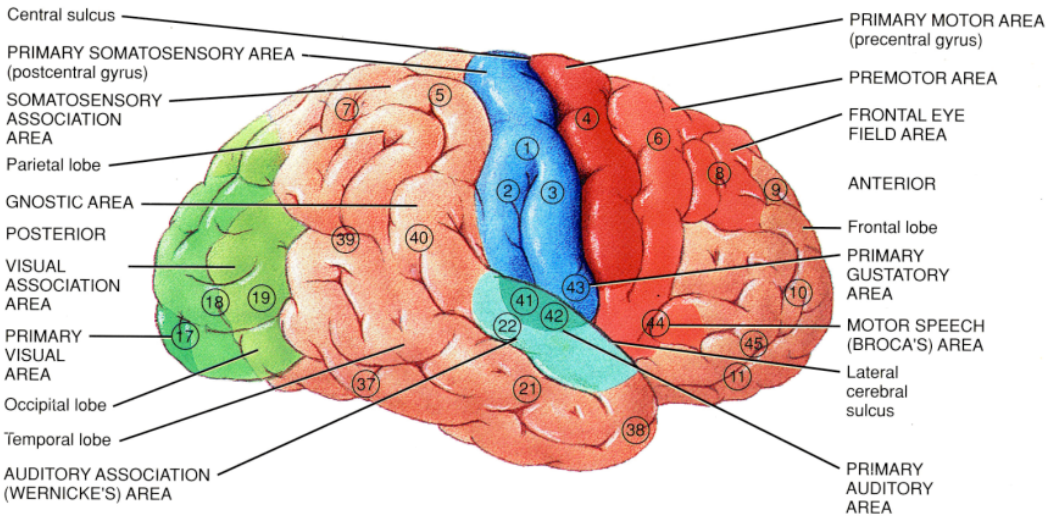

Figure 7.: The different regions of the brain (Photo: [18])

THE SENSORIMOTOR CORTEX AND THE *mu* RHYTHM

Supplementary to the *alpha* rhythms of the visual cortex, other oscillatory activities within almost the same frequency range (8-13 Hz) have been found to occur in the *sensorimotor cortex* [14]. These rhythms have been termed *rolandic mu* (μ) rhythms. The sensorimotor cortex encapsulates the primary motor cortex in addition to the somatosensory processes that occur in different parts of the cerebral cortex. The *mu* rhythms are most prominent when the body is at rest, and they have found to be suppressed when performing a motor-related task. This suppression is referred to as *desynchronization*, and will be discussed in detail in Section 2.4.

2.4 INDUCED ACTIVITIES

Most types of events, for instance, motor-related or sensory stimuli, can produce changes in the neuronal populations. These types of activities can be *time-locked* to the event, meaning that the measured activity will appear with a fixed time-delay in response to the onset of a given stimulus. This phenomenon is referred to as evoked *event-related potential* (ERP), and the most common way to obtain ERPs is to apply averaging techniques [43]. Hans Berger discovered in 1930 that some stimuli can block the alpha rhythms, indicated by a decrease in the signal energy within the alpha frequency band [8]. While the averaging techniques used to obtain ERPs represent the evoked time-locked potentials, these techniques will not be able to detect potentials that are not *phase-locked*. These types of activities are referred to as *induced* activities, and the blocking of the alpha rhythms is an example of such an activity. Since a change in the signal energy within a given rhythm is frequency specific, induced activities can be detected using *time-frequency analysis*. The decrease and increase in the signal energy are referred to as *event-related desynchronization* (ERD) and *event-related synchronization* (ERS), respectively. There exist several standards for computing the ERD/ERS, each resulting in different representations. The method used throughout this research is, as in [43], given as:

$$ERD/ERS = 100 \cdot \frac{A - R}{R} \quad [\%] \tag{1}$$

, where A is the signal power and R is the average power in the *baseline* or *reference* period. The baseline is typically set to somewhere between 100-500ms before stimulus onset. When the ERD/ERS is computed using (1), desynchronized activities will appear as negative values because the signal power is lower compared to the average baseline power. Likewise, for synchronized activities, a higher signal power compared to the average baseline power is represented by a positive value.

This thesis studies the brain activity related to hand movements, and these types of motor-related events can be seen in the sensorimotor cortex. Previous research has shown desynchronized induced activities in the *mu* rhythm occuring both before and during a motor-related event, in addition to both synchronized and desynchronized beta and gamma activities [43]. On the basis of these findings, all the rhythms occuring in the sensorimotor cortex will be analyzed in order to find the ERD/ERS corresponding to a hand movement, more specifically, the clenching of either the left or right fist. Hence, the electrode locations of interest, specified using the *10/5* system, are the following:

- *Left hemisphere*: $FCC3 - FC3 - C1 - C3 - C5 - CP1 - CP3$
- *Right hemisphere*: $FCC4 - FC4 - C2 - C4 - C6 - CP2 - CP4$
- *Midline*: $FCCz - Cz$

3

ARTIFACT DETECTION AND CORRECTION

The main topics explored in this chapter concern the different artifacts that can contaminate the recorded EEG data. Various artifacts are illustrated, and because of their characteristic waveforms, they can be detected and corrected. An artifact correction algorithm combining independent component analysis (ICA) and a statistical property (kurtosis) is presented, and its usefulness examined. Furthermore, a rejection-based EEG analysis approach is investigated.

3.1 External and internal artifacts

EEG signals are contaminated by both external and internal artifacts. EEG acquisition equipment is, like any other sensor measuring unit, vulnerable to environmental sources of disturbance. EEG signals that do not originate from brain activity are considered an artifact. The source of contamination may occur sporadically or frequently, depending on the artifact type. External artifacts originate from outer sources that can corrupt the electrode measurements. Examples of an external artifact include electromagnetic fields, poor electrode contact or electrical line noise. For example, electrical line noise may appear as a 50 Hz component in the EEG signal, which in turn can be removed by applying a notch filter [49].

A variety of internal artifacts can be seen in the EEG data. Some may be easily distinguished by an algorithm, while others may closely mimic brain activity, and are therefore difficult to correct for [51].

Internal artifacts are related to body activities, which can be movement or bioelectrical potentials, such as

- electrooculogram (EOG) activities, which are related to eyeball movements, and includes
  - horizontal eyeball movement (HEOG),
  - vertical eyeball movement (VEOG),
  - eye blinks;
- electromyogram (EMG) activities, the electrical activity produced by muscles;
- electrocardiogram (EKG) activities, the electrical activity of the heart.

Brain-related EEG measurements are usually within the 1-100 μV range. Artifact-related activities can be ten times or hundred times stronger in magnitude, resulting in distorted EEG signals [40]. Detecting and correcting such peaks and magnitude differences can be computationally inexpensive. However, such a correction could potentially give rise to the loss of valuable brain-related EEG data. Eye movements produce electrical activities with a distinct waveform that is clearly seen in EEG recordings. The activities reflect the electrical potential between the cornea and the retina, which changes during an eye movement [51]. This represents a dipole in which the eyeball is negatively charged at the cornea and positively charged at the retina. The strength of an EOG artifact may vary in terms of how the eyeball moves, vertically (VEOG) or horizontally (HEOG). The waveforms can represent low-frequency oscillations and therefore dominate the theta and delta bands. Eye blinks, on the other hand, are characterized by considerable variations in amplitude, and they are observed in almost the same shape and pattern across subjects and sessions. Consequently, they can be automatically detected and corrected using, for example, an *independent component analysis* (ICA) algorithm. ICA will be described in the following section. Figure 8 shows the characteristic waveform of three eye blinks in recorded EEG data.

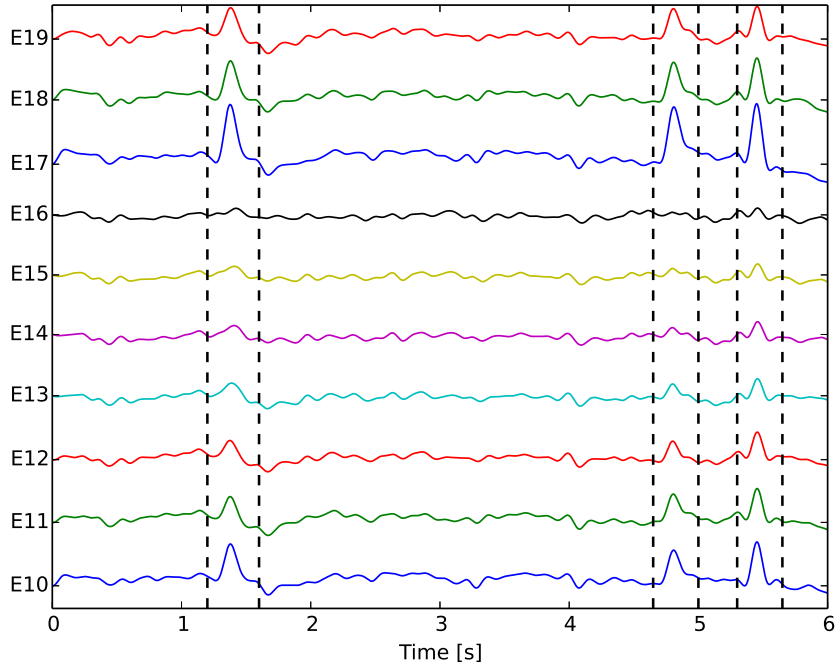

Figure 8.: Ten electrodes recorded from frontal brain regions with data acquired from one of the experiments is this thesis. The vertical lines encapsulate the characteristic waveform of an eye blink. The recorded data is low-pass filtered at 8 Hz for visualization purposes.

EKG artifacts are recognized as a periodic waveform with a rapid variation in amplitude. If an electrode is placed on top of a scalp vein, the pulse will be recorded and appear in the EEG signals. Figure 9 illustrates EKG artifacts in four seconds of recorded EEG data.

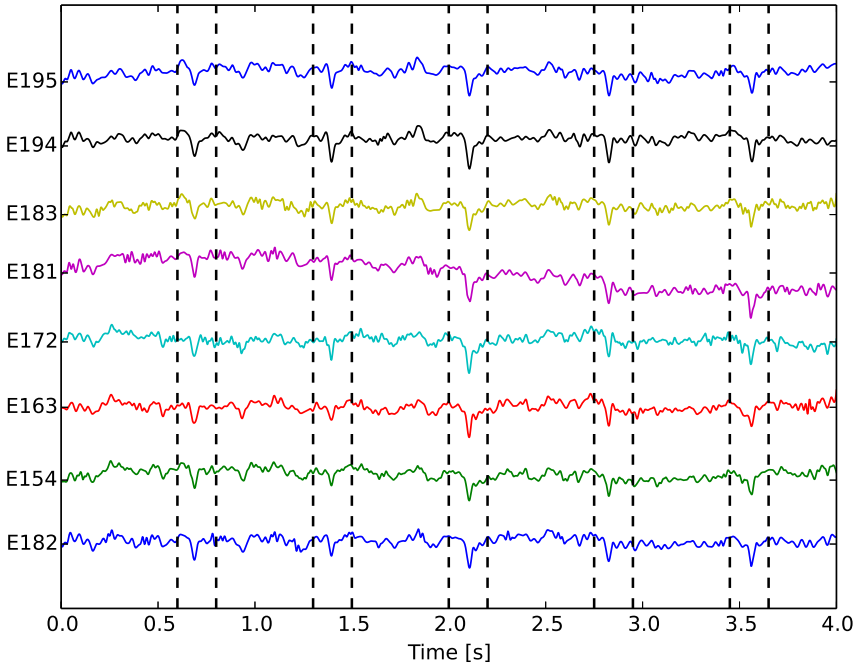

Figure 9.: Eight electrodes recorded from the sensorimotor cortex. The data were extracted from an experiment session and plotted using the platform implemented in this thesis. The vertical lines encapsulate the characteristic waveform of an EKG artifact.

3.2 INDEPENDENT COMPONENT ANALYSIS

One major objective of EEG signal analysis is to extract meaningful signals from a complex set of signals. One possible solution can be obtained by applying independent component analysis (ICA). ICA attempts to recover the original signal by estimating a linear transformation in which statistical independence between the components are provided. This independence is under the assumption that the data are non-Gaussian [54], and this is a fundamental restriction of ICA. The amplitude of an artifact can be of a magnitude much greater than brain-related activities. Moreover, independent components (ICs) reflect isolated sources with high variance in amplitude. Hence, ICA is suitable for extracting artifacts when applied to EEG data. Previous studies show that ICA applied to continuous EEG data has proven itself as an efficient solution for separating distinct artifacts, such as eye blinks and muscular activities [5].

Figure 10 shows ten ICs obtained from EEG data that have been low-pass filtered at 8 Hz, where the fourth IC (*IC4*) contains the typical characteristics of eye blinks.

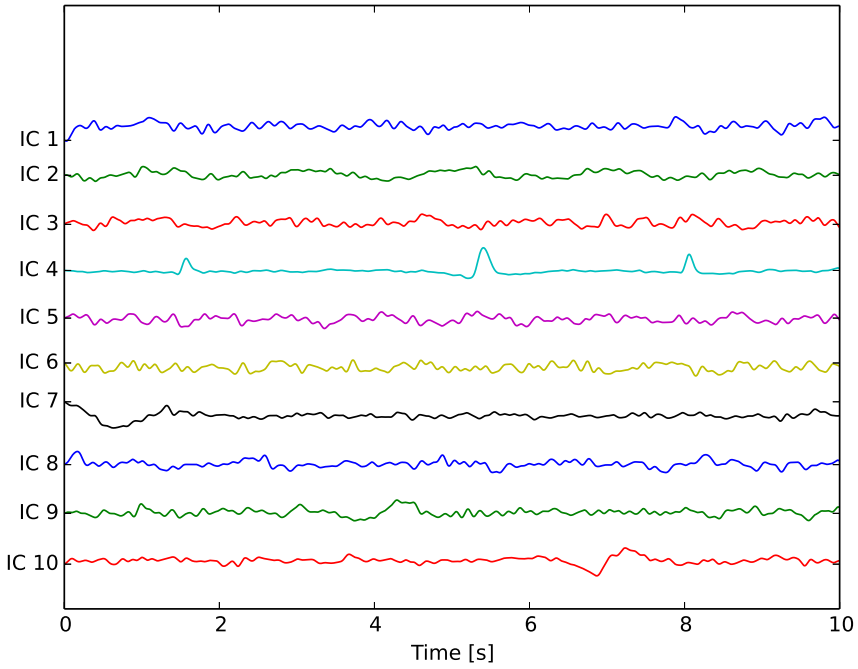

Figure 10.: ICA was applied to recorded EEG data of a single trial. The number of components extracted was equal to the number electrodes. IC4 shows three eye blink artifacts.

The process of isolating artifacts can be achieved through blind source separation in ICs. Blind source separation consists of dividing independent sources from signals based on limited information about the different sources. Independent sources can be obtained when large variations in EEG data are present. Statistical properties of a given IC can be used to find attributes that are significant from other independent sources. The measure of Kurtosis in ICs has been employed in other studies of artifact detection and correction [17]. The Kurtosis value is formally defined as the standardized fourth population moment about the mean [16], and is given as

$$\beta = \frac{E(X-\mu)^4}{(E(X-\mu)^2)^2} = \frac{\mu_4}{\sigma^4} \tag{2}$$

, where E is the expectation operator, μ is the mean, μ^4 is the fourth moment about the mean, and σ is the standard deviation. The kurtosis can reflect highly unusual peak distributions in the ICs. For instance, if an IC shows large positive values, then

the IC could represent an artifact component. As seen in Table 2, IC4 shows a high kurtosis value compared to the other ICs, and as Figure 10 illustrates, the extracted component includes the eye blink artifacts. Hence, ICA, together with the principles of blind source separation with kurtosis, can detect an eye blink artifact in EEG data. Additionally, high negative kurtosis reflects alternating current or direct current artifacts, which can for instance indicate line noise or poor electrode contact [17].

Component	**Kurtosis**
1	1.80825424657
2	-0.985102923753
3	-0.739093441635
4	**18.9327967892**
5	-0.638761790311
6	-0.843447849726
7	4.03093870912
8	1.27907123068
9	1.31731937416
10	4.61159763441

Table 2.: Kurtosis values for ten ICs obtained from the data in Figure 10.

An important task is to define a kurtosis threshold that can detect artifacts such as eye blinks. Configuring thresholds that are positive and large enough not to include ICs that can contain brain-related activities must be carefully considered. ICs that clearly show unusually high peaks can also contain brain-related activities and should, therefore, be studied with caution. If the IC is recognized as an artifact, for instance by possessing a kurtosis value similar to that of IC4 in Table 2, the IC can be excluded from the component matrix and not be considered in the inverse remixing of ICs. Still, the possibility of removing EEG data relevant to the purpose of the study may occur. Other artifacts, such as muscular activities (EMG), are better detected in the frequency domain [17]. EMG activities produce large, relatively high-frequency waveforms in the temporal region. There exist a variety of methods to reject artifacts in EEG data entirely, and the mentioned property of EMG shows that different methods should be used to properly correct for non-related brain activity.

3.3 Rejecting data contaminated by artifacts

When analyzing EEG data, it is necessary to retain as much of the data as possible to get a sufficiently meaningful result. Hence, there should be found a balance between accepted and rejected data. However, the optimal result would be achieved if all of the artifact data were completely removed from the EEG data, yielding only the brain activities related to the study left to analyze. Many detection and correction methods have been utilized in other research of EEG analysis [40][17]. Despite this fact, researchers have not yet come to an agreement regarding any optimal methods for improving the quality of EEG signals, and have not found an optimal way of removing all EEG artifacts while not distorting the brain activities related to the experiment. The latter is the reason why rejection-based EEG analysis is used in this research, resulting in data that are not altered, completely representing the measured EEG signals. Several rejection-based methods have been proposed [32][52], while in this thesis, rejection relies on calculating the variance and the gradient of continuous EEG data. If a significant amplitude variance is detected in a portion of the data, the portion is discarded from further analysis.

4

ADAPTIVE SIGNAL ANALYSIS

One important aspect when performing time-frequency analysis on nonlinear data is to end up with meaningful results. This chapter will present a detailed description of the normalized Hilbert transform (NHT). For the Hilbert transform to extract a meaningful instantaneous frequency, the input signal should be restricted to being an intrinsic mode function (IMF). IMFs are produced by the empirical mode decomposition (EMD), which procedural approach is explained. The theoretical and practical difficulties of EMD are presented, along with their respective solutions. These concern the sifting process, the stopping criterion, the end effects of EMD, and the Bedrosian theorem. Furthermore, a method for visualizing the time-frequency solutions is discussed, namely the Hilbert spectrum.

4.1 INSTANTANEOUS FREQUENCY

In the interest of analyzing nonlinear dynamical systems, e.g., the electrical activity of the brain, a proper definition of instantaneous frequency is required. If the signal of interest is a regular sine wave, which phase characteristics are linear, the instantaneous frequency is equal to the zero-crossing frequency, $f = \frac{1}{T}$, where T is the period of the signal. This is also referred to as the mean frequency, but zero-crossing is the common name simply because the most prominent way to compute the frequency is to measure the time interval between two consecutive zero-crossings [30].

The Fourier transform is another technique used to define the frequency. This transform is applied in order to compute the frequency for a time series, $x(t)$. As explained in [30], the transformed signal is expressed as a sum of N real-valued sine- and cosine components with a given amplitude, a, and frequency, ω, given as

$$x(t) = Re\Big\{ \sum_{j=1}^{N} a_j e^{-i\omega_j t} \Big\} \tag{3}$$

, where

$$a_j = \int_{t=0}^{T} x(t) e^{i\omega_j t} dt \tag{4}$$

, which implies that, during the entire integral interval, the frequency value must be constant. In other words, the frequency as defined by the Fourier transform will only be meaningful if the process represented by the data is *linear* and *stationary*. That is, linear in the way that superposition is allowed and stationary meaning that the frequency is time-independent [30]. This will be further discussed in Chapter 5, in addition to the properties of nonlinear and nonstationary signals. In dynamical processes, for instance the electrical activity of the brain, the data may have multiple extrema between consecutive zero-crossings. This indicates that multiple frequencies could exist at the same time, which implies the necessity of an instantaneous frequency. Figure 11 illustrates this phenomenon. The integral in (4) restricts the Fourier transform in such a way that it does not yield a meaningful frequency when applied to signals with multiple extrema between two zero-crossings, causing the transform to be limited by the uncertainty principle. Heisenberg's uncertainty principle states that two complementary variables, such as position and momentum, must have a well-defined limit to their precision when measured simultaneously. In signal analysis, the principle resembles that of Heisenberg, and states that the precision with which we can measure the frequency at a given time has a finite value. As asserted in [30], the product of frequency resolution, $\Delta\omega$, and the time interval in which it is defined, ΔT, shall not be less than $1/2$, that is

$$\Delta\omega \cdot \Delta T \geq \frac{1}{2} \tag{5}$$

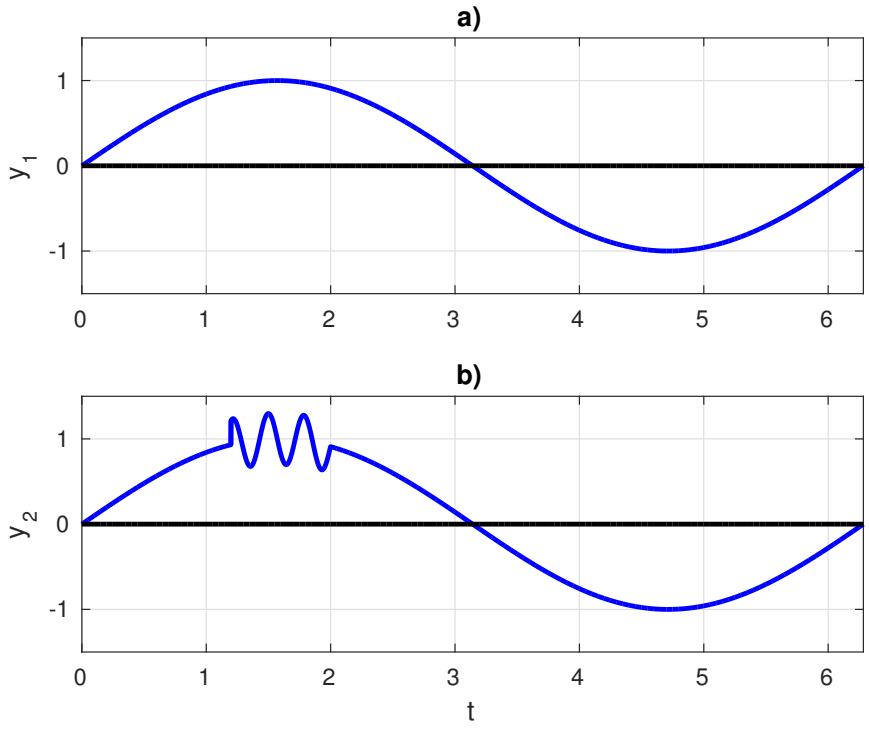

Figure 11.: **a)** Signal with one extremum between two consecutive zero-crossings.
b) Signal with multiple extrema between two consecutive zero-crossings.

When considering the complex signals acquired from EEG, the need for both high frequency resolution and high temporal resolution arise. The solution that has been implemented and discussed in the following research is based on an application of the *Hilbert transform* (HT) on *analytic signals* to obtain time-phase series. Furthermore, the *instantaneous frequency* is obtained by differentiating the phase with respect to time.

4.2 THE HILBERT TRANSFORM

In 1946, *Dennis Gabor* introduced a method to generate an analytic signal by applying the Hilbert transform (HT) [23]. The transformation $\mathscr{H}\{\cdot\}$ for a signal $x(t)$ is defined as

$$y(t) = \mathscr{H}\{x(t)\} = \frac{1}{\pi} P \int_{\tau=-\infty}^{\infty} \frac{x(\tau)}{t-\tau} d\tau \tag{6}$$

, where P is the Cauchy principal value of the complex integral. Moreover, the imaginary part of the transformed signal, $\mathscr{I}m\{y(t)\}$, together with the original signal, $x(t)$, form the analytic signal

$$z(t) = x(t) + i \cdot y(t) = a(t)e^{i\phi(t)} \tag{7}$$

,where

$$a(t) = \sqrt{x(t)^2 + y(t)^2} \tag{8}$$

$$\phi(t) = \tan^{-1}\left(\frac{y(t)}{x(t)}\right) \tag{9}$$

$$\omega(t) = \frac{d\phi(t)}{dt} \tag{10}$$

is the instantaneous amplitude, instantaneous phase and instantaneous frequency of the signal, respectively.

From (6) and (7), it is clear that the existence of the Hilbert transform is general and global. The limitation lies in the existence of a meaningful instantaneous frequency, which is restrictive and local [20]. As seen in Figure 11 a), the signal needs to be represented as a function that is symmetric locally to the zero mean level. These kinds of functions are called *monocomponent* functions and, at any given time, they possess only a single, well-defined, positive, instantaneous frequency represented by the derivative of its phase function [28]. Figure 11 b) illustrates the opposite, which is referred to as a multicomponent function. In the interest of applying the Hilbert transform to complex data, the *intrinsic mode function* (IMF) was defined by Professor Norden Huang in [20]. This function is symmetric locally to the zero mean level, hence a meaningful instantaneous frequency can be defined anywhere in the function.

4.3 INTRINSIC MODE FUNCTION

An intrinsic mode function (IMF) has no mathematical definition, but is defined as a function that satisfies two conditions [20]:

- **C1**: In the whole data set, the number of extrema and the number of zero-crossings must either be equal or differ by no more than one.
- **C2**: At any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero.

The first condition is similar to the narrow band requirements for a stationary Gaussian process, while the second condition is necessary in order to obtain a meaningful instantaneous frequency. If the second condition is not satisfied, the signal is no longer symmetric with respect to the zero local mean. This will introduce unwanted fluctuations in the instantaneous frequency. Figure 12 shows an IMF, a function that satisfies both of these conditions.

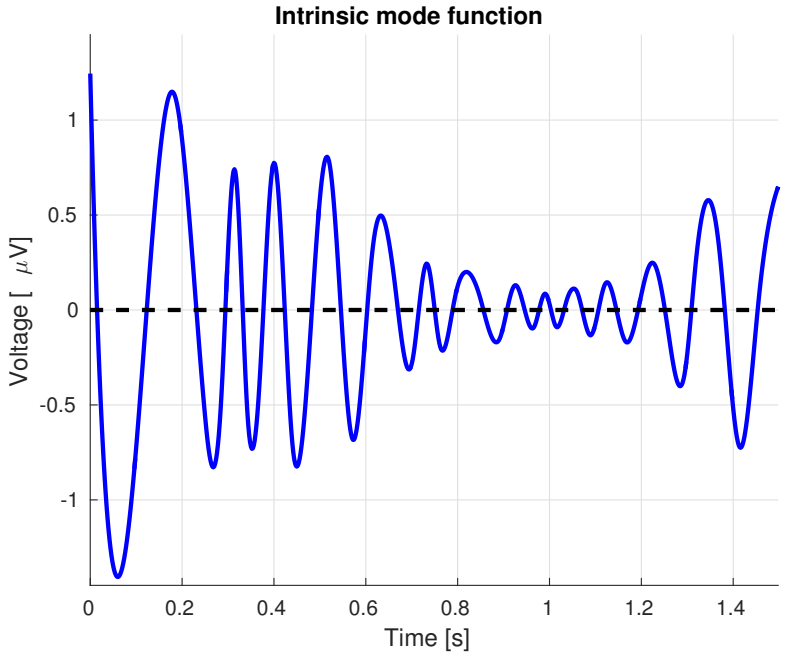

Figure 12.: Example of an intrinsic mode function (IMF). It is a monocomponent, i.e., there is only one extremum between two consecutive zero-crossings (**C1**).
The IMF is symmetric with respect to the zero local mean (**C2**).

With this definition of the IMF, the function involves only *one* mode of oscillation, that is, it represents one of the oscillatory modes embedded in the data [20]. This mode is not restricted to a narrow band signal, it can be both amplitude and frequency modulated.

The previous section described that in order to obtain a meaningful instantaneous frequency, a well-behaved Hilbert transform (HT) is of great importance. The IMF, which satisfies the two conditions **C1** and **C2**, possesses this quality. IMFs are acquired through the use of an adaptive method called *empirical mode decomposition* (EMD). This method will be described in detail in Section 4.4. The procedure of employing EMD together with the Hilbert transform is referred to as the *Hilbert-Huang transform* (HHT) [20].

4.4 Empirical mode decomposition

Empirical mode decomposition (EMD) is based on the direct extraction of the energy associated with various intrinsic time scales [20]. It breaks a signal down into several IMFs without leaving the time-domain. This is accomplished through a process called *sifting*. As described in [20] and in [28], the sifting process will, in addition to eliminate riding waves, make the wave-profiles more symmetric. A detailed description of the sifting process will now be outlined.

The sifting process

The sifting process is an iterative procedure that involves the use of interpolation methods in order to compute the local mean of a signal. The output of the sifting process is an IMF.

Considering a time-dependent signal $x(t)$, and letting the envelope $e_{up}(t)$ be defined as a cubic spline interpolation of all the maxima in the signal. Moreover, let the envelope $e_{low}(t)$ be defined as a cubic spline interpolation of all the minima in the signal. The local mean $m_{1,1}(t)$ is then defined by computing the average of the lower envelope and the upper envelope:

$$m_{1,1}(t) = 0.5 \cdot (e_{up}(t) + e_{low}(t)) \tag{11}$$

The first subscript index indicates the IMF number, whereas the second subscript indicates the current sifting iteration that is being performed on that specific IMF. Figure 13 illustrates this procedure, which is the first step in the sifting process.

The next step is defining the *mode* $h_{1,1}(t)$, which is obtained by subtracting the mean $m_{1,1}(t)$ from the original signal $x(t)$, that is:

$$h_{1,1}(t) = x(t) - m_{1,1}(t) \tag{12}$$

This describes one sifting iteration. In the following iteration, the previous mode, $h_{1,1}(t)$, is treated as the input signal, instead of $x(t)$. This means that a new pair of upper and lower envelopes, $e_{up}(t)$ and $e_{low}(t)$, are obtained by interpolating the maxima and minima of the previous mode, resulting in a new local mean:

$$m_{1,2}(t) = 0.5 \cdot (e_{up}(t) + e_{low}(t)) \tag{13}$$

The next mode is then defined by subtracting the new local mean from the previous mode:

$$h_{1,2}(t) = h_{1,1}(t) - m_{1,2}(t) \tag{14}$$

These steps are to be repeated until the mode satisfies the conditions to become an IMF. These two conditions, **C1** and **C2**, are stated in Section 4.3. If, after k sifting iterations, the mode satisfies these conditions, it is defined as an IMF, $c(t)$, that is:

$$m_{1,k}(t) = 0.5 \cdot (e_{up,k}(t) + e_{low,k}(t)) \tag{15}$$

$$h_{1,k}(t) = h_{1,k-1}(t) - m_{1,k}(t) \tag{16}$$

$$c_1(t) = h_{1,k}(t) \tag{17}$$

Now that the first IMF is obtained, the IMF is subtracted from the original signal, thus removing one oscillatory mode from the signal:

$$x_2(t) = x(t) - c_1(t) \tag{18}$$

The entire sifting process (15-18) is performed on the new signal, resulting in a new IMF, $c_2(t)$. This is repeated until there are no more oscillatory modes in the data. In other words, EMD is finished when the last mode $h_{i,k}(t)$ analyzed is a monotonic function: a function with no extrema. This is referred to as the *residue*, $r(t)$, or the *trend*. If the decomposition is finished after obtaining n IMFs, the original data can be expressed as

$$x(t) = \sum_{i=1}^{n} c_i(t) + r_n(t) \tag{19}$$

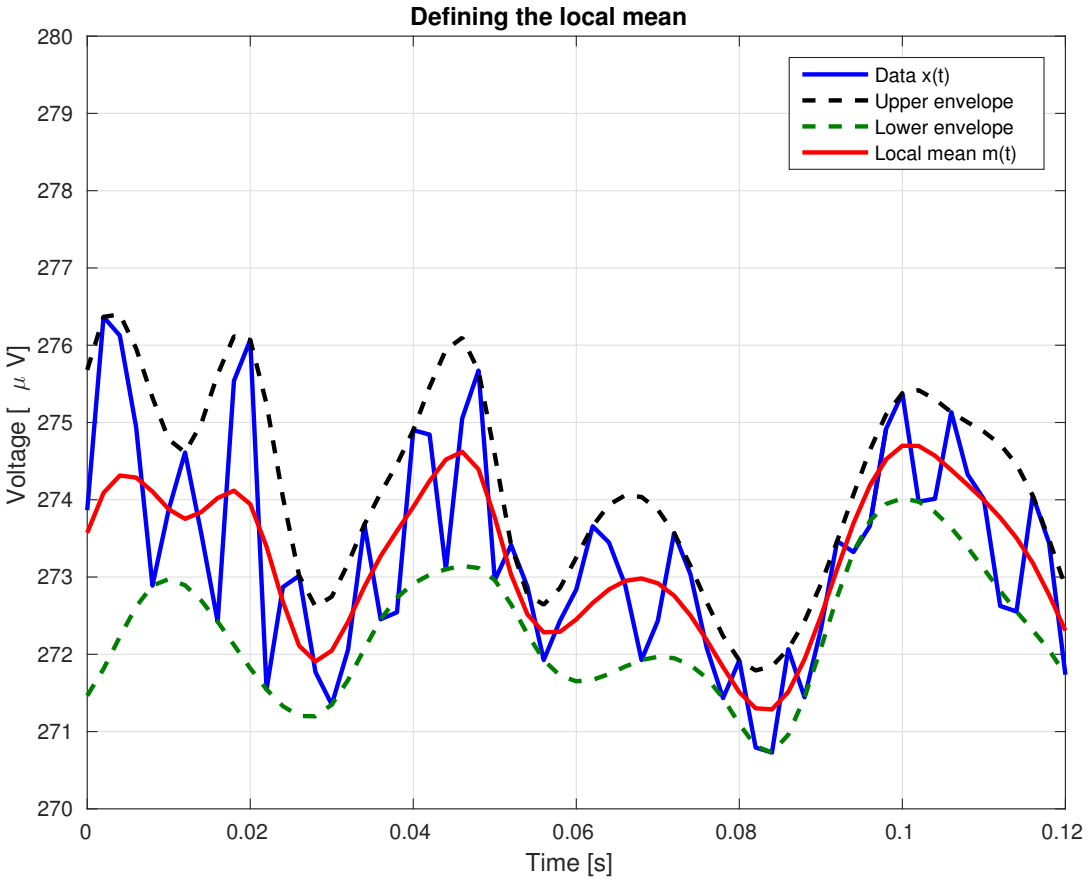

Figure 13.: Example illustrating the interpolation between the maxima and minima so that the local mean can be defined.

As an example, a signal $x(t)$ containing $600ms$ of EEG data has been decomposed using EMD. This is shown in Figure 14, along with *five* IMFs $c_1(t) - c_5(t)$, and the final residue, $r(t)$. The maximum number of siftings was set to 400, and all of the IMFs were extracted without reaching this number.

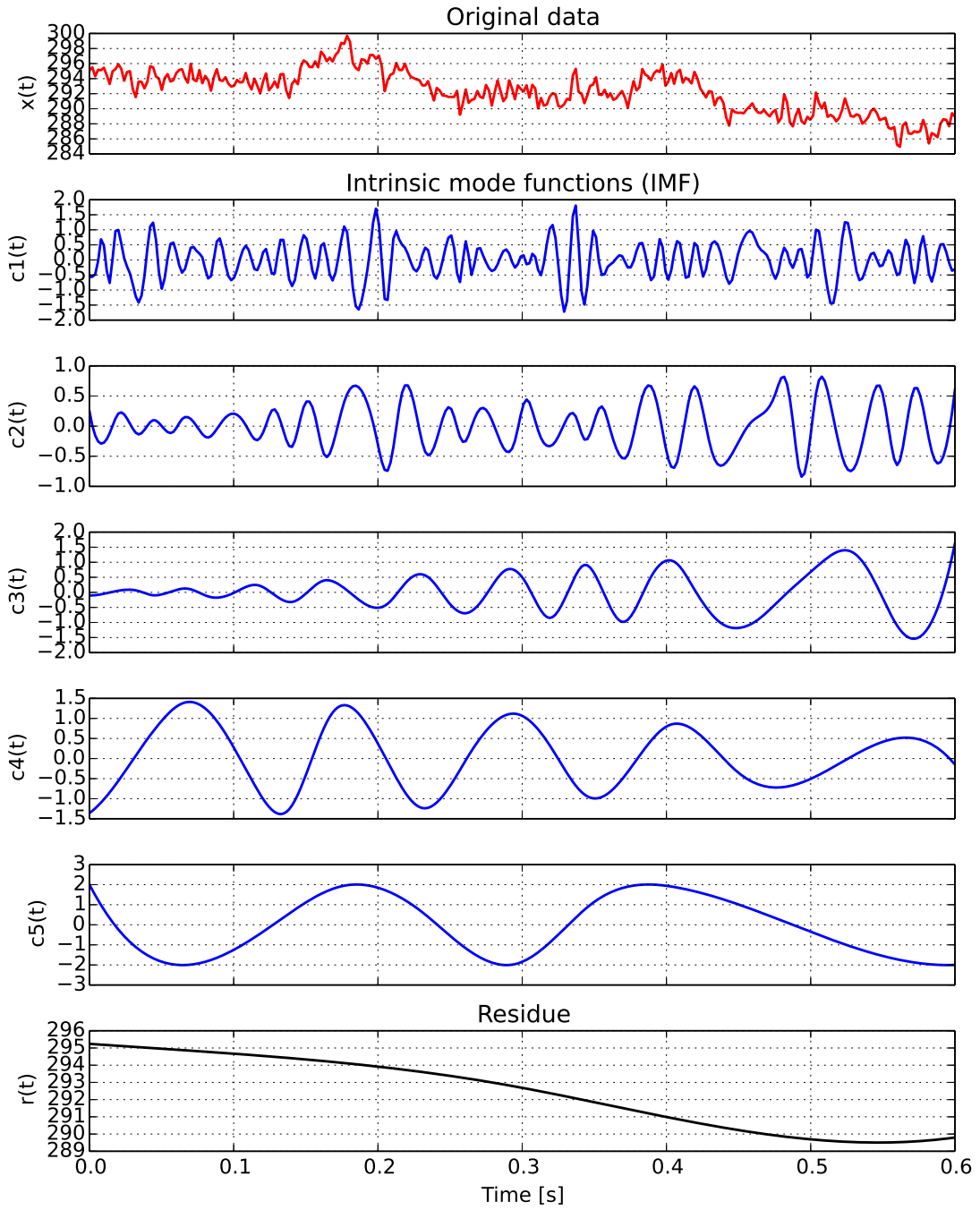

Figure 14.: Example illustrating the empirical mode decomposition of an EEG signal. The signal is 600ms long and only meant for visualizing the IMFs.

STOPPING CRITERION

As explained in [20], too many sifting iterations may produce an IMF that is purely frequency modulated with constant amplitude. After all, the desired result is for the IMF to encompass a physically meaningful oscillatory mode. To ensure that this is satisfied, a *stopping criterion* must be implemented. The inventor of EMD, Norden E. Huang, proposed two possible criteria in [2]. The first criterion was originally stated in [20] and is referred to as the Cauchy-type convergence criterion. When applying this criterion, the sifting process stops when the difference between two successive sifting iterations is smaller than a given threshold, Δh, typically set between 0.2 and 0.3. The criterion can be expressed as

$$\Delta h_i = \sum_{t=0}^{T} \frac{|h_{i,k}(t) - h_{i,k-1}(t)|^2}{h_{i,k-1}^2(t)} \tag{20}$$

, where i is the IMF candidate number and k is the number of siftings performed on the given candidate.

However, as stated in [28], this stopping criterion should be used with caution, because it does not depend on the definition of the IMFs. Even though the difference between two siftings can be small, it is not guaranteed that the function will have the same numbers of extrema and zero-crossings. Consequently, an improved and simplified stopping criterion was defined by terminating the sifting procedure when the number of zero crossings and extrema is the same for S consecutive sifting steps. A value of $3 \leq S \leq 5$ has turned out to be acceptable as the default stopping criterion for the sifting process [2].

END EFFECTS

A theoretical difficulty tends to occur if the data is nonlinear. The envelope mean obtained from the interpolation of extrema will experience some deviation from the true mean, resulting in some asymmetric waveforms. This can be reduced by ensuring that enough sifting iterations are performed, but since the spline interpolation is an approximation, the envelope mean will still experience some minor deviation from the true mean.

Another difficulty occurs near the ends of the data, where the spline interpolation can produce large swings. This is referred to as the end effects of EMD and is, even today, a known challenge. There is a need to determine the end point of the spline curve of both the upper and the lower envelope. One solution is presented in [20], where the last two maxima are used to create a straight line. The endpoint of this line is then compared to the endpoint of the data itself, and if the endpoint of the straight line has

a *higher* value than the endpoint of the data, the last maximum in the upper envelope is chosen as the endpoint of the straight line. If not, the data point itself is chosen as the last maximum. Similarly, for the lower envelope, the last two minima are used to create a straight line. The endpoint of the line is compared to the endpoint of the data itself, and if the endpoint of the straight line has a *lower* value than the endpoint of the data, the last minimum in the lower envelope is chosen as the endpoint of the straight line. If not, the data point itself is chosen as the last minimum.

Figure 15 shows the end effects of the empirical mode decomposition. The envelope mean is deviating from the true mean due to the upwards swing resulting from the spline interpolation of the maxima. A deviation in the mean would also occur if the interpolation of the minima resulted in a downwards swing, though in the opposite direction of the one illustrated here. Figure 16 shows the results when applying the technique described above. The upper and lower envelopes are adjusted by straight lines constructed by the last two extrema, thus removing the large swings produced by the spline interpolation. Hence, the end-effect-corrected envelope mean will be closer to the true mean.

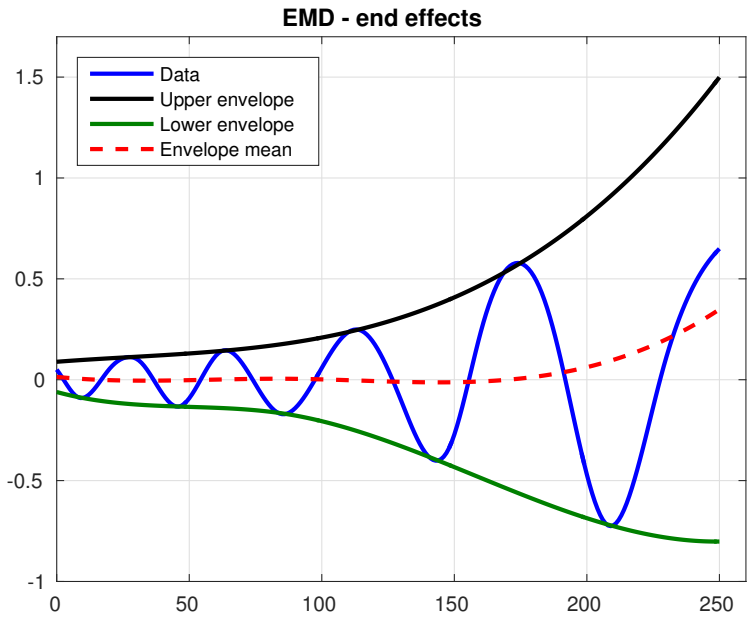

Figure 15.: Example illustrating the end effects produced by EMD.

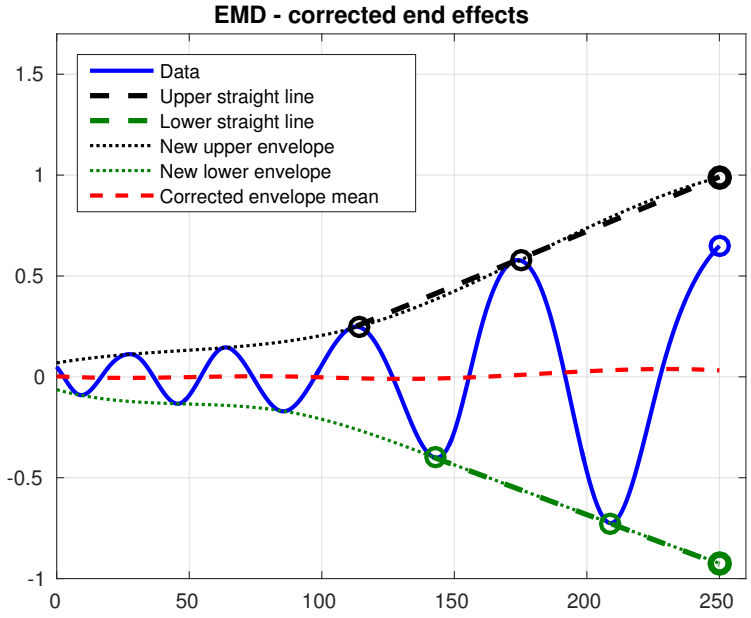

Figure 16.: Example illustrating the corrected end effects by using the endpoints of two straight lines made by the two last maxima and minima, respectively.

4.5 THE BEDROSIAN THEOREM

Already in 1962, mathematicians stated inquiries regarding the computation of instantaneous frequency based on analytic signals. *E. Bedrosian* discovered a necessary condition for the instantaneous frequency to be meaningful. It is based on the separation of the Hilbert transform of the carrier signal from its amplitude envelope and is referred to as the *product theorem*, or simply the Bedrosian theorem, and is given as

$$\mathscr{H}\{a(t)cos(\phi(t))\} = a(t)\mathscr{H}\{cos(\phi(t))\} \tag{21}$$

For (21) to hold, the amplitude has to change very slowly as a function of time. In fact, the frequency domains of $a(t)$ and $\phi(t)$ must be disjoint [28]. Figure 17 illustrates a 5 Hz sine wave with constant amplitude. The instantaneous frequency is obtained from the Hilbert-Huang transform (HHT) and shows acceptable values compared to the true frequency. On the other hand, Figure 18 pictures a 5 Hz sine wave that experiences a jump in the amplitude by a factor of 10 at $t = 1.5s$. Because of this large variation, the associated IMF does not satisfy the Bedrosian theorem stated in (21), resulting in fluctuations in the instantaneous frequency as depicted in Figure 18.

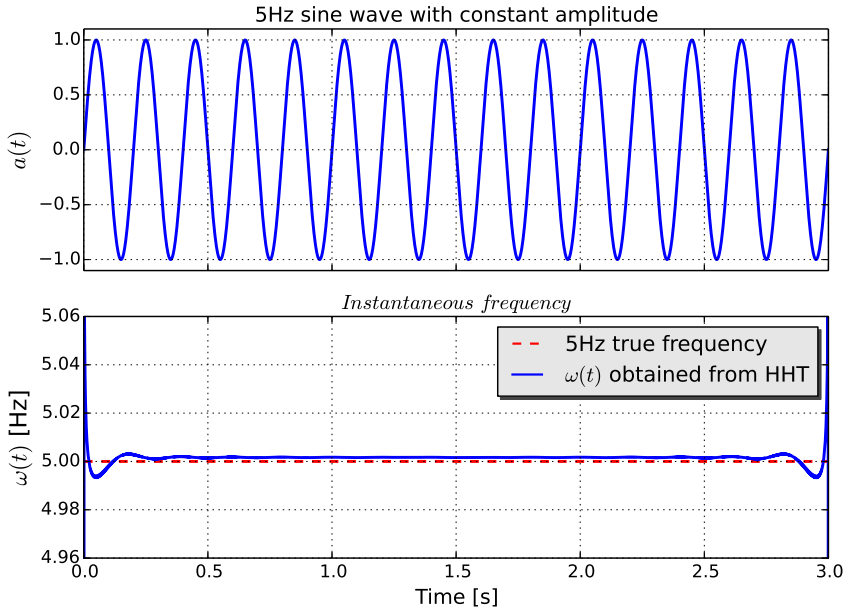

Figure 17.: A 5 Hz sine wave with constant amplitude, $a(t) = 1$, and its instantaneous frequency

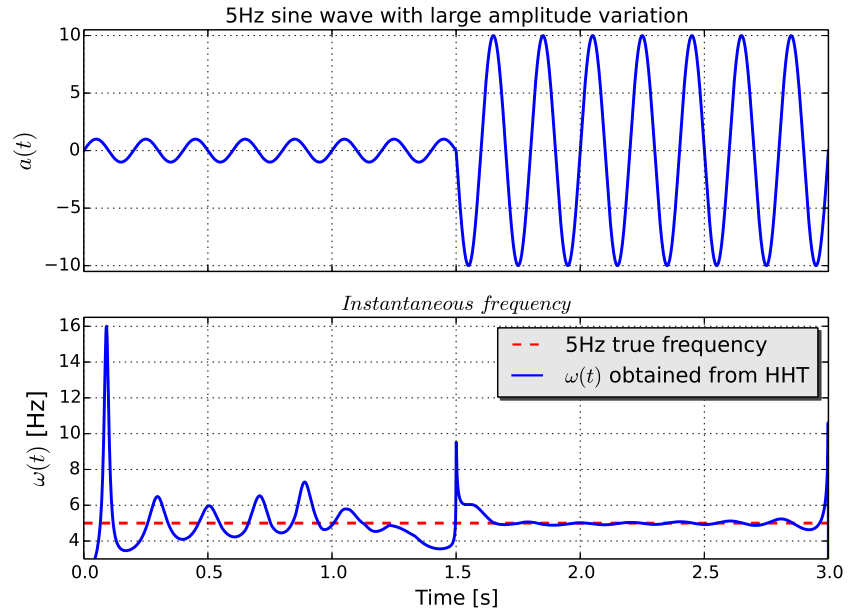

Figure 18.: A 5 Hz sine wave with a jump in the amplitude by a factor of 10 at $t = 1.5s$ and its instantaneous frequency.

In order to satisfy the condition set by the Bedrosian theorem, Huang and Long (2003) proposed a solution referred to as the *normalized Hilbert transform* [27]. This method will be described in detail in the next section.

4.6 THE NORMALIZED HILBERT TRANSFORM

The normalization scheme described in this section will eliminate the limitation stated by Bedrosian. By performing this normalization method, the IMFs will have unity amplitude, hence, naturally satisfying the Bedrosian theorem stated in Section 4.5, Equation 21. The steps in the normalization method are given as:

1. Obtaining the IMFs from EMD, $c_i(t), i = 1, ...n$, as explained in Section 4.4.

   For each of the n IMFs, the following steps should be applied:

2. Defining the data to be normalized as the IMF of interest, $x(t) = c_i(t)$

3. Defining the absolute value of the IMF as $x_{abs}(t) = |x(t)|$.

4. Constructing an envelope, $e_1(t)$, defined by the cubic spline interpolation of all the maxima in $x_{abs}(t)$.

5. Scaling the IMF by dividing the IMF by its associated maxima envelope, so that

$$y_1(t) = \frac{x(t)}{e_1(t)} \tag{22}$$

   Since the spline interpolation is an approximation, this step should be repeated a couple of times to ensure unity amplitude in the normalized signal $y(t)$. When this is ensured, after k repeated iterations (typically $k = 4$), the normalized IMF is given as

$$\begin{aligned} y_2(t) &= \frac{y_1(t)}{e_2(t)} \\ &\vdots \\ y_k(t) &= \frac{y_{k-1}(t)}{e_k(t)} \end{aligned} \tag{23}$$

The instantaneous amplitude is now given as the product of all the maxima envelopes

$$a(t) = e_1(t) \cdot e_2(t) \cdots e_k(t) \tag{24}$$

The normalized IMFs are now strictly frequency modulated (FM), because of the unity amplitude property. They are now satisfying the Bedrosian theorem, in addition to being monotonic functions. Hence, the IMFs should experience well-behaved Hilbert transform, from which a meaningful instantaneous frequency can be computed. The normalization scheme, as explained through these steps, is illustrated in Figure 19 and Figure 20. The whole adaptive signal analysis method is referred to as the *normalized Hilbert-Huang transform* (NHHT), or simply the *normalized Hilbert transform* (NHT), and can be summarized as:

1. Obtaining the IMFs from EMD
2. Applying the normalization scheme to each of the IMFs, ensuring the following signal properties:
    a) unity amplitude
    b) strictly frequency modulated (FM)
3. Obtaining the Hilbert transform of each IMF
4. Calculating the instantaneous frequency of each IMF

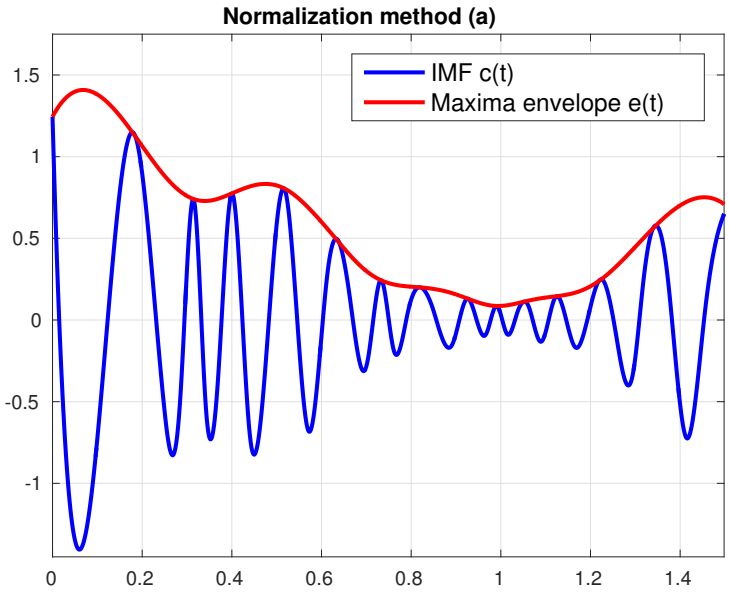

Figure 19.: Example illustrating the first steps in the normalization method. The maxima envelope is constructed by a cubic spline interpolation of the absolute value of the IMF.

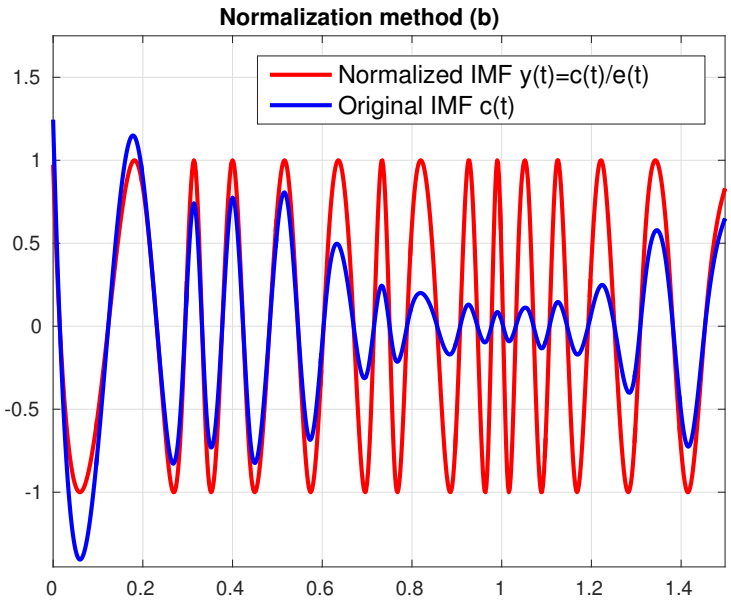

Figure 20.: Example illustrating the normalized IMF compared to the original IMF.

Figure 21 shows the instantaneous frequency obtained from both the Hilbert-Huang Transform (HHT) and the normalized Hilbert Transform (NHT). The input signal is a 25 Hz sine wave which amplitude is increasing as an exponential function, $a(t) = e^{2.25t-1}$. The obtained IMFs are not satisfying the Bedrosian theorem, hence HHT results in large fluctuations in the frequency, especially for the low amplitude values from $t \in [0,1]$. After applying the normalization scheme, the Bedrosian theorem is satisfied and, therefore, resulting in a much more stable and accurate instantaneous frequency.

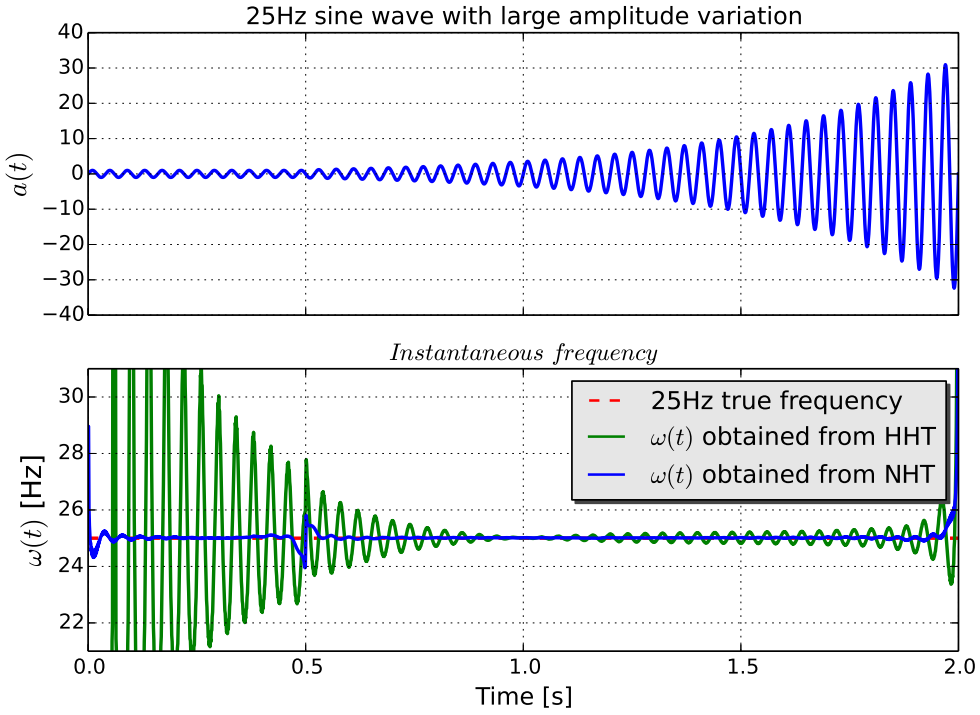

Figure 21.: Comparison of the instantaneous frequency of a 25Hz sine wave with increasing amplitude. Showing the normalized Hilbert transform (NHT) in blue and the original Hilbert-Huang transform (HHT) in green.

After EMD has decomposed the data into several IMFs, the Hilbert transform can be applied to the each of these. As explained in Section 4.3, the IMFs experience well-behaved Hilbert transform, from which a meaningful instantaneous frequency can be extracted. The instantaneous frequency of the IMFs and their respective instantaneous amplitude can now be presented in the time-frequency domain. This is referred to as *Hilbert spectral analysis* (HSA), and will be discussed in the next section.

4.7 HILBERT SPECTRAL ANALYSIS

One way of studying the instantaneous frequency and amplitude, including how they change as functions of time, is by visualizing the three dimensions in a two-dimensional space. The amplitude changes will be represented by colors where red is, typically, assigned to high amplitude values and blue to low amplitude values. Power can, of course, be used instead of amplitude. The same representation will apply in that case. Moreover, the amplitude changes will be expressed in the time-frequency domain. Figure 22 shows the instantaneous frequency from four IMFs obtained by applying NHT to EEG data. The corresponding instantaneous amplitude values are normalized and squared, resulting in the normalized power visualized through the use of colors.

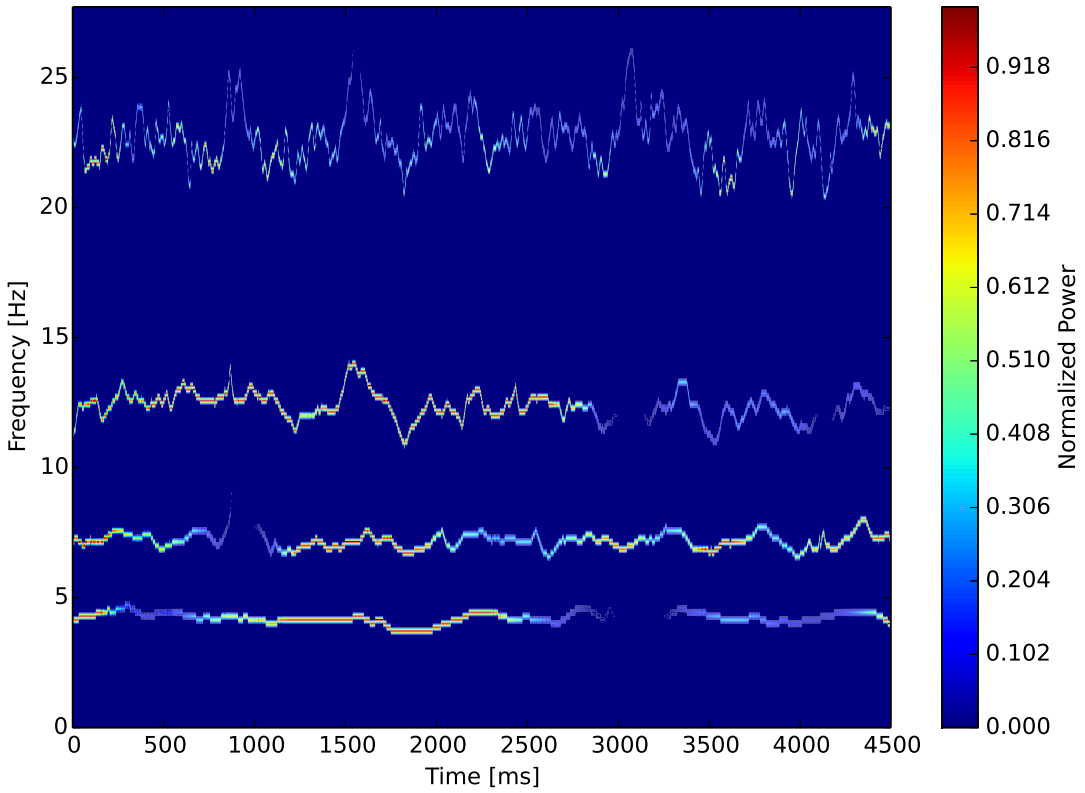

Figure 22.: Example illustrating the Hilbert spectrum (HS) of several IMFs obtained by applying NHT to EEG data.

5

FILTERING TECHNIQUES AND NONLINEAR DATA

This chapter states some of the challenges occurring when dealing with nonlinear and nonstationary data, in addition to the conditions that have to be met for the data to be successfully represented. The differences between signal processing and signal analysis are clarified. Moreover, examples illustrating how some filtering techniques deform a signal are discussed. A method for increasing the low spatial resolution associated with EEG is explained, namely the Laplacian spatial filter (LSF).

5.1 NONLINEAR AND NONSTATIONARY SIGNALS

One important aspect of data analysis is to give a meaningful representation of the underlying physics of a process. Models are one way of representing different physical processes. Different linearization techniques can be applied to get an approximation of a phenomenon, constructing a model of the natural process. This is done to overcome the problems of nonlinear and nonstationary signals. What separates a nonlinear signal from a linear one is the fact that superposition principle does not hold. For the superposition principle to hold, the two following properties must be satisfied:

$$F(x_1 + x_2) = F(x_1) + F(x_2) \tag{25}$$

$$F(a \cdot x) = a \cdot F(x) \tag{26}$$

, where (25) is called the additivity property and (26) is referred to as the homogeneity property. Moreover, a nonstationary signal is, simply put, a signal in which the frequency and the amplitude are time-dependent functions. Although linearized systems can elicit acceptable representations of natural processes, the linear approximation is only valid as long as its constraints are preserved.

As discussed in [20], the necessary conditions for representing a nonlinear and nonstationary signal are:

- completeness,
- orthogonality,
- locality,
- adaptivity.

The first *two* requirements are standard for all linear expansion methods, for instance, the Fourier transform. The first condition guarantees the degree of precision of the expansion, meaning how precise the approximation is. The second condition indicates the statistical independence of the different components used to represent the nonlinear and nonstationary data. For nonstationary data, the frequency and amplitude are time-dependent functions. This requires the locality condition to be satisfied, which implicate that all events have to be identified by their time of occurance. Moreover, the adaptivity condition is based on the fact that the model representing the nonlinear and nonstationary data must account for the local variations of the data without any prior knowledge [20].

In general, time-frequency analysis is widely used to provide both temporal and spectral information of any given data. Here, Fourier spectral analysis has been dominating soon after its introduction. This is a general method, and because of its simplicity, it can obtain the time-frequency solution of several types of data. After applying the Fourier transform given in Equation 3, the transformed signal can be seen as a sum of sinusoidal waves, each with constant amplitude and frequency. However, the transform is restricted to the fact that the signal must be linear and stationary. Otherwise, the resulting spectrum makes little physical sense [20].

An important method implemented and practiced throughout this thesis is the *normalized Hilbert transform*, which enables the frequency and amplitude to be functions of time. The nonlinear processes in EEG can now be fully extracted and perhaps give a meaningful representation of brain activity. A full informational representation of nonlinear and nonstationary data has been presented in [27] and in [30].

It is important to understand the differences between signal processing and signal analysis. The former extracts relevant information from the data, and can result in a simpler representation of the data, while the latter reflects the information obtained from the data, resulting in an increased understanding. In signal processing, no meaning is added to the data. An example of a signal processing method is the notch filter, in which a single operation is performed. Signal analysis, on the other

hand, will provide increased understanding of the data, for instance time-frequency analysis or the empirical mode decomposition (EMD). In this project, the EMD was used as a signal analysis method to extract meaningful oscillatory modes from EEG data.

5.2 ALTERING THE DATA

In EEG analysis, notch filters and band-pass filters are often applied in the process of analyzing high-density EEG data, that is, EEG recorded by a large number of electrodes. A notch filter can be used to remove a specific frequency component in EEG signals, whereas band-pass filters allow for reclaiming signals within specific frequency bands. As shown in Figure 23, a 5 Hz sinusoidal signal with unity amplitude is contaminated by a 50 Hz sinusoidal signal with ten times the amplitude. This is a typical situation in EEG analysis, where the signal could be corrupted by electrical line noise. To retrieve the 5 Hz sinusoidal signal without the line noise, three different approaches have been tested. Figure 24 shows how a filter alters the data of interest. The notch-filtered signal will eventually resemble the true sine wave, but large fluctuations are observed during 25% of the total length of the signal. Moreover, when applying a band-pass filter, the amplitude of the 5 Hz sine wave is significantly damped. The band-pass filter's low-cutoff and high-cutoff frequencies were 0.001 Hz and 5.5 Hz, respectively. EMD, on the other hand, has successfully retrieved the real 5 Hz sine wave. The IMFs obtained using EMD reflect the sinusoidal components in both the true sine wave and the line noise component. Hence, the IMF that represents the 50 Hz component could be excluded when reconstructing the original signal, which in turn yields the sinusoidal signal of interest.

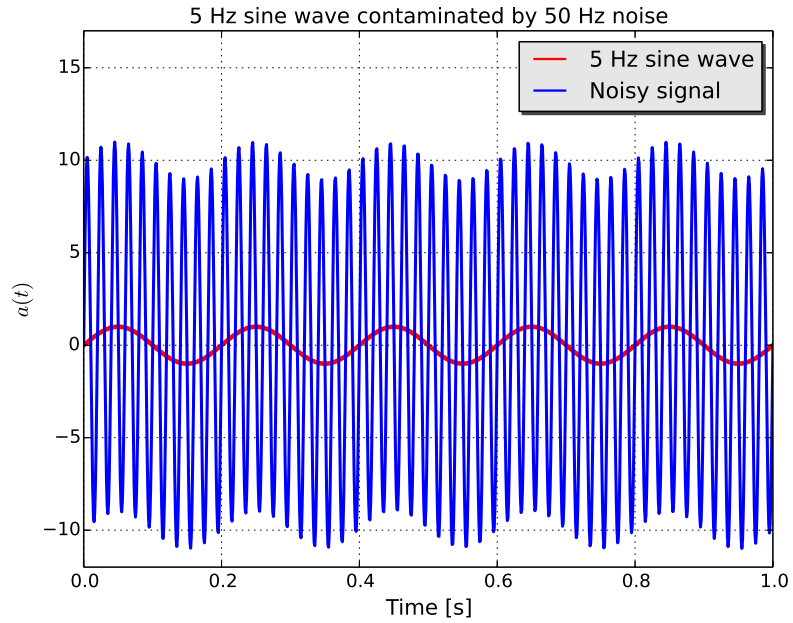

Figure 23.: Simulated line noise (50 Hz) contaminating a 5 Hz sinusoidal signal. The red signal shows the true 5 Hz sine wave.

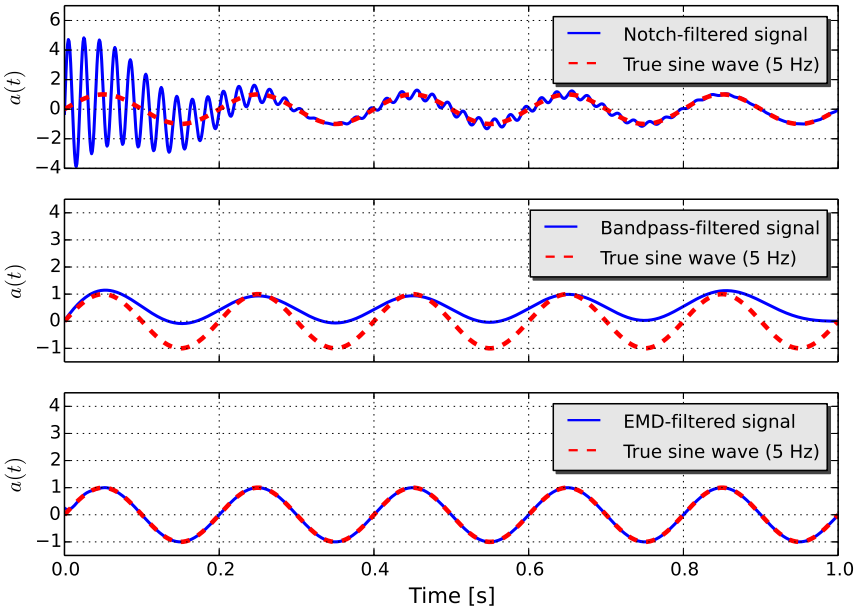

Figure 24.: The three different filtering methods applied to the contaminated sine wave.

This illustrates how the two signal processing method alters the data. How can results from EEG data based on notch or band-pass filters be justified when filtering simple sine waves deforms the signal? These results have motivated the researchers to fully analyze EEG data without any filtering methods, but resorting to the EMD, namely a data analysis method.

5.3 Surface Laplacian spatial filter

Conventional EEG studies have expressed the need for both high temporal and spatial resolution. However, as mentioned in Chapter 2, the conductivity of the brain makes the resulting spatial resolution low. The need for improving the spatial resolution is motivated by the fact that EEG researchers study the electrical activity from an increasingly specific cortical area. The surface Laplacian method has become popular among EEG researchers because of its influential initial work in 1973 [39], in addition to the further enhancement in 1975 [26], yielding an improvement of the spatial resolution. The method can be beneficial, since it does not require a volume conductor model of the head, or any detailed specification of neural sources. As stated in [12], the motivation for using the surface Laplacian technique is based on Ohms's law. Here, a relationship is established concerning the surface Laplacian of scalp potentials, and the underlying flow of electrical current caused by brain activity.

The surface Laplacian combines the volume Laplacian and the physics behind electrical fields and potentials. The former is calculated as the divergence of the gradient of the volume, given in Equation (27), whereas the latter reflects the relationship between the electrical potential of the field (28) [12]. Both equations are expressed in Cartesian coordinates, that is *x, y,* and *z*.

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = 0 \tag{27}$$

$$\frac{\partial^2 E}{\partial x^2} + \frac{\partial^2 E}{\partial y^2} + \frac{\partial^2 E}{\partial z^2} = 4\pi\rho_Q \tag{28}$$

, where E represents the electric field, and V the electric potential of the corresponding field. Moreover, ρ_Q represents the charge density, meaning the electric charge per unit volume.

By projecting the scalp coordinates onto the xy-plane, the equation for the 2D surface Laplacian of V is defined as

$$\mathscr{L}(V) = \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} \tag{29}$$

Further information regarding the surface Laplacian can be found in [26]. The Application of the surface Laplacian spatial filter (LSF) aids in removing broad features shared among EEG recording electrodes, yielding enhanced spatial resolution. The surface Laplacian is utilized by computing the average activity of the neighboring electrodes, and subtracting this activity from the specific electrode. This results in increased spatial resolution when re-referencing the recorded activity from electrodes placed on the scalp. The computational iterative procedure for achieving this can be summaraized in the equation below.

$$y_{LSF}(t) = y(t) - \frac{1}{N}\sum_{i=1}^{N} x_i(t), \quad i = 1,2,3...,N \tag{30}$$

, where N is the number of neighbors. The distance from the reference electrode $y(t)$ to the neighboring electrodes can be specified by either small LSF, using the adjacent electrodes or large LSF, meaning electrodes further away.

6

CLASSIFICATION SCHEME

This chapter embraces important aspects of classification in BCI applications. Feature extraction and selection are outlined to get an improved understanding of feature properties such that the most appropriate classifier can be chosen. Moreover, classification algorithms are discussed to comply with the major problems related to the curse-of-dimensionality and the bias-variance tradeoff. When evaluating the performance and optimization of the classifier output, the area under the receiver operating characteristic (ROC) curve and the grid search are explored.

6.1 FEATURE EXTRACTION AND SELECTION

If two or more classes are introduced to a classifier model, feature extraction consists of choosing those features which are most effective for preserving class separability [53]. To separate samples in variables, the proper design of any classifier model is necessary. This can be achieved by selecting optimal features to use in the classifier. Feature selection can be seen as a process where the original signal is mapped into more suitable features. It is essential to understand how features are used, and simultaneously know what their properties are. Both the former and the latter are necessary to select the most appropriate classifier for a given brain-computer interface (BCI) system. As reviewed in [37], a wide collection of features have been tested in the design of BCI classification systems, such as:

- Amplitude values
- Power spectral density (PSD)
- Band powers
- Autoregressive and adaptive autoregressive parameters
- Time-frequency features of EEG signals

Further, some critical properties of features need to be examined when designing classifiers for a BCI system, and these will now be explained.

CRITICAL FEATURE PROPERTIES

Features can be noisy and contain outliers because of poor signal-to-noise ratio in EEG data. It is, therefore, important to increase that property to improve the design of a specific classifier. Moreover, concatenating feature vectors from multiple time segments give rise to high dimensionality. The amount of data needed to sufficiently describe the different classes increase exponentially with the dimensionality. This is known as the curse-of-dimensionality, and is a common problem when dealing with classification problems in BCI research [33]. The classifier will probably give unsatisfactory results if the size of the training data is small compared to the dimensionality of the feature vector. It has been suggested to use five to ten times as many training samples per class as the dimensionality [45]. The curse is therefore that this is not always suitable for all BCI systems. Another major aspect to contemplate is the fact that features should contain temporal information, as brain activity patterns are related to specific time variations in EEG signals. Regarding nonstationarity, features should consider that EEG signals may rapidly vary over time and of course over sessions. The process of training a classifier is time-consuming and demanding for participants. That is why training sets are often kept small, which in turn could cause problems for the classifier performance.

6.2 CLASSIFICATION ALGORITHMS

Classification in a BCI system deals with identifying to which category brain patterns belong. Consequently, an important aspect concerns the choice of classifier used for a given set of features. To properly discuss this, two main problems should be taken into consideration, namely the *curse-of-dimensionality*, as mentioned in Section 6.1, and the *bias-variance tradeoff* [37]. Classification errors can be decomposed into three terms, which in turn explain three possible sources of classification errors [10]:

- Noise represented in the system, which is an immutable error.
- Bias between estimated mapping and the best mapping.
- The variance, which reflects the sensitivity to the training set used.

Keeping these sources of errors low will result in reduced classification errors. As stated in [37], stable classifiers tend to have high bias, but low variance in training sets, while unstable classifiers have low bias and high variance.

SUPPORT VECTOR MACHINE

Support vector machine (SVM) is a group of supervised learning methods that can be applied to classification or regression. It is commonly associated with a set of learning algorithms that analyze and recognize patterns. *Supervised learning* means that the computer is presented with a set of trained data as input, along with their respective desired output, and the objective is to find a general rule in which the computer maps input to output. Participants should go through one or several recording sessions. The trained data could be a motor-related event, for instance, a left hand movement. After extracting the features of that event, the features are fed as input to the SVM, and stored as a support vector. An SVM training algorithm will build a model that assigns the variables into a specific category, and this is referred to as the *trained set*. Machine learning revolves around learning some properties which then can be applied to new data. Both off-line data and real-time data can be applied as a *test set*, in which the predictive properties of the SVM are tested.

SVMs are discriminative algorithms that use different kernel functions to distinguish classes, such as linear, radial basis, sigmoid, or a polynomial function. Discriminant means that the classifier focuses on the class boundaries or the class membership probabilities directly [48]. In BCI research, SVM is probably the most popular algorithm used so far. It applies a discriminative hyperplane to identify classes [11], and the chosen hyperplane will be the one that maximizes the margins, meaning the distance between training points. Figure 25 shows the optimal hyperplane that gives best separation of the support vectors. The kernel function of the SVM illustrated in Figure 25 is linear, but the kernel function can, however, be of a higher complexity. Such a kernel function will implicitly map the feature vectors to another space of higher dimensionality. Kernels can be customized for a given SVM to correctly fit the feature vectors that have been extracted, before designing the classifier model. The radial basis function (RBF) Kernel is shown in Equation 31.

$$k(x, c) = e^{-\gamma|x-c|^2} \tag{31}$$

, where x is the feature vector points, and c is the distance to a center point, which can for instance be located at the origin. As seen in the RBF kernel (31), the function depends on the distance to the point c, in addition to the variable γ, which is one of the hyperparameters of an SVM algorithm. The other hyperparameter is the regularization parameter C, which accommodates for outliers and allows for errors in the training set. The fact that the SVM employs the maximization of margins and the regularization hyperparameter C give rise to acceptable generalization properties [6].

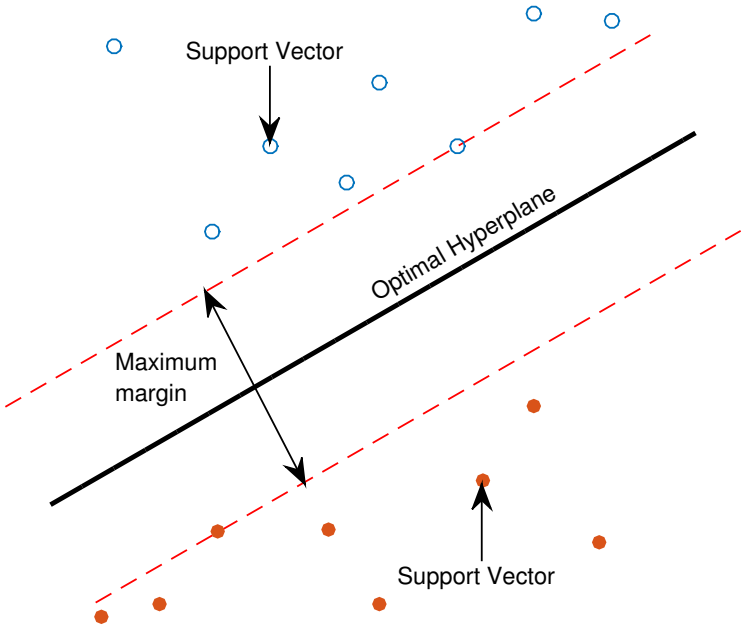

Figure 25.: The SVM is finding the optimal hyperplane by maximizing the margins, meaning the distances between training points. The hyperplane is then used for separating the support vectors. The blue and the red dots represent two different classes.

6.3 MODEL SELECTION

Model selection is the part in which the goal is to improve the accuracy of the classifier output. Acceptable classification performance can be achieved by comparing, validating and choosing optimal parameters for classification models. It comprehends the study of the behavior of different models and how to optimize their performance, for instance, by using a *grid search*. Two important aspects of any given classification scheme are the *validation* and *optimization* of the classifier model. The concept of the former and the latter is outlined in the following sections, and can be further addressed in [38]. Their properties will reflect how well the model will perform on new data. When the training set is chosen and given as input to the classifier model, an evaluation of the classifier performance should be done with validation.

EVALUATION OF CLASSIFIER PERFORMANCE

In many classifications problems, evaluation of the performance, meaning quantifying the quality of the predictions, and the evalution of the final validation accuracy of a classifier is important. The latter estimates the future performance of a classifier and it is imperative that the estimator method provides low bias and variance [36]. Cross-validation is a way of evaluating the estimation of future predictions. The idea behind cross-validation is to split the data into one or several sets. One part of the data is used to train the model, and the remaining portion is employed to estimate the accuracy of the model [4]. This procedure is repeated several times, and the reported measure is an average over the values computed in the validation steps. Another evaluation property is the ROC value, which reflects the area under the receiver operating characteristic curve, and is an assessment of the quality of the output of the classifier. The area corresponds to the probability of correct identification of classes represented in the classification model [25]. Other performance measures can be used for evaluation, such as the average precision scores, which are computed from the prediction scores and reflect the area under the precision-recall curve [15].

OPTIMALIZATION OF CLASSIFIER PERFORMANCE

Optimizing the performance of an SVM is accomplished by testing the predictive abilities when changing the hyperparameters γ and C. A grid search is an excellent tool to find the optimal hyperparameters that result in the best output from the classifier. It consists of giving the model a set of different parameters of both γ and C and then evaluating the training set with the parameters specified in the parameters set. The parameter combination that contributed to the best performance on the training set will, in turn, provide the best performance when predicting new data. Evaluation of the training set means that a subset is extracted from the trained set and used as a test set. The predictive abilities are then tested with the subset and the optimal scores are used to select the classifier hyperparameters.

Part II

EMPIRICAL RESEARCH AND EVIDENCE

7

METHODS

In this chapter, the experimental paradigm is described, including information concerning the number of participants, the EEG acquisition equipment, and how a typical EEG recording session was carried out. Furthermore, the whole EEG analysis procedure is explained, all the steps from the segmentation of raw EEG data to the classification of hand movements. Some important aspects regarding the software implementation are mentioned, illustrating a few of the different functionalities that have been developed during this research.

7.1 Experimental paradigm

Participants

Eight healthy participants were recruited, seven male and one female. All participants were right-handed and of similar age (25 years, SD = 1 year). The EEG recordings caused no physical harm to the participants. Moreover, they had given their informed consent and could withdraw from the testing at any time without any consequences.

Paradigm

The stimuli were projected with a standard 60 Hz refresh rate onto a display by a *Texas Instruments M-Vision Cine 3D*. The width and height of the display were 244 cm and 224 cm, respectively, and placed at the constant distance of 80 cm in front of the participants. The stimuli consisted of black arrows pointing either to the left or the right, as depicted in Figure 26.

E-prime 2.0 software tool was used to provide timing services and to establish the stimulus type. The experimental paradigm was designed using E-prime's graphical

interface as well as the E-Basic scripting language. The former is simply a drag and drop approach while the latter is similar to the Visual Basic programming language.

The recording session started by presenting a few explanatory sentences regarding the tasks assigned to the participant. The beginning of each trial was indicated by a fixation cross in the middle of the display, appearing for 750 ms. Then the cross disappeared for 250 ms, and a randomly chosen stimulus was presented for 3000 ms. Afterwards there was an inter-trial period spanning for 2000 ms, resulting in a total trial duration of 6 seconds. These sequential steps are illustrated in Figure 27.

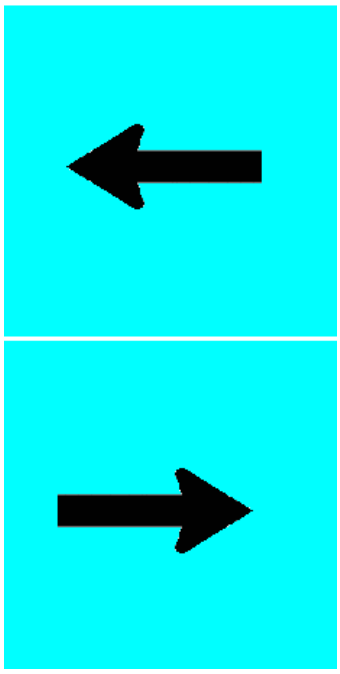

Figure 26.: The stimuli used throughout this research.

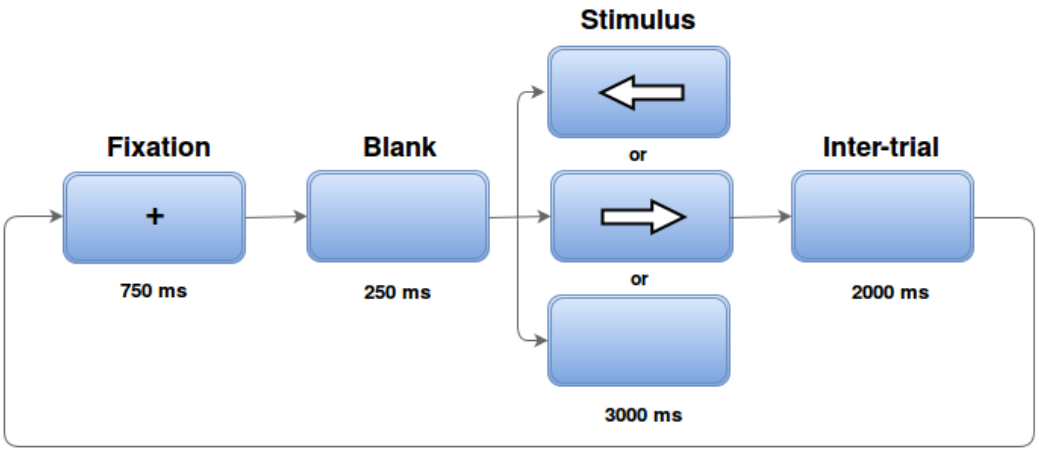

Figure 27.: The experimental steps for each trial during a recording session.

DATA ACQUISITION

The electrical activity of the brain was recorded using a 256-electrode Geodesic Sensor Net (GSN), as depicted in Figure 28. Furthermore, this activity (EEG) was amplified using an *Electrical Geodesics, Inc.* (EGI) amplifier, and the maximum electrode impedance was set to 100 $k\Omega$. The amplified EEG activity was recorded at a sampling rate of 500 Hz using *Net Station* software on a Macintosh computer. The experimental paradigm designed in E-prime interacted with the Net Station software, making sure that a particular stimulus type was connected to its corresponding EEG data segment.

Figure 28.: The 256-electrode GSN net when mounted on the head of a participant.

As stated in Section 2.4, the 16 electrodes of interest are placed above the sensorimotor cortex. These are marked *green* in Figure 29, which shows all the 256 electrodes of

the GSN net used throughout this research. In accordance with the *10/5* system, the naming of the electrodes are:

FCCz, Cz, FCC3, FCC4, FC3, FC4, C1, C2, C3, C4, C5, C6, CP1, CP2, CP3, CP4

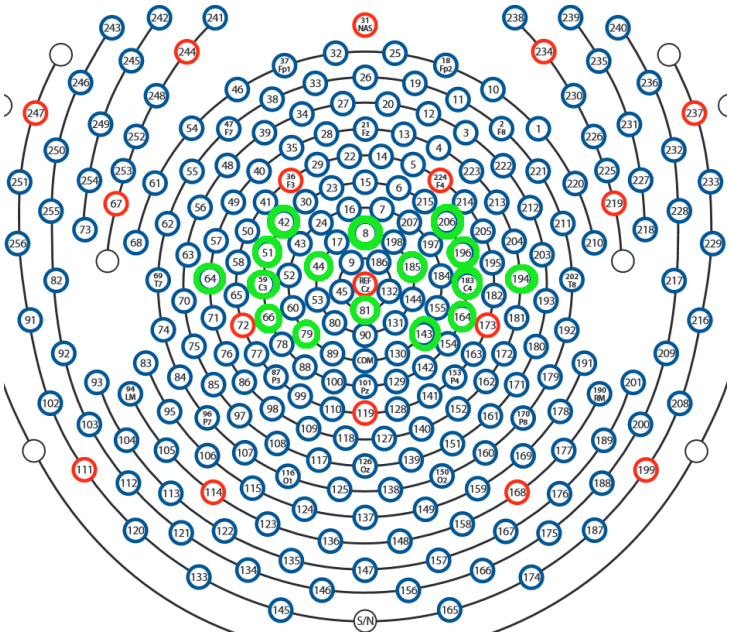

Figure 29.: The electrode numbering and location of the 256-electrode Geodesic Sensor Net (GSN) used in this research. The electrodes of interest when analyzing motor-related events are marked green. (Photo: [31])

PROCEDURE

After signing the consent form, the participant's head circumference was measured so as to select the most suitable net size. Moreover, the net was soaked in a saline solution to optimize the electrical conductivity. After being mounted on the participant's head, the net was plugged into the amplifier. If the impedance of an electrode was too high, contact quality was improved by adding saline solution to the electrode or by adjusting its position. To reduce external interference with the EEG signal, a grounding strap was connected around the ankle of the participant. This is depicted in Figure 30.

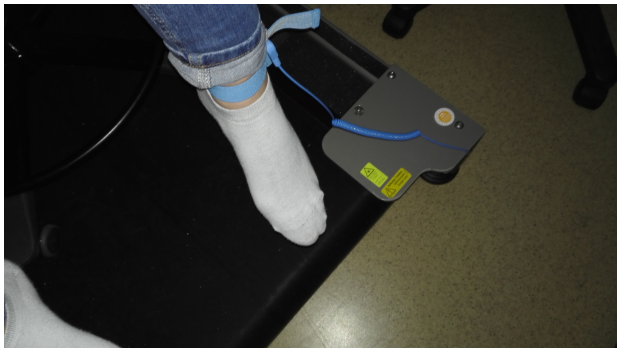

Figure 30.: An ankle strap was used in order to ground the participant, hence reducing interference with the recorded EEG data.

After making sure that the task assigned to the participant was fully understood, the testing session began. The three different tasks involved clenching either the left or right fist, or doing nothing at all. The participant was asked to open and close the left or right fist for as long as the corresponding arrow was showing, which meant the participant were to clench repeatedly for *three* seconds. Moreover, the stimuli were presented in random order. For the first *four* participants, the number of trials were 80 for both left and right hand movements, in addition to 10 trials of a neutral condition where no arrows were shown. To obtain more recordings of the neutral conditions, the number of trials for the last *four* participants were 70 for all three conditions.

7.2 DATA ANALYSIS

This section will provide a description of the procedural approach from the obtaining of the raw EEG recording to the classification of hand movements.

After obtaining the raw EEG recording in Net Station, the data were exported as a binary MATLAB format file, called a *MAT*-file. This format is compatible with the Python software implemented and used throughout this thesis. After loading the exported data into the Python program, the data were segmented, meaning that the correct stimulus type gets connected to its corresponding trial. A trial, in this context, refers to a segment of EEG data containing the activity related to the onset of a stimulus.

Since scalp EEG is prone to disturbances, the voltage potential measured from an electrode is often disturbed by line noise from the mains electricity. Moreover, the line noise can be detected in the first *two* intrinsic mode functions (IMFs) obtained after applying empirical mode decomposition (EMD). Consequently, this method was

employed to eliminate the line noise. A detailed explanation regarding EMD is given in Chapter 4. The last IMF, also referred to as the *trend*, was subtracted from the data. This was done to remove slowly varying drift in the measured potential.

After obtaining the EMD-filtered EEG data, an automatic trial rejection procedure was applied. This was done because some trials could possibly be contaminated by artifacts and noisy disturbances. Furthermore, the normalized Hilbert-Huang transform (NHT), as explained in 4.6, was applied to the accepted trials and the resulting frequencies and amplitudes from *four* IMFs were stored. Finally, feature vectors were extracted from this data, which provided the basis for the classification scheme. Features were calculated using the instantaneous amplitude from the IMFs as input to the equation for *event-related desynchronization*, described in Chapter 2 (Equation 1). Moreover, the features obtained from a given number of trials were employed as support vectors, whereas features obtained from the remaining trials were utilized for predicting if the hand movements were executed using the left or right hand. The flow of this entire approach is depicted in Figure 31 and can be summarized as:

1. *Export*: Exporting the raw EEG recording from Net Station as a *MAT*-file.
   - Input: raw data [EGI's Metafile format (MFF)]
   - Output: raw data [Binary MATLAB file format (MAT)]
2. *Segmentation*: assigning the correct stimulus type (event) to the corresponding trial.
   - Input: raw data [(MAT)]
   - Output: $X_{raw}(t)$ [Matrix] , $Events$ [Array]
3. *EMD-filtering*: subtracting the first two IMFs to remove line noise as well as subtracting the trend from the data to remove slow drift in the data.
   - Input: $X_{raw}(t)$ [Matrix]
   - Output: $X_{filtered}(t)$ [Matrix] , $X_{linenoise}(t)$ [Matrix] , $X_{trend}(t)$ [Matrix]
4. *Trial rejection*: discarding trials contaminated by artifacts and external interference.
   - Input: $X_{filtered}(t)$ [Matrix]
   - Output: $X_{cleaned}(t)$ [Matrix]

5. *Adaptive analysis*: applying the normalized Hilbert-Huang Transform (NHHT), resulting in the instantaneous frequency, $\omega(t)$, and amplitude, $a(t)$, of the first *four* IMFs.

   - Input: $X_{cleaned}(t)$ [Matrix]
   - Output: $a_j(t)$ [Matrix] , $\omega_j(t)$ [Matrix] , $j = 1...4$ (Number of IMFs)

6. *Feature extraction*: calculating the mean event-related desynchronization (ERD) of the instantaneous amplitude.

   - Input: $a_j(t)$ [Matrix] , $j = 1...4$ (Number of IMFs)
   - Output: $\overline{ERD_j} = \frac{1}{N}\sum_{t=1}^{N} 100 \cdot \frac{a_j(t) - R_j}{R_j}$, $j = 1...4$ (Number of IMFs)

7. *Classification*: differentiating between the three conditions, namely, left hand movement, right hand movement, and no hand movement.

   - Input: $\overline{ERD_j}$, $Events$ [Array]
   - Output: Classification results

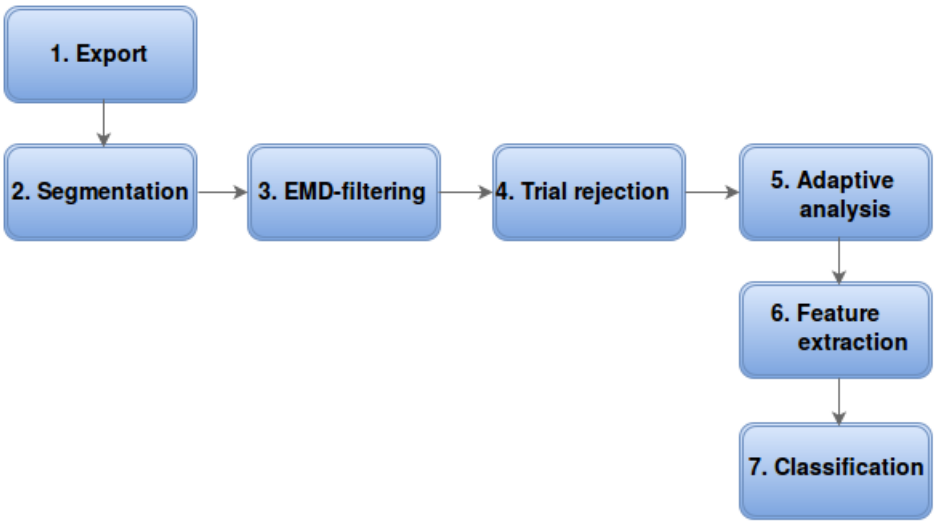

Figure 31.: The sequential steps in the data analysis, from raw EEG data to the classification of hand movements.

7.3 SOFTWARE IMPLEMENTATION

The raw EEG data acquired from Net Station software were imported to the Python program implemented in this thesis. The primary software module implementation encapsulates the automatic procedure explained in Section 7.2 and illustrated in Figure 31.

DEPENDENCIES

The project implementation dependencies consist of NumPy, SciPy, Matplotlib, Scikit-Learn, and a separate module, implemented by the researchers, called Pysa. SciPy is a Python-based ecosystem of open-source software for mathematics, science and engineering. Numpy is a core package in SciPy and is fundamental for creating N-dimensional arrays. Matplotlib is a python 2D and 3D plotting library which produces publication quality figures, and with the Pyplot interface, plots can be provided with a Matlab-like interface. Scikit-Learn is an open-source machine learning library in Python built on NumPy, SciPy, and Matplotlib. The library is a straightforward and efficient tool for data mining and data analysis. The implemented Pysa module is based on the theory explained in Chapter 4. Pysa stands for Python signal analysis, and is developed as an open-source signal analysis toolbox for neuroscientists.

PROCEDURE

As an initial pre-processing step, the raw EEG data loaded from the Matlab-file were segmented and stored locally. Figure 32 depicts the segmented raw EEG data in a sliding plot, where the window reflects one trial. The segmentation process loops through the data and assigns a segmented data matrix to the corresponding event.

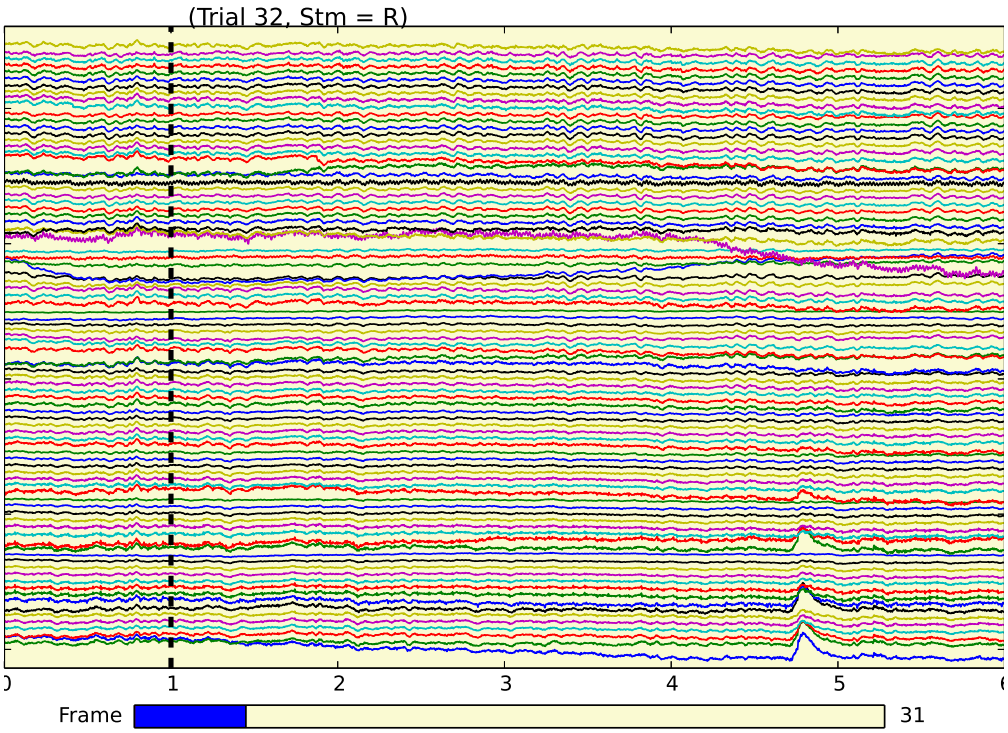

Figure 32.: Raw EEG data (90 electrodes) segmented using the platform implemented in Python.

Furthermore, the segmented data were filtered using the EMD method implemented in the Pysa module. Other studies have shown that EMD acts primarily as a dyadic filter bank [21], meaning that the IMFs that reflect the electrical line noise could be rejected when reconstructing the original signal, or subtracted from the data. Figure 14 in Chapter 4 shows the IMFs extracted from one electrode using the Matplotlib interface for plotting. To begin with, the EMD implementation will normalize the signal to be decomposed. After EMD is applied to the normalized data, the IMFs are returned to the main module. The first *two* IMFs extracted from one electrode were then subtracted from the signal. Figure 33 shows how the dyadic EMD filter was applied on multiple electrodes.

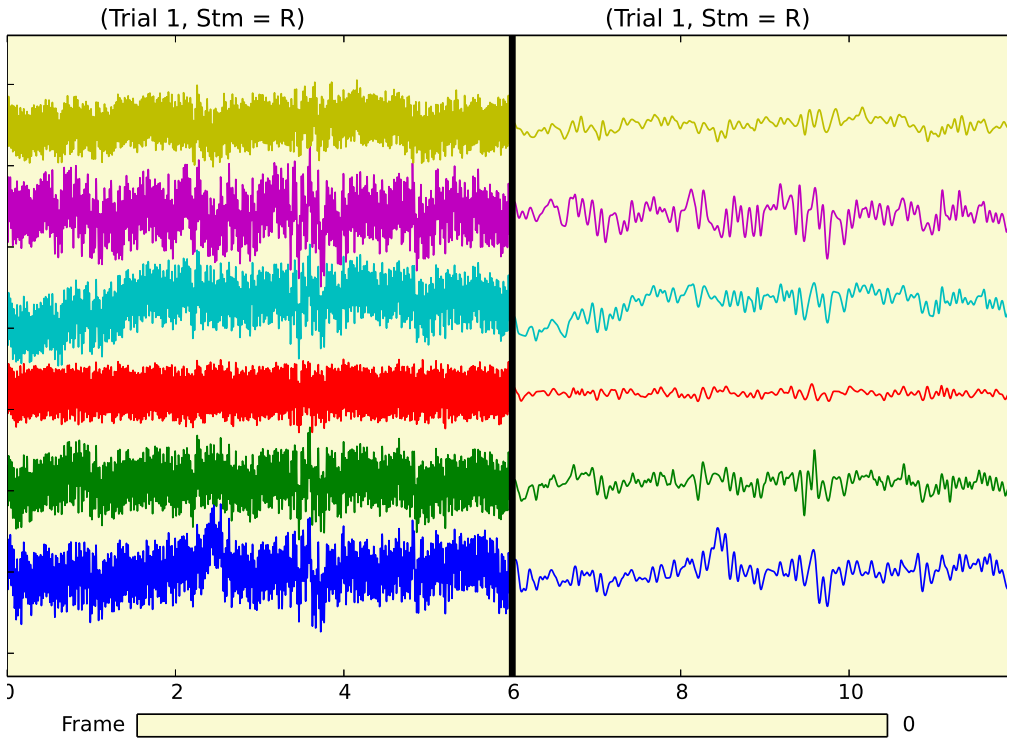

Figure 33.: Six electrodes particularly contaminated by electrical line noise. The black vertical line separates a specific trial before and after the EMD-filtering procedure was applied.

The trial rejection process consists of detecting significant amplitude variations in the measured scalp EEG waveforms based on predefined and tuned thresholds within a short period, for instance 100 ms. The variance and the gradient of one electrode over the preceding period were calculated. The variation in the measured data was participant specific, hence, the chosen thresholds varied from subject to subject. The most frequenctly used time period to search for too large amplitude variations was 100 ms. When detecting too large fluctuations from a specific electrode within this time period, the electrode in question was marked as bad, meaning that the whole trial recorded from the bad electrode was rejected from further analysis. Figure 34 shows how eye blinks are detected in a particular trial and discarded.

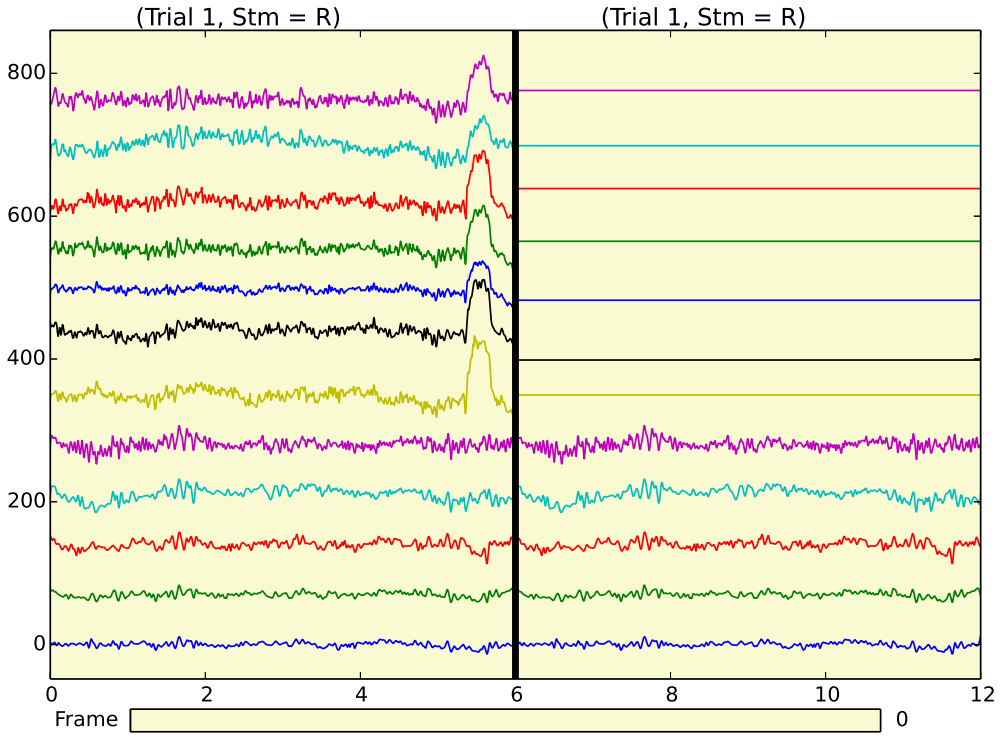

Figure 34.: 12 electrodes after EMD-filtering procedure was applied. The black line separates a specific trial before and after the trial rejection procedure was performed.

After cleaning the data using both EMD filtering and trial rejection, the normalized Hilbert transform (NHT) was applied using the Pysa module to get the instantaneous frequencies and amplitudes. The NHT method takes the normalized IMFs as input. To get the maxima envelope of an IMF as described in Section 4.6, a Pchipinterpolator was utilized instead of a cubic spline interpolator. The Scipy library was used to get the Hilbert transform of the normalized IMFs. After the instantaneous frequency was calculated by differentiating the instantaneous phase with respect to time, detection and correction of rapid unrealistic changes in the frequency were applied, using a threshold-based detection method. The threshold was set to 50, meaning that the difference between two consecutive samples in the computed frequency should be less than this value. If any changes were detected, a correction was comprehended by setting the current data point to be equal to the previous one. The returned frequency and amplitude modulations were stored locally using the main module.

Even though the method for correcting the end effect produced by EMD was implemented, some end effects were still observed. In addition to this, the Hilbert transform algorithm also introduced some small fluctuations near the ends. To overcome

these issues, the trial was reduced in size. After EMD-filtering had been executed, the trial was reduced by 200 ms from the start and 400 ms from the end, yielding a total of 5.4 seconds of recorded data left to analyze. Moreover, after the instantaneous frequency was calculated by the Hilbert transform, the trial duration was reduced by exluding the first 300 ms and the last 600 ms. This resulted in a final trial duration of 4.5 seconds, in which the stimulus onset was after 500 ms.

Feature extraction was performed by the primary module by the procedure explained in the previous section. For classification, a support vector machine with radial basis function (RBF) kernel was applied. A grid search was used to optimize the classifier output. The following values for the hyperparameters C and γ were tested:

- C: 1, 2, 4, 6, 8 , and 10;
- γ: 0.0001, 0.001, 0.01, 0.1, 1.0, and 2.0.

Furthermore, an optimization method similar to the grid search was implemented and used to find the combination of baseline and ERD duration parameters that provided the highest accuracy output from the classifier. A grid of different baseline and ERD duration parameters was set up with the following values, in milliseconds:

- baseline: 150, 175, 200, 250, 300, 350, 400, 450, 500, and 550;
- ERD duration: 1100, 1150, 1200, 1250, 1300, 1400, 1450, 1500, 1550, 1600, 1700, 1750, 1800, 1900, and 2000.

The ratio between the number of trials used in the training set and the testing set was selected such that the total number of predictions were within the interval from 18 to 24 predictions. Moreover, after the classification outcome, the area under the receiver operating characteristics (ROC) curve was used to evaluate the classifier performance.

8

RESULTS

This chapter starts by presenting an average time-frequency analysis of all the trials, comparing left and right hand movements to the neutral condition. Event-related desynchronizations (ERDs) are clearly observed during fist clenching, both when averaging multiple trials and when investigating a single trial. These results are visualized in multiple ERD plots and Hilbert spectra, indicating the differences between the conditions. The classification results are presented, both for single subjects and a grand average for all eight participants. Classification accuracies of up to 95% are reported, including an average accuracy of 87.86% when differentiating between left and right hand movements. The evaluation of the classifier and its hyperparameters are examined, in addition to the optimization of the feature parameters.

8.1 TIME-FREQUENCY ANALYSIS

AVERAGE

After obtaining the instantaneous frequency and amplitude by applying the normalized Hilbert transform (NHT), the average results from each condition were computed. That is, the average time-frequency solutions for all the accepted left hand trials, right hand trials, and neutral trials were calculated, separately, for each of the participants. Using the recorded EEG data from one specific subject, the average normalized power in the *first* IMF was computed from 62 left hand trials, recorded from the Cz electrode. The results are illustrated in the Hilbert spectrum (HS) in Figure 35. Moreover, the same property is depicted in Figure 36, but here the average power is calculated from 61 neutral conditions, meaning trials in which no hand movements occurred.

The results show low-gamma ($\omega \simeq 35$ Hz) desynchronization during left fist clenching, starting shortly after stimulus onset and lasting for the entire three-second execution period. Because of the removal of the first *two* IMFs in the EMD-filtering procedure, the first IMF is, strictly speaking, the third IMF. The numbering of the IMFs will, however, not account for the IMFs removed during this procedure, because they are not a part of the time-frequency analysis. The frequency content in the IMFs were, after averaging all trials, approximately divided into the following frequency ranges:

- IMF 1: $\omega_1 \in (20, 40)$ [Hz]
- IMF 2: $\omega_2 \in (10, 20)$ [Hz]
- IMF 3: $\omega_3 \in (7, 15)$ [Hz]
- IMF 4: $\omega_4 \in (3, 8)$ [Hz]

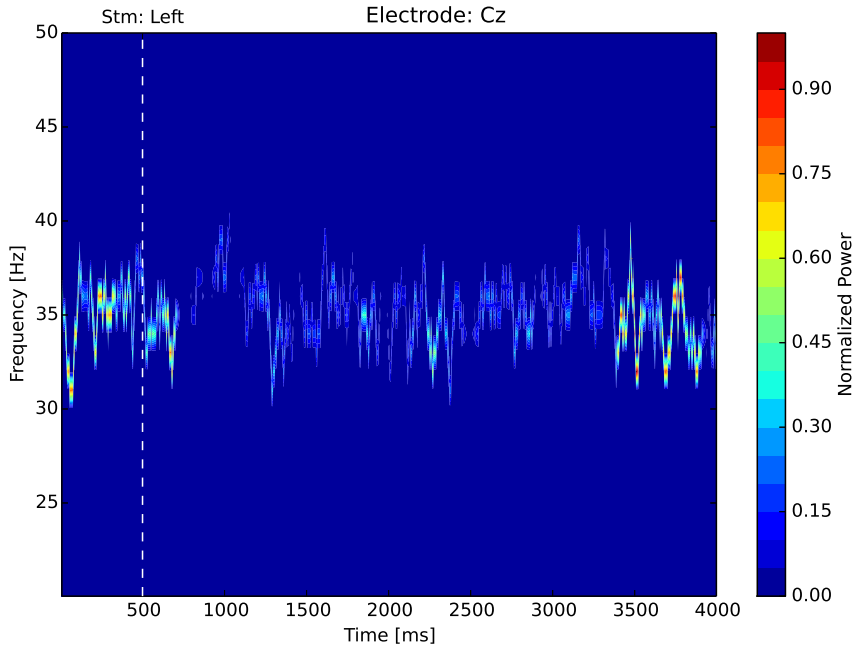

Figure 35.: Normalized power in the first IMF averaged over 62 left hand movements, recorded from *Subject 1*.

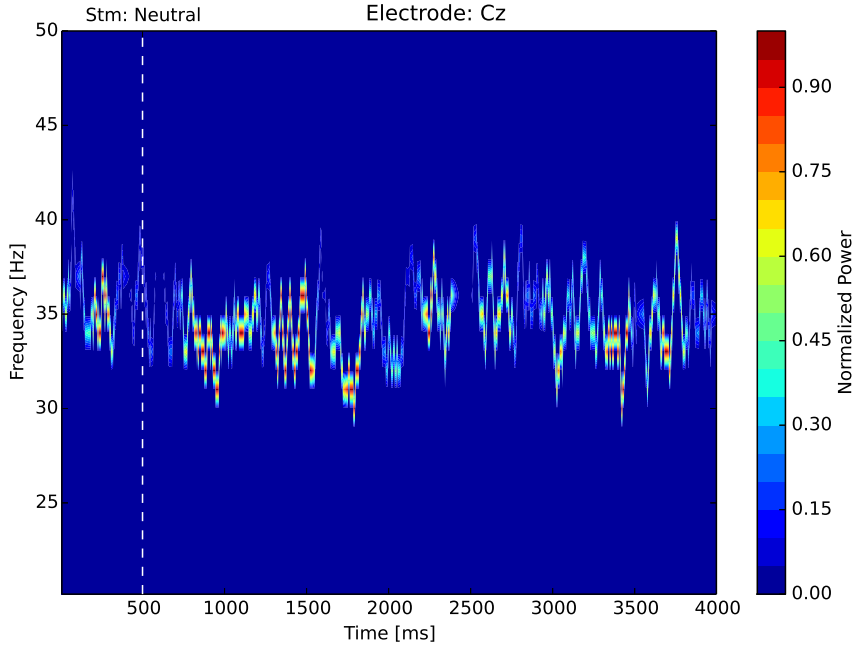

Figure 36.: Normalized power in the first IMF averaged over 61 trials in which no hand movements were executed, recorded from *Subject 1*.

By using the equation for event-related desynchronization, as described in Chapter 2 (Equation 1), both the average ERD from multiple trials and the ERD from single trials were computed. Even though the equation yields both positive and negative values, which in turn signifies synchronized and desynchronized activities, the acronym ERD is used instead of ERD/ERS when referring to event-related (de)synchronization. The average ERD resulting from 62 left hand trials and 61 neutral trials are depicted in Figure 37 and Figure 38, respectively. These results are obtained using the recorded EEG data from the same subject as before, referred to as *Subject 1*. Similarly to the Hilbert spectra in Figure 35 and 36, the calculations were applied to the first IMF, and the electrode location was Cz. The mean baseline power was calculated as the mean power in the 500 ms prior to the stimulus onset. The results in Figure 37 show that, during the approximately three-second long hand movement, the power in the low-gamma IMF is significantly lower compared to the mean baseline power. Moreover, the signal power is increasing shortly after the fist clenching task is executed. The results from the neutral trials, as illustrated in Figure 36 and Figure 38, show no significant change in power relative to the stimulus onset.

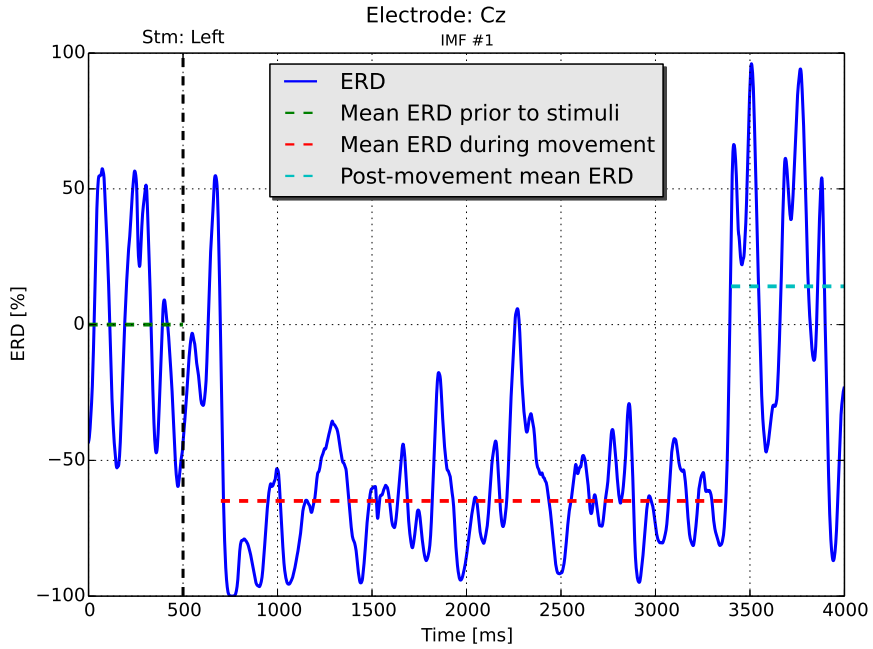

Figure 37.: Average ERD calculated using the first IMF from 62 left hand movements, recorded from *Subject 1*.

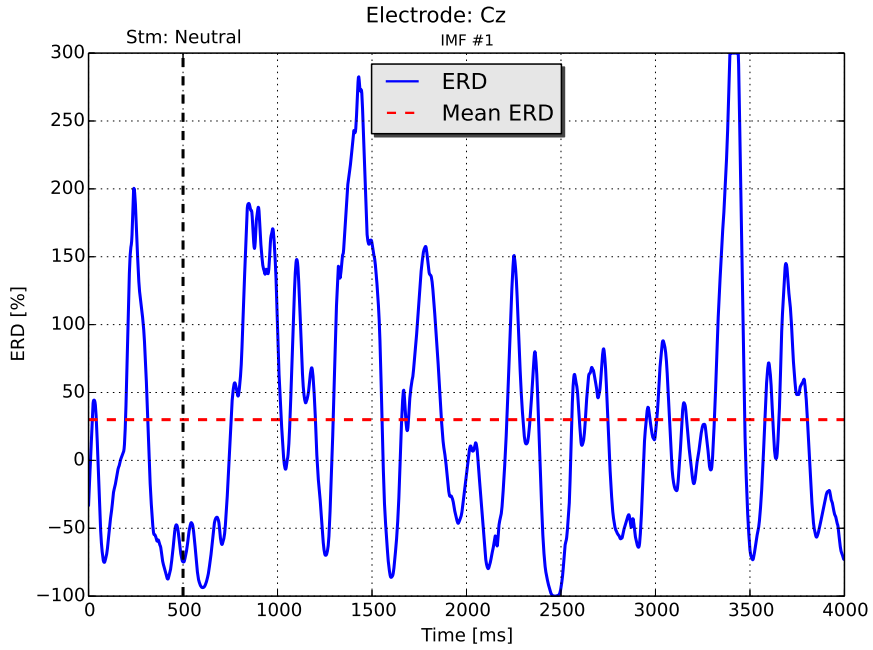

Figure 38.: Average ERD calculated using the first IMF from 61 trials with no hand movements, recorded from *Subject 1*.

Similar results were gathered from other subjects, particularly the low-gamma desynchronization during a motor-related event. Figure 39 shows the results obtained from the $CP3$ electrode, from *Subject 2*, when calculating the normalized power in the first IMF averaged over 65 right hand movements. In comparison, Figure 40 shows the normalized power in the first IMF averaged over 66 neutral trials. The frequency of the IMF is centered around the border between the high-beta and low-gamma brain rhythms ($\omega \simeq 30$ Hz). The signal power experiences a distinct suppression during the right fist clenching when compared to the mean baseline power. This suppression is not found in the neutral trials, that is, when the subject was physically at rest. These results are also illustrated by computing the event-related (de)synchroniation, both for right hand movements and neutral trials, in Figure 41 and in Figure 42. Figure 41 shows a mean ERD of 90% during movement when compared to the mean baseline power.

In addition to the 30 Hz desynchronized activity found in *Subject 2*, a clear *mu* rhythm desynchronization ($\omega \simeq 9$ Hz) was found when the subject was executing a left hand movement. The induced activity was recorded from the $CP4$ electrode, and the normalized power in the *third* IMF averaged over 65 left hand movements is depicted in Figure 43. The *mu*-suppression is also seen in the average ERD calculated from the same trials, illustrated in Figure 45. In contrast to the desynchronized activity, Figure 44 and Figure 46 show the average power and the average ERD in the third IMF obtained from 56 neutral trials, respectively.

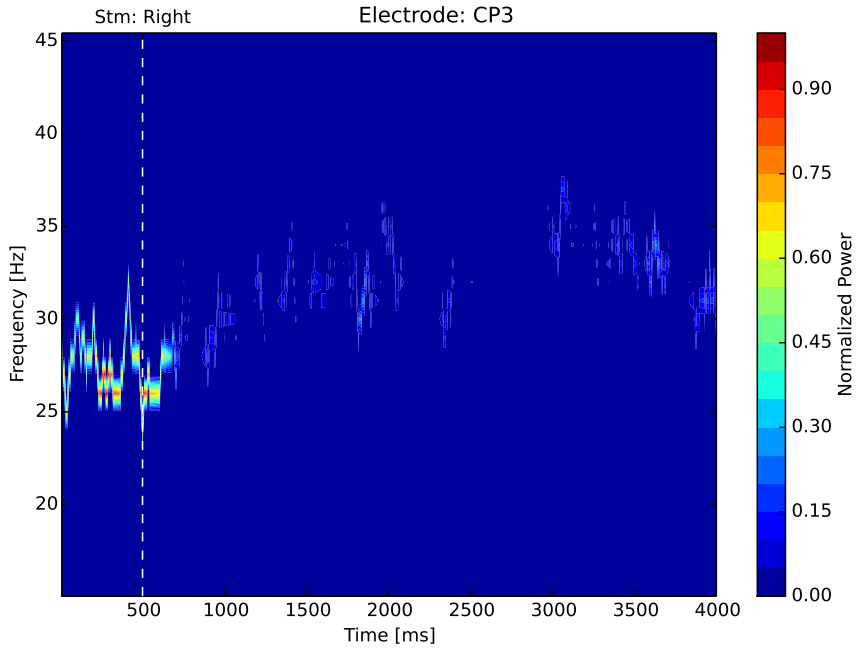

Figure 39.: Normalized power in the first IMF averaged over 65 right hand movements, recorded from *Subject 2*.

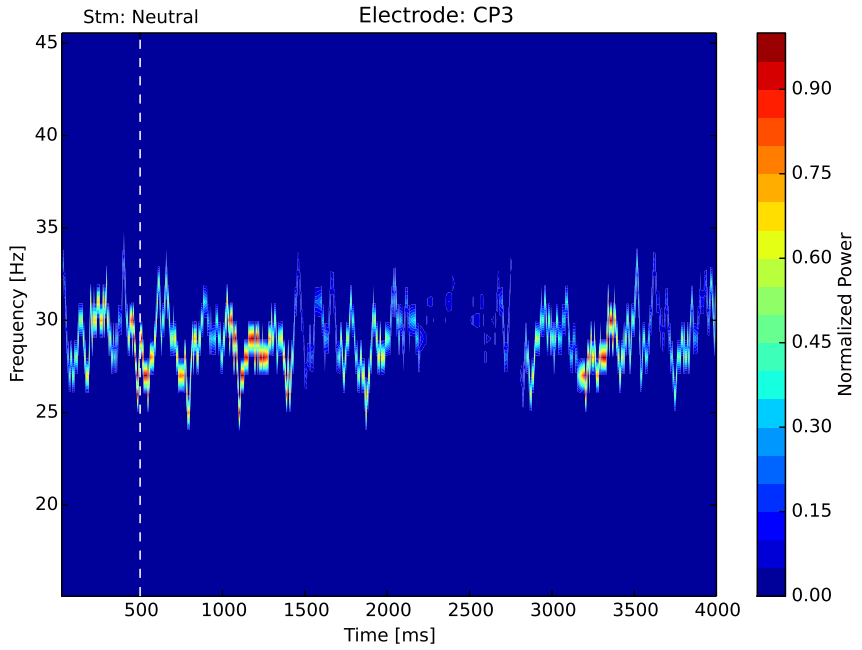

Figure 40.: Normalized power in the first IMF averaged over 66 trials in which no hand movements occurred, recorded from *Subject 2*.

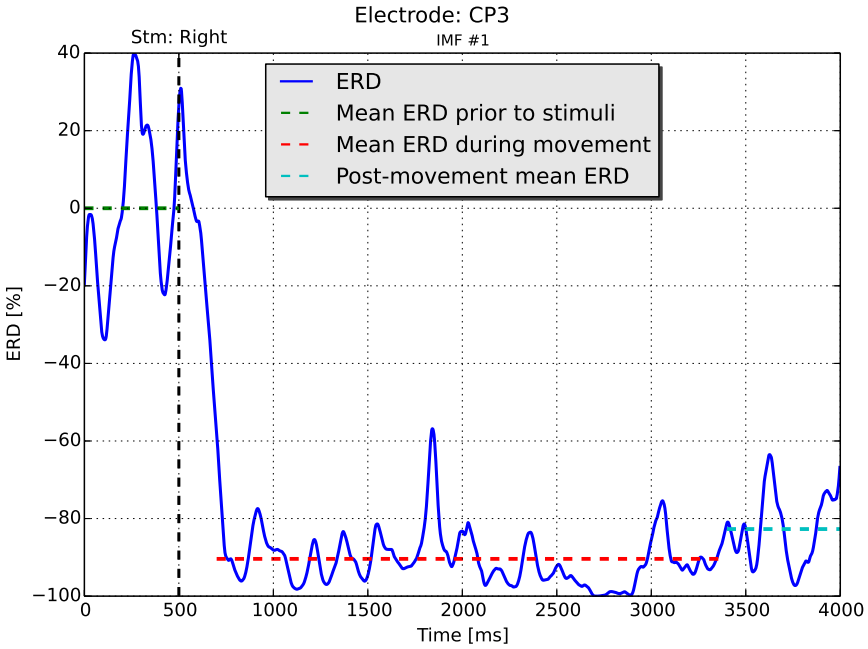

Figure 41.: Average ERD calculated using the first IMF from 65 right hand movements, recorded from *Subject 2*.

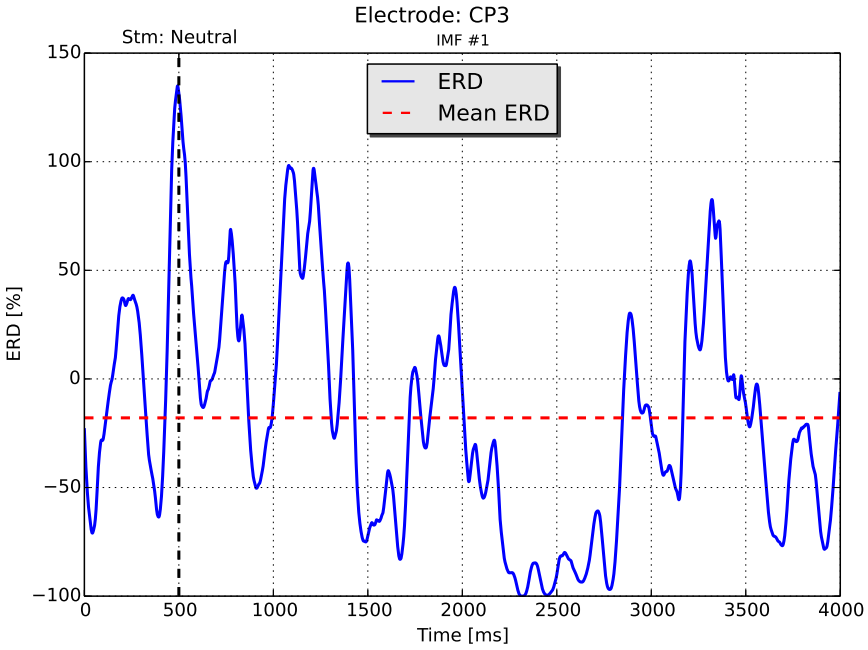

Figure 42.: Average ERD calculated using the first IMF from 66 trials with no hand movements, recorded from *Subject 2*.

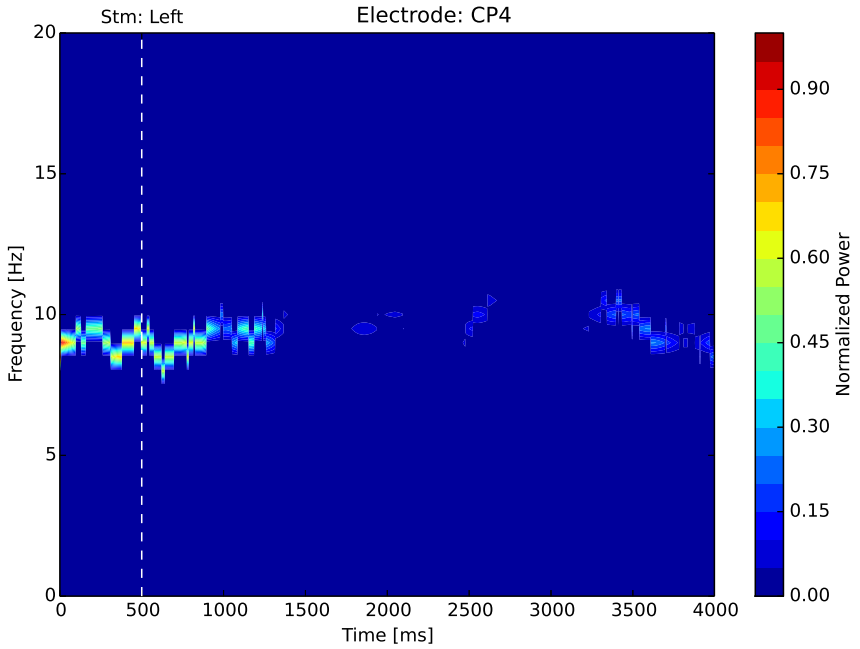

Figure 43.: Normalized power in the third IMF averaged over 65 left hand movements, recorded from *Subject 2*.

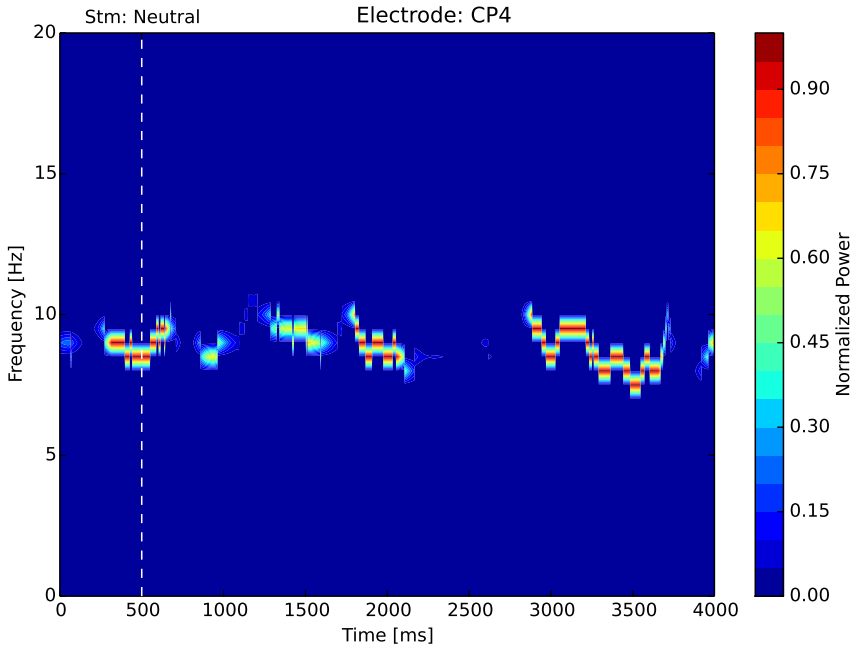

Figure 44.: Normalized power in the third IMF averaged over 56 trials in which no hand movements occurred, recorded from *Subject 2*.

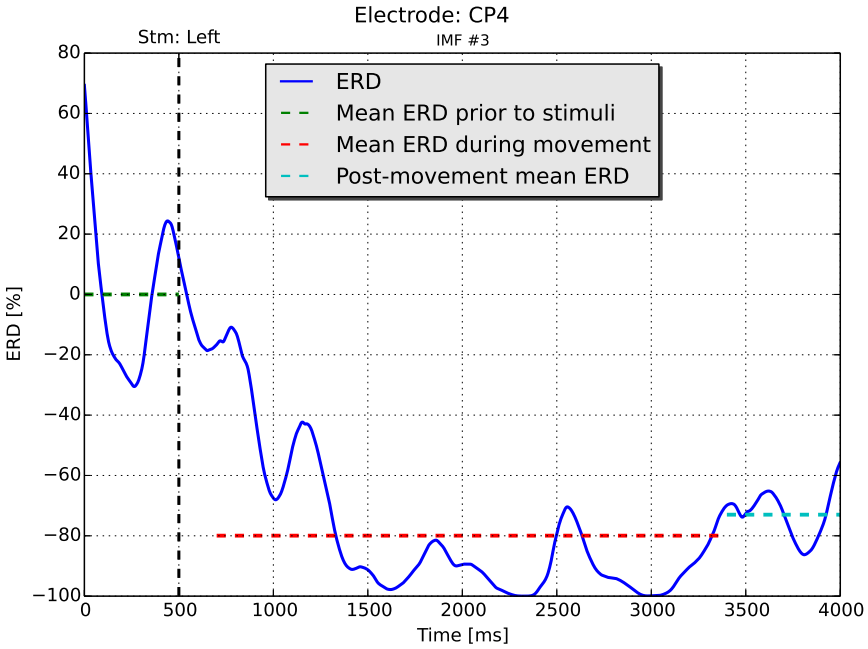

Figure 45.: Average ERD calculated using the third IMF from 65 left hand movements, recorded from *Subject 2*.

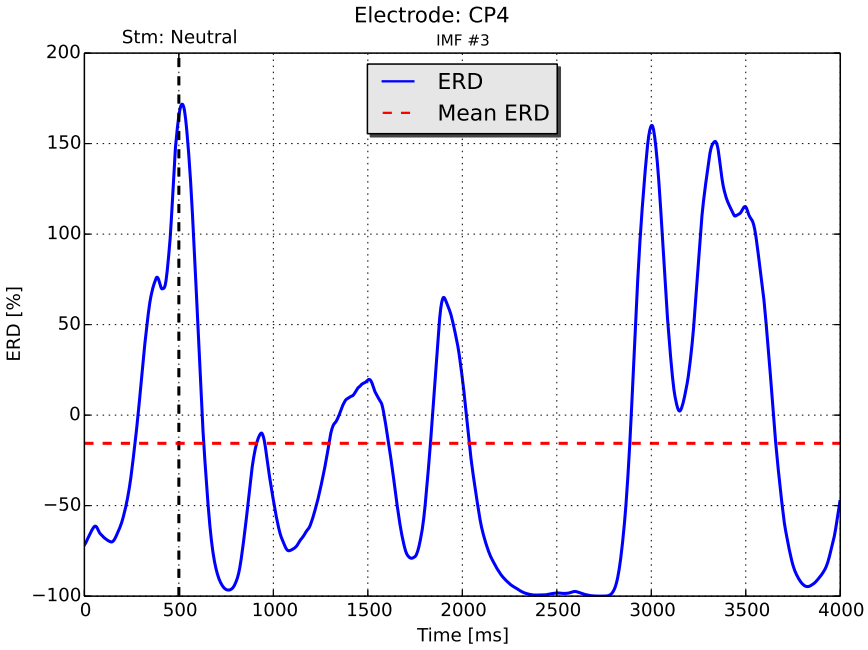

Figure 46.: Average ERD calculated using the third IMF from 56 trials with no hand movements, recorded from *Subject 2*.

SINGLE TRIAL

The normalized power in the *first* IMF from a single left hand trial and a single neutral trial, acquired from *Subject 1*, are illustrated in the Hilbert spectra in Figure 47 and Figure 48. The low-gamma desynchronization during left fist clenching is recognized both in a single trial, as well as in the average results of all trials. Furthermore, the ERD calculated from the same trials are illustrated in Figure 49 and Figure 50. The decreased power compared to the mean baseline power is prominent during the execution of a hand movement, which is similar to the average results depicted in Figure 37. For the neutral condition in Figure 50, there is shown a relative increase in the signal power, however, a much higher variance is seen when compared to the hand movement condition.

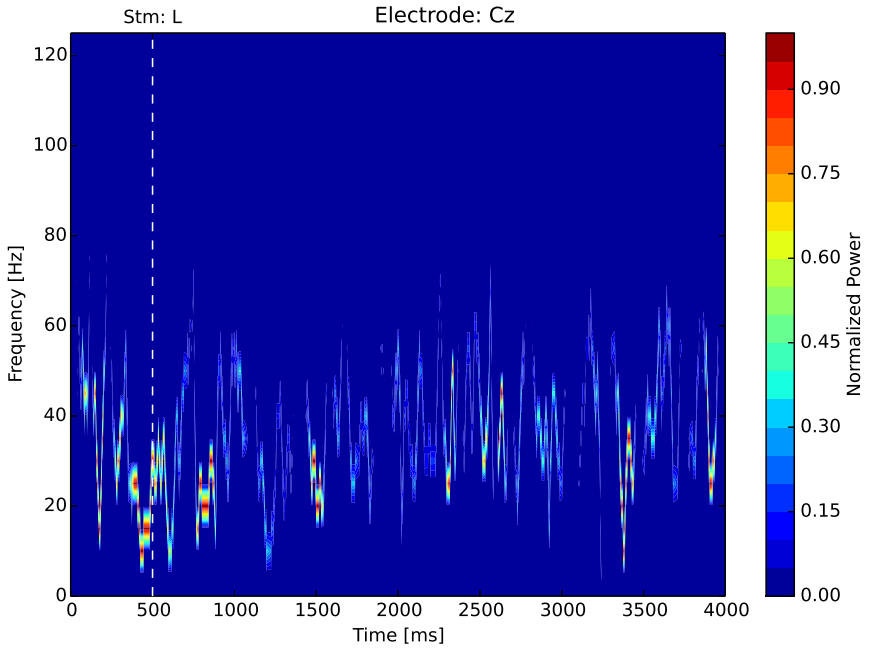

Figure 47.: Hilbert spectrum of the normalized power in the first IMF from a single left hand trial, recorded from *Subject 1*.

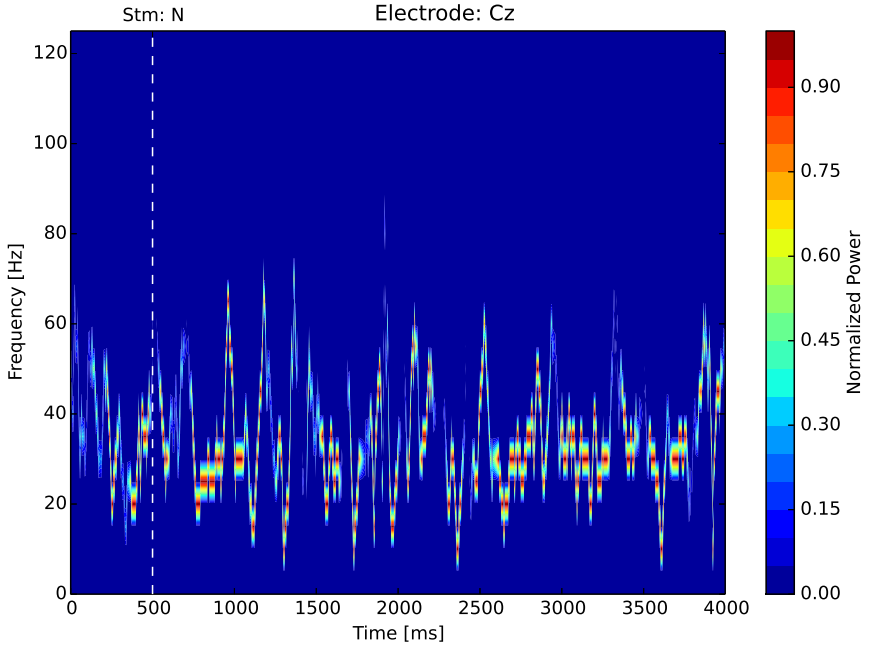

Figure 48.: Hilbert spectrum of the normalized power in the first IMF from a single trial with no hand movements, recorded from *Subject 1*.

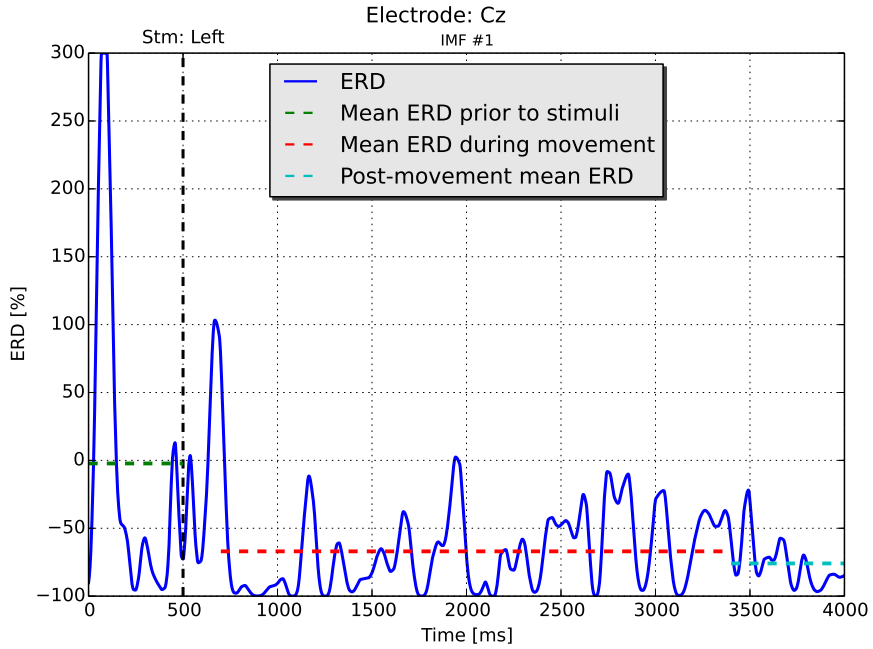

Figure 49.: ERD calculated using the first IMF from a specific left hand trial, recorded from *Subject 1*.

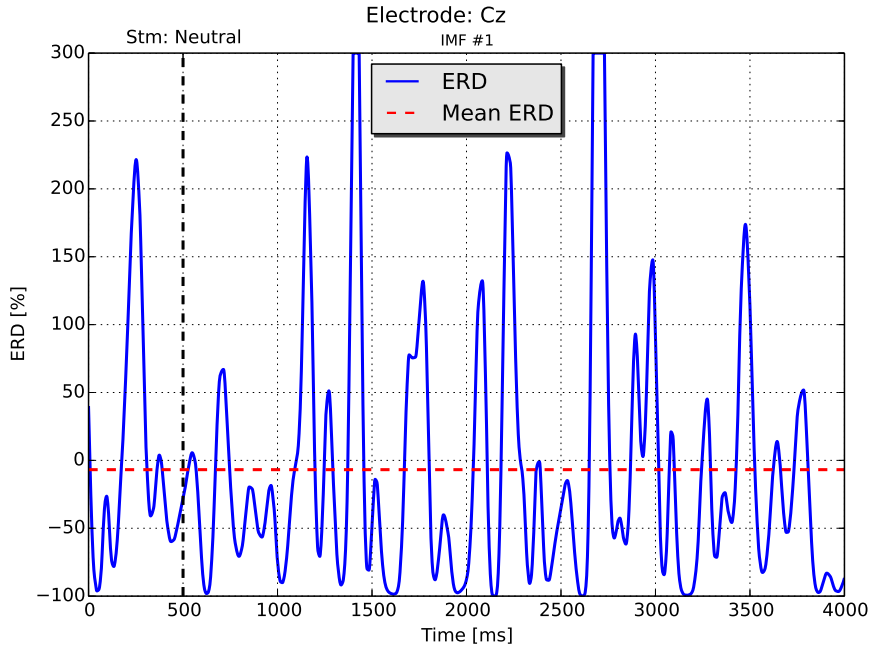

Figure 50.: ERD calculated using the first IMF from a single trial with no hand movements, recorded from *Subject 1*.

RIGHT VERSUS LEFT

Subject 5 experienced a clear suppression of the *mu* rhythm ($\omega \simeq 13$ Hz) during right hand movements. Figure 51 illustrates the normalized power in the *second* IMF averaged over 69 right hand trials, recorded at the $C1$ electrode. A similar suppression was *not* evident during left hand movements, as seen in Figure 52, which depicts the normalized power in the same IMF averaged over 68 left hand trials. To make the differentiation even more distinct, Figure 53 and Figure 54 illustrate the ERD from the same right hand trials and left hand trials, respectively. An additional comparison of right and left hand movements can be found in Appendix A, showing the Hilbert spectra and ERD plots of the *second* IMF recorded from C2, using the data acquired from *Subject 7*.

All the results presented in this section show a distinct difference in signal power before and during a motor-related event, both for single trials, and when averaging the results obtained from multiple trials. The electrode location and the IMF number that provided the most prominent results were subject specific, and this will become even more apparent in the next section. However, after acquiring the optimal electrode and IMF, the mean ERD proves to be an acceptable feature for differentiating between the three conditions, namely, left hand, right hand, and neutral. The next section will provide the classification results obtained from multiple subjects and a grand average classification success rate, in addition to an evaluation of the classifier.

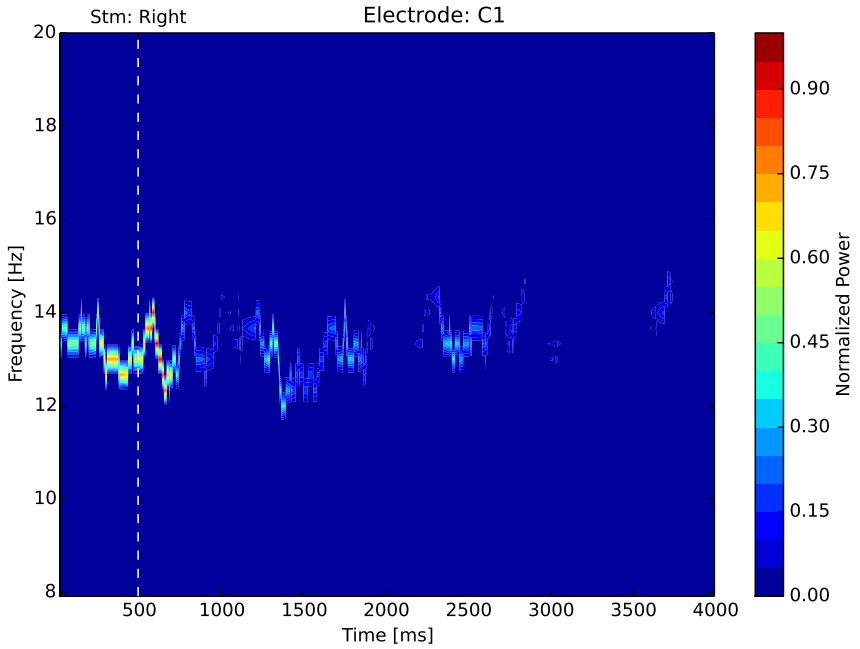

Figure 51.: Normalized power in the second IMF averaged over 69 right hand movements, recorded from *Subject 5*.

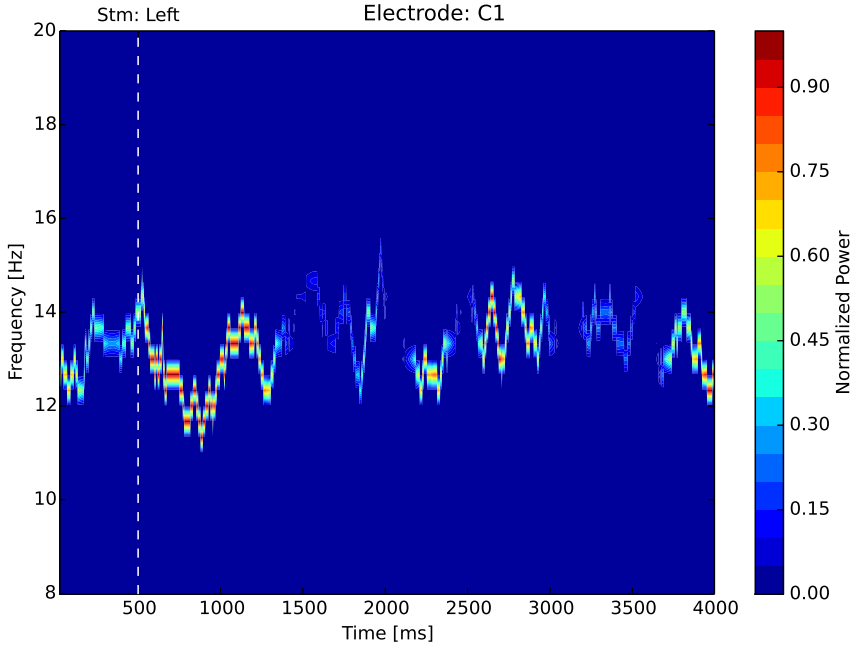

Figure 52.: Normalized power in the second IMF averaged over 68 left hand movements, recorded from *Subject 5*.

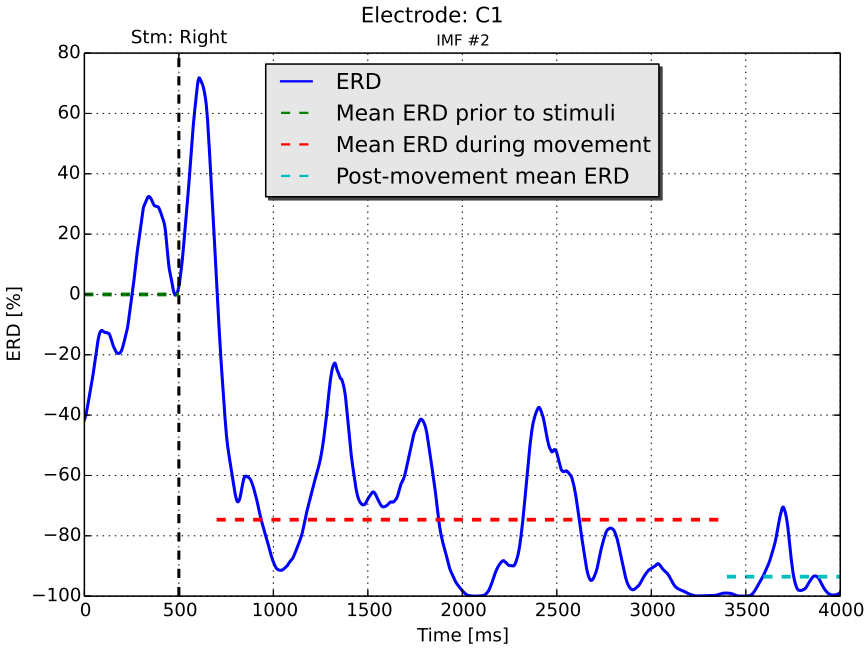

Figure 53.: Average ERD calculated using the second IMF from 69 right hand movements, recorded from *Subject 5*.

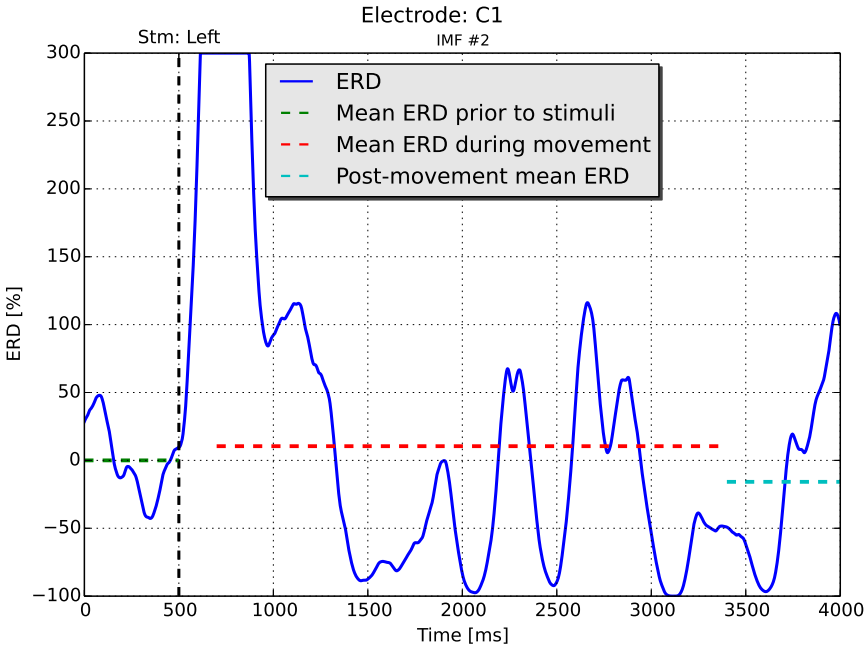

Figure 54.: Average ERD calculated using the second IMF from 68 left hand movements, recorded from *Subject 5*.

8.2 Classification

The classification performances were studied using standard measures including overall accuracy and the *ROC* curve. The former is the percentage of conditions that were correctly predicted, while the latter is the area under the receiver operating characteristic curve. A support vector machine (SVM) classifier was applied, with an RBF kernel optimized by a grid search. For the eight subjects, *three* different conditions were considered. The conditions are from now on referred to as:

- **Condition 1**: Right hand movement.
- **Condition 2**: Left hand movement.
- **Condition 3**: No hand movement (neutral).

Moreover, the conditions resulted in three binary classification problems, which are also known as *two-class problems*. The three classification problems are from now on referred to as:

- **Problem I**: Condition 1 against Condition 2 (right versus left).
- **Problem II**: Condition 1 against Condition 3 (right versus neutral).
- **Problem III**: Condition 2 against Condition 3 (left versus neutral).

Because of the change in the experimental paradigm, the classification scheme for the first *four* participants could only consider *Problem I*, namely *Condition 1* against *Condition 2*. This was because these subjects only had 10 trials of *Condition 3*, instead of 70. As for the last four subjects, all three classification problems were investigated.

Within-subject classification scores were based on various settings. Consequently, the results will include every important parameter in addition to the classification scores. These parameters are:

- the recording electrode
- the number of accepted trials for the different conditions
- the ratio between the trials employed as support vectors and the trials used for prediction
- the IMF number used for computing the ERD feature
- the baseline duration in milliseconds, used for obtaining the reference power when calculating the ERD
- the ERD duration in milliseconds. This determines the period in which the ERD feature should be calculated

A method for optimizing the parameters regarding baseline and ERD duration was applied. The method used a technique similar to a grid search, and found the optimal combination of the baseline and ERD duration that provided the highest classification accuracy.

PERFORMANCE MEASURES

The optimal classification scores were calculated at different electrodes above the sensorimotor cortex, with various subject-specific settings. Table 3 shows the classification results of *Problem I* for *Subject 2*, from each of the 16 electrodes. The best result was obtained from the *first* IMF, recorded from the FCCz electrode, yielding a classification accuracy of **90%**. As seen in the table, 18 correct predictions out of 20 were the basis for the score. The mean number of predictions based on all electrodes was 20, with a standard deviation (SD) of 0.7.

Classification: Problem I

	Parameters						Scores	
Electrode	**Accepted trials left**	**Accepted trials right**	**Ratio trained**	**IMF number**	**Baseline [ms]**	**ERD duration [ms]**	**Accuracy [%]**	**Correct Predictions**
Cz	70	70	0.86	4	500	3000	**85**	17 / 20
FCCz	68	67	0.86	1	400	3100	**90**	18 / 20
FC3	69	69	0.86	2	350	3200	**85**	17 / 20
FCC3	69	69	0.86	1	300	2400	**75**	15 / 20
C1	69	69	0.86	1	1000	2800	**80**	16 / 20
C3	69	70	0.86	1	600	2900	**75**	15 / 20
C5	66	60	0.84	3	900	2200	**85.71**	18 / 21
CP1	68	69	0.86	1	500	3600	**80**	16 / 20
CP3	67	66	0.86	1	350	2300	**75**	15 / 20
FC4	70	67	0.86	2	500	3400	**80**	16 / 20
FCC4	70	67	0.86	2	500	3500	**75**	15 / 20
C2	70	70	0.88	1	600	3200	**83.33**	15 / 18
C4	70	67	0.86	2	350	3000	**85**	17 / 20
C6	70	68	0.86	1	350	3600	**80**	16 / 20
CP2	68	64	0.85	4	600	2400	**80.95**	17 / 21
CP4	68	61	0.85	1	1100	4000	**85.71**	18 / 21

Table 3.: Classification results for *Subject 2*, using left and right hand movements as conditions. The electrode with the highest accuracy is marked with green.

The classification results of *Problem II*, also for *Subject 2*, are shown in Table 4. Similar to the previous classification problem, the best score was also **90%** and obtained from FCCz, but from the *fourth* IMF. In addition to this, a **90%** score was obtained both from the electrodes C1 and FCC4, using IMF number two and three, respectively. The mean number of predictions based on all 16 electrodes was 20 (SD = 2.2).

Classification: Problem II

	Parameters						Scores	
Electrode	**Accepted trials right**	**Accepted trials neutral**	**Ratio trained**	**IMF number**	**Baseline [ms]**	**ERD duration [ms]**	**Accuracy [%]**	**Correct Predictions**
Cz	70	69	0.86	4	300	2300	**80**	16 / 20
FCCz	66	55	0.84	4	500	3000	**90.00**	18 / 20
FC3	67	63	0.91	2	500	3000	**84.62**	11 / 13
FCC3	69	67	0.86	4	500	2500	**75.00**	15 / 20
C1	70	69	0.86	2	500	3800	**90.00**	18 / 20
C3	70	69	0.86	3	800	3600	**85.00**	17 / 20
C5	60	69	0.86	2	700	2400	**78.94**	15 / 19
CP1	69	69	0.86	3	900	3000	**85.00**	17 / 20
CP3	60	63	0.82	1	500	3000	**82.61**	19 / 23
FC4	67	67	0.86	1	350	3000	**80.00**	16 / 20
FCC4	67	66	0.86	3	1000	3000	**90.00**	18 / 20
C2	70	68	0.84	1	380	2500	**82.61**	19 / 23
C4	67	69	0.86	3	700	2500	**80.00**	16 / 20
C6	68	67	0.86	4	350	3800	**75.00**	15 / 20
CP2	57	48	0.8	1	500	3800	**77.27**	17 / 22
CP4	54	41	0.8	1	450	2400	**75.00**	15 / 20

Table 4.: Classification results for *Subject 2*, using right hand movements and neutral trials as conditions. The electrodes with the highest accuracy are marked with green.

The results from the third classification problem, still using the data from *Subject 2*, are presented in Table 5. The two best performing electrodes were CP1 and CP2, both with features calculated using the first IMF, resulting in 18 correct predictions out of 20, which coincides with an accuracy of **90%**. The mean number of predictions for all the electrodes was 20 (SD = 1.1).

Classification: Problem III

	Parameters						Scores	
Electrode	**Accepted trials left**	**Accepted trials neutral**	**Ratio trained**	**IMF number**	**Baseline [ms]**	**ERD duration [ms]**	**Accuracy [%]**	**Correct Predictions**
Cz	70	69	0.86	4	300	2300	**80**	16 / 20
FCCz	62	44	0.78	4	500	3000	**83.33**	20 / 24
FC3	69	67	0.86	3	700	2900	**85**	17 / 20
FCC3	69	69	0.86	2	800	2300	**85**	17 / 20
C1	70	69	0.86	4	300	2300	**80**	16 / 20
C3	69	69	0.86	2	700	2500	**85**	17 / 20
C5	66	69	0.86	2	700	2900	**80**	16 / 20
CP1	68	69	0.86	1	800	2900	**90**	18 / 20
CP3	67	68	0.86	1	600	2800	**85**	17 / 20
FC4	70	67	0.86	2	300	3200	**75**	15 / 20
FCC4	65	61	0.85	2	400	2800	**80**	16 / 20
C2	70	68	0.85	1	500	3000	**81.82**	18 / 22
C4	70	69	0.86	1	1100	3800	**85**	17 / 20
C6	70	67	0.86	4	500	2900	**75**	15 / 20
CP2	68	59	0.85	1	600	2200	**90**	18 / 20
CP4	68	58	0.85	3	300	2200	**85**	17 / 20

Table 5.: Classification results for *Subject 2*, using left hand movements and neutral trials as conditions. The electrodes with the highest accuracy are marked with green.

Similar classification results from *Problem II* and *Problem III* were obtained for the other subjects. These results are presented in Appendix B. Given the objective of this research, the remaining classification results will focus on *Problem I*, namely right and left hand movements. Table 6 shows the classification results from *Problem I*, using the data from *Subject 1*. The best score, **85.71%**, was obtained from C4 and C5, using IMF number three and four, respectively. In both cases, 18 out of 21 trials were correctly predicted. When considering all of the 16 electrodes, the mean number of trials used in the testing set was 20 (SD = 1).

Classification: Problem I

	Parameters						Scores	
Electrode	**Accepted trials left**	**Accepted trials right**	**Ratio trained**	**IMF number**	**Baseline [ms]**	**ERD duration [ms]**	**Accuracy [%]**	**Correct Predictions**
Cz	70	70	0.86	3	400	4000	**80**	16 / 20
FCCz	68	66	0.86	1	900	3800	**85**	17 / 20
FC3	55	59	0.84	4	700	3100	**84.21**	16 / 19
FCC3	54	60	0.84	2	300	2800	**84.21**	16 / 19
C1	69	69	0.86	3	1100	2600	**80**	16 / 20
C3	50	54	0.82	1	700	3000	**78.95**	15 / 19
C5	30	37	0.7	4	380	2100	**85.71**	18 / 21
CP1	56	58	0.84	1	300	4000	**84.21**	16 / 19
CP3	42	48	0.78	3	800	3000	**76.19**	16 / 21
FC4	23	18	0.58	3	600	3600	**83.33**	15 / 18
FCC4	49	51	0.8	1	300	3400	**71.43**	15 / 21
C2	65	63	0.85	1	700	3200	**80**	16 / 20
C4	56	59	0.83	3	1100	2400	**85.71**	18 / 21
C6	52	53	0.82	3	800	2400	**85**	17 / 20
CP2	59	51	0.82	1	400	2800	**77.27**	17 / 22
CP4	44	41	0.78	1	600	3400	**75**	15 / 20

Table 6.: Classification results for *Subject 1*, using right and left hand movements as conditions. The electrodes with the highest accuracy are marked with green.

When considering *Problem I*, the highest classification accuracy was obtained using the data from *Subject 5*, and the corresponding results are presented in Table 7. The best electrodes were C1 and C3, both yielding an accuracy of **95%**. This coincides to 19 correct predictions out of 20. The IMFs utilized when computing the ERD features were IMF number four and number one, for C1 and C3, in that order. The mean number of predictions based on all 16 electrodes was 20 (SD = 1.2).

Classification: Problem I

	Parameters						Scores	
Electrode	**Accepted trials left**	**Accepted trials right**	**Ratio trained**	**IMF number**	**Baseline [ms]**	**ERD duration [ms]**	**Accuracy [%]**	**Correct Predictions**
Cz	76	78	0.9	1	400	2400	**81.25**	13 / 16
FCCz	73	79	0.87	3	400	3200	**76.19**	16 / 21
FC3	67	59	0.85	3	600	3000	**85**	17 / 20
FCC3	73	73	0.87	4	1100	2900	**80**	16 / 20
C1	78	79	0.88	4	540	3550	**95**	19 / 20
C3	76	77	0.88	1	900	2300	**95**	19 / 20
C5	75	78	0.87	4	350	3000	**80.95**	17 / 21
CP1	75	73	0.87	2	500	4000	**80**	16 / 20
CP3	78	76	0.88	4	350	2400	**85**	17 / 20
FC4	61	59	0.84	1	800	3200	**80**	16 / 20
FCC4	73	75	0.87	3	700	2600	**75**	15 / 20
C2	76	78	0.88	3	1000	3500	**80**	16 / 20
C4	77	79	0.88	2	400	2560	**75**	15 / 20
C6	67	71	0.86	4	1000	3500	**75**	15 / 20
CP2	73	78	0.86	1	600	2300	**77.27**	17 / 22
CP4	76	78	0.88	3	900	2200	**85**	17 / 20

Table 7.: Classification results for *Subject 5*, using right and left hand movements as conditions. The electrodes with the highest accuracy are marked with green.

The rest of the classification results from *Problem I* are presented in Appendix B, and the highest scores attained from each of the remaining subject were:

- *Subject 3*: **85%** from Cz, FCC3, and CP4 (17/20 correct predictions)
- *Subject 4*: **86.36%** from C1 (19/22 correct predictions)
- *Subject 6*: **90%** from FC4 (18/20 correct predictions)
- *Subject 7*: **83.33%** from both Cz and C2 (20/24 correct predictions)
- *Subject 8*: **87.5%** from both FCC3 and FC4 (21/24 correct predictions)

GRAND AVERAGE

Even though the optimal electrode was subject specific, the average classification scores for each electrode and all of the classification problems are presented in Table 8, along with their respective standard deviation (SD). The electrodes that provided the best overall classification accuracy for each problem are the highlighted ones, namely, C5 for *Problem I*, FCCz for *Problem II*, and FCC3 for *Problem III*. The results from all eight subjects were used for the average result of *Problem I*. Considering that four of the subjects only had 10 neutral trials, the average results for *Problem II* and *Problem III* were from the four subjects that had 70 neutral trials. Furthermore, for all classification problems, the average percentage of accepted trials was 91.4%.

Classification accuracy - grand average (1)

Electrode	Problem I	Problem II	Problem III
Cz	81.53 (SD=2.8)	81.25 (SD=2.5)	79.79 (SD=0.4)
FCCz	79.09 (SD=6.3)	**84.87 (SD=6.9)**	81.88 (SD=2.8)
FC3	81.92 (SD=3.8)	82.41 (SD=2.8)	80.08 (SD=8.6)
FCC3	80.51 (SD=5.1)	81.25 (SD=9.5)	**84.56 (SD=1.0)**
C1	81.45 (SD=6.4)	82.50 (SD=6.5)	82.65 (SD=5.5)
C3	80.40 (SD=6.3)	79.32 (SD=4.3)	81.68 (SD=4.8)
C5	**83.22 (SD=3.6)**	79.15 (SD=3.6)	80.81 (SD=3.2)
CP1	79.94 (SD=2.1)	79.79 (SD=4.1)	83.66 (SD=6.9)
CP3	79.63 (SD=3.2)	80.46 (SD=4.0)	82.48 (SD=3.0)
FC4	82.50 (SD=4.9)	83.20 (SD=4.5)	81.82 (SD=6.9)
FCC4	77.37 (SD=3.8)	84.05 (SD=5.7)	78.75 (SD=4.8)
C2	80.06 (SD=2.4)	81.11 (SD=1.3)	79.57 (SD=1.9)
C4	79.15 (SD=5.9)	81.43 (SD=5.0)	79.56 (SD=4.4)
C6	78.38 (SD=4.4)	80.24 (SD=4.1)	77.00 (SD=3.3)
CP2	78.79 (SD=3.8)	80.01 (SD=1.9)	84.34 (SD=5.2)
CP4	79.28 (SD=5.8)	81.05 (SD=2.0)	80.83 (SD=4.4)

Table 8.: Grand average classification scores for the three classification problems, for each electrode.

To get an assessment of the overall classification performance, the highest accuracy obtained from each subject, independent of the electrode, was examined. The results are presented in Table 9, showing an average classification accuracy of:

- **87.86%** when differentiating between right and left hand movements.
- **90.05%** when differentiating between right hand movements and no hand movements.
- **88.3%** when differentiating between left hand movements and no hand movements.

The mean number of predictions used when calculating the average values was 20 for all of the three classification problems.

Classification accuracy - grand average (2)

Classification	Accuracy [%]
Problem I	**87.86** (SD=3.7)
Problem II	**90.05** (SD=3.8)
Problem III	**88.30** (SD=2.1)

Table 9.: Mean classification accuracy for all three classification problems.

EVALUATION AND OPTIMIZATION

Table 10 shows the evaluation of the classification results, expressing a performance measure of the classifier output using the ROC area. The classification problem was *Problem I*, recorded from all subjects, meaning that the evaluations are associated with the results observed in the above tables. The evaluation of the classifier performance of Problem II and III are reported in Appendix B, regarding *Subject* 1 to 4. The highest performance measure was observed in electrode C1 and C3 for *Subject* 5 both yielding a 95% score.

ROC area: Problem I

Electrode	Subject 1	Subject 2	Subject 3	Subject 4	Subject 5	Subject 6	Subject 7	Subject 8
Cz	0.8	0.85	0.85	0.8	0.81	0.81	0.83	0.77
FCCz	0.85	0.9	0.76	0.81	0.75	0.77	0.78	0.69
FC3	0.84	0.85	0.8	0.75	0.83	0.8	0.78	0.86
FCC3	0.84	0.75	0.85	0.73	0.8	0.78	0.82	0.87
C1	0.8	0.8	0.75	0.86	**0.95**	0.79	0.79	0.77
C3	0.79	0.75	0.8	0.82	**0.95**	0.77	0.79	0.75
C5	0.87	0.85	0.8	0.85	0.81	0.89	0.78	0.82
CP1	0.84	0.8	0.81	0.8	0.8	0.79	0.78	0.77
CP3	0.75	0.75	0.8	0.8	0.85	0.82	0.78	0.79
FC4	0.81	0.8	0.75	0.85	0.8	0.9	0.79	0.87
FCC4	0.7	0.75	0.8	0.75	0.75	0.83	0.79	0.81
C2	0.8	0.83	0.8	0.8	0.8	0.77	0.83	0.77
C4	0.85	0.85	0.8	0.8	0.75	0.71	0.71	0.85
C6	0.85	0.8	0.75	0.8	0.75	0.82	0.79	0.7
CP2	0.77	0.8	0.8	0.85	0.77	0.81	0.75	0.73
CP4	0.75	0.86	0.83	0.8	0.85	0.77	0.78	0.69

Table 10.: Evaluation of classification performance using ROC area for *Problem I*, using the data from all subjects.

The optimized hyperparameters, C and *gamma* (γ), are presented in Table 11, revealing how the hyperparameters were changed by testing the predictive abilities in all classification problems and within subjects. After applying the grid search, the most frequently observed C and γ were 1 and 0.1, respectively. The optimized hyperparameters of Problem II and III, regarding *Subject* 1 to 4, are summarized in Appendix B.

C/γ: Problem I

Electrode	Subject 1	Subject 2	Subject 3	Subject 4	Subject 5	Subject 6	Subject 7	Subject 8
Cz	8 / 0.001	1 / 0.001	1/0.01	1 / 0.01	2 / 1.0	4/0.001	2/0.0001	1/0.1
FCCz	1 / 0.1	1 / 2.0	10/0.01	1 / 0.0001	1 / 2.0	2/0.0001	10/0.001	8/0.1
FC3	4 / 0.0001	10 / 0.0001	10/0.01	10 / 0.1	2 / 1.0	4/0.0001	1/1.0	8/0.0001
FCC3	6 / 0.0001	4 / 0.1	1/0.01	1 / 0.1	8 / 0.001	10/0.01	4/0.1	1/0.1
C1	4 / 0.01	1 / 0.01	2/1.0	6 / 0.0001	10 / 0.001	1/0.001	4/1.0	8/0.01
C3	10 / 0.01	1 / 0.01	1/2.0	1 / 0.001	8 / 0.001	1/0.1	1/0.1	1/0.1
C5	1 / 0.01	1 / 2.0	1/0.001	4 / 1.0	6 / 0.001	1/0.01	1/2.0	2/0.01
CP1	8 / 0.1	4 / 0.1	4/0.1	1 / 0.1	10 / 2.0	4/0.1	2/0.1	1/0.001
CP3	2 / 2.0	2 / 0.1	4/1.0	10 / 0.01	10 / 0.1	10/0.001	4/0.001	1/1.0
FC4	2 / 0.0001	1 / 2.0	6/0.1	6 / 0.001	1 / 1.0	8/0.001	6/0.01	8/0.1
FCC4	10 / 0.1	1 / 2.0	1/1.0	1 / 2.0	1 / 2.0	6/1.0	1/0.1	1/0.1
C2	1 / 0.001	2 / 0.001	1/1.0	2 / 2.0	1 / 0.1	1/0.0001	8/1.0	2/0.001
C4	1 / 0.0001	8 / 0.0001	1/0.001	6 / 0.01	6 / 0.1	1/1.0	1/0.1	6/0.01
C6	8 / 0.001	8 / 0.01	8/0.01	4 / 0.1	6 / 0.1	8/0.1	4/0.1	2/0.0001
CP2	2 / 0.01	4 / 0.001	1/2.0	6 / 0.001	2 / 0.1	6/0.01	2/0.01	1/2.0
CP4	1 / 0.1	6 / 0.001	1/1.0	8 / 0.0001	1 / 0.0001	1/0.0001	4/0.001	1/1.0

Table 11.: The optimal hyperparameters for *Problem I*, using the data from all subjects.

Part III

DISCUSSION, CONCLUSION, AND IMPLICATIONS

9

DISCUSSION

As argued in Chapter 4, the use of notch filters or band-pass filters will deform the data, illustrated through the use of two sine waves. After learning that the band-pass filter was not able to attain a 5 Hz sine wave contaminated by 50 Hz noise, it was decided that EMD alone should be applied to remove the electrical line noise identified in the EEG data. The mean frequencies in the first two IMFs were usually 150 Hz and 50 Hz, respectively. By excluding these IMFs from the rest of the EEG data, all occurrences of the line noise were removed, as revealed in Figure 33.

Even though an artifact correction tool using independent component analysis (ICA) was implemented, it was not utilized in any of the analysis methods. Components containing eye blinks and other artifacts were recognized when applying ICA, but the components did, however, contain non-artifact data as well. Additionally, in order to detect artifacts, ICA was dependent on the data being low-pass filtered, preferably below 10 Hz. Consequently, it was chosen *not* to apply ICA, since the artifact component could also hold valuable content. With this in mind, the automatic trial rejection tool was implemented. If a trial was particularly contaminated by artifacts, the whole trial was discarded. This was done to ensure that no unacceptable trials were employed as support vectors in the classification scheme. As mentioned in Chapter 8, the average percentage of accepted trials was 91.4%, meaning that when the number of trials was 70, approximately *six* trials would be discarded.

As described in Section 5.3, a spatial filtering process has been implemented, using the surface Laplacian. When applied, it can enhance the localized brain activity, resulting in an increased spatial resolution. The Laplacian spatial filter process has, however, *not* been utilized in any of the EEG analysis performed in this research. Any filtering techniques have deliberately been avoided, except for the EMD method. As explained in Section 4.4, EMD extracts the energy associated with the different oscillatory modes of a signal, resulting in IMFs that can incorporate meaningful content. Therefore, EMD should be able to extract informational data from a single

electrode, without having to use any form of spatial filtering method. By studying the results presented in Chapter 8, it is clear that EMD is able to acquire the motor-related EEG activity, from which the Hilbert transform can extract a meaningful instantaneous frequency. Moreover, as reported in Section 8.1, the different frequency domains of the IMFs are approximately divided into the internationally recognized brain rhythms, that is, *theta, alpha (mu), beta,* and *gamma*. Hence, the use of EMD is once again justified.

The findings of Pfurtscheller and Lopes da Silva in [43] was the main reason for choosing the ERDs as features. Induced desynchronized activities when executing the fist clenching were found in every subject, in either one or several of the four IMFs, depending on the subject. One occurrence illustrated in Figure 35 shows that the normalized power in the *first* IMF, recorded from the Cz electrode, is experiencing a clear suppression from $t \simeq 700$ ms to $t \simeq 3400$ ms. This is even more evident in the ERD plot in Figure 37. The stimulus onset was at $t = 500$ ms, and the participant's task was to clench the left fist for as long as the visual stimulus was presented (three seconds). The average reaction time from stimulus onset to movement execution can be interpreted as the 200 ms period before the drop in the average ERD, that is, from $t = 500$ ms to $t = 700$ ms. In the same figure, the mean signal power 500 ms prior to stimlus onset is employed as the reference, and the signal power during the three seconds of movement is significantly lower compared to the reference power. Because of the repetetive clenching, some synchronized activities are found during the movement period, for instance the peak at $t \simeq 2250$ ms. This could indicate the time in between the movements. Moreover, induced synchronized activities were observed following the end of the fist clenching, indicating that the participant had returned to a state of physical rest.

The fact that desynchronized activities were found not only when averaging multiple trials, but also during single trials, was a valid indication that ERDs would produce desirable features for detecting and differentiating hand movements. Figure 47 and Figure 49 show that, for a single left hand movement, desynchronized activity occurs in the *gamma* rhythm. Even though desynchronizations were observed in different IMFs, the most prominent desynchronized activities were found in this frequency range. This shows that EMD can be used for extracting the features within all the different brain rhythms, without the need for any band-pass filtering methods. Because of EMD's adaptive property, there is no reason to divide a signal into some predefined frequency regions. On the contrary, EMD will only produce IMFs with a particular frequency-specific information if such content exists in the signal. This means that EMD can decompose any signal, without the need for any prior knowledge regarding the temporal or spectral content of the signal.

As presented in the previous chapter, the scores obtained from each of the three classification problems were approximately **90%**. Moreover, the average number of predictions was 20, meaning that *18 out of 20* trials were correctly predicted. The IMFs used when calculating the features were dependent on the subject. However, when considering the best scores from all subjects, for all of the classification problems, the most prominent IMF was the *first* one, occurring as the optimal IMF eight times. In comparison, each of the three remaining IMFs appeared as the optimal IMF six times. This indicates that when detecting and differentiating left and right fist clenching, the brain activity in the high-beta and low-gamma frequency bands will experience the most significant change in ERD during a motor-related event. Still, clear changes in ERD are observed in the other brain rhythms, that is, *theta* and *mu*, found in IMFs number two, three, and four.

The Hilbert spectra in Figure 39 and Figure 40 demonstrated a distinct difference in the normalized power at the CP3 electrode when comparing right hand movements to the neutral condition, using the data from *Subject 2*. This characteristic suppression of the energy, indicating induced desynchronized activity in the first IMF, resulted in the classification accuracy presented in Table 4, at electrode CP3. The accuracy shows 82.61%, which corresponds to 19 correct predictions out of 23. As seen in the table, the baseline and ERD durations that provided this result were 500 ms and 3000 ms, respectively. The fact that the stimulus onset was at 500 ms, and the total duration of the right hand movement was 3000 ms, explains why these exact values would yield such a high classification accuracy.

As shown in Table 8, the optimal electrode was C5 for *Problem I*, with an average accuracy of 83.22%, which corresponds to approximately 20 out of 24 correctly predicted trials. This result is only slightly higher than for the electrodes Cz, FC3, and FC4, meaning that no significant effect between the specific electrode location and the best classification features can be reported, that is, all 16 electrodes show acceptable classification results. Similarly, for the two remaining problems, the classification scores from all 16 electrodes are not significantly different, but each electrode yields a promising and robust accuracy.

When considering real-time applications for detecting and differentiating hand movements, the process time is of high importance. Using the procedural approach outlined in Chapter 7, illustrated in Figure 31, the average process time for *one* trial (6 seconds of EEG data) recorded at *one* electrode, was found to be 1003 ms (SD=6.2 ms). This indicates the possibility of using this platform in a real-time application, because of the low computation time compared to the duration of the trial.

The use of an SVM in trying to differentiate between right and left hand movements can be justified by the fact that it has several advantages when applied in BCI research, giving satisfactory results [35]. As stated in [11], the maximization of margins and the regularization property provides proper generalization properties. Further, the SVM is not very sensitive to overtraining and to the curse-of-dimensionality, which often occurs when training classifiers in BCI experiments. The kernel function is used to map the data to another space of higher dimensionality. This will, in addition to slightly increasing the complexity of the classifier, allow for the expression of nonlinear boundaries. As a result, the classifier output may experience increased performance when applied to nonlinear EEG data. Moreover, the fact that the SVM can be optimized using a grid search by defining the regularization parameter C and the kernel parameter γ, provide the advantages of getting higher classification output. However, these benefits are gained at the expense of an increased computational time. Several different SVM kernels have been tested, including *linear*, *sigmoid*, *polynomial*, and *rbf*, where the rbf kernel provided the most accurate and robust results. The selected kernel, along with the optimized properties of the SVM, are reflected in the results, yielding the optimal classifier model used for the three classification problems. Furthermore, as discussed above, all of the 16 electrodes show acceptable classification accuracy, indicating that the SVM has proven to be stable and robust for different electrode locations using the ERDs as features.

To quantify the predictive abilities of the classifier used in this thesis, the area under the receiver operating curve was used. The ROC area, which varies between 0 and 1, provides a measure of the probability in the predicted binary classifications. As perceived in Table 10, the separability of *Condition 1* and *Condition 2* was prominent in both the C1 and C2 electrodes for *Subject 5*, reflected in the exceptional classification accuracy of 95%. Moreover, a similar evaluation of the classification performance was observed at the electrode C5, for *Subject 1* and *Subject 6*, which also gave high classification scores. From the ROC evaluation, it seems evident that the chosen procedure is applicable for detecting and differentiating motor-related events. Since the area under the ROC has provided continuous evaluation of the classifier output, without heavy computations, the ROC-based evaluation could be suitable for BCI applications, yielding a real-time assessment of the performance.

As observed in Table 11, a typical value of the regularization parameter C is 1, meaning suppressed penalizations of the classification errors in the training set. This indicates that the selected features are useful for preserving class separability, resulting in a classifier that can accommodate outliers and obtain better generalization properties. Even though outliers are common in EEG data, the regularized classifier used in this thesis has provided valuable classification results, which conforms with the research in [9] and [38]. Although different combinations of the hyperparameter pair C and γ were observed, the most frequently occurring parameters were $C = 1$ and $\gamma = 0.1$, yielding a reduced penalty of the errors in the training set and a moderate width in the kernel. Still, the optimal parameter pair for *Subject 5* was $C = 10$ and $\gamma = 0.001$ when obtaining the 95% classification accuracy, indicating the fact that the classification is very subject-specific and that different parameters must be optimized to yield the best results.

Adjustments of the baseline and ERD durations were optimized similar to a grid search, to acquire desirable features from single trials. This method can be regarded as a regularized feature selection technique, meaning that poor classification scores will be reconsidered with a new set of parameters defining the ERD onset and duration. However, if the optimization search did not yield acceptable results, it could indicate that the two conditions are not separable. One possible reason for this may be that the support vectors contain outliers and, hence, corrupting the trained set. Nonetheless, the regularized feature selection technique implemented in this thesis provides the ability for researchers to improve the feature selection in other classification problems.

10

CONCLUSION

The entire platform implemented in this thesis contains several modules combining EEG data analysis and machine learning. The steps include filtering of electrical line noise, automatic trial rejection, adaptive time-frequency analysis, and classification of motor-related events. A valuable library, *Pysa*, has been implemented. The most important functionality in the library is the empirical mode decomposition (EMD) utilized in the normalized Hilbert transform (NHT). The unified feature extraction and classification scheme, in addition to the Pysa library, may be useful for neuroscientists, both beginners, and experienced ones.

Some important advantages of the EMD have been outlined, in which the most important feature concerns the adaptiveness of the method. There is no need for any prior knowledge regarding the signal, neither temporal nor spectral, meaning that just about any signal can be decomposed by EMD, resulting in IMFs that hold meaningful content. By deliberately avoiding the use of any filtering techniques other than EMD, the researchers have ensured that the EEG signals would not experience any unwanted alteration or deformation. Disturbance from the mains electricity is often related to EEG acquisition. By using EMD as a dyadic filter, the electrical line noise can be completely separated from the brain-related EEG signals. Furthermore, the frequency spectra of the intrinsic mode functions (IMFs), produced by applying EMD to the filtered EEG data, are comparable to the internationally recognized brain rhythms. The energy associated with the different oscillatory modes of the EEG signal will be incorporated in each of the IMFs, which makes EMD a feasible method for the extraction of features related to hand movements. By utilizing the standard equation for event-related desynchronization (ERD), the average ERD values for a given number of trials are calculated and employed as support vectors. Moreover, the features of the remaining trials are predicted using an RBF support vector machine (SVM) classifier, yielding a classification accuracy of up to 95% when differentiating between left and right hand movements.

The normalized Hilbert transform (NHT) combines primarily three methods, namely the EMD, a normalization scheme, and the Hilbert transform. NHT has proven itself as a reliable and accurate resource when analyzing the time-frequency solutions of nonlinear data, such as the electrical activity of the brain. The instantaneous frequencies and amplitudes extracted from the normalized IMFs can be represented in a Hilbert spectrum (HS). The cortical activity related to a hand movement can, when presented in a HS, clearly be distinguished from the activity recorded from a subject being physically at rest.

When considering actuation purposes, for instance the control of an electric wheelchair using EEG, minimizing the time delay from control signal to actuator is a high priority. The computational effort with respect to the proposed procedure is concluded by off-line simulations, and found to be within the real-time demands of a brain-computer interface (BCI).

11

IMPLICATIONS

The findings in this research have motivated the researchers to work further on applying the procedure in real-time BCI applications. The proposed procedure could, for instance, be used for actuation purposes, controlling quadcopters or perhaps an advanced prosthetic hand. Hopefully it is a pathway of reading brain activity that relates the cognitive activation of muscular movement. The fact that no prior knowledge regarding the signal is required when applying EMD, and that the resulting IMFs consolidate meaningful information of the underlying physics of a process, provides the ability to let the electrical activities of the brain communicate to us, and not the other way around. This path seems to be promising concerning the different aspects of analyzing brain activity.

The proposed procedure of identifying motor-related activities have inspired the researchers to follow the recent research of Dr. Norden Huang, in which the Holo-Hilbert spectral analysis (HHSA) has been developed [29]. HHSA utilizes a nested empirical mode decomposition to extract intrinsic frequency and amplitude modulations represented in nonlinear systems, both inter and intra-wave. This means that additional dimensions in the spectrum can be analyzed to account for the variations within the frequency and the amplitude modulations, respectively. Furthermore, for real-time BCI application purposes, the EMD sifting procedure could perhaps be optimized in the sense of computational effort, and concurrent programming can be introduced to process heavy tasks in parallel.

The platform implemented in this thesis may be applied as a toolbox for analyzing brain activities, for which other interested researchers can exploit the rhythms of the brain, and hopefully, utilize the offered methods in other real-time BCI purposes. The idea behind letting an EEG signal speak for itself, not looking for predefined rhythms of interest by applying signal processing methods, is a concept that provides close relations to how our cognitive abilities works.

BIBLIOGRAPHY

[1] *10/20 System Position*. https://www.trans-cranial.com/local/manuals/10_20_pos_man_v1_0_pdf.pdf. Accessed: 2016-04-26.

[2] Norden E. Huang et al. "A confidence limit for the empirical mode decomposition and Hilbert spectral analysis". In: *The Royal Society* 459 (2003), 2317–2345.

[3] Kai Keng Ang et al. "Clinical study of neurorehabilitation in stroke using EEG-based motor imagery brain-computer interface with robotic feedback". In: *32nd Annual International Conference of the IEEE EMBS* (2010), pp. 5549–5552.

[4] Sylvain Arlot and Alain Celisse. "A survey of cross-validation procedures for model selection". In: *Statistics surveys* 4 (2010), pp. 40–79.

[5] G. Barbati et al. "Optimization of an independent component analysis approach for artifact identification and removal in magnetoencephalographic signals". In: *Clinical Neurophysiology* 115 (2004), 1220–1232.

[6] Kristin P. Bennett and Colin Campbell. "Support Vector Machines: Hype or Hallelujah?" In: *Explor. Newslett.* 2 (2000), pp. 1–13.

[7] Hans Berger. "Ueber das Elektroenkephalogramm des Menschen". In: *Arch Psychiatr Nervenkrankh* 87 (1929), pp. 527–570.

[8] Hans Berger. "Ueber das Elektroenkephalogramm des Menschen II". In: *Psychol Neurol* 40 (1930), pp. 160–179.

[9] Blankertz et al. "Classifying single trial EEG: Towards brain computer interfacing". In: *Advances in neural information processing systems* 1 (2002), pp. 157–164.

[10] Leo Braiman. "Theory of Communication". In: *The Annals of Statistics* 26.3 (1998), pp. 801–849.

[11] Christopher J.C. Burgers. "A Tutorial on Support Vector Machines for Pattern Recognition". In: *Data Mining and Knowledge Discovery* 2 (1998), 121–167.

[12] Claudio Carvalhaes and J. Acacio de Barros. "The surface Laplacian technique in EEG: Theory and methods". In: *International Journal of Psychophysiology* 97.3 (2015), pp. 174–188.

[13] Duo Chen, Suiren Wan, and Forrest Sheng Bao. “EEG-based seizure detection using discrete wavelet transform through full-level decomposition”. In: *Bioinformatics and Biomedicine, IEEE International Conference* (2015), pp. 1596–1602.

[14] Louis Lemieux Christoph Mulert, ed. *EEG-fMRI Physiological Basis, Technique and Applications*. Springer Heidelberg Dordrecht London New York, 2010. Chap. 2. ISBN: 9783540879183.

[15] Jesse Davis and Mark Goadrich. “The relationship between Precision-Recall and ROC curves”. In: *Proceedings of the 23rd international conference on Machine learning* (2006), pp. 233–240.

[16] Lawrence T. DeCarlo. “On the Meaning and Use of Kurtosis”. In: *Psychological Methods* 2.3 (1997), pp. 292–307.

[17] Arnaud Delorme, Terrence Sejnowski, and Scott Makeig. “Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis”. In: *NeuroImage* 34 (4 2007), 1443–1449.

[18] *Divisions of the cerebral cortex, functions areas of the brain*. `https://classconnection.s3.amazonaws.com/67582/flashcards/670169/png/divisions-of-the-cerebral-cortex,-functions-areas-of-the-brain.png`. Accessed: 2016-04-25.

[19] *Electric dipole field lines*. `https://commons.wikimedia.org/wiki/File:Electric_dipole_field_lines.svg`. Accessed: 2016-04-26.

[20] Norden E. Huang et.al. “The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis”. In: *The Royal Society* 454 (1998), pp. 903–995.

[21] Patrick Flandrin, Babriel Rilling, and Paulo Gonçalvés. “Empirical Mode Decomposition as a Filter Bank”. In: *IEEE Signal Processing Letters* 11.2 (2004).

[22] Walter J. Freeman and Bob W. van Dijk. “Spatial patterns of visual cortical fast EEG during conditioned reflex in a rhesus monkey”. In: *Brain Research* 422 (1987), pp. 267–276.

[23] Dennis Gabor. “Theory of Communication”. In: *J. IEEE* 93 (1946), pp. 429–457.

[24] Alan Gevins et al. “High-resolution EEG Mapping of Cortical Activation Related to Working Memory: Effects of Task Difficulty, Type of Processing, and Practive”. In: *Cereb. Cortex* 7 (1997), pp. 374–385.

[25] James A. Hanley and Barbara J. McNeil. “The Meaning and Use of the Area under a Receiver Operating Characteristic (ROC) Curve”. In: *Radiology* 143 (1982), pp. 29–36.

[26] B. Hjorth. “An on-line transformation of EEG scalp potentials into orthogonal source derivations”. In: *Electroencephalogr. Clin. Neurophysiol.* 39 (1975), pp. 526–530.

[27] Norden E. Huang and Steven R. Long. “Normalized Hilbert transform and instantaneous frequency”. In: *NASA Patent Pending GSC 14 673-1* (2003).

[28] Norden E. Huang and Samuel S. P. Shen, eds. *Hilbert-Huang Transform and Its Applications*. Second Edition. World Scientific Publishing Co. Pte. Ltd., 2014. ISBN: 9789814508230.

[29] Norden E. Huang et al. “On Holo-Hilbert spectral analysis: a full informational spectral representation for nonlinear and non-stationary data”. In: *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 374.2065 (2016).

[30] Norden E. Huang et al. “On instantaneous frequency”. In: *Advances in Adaptive Data Analysis* 1.2 (2009), pp. 177–229.

[31] *HydroCel Geodesic Sensor Net 256-Channel Map*. `https://www.egi.com`. Accessed: 2016-05-22.

[32] Croft R J et al. “EOG correction: a comparison of four methods”. In: *Psychophysiology* 42 (2005), 16–24.

[33] Anil K. Jain, Ropert P. W. Duin, and Jianchang Mao. “Statistical Pattern Recognition: A Review”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22.1 (2000), pp. 4–37.

[34] Valer Jurcak, Daisuke Tsuzuki, and Ippeita Dan. “10/20, 10/10, and 10/5 systems revisited: Their validity as relative head-surface-based positioning systems”. In: *NeuroImage* 34 (2007), pp. 1600–1611.

[35] M. Kaper et al. “BCI competition 2003-data set IIb: support vector machines for the P300 speller paradigm”. In: *Biomedical Engineering, IEEE Transactions* 51 (2004), pp. 1073–1076.

[36] Ron Kohavi. “A study of cross-validation and bootstrap for accuracy estimation and model selection”. In: *International Joint Conference on Artificial Intelligence* 14.2 (1995), pp. 1137–1145.

[37] Fabien Lotte et al. “A review of classification algorithms for EEG-based brain–computer interfaces”. In: *Journal of Neural Engineering* 4.2 (2007), R1.

[38] Klaus-Robert Müller et al. “Machine learning techniques for brain-computer interfaces”. In: *Biomed. Tech* 49.1 (2004), pp. 11–22.

[39] C. Nicholson. “Theoretical analysis of field potentials in anisotropic ensembles of neuronal elements”. In: *IEEE Trans. Biomed. Eng.* 20 (1973), pp. 278–288.

[40] H. Nolan, R. Whelan, and R.B. Reilly. “FASTER: Fully Automated Statistical Thresholding for EEG artifact Rejection”. In: *Journal of Neuroscience Methods* 192 (2010), pp. 152–162.

[41] Paul L. Nunez and Ramesh Srinivasan. *Electric Fields of the Brain: The Neurophysis of EEG*. 2nd ed. Oxford University Press, Inc., 2006. ISBN: 9780195050387.

[42] Rafael Lorente De Nó. “Action potential of the motoneurons of the hypoglossus nucleus”. In: *Journal of Cellular and Comparative Physiology* 29.3 (1947), pp. 207–287.

[43] G. Pfurtscheller and F.H. Lopes da Silva. “Event-related EEG/MEG synchronization and desynchronization: basic principles”. In: *Clinical Neurophysiology* 110 (1999), pp. 1842–1857.

[44] Kemal Polat and Salih Günes. “Classification of epileptiform EEG using a hybrid system based on decision tree classifier and fast Fourier transform”. In: *Applied Mathematics and Computation* 187.2 (2007), pp. 1017–1026.

[45] S.J. Raudys and A.K. Jain. “Small Sample Size Effects in Statistical Pattern Recognition: Recommendations for Practitioners”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13.3 (1991), pp. 252–264.

[46] Farhan Riaz et al. “Detecting and classifying movement-related cortical potentials associated with hand movements in healthy subjects and stroke patients from single-electrode, single-trial EEG”. In: *Journal of Neural Engineering* 12.5 (2015).

[47] Farhan Riaz et al. “EMD-Based Temporal and Spectral Features for the Classification of EEG Signals Using Supervised Learning”. In: *IEEE Transactions on neural systems and rehabilation engineering* 24.1 (2016), pp. 28–35.

[48] Y. Dan Rubinstein and Trevor Hastie. “Discriminative vs Informative Learning”. In: *KDD-97 Proceedings* (1997).

[49] Antti Savelainen. “An introduction to EEG artifacts”. Independent research projects in applied mathematics. 2010.

[50] F.H. Lopes da Silva et al. “Dynamic characteristics of visual evoked potentials in the dog. II. Beta frequency selectivity in evoked potentials and background activity”. In: *Electroencephalography and Clinical Neurophysiology* 29.3 (1970), pp. 260–268.

[51] Leif Sornmo and Pablo Laguna. *Bioelectrical Signal Processing in Cardiac and Neurological Applications*. Revised. Academic Press, 2005. ISBN: 0080527922, 9780080527925.

[52] Pham T T et al. “A test of four EOG correction methods using an improved validation technique”. In: *Int. J. Psychophysiol.* 79 (2011), 203–210.

[53] Andrew R. Webb. *Statistical Pattern Recognition*. Second Edition. John Wiley & Sons Ltd., 2002. ISBN: 9780198520115.

[54] Frank Westad and Martin Kermit. “Cross validation and uncertainty estimates in independant component analysis”. In: *Analytica Chimica Acta* 490 (2003), pp. 341–354.

[55] Diane J. Whitmer. “Interpretation of mammalian brain rhythms of sensorimotor processing”. PhD thesis. UC San Diego Electronic Theses and Dissertations, 2008.

[56] C.H. Wolters et al. “Influence of tissue conductivity anisotropy on EEG/MEG field and return current computation in a realistic head model: A simulation and visualization study using high-resolution finite element modeling”. In: *NeuroImage* 30 (2006), pp. 813–826.

A

EVENT-RELATED (DE)SYNCHRONIZATION

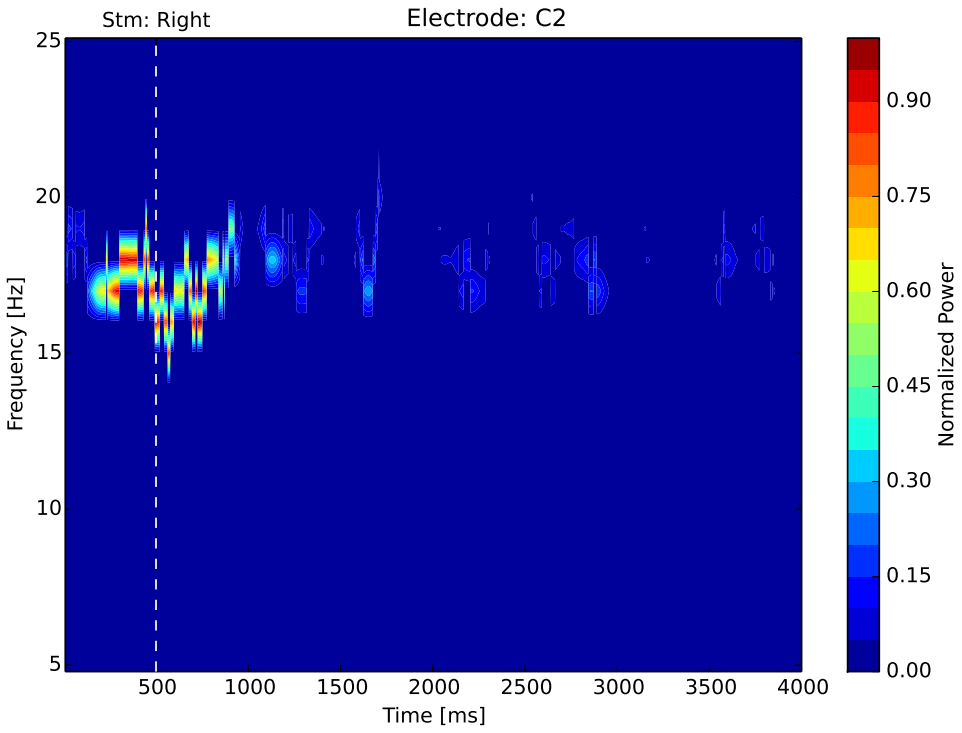

Figure 55.: Normalized power in the second IMF averaged over 80 right hand movements, recorded from *Subject 7*.

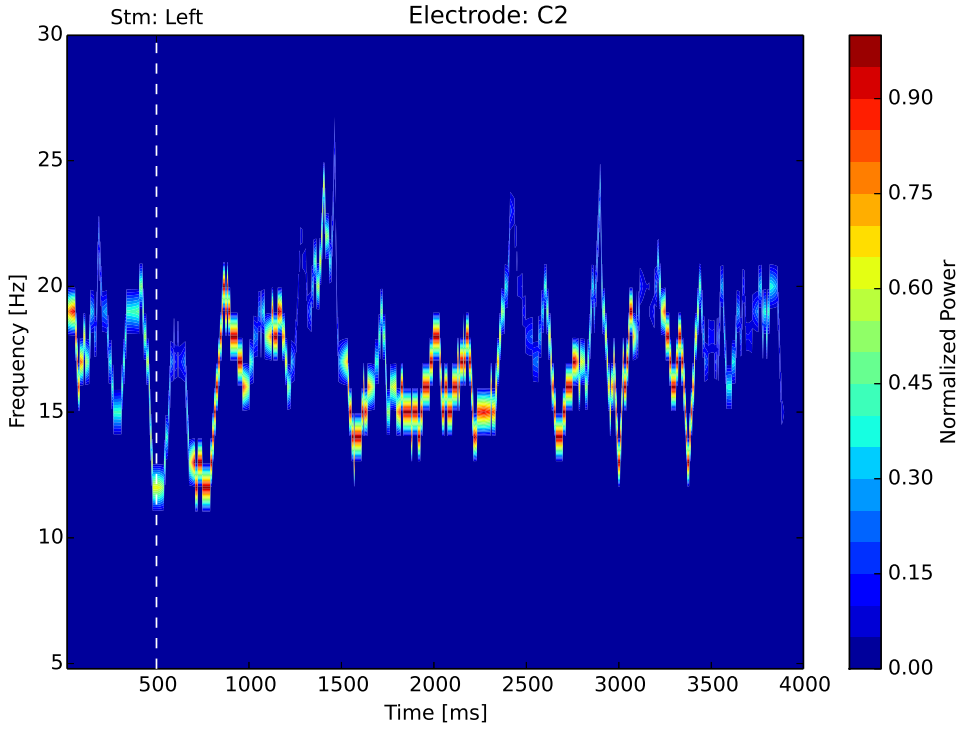

Figure 56.: Normalized power in the second IMF averaged over 79 left hand movements, recorded from *Subject 7*.

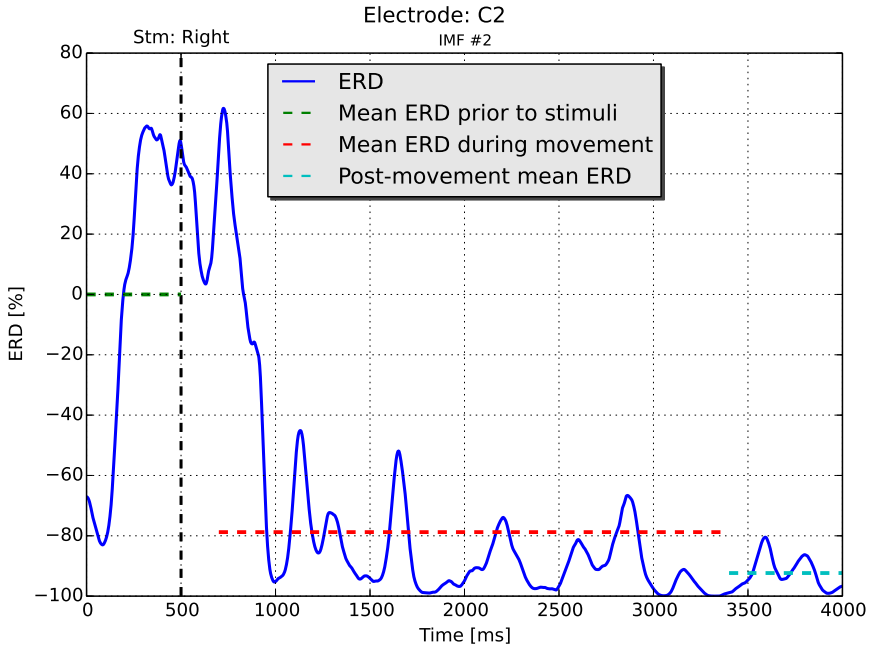

Figure 57.: Average ERD calculated using the second IMF from 80 right hand movements, recorded from *Subject 7*.

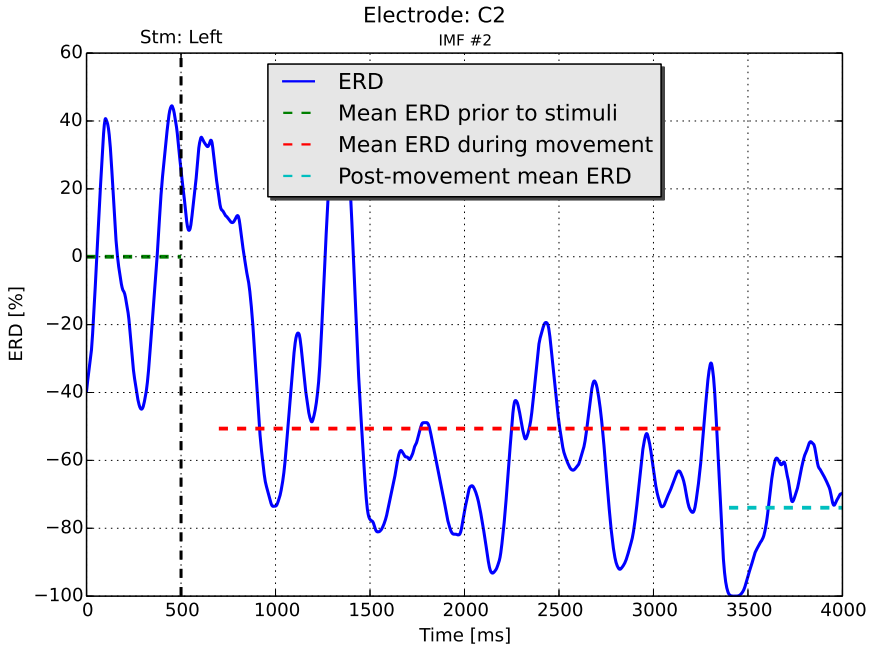

Figure 58.: Average ERD calculated using the second IMF from 79 left hand movements, recorded from *Subject 7*.

B

CLASSIFICATION PERFORMANCE AND EVALUATION

Classification: Problem I

	Parameters						Scores	
Electrode	**Accepted trials left**	**Accepted trials right**	**Ratio trained**	**IMF number**	**Baseline [ms]**	**ERD duration [ms]**	**Accuracy [%]**	**Correct Predictions**
Cz	70	70	0.86	1	400	4000	**85**	17 / 20
FCCz	60	65	0.84	2	500	3400	**76.19**	16 / 21
FC3	70	70	0.86	2	600	3200	**80**	16 / 20
FCC3	70	70	0.86	3	400	2300	**85**	17 / 20
C1	70	70	0.86	1	1100	2300	**75**	15 / 20
C3	69	70	0.86	2	600	3500	**80**	16 / 20
C5	68	67	0.86	1	600	3400	**80**	16 / 20
CP1	60	64	0.84	4	300	2200	**80.95**	17 / 21
CP3	68	68	0.86	4	900	2400	**80**	16 / 20
FC4	66	67	0.86	1	400	2800	**75**	15 / 20
FCC4	70	70	0.86	1	300	2300	**80**	16 / 20
C2	70	70	0.86	1	500	2200	**80**	16 / 20
C4	67	69	0.86	2	600	2300	**80**	16 / 20
C6	53	65	0.84	1	400	3400	**75**	15 / 20
CP2	59	66	0.84	1	800	3800	**80.95**	17 / 21
CP4	54	63	0.84	2	500	3200	**85**	17 / 20

Table 12.: Classification results for *Subject 3*, using right and left hand movements as conditions. The electrodes with the highest accuracy are marked with green.

Classification: Problem I

	Parameters						Scores	
Electrode	**Accepted trials left**	**Accepted trials right**	**Ratio trained**	**IMF number**	**Baseline [ms]**	**ERD duration [ms]**	**Accuracy [%]**	**Correct Predictions**
Cz	70	70	0.86	3	500	2200	**80**	16 / 20
FCCz	34	29	0.68	1	350	2300	**80.95**	17 / 21
FC3	62	54	0.83	3	300	2200	**76.19**	16 / 21
FCC3	68	69	0.85	1	500	3500	**72.73**	16 / 22
C1	70	70	0.85	4	400	2400	**86.36**	19 / 22
C3	67	70	0.85	4	800	4000	**81.82**	18 / 22
C5	63	66	0.85	1	400	3400	**85**	17 / 20
CP1	69	70	0.86	4	400	2300	**80**	16 / 20
CP3	67	69	0.86	4	400	2800	**80**	16 / 20
FC4	70	69	0.86	4	700	3100	**85**	17 / 20
FCC4	70	70	0.86	1	300	2300	**75**	15 / 20
C2	70	70	0.86	1	600	2600	**80**	16 / 20
C4	70	70	0.86	4	350	2400	**80**	16 / 20
C6	70	67	0.86	1	350	3200	**80**	16 / 20
CP2	70	70	0.86	3	900	3400	**85**	17 / 20
CP4	68	70	0.86	4	300	4000	**80**	16 / 20

Table 13.: Classification results for *Subject 4*, using right and left hand movements as conditions. The electrode with the highest accuracy is marked with green.

Classification: Problem I

	Parameters						Scores	
Electrode	**Accepted trials left**	**Accepted trials right**	**Ratio trained**	**IMF number**	**Baseline [ms]**	**ERD duration [ms]**	**Accuracy [%]**	**Correct Predictions**
Cz	80	80	0.84	4	300	3200	**80.77**	21 / 26
FCCz	79	79	0.84	4	500	3400	**76.92**	20 / 26
FC3	48	50	0.8	4	400	3500	**80**	16 / 20
FCC3	67	64	0.8	3	900	2300	**77.78**	21 / 27
C1	80	80	0.86	4	400	3800	**79.17**	20 / 24
C3	77	72	0.86	1	600	2200	**77.27**	17 / 22
C5	30	28	0.7	4	900	2400	**88.89**	16 / 18
CP1	80	80	0.86	2	900	2300	**79.17**	19 / 24
CP3	79	73	0.86	4	1100	2500	**82.61**	19 / 23
FC4	57	58	0.84	4	1100	2800	**90**	18 / 20
FCC4	79	77	0.86	1	300	3500	**82.61**	19 / 23
C2	80	80	0.84	4	500	2200	**76.92**	20 / 26
C4	78	71	0.84	1	500	2600	**72**	18 / 25
C6	40	43	0.8	1	600	3100	**82.35**	14 / 17
CP2	80	80	0.84	2	300	3400	**80.77**	21 / 26
CP4	75	76	0.84	2	400	2400	**76**	19 / 25

Table 14.: Classification results for *Subject 6*, using right and left hand movements as conditions. The electrode with the highest accuracy is marked with green.

Classification: Problem I

	Parameters						Scores	
Electrode	**Accepted trials left**	**Accepted trials right**	**Ratio trained**	**IMF number**	**Baseline [ms]**	**ERD duration [ms]**	**Accuracy [%]**	**Correct Predictions**
Cz	80	79	0.86	3	500	2400	**83.33**	20 / 24
FCCz	79	78	0.86	2	1100	2400	**78.26**	18 / 23
FC3	80	72	0.86	3	400	2300	**78.26**	18 / 23
FCC3	78	77	0.86	1	500	3600	**81.82**	18 / 22
C1	79	79	0.86	4	500	2600	**79.17**	19 / 24
C3	80	79	0.86	4	500	3100	**79.17**	19 / 24
C5	78	79	0.86	3	700	4000	**78.26**	18 / 23
CP1	80	78	0.86	4	600	2800	**78.26**	18 / 23
CP3	79	76	0.86	1	350	2800	**78.26**	18 / 23
FC4	79	79	0.86	3	700	3500	**79.17**	19 / 24
FCC4	80	79	0.86	3	1000	3100	**79.17**	19 / 24
C2	80	79	0.86	2	1100	4000	**83.33**	20 / 24
C4	80	79	0.86	3	350	2300	**70.83**	17 / 24
C6	79	75	0.86	2	900	2300	**78.26**	18 / 23
CP2	80	79	0.86	3	350	3500	**75**	18 / 24
CP4	79	78	0.86	3	300	2900	**78.26**	18 / 23

Table 15.: Classification results for *Subject 7*, using right and left hand movements as conditions. The electrodes with the highest accuracy are marked with green.

Classification: Problem I

	Parameters						Scores	
Electrode	**Accepted trials left**	**Accepted trials right**	**Ratio trained**	**IMF number**	**Baseline [ms]**	**ERD duration [ms]**	**Accuracy [%]**	**Correct Predictions**
Cz	80	80	0.84	3	600	2800	**76.92**	20 / 26
FCCz	80	80	0.84	1	700	3500	**69.23**	18 / 26
FC3	45	43	0.84	1	1000	2500	**86.67**	13 / 15
FCC3	75	74	0.84	1	400	3400	**87.5**	21 / 24
C1	80	80	0.84	2	1100	2600	**76.92**	20 / 26
C3	77	69	0.84	4	500	3600	**76**	19 / 25
C5	51	41	0.84	2	500	3400	**81.25**	13 / 16
CP1	80	80	0.84	3	1100	3000	**76.92**	20 / 26
CP3	76	73	0.84	2	400	2500	**80**	20 / 25
FC4	69	69	0.84	2	800	3800	**87.5**	21 / 24
FCC4	79	79	0.84	2	800	3800	**80.77**	21 / 26
C2	80	80	0.84	3	1000	2300	**76.92**	20 / 26
C4	80	79	0.84	4	500	4000	**84.62**	22 / 26
C6	68	57	0.84	1	350	2600	**71.43**	15 / 21
CP2	80	80	0.84	3	1100	3400	**73.08**	19 / 26
CP4	79	79	0.84	3	600	2400	**69.23**	18 / 26

Table 16.: Classification results for *Subject 8*, using right and left hand movements as conditions. The electrodes with the highest accuracy are marked with green.

Classification: Problem II

	Parameters						Scores	
Electrode	**Accepted trials right**	**Accepted trials neutral**	**Ratio trained**	**IMF number**	**Baseline [ms]**	**ERD duration [ms]**	**Accuracy [%]**	**Correct Predictions**
Cz	70	70	0.86	3	500	2600	**85**	17 / 20
FCCz	66	67	0.86	3	400	2200	**75**	15 / 20
FC3	59	59	0.84	3	500	3400	**85**	17 / 20
FCC3	60	59	0.84	3	350	2400	**75**	15 / 20
C1	69	69	0.86	4	800	2800	**80**	16 / 20
C3	54	57	0.81	1	700	3100	**77.27**	17 / 22
C5	26	41	0.7	3	300	2500	**76.19**	16 / 21
CP1	58	62	0.84	1	300	3600	**75**	15 / 20
CP3	48	48	0.8	2	500	3600	**75**	15 / 20
FC4	18	25	0.6	4	1000	3100	**83.33**	15 / 18
FCC4	51	46	0.8	1	800	2500	**76.19**	16 / 21
C2	63	63	0.84	2	700	3100	**81.82**	18 / 22
C4	59	55	0.83	3	1000	3400	**85.71**	18 / 21
C6	53	48	0.81	2	400	3500	**80.95**	17 / 21
CP2	61	61	0.82	3	400	2400	**81.82**	18 / 22
CP4	41	44	0.78	2	350	3500	**80**	16 / 20

Table 17.: Classification results for *Subject 1*, using right hand movements and neutral trials as conditions. The electrode with the highest accuracy is marked with green.

Classification: Problem II

	Parameters						Scores	
Electrode	**Accepted trials right**	**Accepted trials neutral**	**Ratio trained**	**IMF number**	**Baseline [ms]**	**ERD duration [ms]**	**Accuracy [%]**	**Correct Predictions**
Cz	70	70	0.86	4	900	3500	**80**	16 / 20
FCCz	65	56	0.84	1	300	3200	**85**	17 / 20
FC3	70	70	0.86	4	700	4000	**80**	16 / 20
FCC3	70	69	0.86	2	900	3400	**95**	19 / 20
C1	70	70	0.86	1	900	2400	**75**	15 / 20
C3	70	69	0.86	2	500	2900	**80**	16 / 20
C5	67	60	0.86	1	400	2500	**84.21**	16 / 19
CP1	64	53	0.84	4	1100	4000	**80**	16 / 20
CP3	68	59	0.86	4	1000	2400	**84.21**	16 / 19
FC4	67	63	0.86	4	700	3100	**89.47**	17 / 19
FCC4	70	70	0.86	2	400	3600	**85**	17 / 20
C2	70	70	0.86	3	900	3600	**80**	16 / 20
C4	69	68	0.86	4	600	3600	**85**	17 / 20
C6	65	54	0.84	4	600	4000	**85**	17 / 20
CP2	66	57	0.84	3	500	3000	**80.95**	17 / 21
CP4	63	45	0.84	2	350	3000	**84.21**	16 / 19

Table 18.: Classification results for *Subject 3*, using right hand movements and neutral trials as conditions. The electrode with the highest accuracy is marked with green.

Classification: Problem II

	Parameters						Scores	
Electrode	**Accepted trials right**	**Accepted trials neutral**	**Ratio trained**	**IMF number**	**Baseline [ms]**	**ERD duration [ms]**	**Accuracy [%]**	**Correct Predictions**
Cz	70	70	0.86	3	500	2400	**80**	16 / 20
FCCz	29	61	0.8	1	300	2800	**89.47**	17 / 19
FC3	54	66	0.84	4	700	2400	**80**	16 / 20
FCC3	69	66	0.86	4	800	4000	**80**	16 / 20
C1	70	69	0.86	1	500	2800	**85**	17 / 20
C3	70	65	0.86	1	500	3600	**75**	15 / 20
C5	66	63	0.84	1	700	2600	**77.27**	17 / 22
CP1	70	69	0.84	4	300	3800	**79.17**	19 / 24
CP3	69	65	0.86	3	900	2300	**80**	16 / 20
FC4	69	70	0.86	4	500	3000	**80**	16 / 20
FCC4	70	69	0.86	4	350	2800	**85**	17 / 20
C2	70	70	0.86	1	350	2500	**80**	16 / 20
C4	70	70	0.86	1	600	3200	**75**	15 / 20
C6	67	66	0.86	2	1100	2500	**80**	16 / 20
CP2	70	70	0.86	1	500	2400	**80**	16 / 20
CP4	70	70	0.86	4	300	2500	**85**	17 / 20

Table 19.: Classification results for *Subject 4*, using right hand movements and neutral trials as conditions. The electrode with the highest accuracy is marked with green.

Classification: Problem III

	Parameters						Scores	
Electrode	**Accepted trials left**	**Accepted trials neutral**	**Ratio trained**	**IMF number**	**Baseline [ms]**	**ERD duration [ms]**	**Accuracy [%]**	**Correct Predictions**
Cz	70	70	0.86	4	1100	2400	**80**	16 / 20
FCCz	68	67	0.86	2	400	2200	**85**	17 / 20
FC3	55	59	0.84	1	600	3100	**89.47**	17 / 19
FCC3	54	59	0.84	1	500	2500	**84.21**	16 / 19
C1	69	69	0.86	4	700	3600	**90**	18 / 20
C3	50	57	0.8	2	700	2200	**86.36**	19 / 21
C5	29	41	0.7	2	700	4000	**77.27**	17 / 22
CP1	56	62	0.83	3	800	3400	**85.71**	18 / 21
CP3	42	48	0.8	3	500	3400	**78.95**	15 / 19
FC4	23	25	0.6	3	800	2300	**90**	18 / 20
FCC4	49	46	0.8	4	600	2400	**85**	17 / 20
C2	65	63	0.84	1	1100	2200	**77.27**	17 / 22
C4	56	55	0.82	3	900	3400	**80.95**	17 / 21
C6	52	48	0.8	2	900	2900	**76.19**	16 / 21
CP2	59	61	0.84	1	350	2800	**80**	16 / 20
CP4	44	44	0.8	3	500	3800	**83.33**	15 / 18

Table 20.: Classification results for *Subject 1*, using left hand movements and neutral trials as conditions. The electrodes with the highest accuracy are marked with green.

Classification: Problem III

	Parameters						Scores	
Electrode	**Accepted trials left**	**Accepted trials neutral**	**Ratio trained**	**IMF number**	**Baseline [ms]**	**ERD duration [ms]**	**Accuracy [%]**	**Correct Predictions**
Cz	70	70	0.84	4	600	4000	**79.17**	19 / 24
FCCz	60	56	0.8	1	600	3500	**79.17**	19 / 24
FC3	70	70	0.84	2	1100	4000	**70.83**	17 / 24
FCC3	70	69	0.84	3	1000	3100	**83.33**	20 / 24
C1	70	70	0.84	4	700	3200	**83.33**	20 / 24
C3	69	69	0.84	2	350	2600	**79.17**	19 / 24
C5	68	60	0.84	2	350	2900	**80.95**	17 / 21
CP1	60	53	0.8	4	900	3200	**73.91**	17 / 23
CP3	68	59	0.84	3	1000	3000	**80.95**	17 / 21
FC4	66	63	0.84	4	350	2400	**77.27**	17 / 22
FCC4	70	70	0.84	3	700	4000	**75**	18 / 24
C2	70	70	0.84	3	400	4000	**79.17**	19 / 24
C4	67	68	0.84	3	600	2500	**77.27**	17 / 22
C6	53	54	0.8	3	600	3100	**81.82**	18 / 22
CP2	59	57	0.8	1	600	2200	**87.5**	21 / 24
CP4	54	45	0.76	3	900	3600	**75**	18 / 24

Table 21.: Classification results for *Subject 3*, using left hand movements and neutral trials as conditions. The electrode with the highest accuracy is marked with green.

Classification: Problem III

	Parameters						Scores	
Electrode	**Accepted trials left**	**Accepted trials neutral**	**Ratio trained**	**IMF number**	**Baseline [ms]**	**ERD duration [ms]**	**Accuracy [%]**	**Correct Predictions**
Cz	70	70	0.86	2	900	4000	**80**	16 / 20
FCCz	34	61	0.8	4	700	3600	**80**	16 / 20
FC3	62	66	0.85	4	350	2600	**75**	15 / 20
FCC3	68	66	0.85	2	300	2800	**85.71**	18 / 21
C1	70	69	0.85	2	300	3600	**77.27**	18 / 22
C3	67	65	0.85	1	400	3800	**76.19**	16 / 21
C5	63	63	0.85	4	300	2200	**85**	17 / 20
CP1	69	69	0.86	1	1000	3400	**85**	17 / 20
CP3	67	65	0.86	4	700	2300	**85**	17 / 20
FC4	70	70	0.86	3	400	3400	**85**	17 / 20
FCC4	70	69	0.86	2	800	4000	**75**	15 / 20
C2	70	70	0.86	3	600	2600	**80**	16 / 20
C4	70	70	0.86	2	500	3800	**75**	15 / 20
C6	70	66	0.86	2	1100	4000	**75**	15 / 20
CP2	70	70	0.86	3	900	3200	**80**	16 / 20
CP4	68	70	0.86	1	700	3400	**80**	16 / 20

Table 22.: Classification results for *Subject 4*, using left hand movements and neutral trials as conditions. The electrode with the highest accuracy is marked with green.

ROC area: Problem II

Electrode	Subject 1	Subject 2	Subject 3	Subject 4
Cz	0.85	0.8	0.8	0.8
FCCz	0.75	0.8989	0.85	0.88
FC3	0.85	0.8333	0.8	0.79
FCC3	0.75	0.75	0.95	0.8
C1	0.8	0.9	0.75	0.85
C3	0.77	0.85	0.8	0.75
C5	0.71	0.79	0.84	0.77
CP1	0.75	0.85	0.78	0.79
CP3	0.75	0.83	0.84	0.8
FC4	0.84	0.8	0.89	0.8
FCC4	0.76	0.9	0.85	0.85
C2	0.82	0.82	0.8	0.8
C4	0.85	0.8	0.85	0.75
C6	0.81	0.75	0.85	0.8
CP2	0.82	0.77	0.8	0.8
CP4	0.8	0.72	0.81	0.85

Table 23.: Evaluation of classification performance using ROC area for *Problem II*, using the data from four subjects.

ROC area: Problem III

Electrode	Subject 1	Subject 2	Subject 3	Subject 4
Cz	0.8	0.8	0.79	0.8
FCCz	0.85	0.83	0.79	0.75
FC3	0.89	0.85	0.71	0.75
FCC3	0.84	0.85	0.83	0.86
C1	0.9	0.8	0.83	0.77
C3	0.86	0.85	0.79	0.76
C5	0.79	0.8	0.8	0.85
CP1	0.85	0.9	0.74	0.85
CP3	0.78	0.85	0.81	0.85
FC4	0.9	0.75	0.77	0.85
FCC4	0.85	0.8	0.75	0.75
C2	0.77	0.82	0.79	0.8
C4	0.81	0.85	0.77	0.75
C6	0.76	0.75	0.82	0.75
CP2	0.8	0.89	0.88	0.8
CP4	0.83	0.83	0.74	0.8

Table 24.: Evaluation of classification performance using ROC area for *Problem III*, using the data from four subjects.

C/γ: **Problem II**				
Electrode	**Subject 1**	**Subject 2**	**Subject 3**	**Subject 4**
Cz	2 / 2.0	6 / 1.0	1/0.01	1 / 0.1
FCCz	1 / 1.0	1 / 0.0001	2/0.01	2 / 0.0001
FC3	2 / 0.001	1 / 0.0001	8/2.0	1 / 0.01
FCC3	2 / 0.01	10 / 0.01	2/0.001	2 / 0.1
C1	2 / 0.01	4 / 0.01	4/2.0	4 / 1.0
C3	10 / 0.01	10 / 0.0001	8/1.0	1 / 0.001
C5	1 / 0.01	8 / 0.01	10/0.01	1 / 1.0
CP1	1 / 2.0	4 / 0.1	1/0.0001	1 / 1.0
CP3	2 / 2.0	1 / 0.0001	6/0.01	1 / 2.0
FC4	10 / 0.0001	8 / 0.01	10/0.001	1 / 0.0001
FCC4	1 / 0.0001	1 / 0.0001	6/0.001	1 / 0.01
C2	6 / 0.001	1 / 0.01	1/0.1	2 / 0.1
C4	2 / 0.01	10 / 2.0	1/0.01	8 / 1.0
C6	1 / 0.1	6 / 0.0001	2/0.001	10 / 0.001
CP2	6 / 0.1	4 / 0.01	2/0.01	1 / 2.0
CP4	1 / 0.0001	1 / 0.0001	1/2.0	1 / 2.0

Table 25.: The optimal hyperparameters for *Problem II*, using the data from four subjects.

C/γ: **Problem III**				
Electrode	**Subject 1**	**Subject 2**	**Subject 3**	**Subject 4**
Cz	1 / 0.0001	6 / 1.0	2/0.0001	10 / 0.0001
FCCz	6 / 0.1	2 / 0.0001	6/1.0	1 / 0.001
FC3	1 / 0.1	1 / 0.0001	2/0.01	8 / 0.01
FCC3	2 / 1.0	1 / 0.001	1/0.01	2 / 0.01
C1	6 / 0.001	2 / 0.01	4/0.1	4 / 0.01
C3	4 / 0.001	2 / 0.1	2/0.0001	10 / 0.001
C5	4 / 0.001	1 / 0.1	4/0.1	4 / 0.0001
CP1	1 / 0.0001	2 / 0.001	2/0.01	1 / 0.1
CP3	1 / 0.01	6 / 0.0001	8/0.1	1 / 1.0
FC4	4 / 0.01	1 / 1.0	10/0.001	4 / 0.001
FCC4	2 / 0.001	8 / 0.0001	4/0.001	1 / 0.01
C2	1 / 0.0001	1 / 0.0001	1/0.0001	6 / 0.1
C4	2 / 0.1	1 / 0.001	4/0.001	1 / 0.1
C6	10 / 0.0001	8 / 0.001	8/0.0001	4 / 0.001
CP2	1 / 0.01	1 / 0.1	1/1.0	2 / 0.0001
CP4	1 / 0.0001	4 / 0.001	1/0.0001	1 / 2.0

Table 26.: The optimal hyperparameters for *Problem III*, using the data from four subjects.