

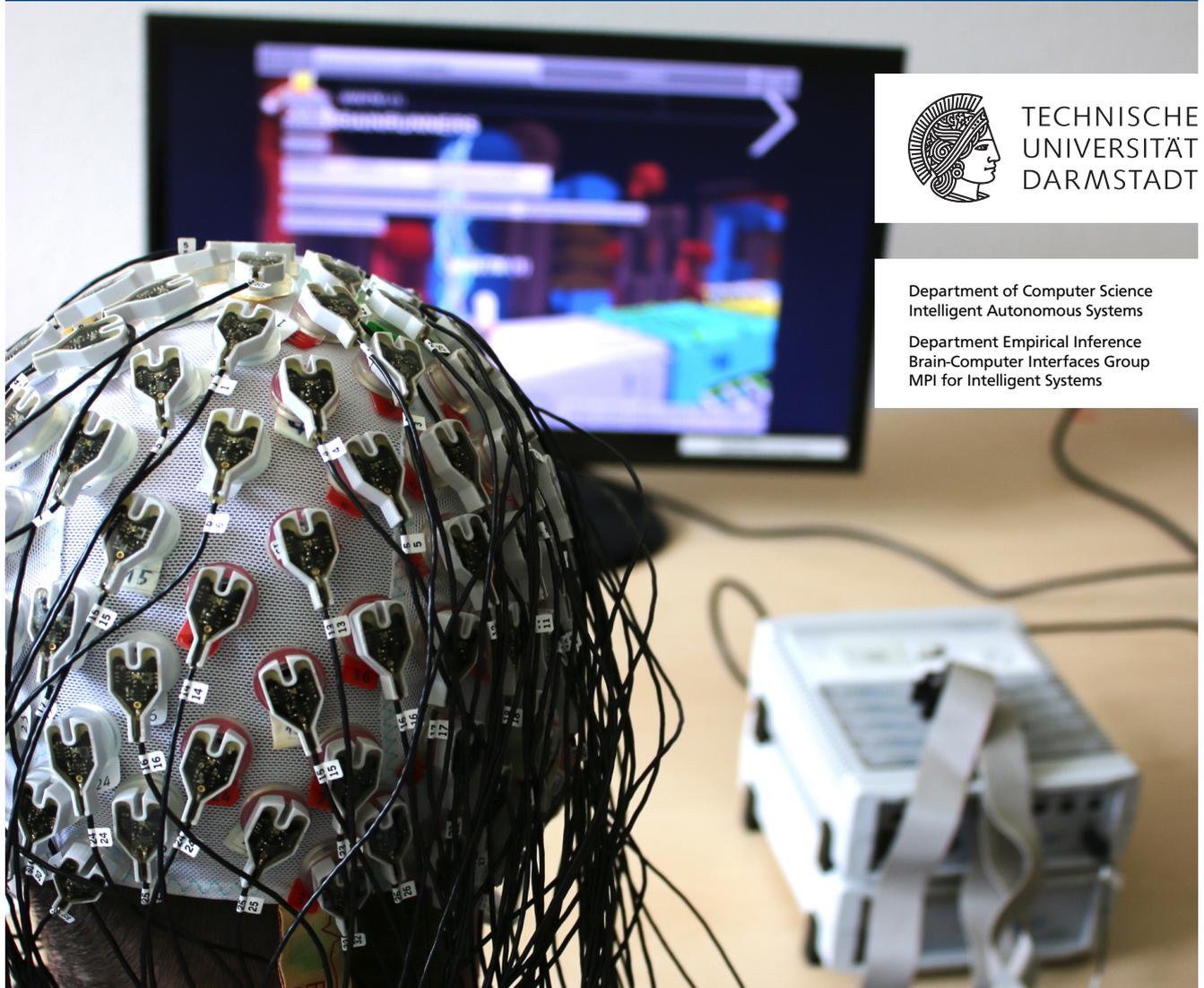
Learning a Filter for Noise Attenuation in EEG Data for Brain-Computer Interfaces

Lernen eines Filters zur Rauschunterdrückung in EEG Daten für Gehirn-Computer Schnittstellen

Master-Thesis von Thomas Hesse aus Frankfurt am Main

Tag der Einreichung:

1. Gutachten: Prof. Dr. Jan Peters
2. Gutachten: Dr.-Ing. Moritz Grosse-Wentrup
3. Gutachten: Vinay Jayaram



TECHNISCHE
UNIVERSITÄT
DARMSTADT

Department of Computer Science
Intelligent Autonomous Systems

Department Empirical Inference
Brain-Computer Interfaces Group
MPI for Intelligent Systems

Learning a Filter for Noise Attenuation in EEG Data for Brain-Computer Interfaces
Lernen eines Filters zur Rauschunterdrückung in EEG Daten für Gehirn-Computer Schnittstellen

Vorgelegte Master-Thesis von Thomas Hesse aus Frankfurt am Main

1. Gutachten: Prof. Dr. Jan Peters
2. Gutachten: Dr.-Ing. Moritz Grosse-Wentrup
3. Gutachten: Vinay Jayaram

Tag der Einreichung:

Erklärung zur Master-Thesis gemäß § 22 Abs. 7 APB der TU Darmstadt

Hiermit versichere ich, Thomas Hesse, die vorliegende Master-Thesis ohne Hilfe Dritter nur mit den angegebenen Quellen und Hilfsmitteln angefertigt zu haben. Alle Stellen, die aus Quellen entnommen wurden, sind als solche kenntlich gemacht. Diese Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

In der abgegebenen Thesis stimmen die schriftliche und elektronische Fassung überein.

Darmstadt, den 29. September 2016

(Thomas Hesse)

Thesis Statement pursuant to § 22 paragraph 7 of APB TU Darmstadt

I herewith formally declare that I have written the submitted thesis independently. I did not use any outside support except for the quoted literature and other sources mentioned in the paper. I clearly marked and separately listed all of the literature and all of the other sources which I employed when producing this academic work, either literally or in content. This thesis has not been handed in or published before in the same or similar form.

In the submitted thesis the written copies and the electronic version are identical in content.

Darmstadt, September 29, 2016

(Thomas Hesse)

Table of Contents

1	Background	3
1.1	Foundations of electroencephalogram characteristics	4
1.2	Artefactual signal morphology in electroencephalogram	5
1.3	Non-Invasive Brain-Computer Interface Pipeline	6
2	Motivation and related work	7
2.1	Thesis Outline	7
3	Spatial Filter Methods used for Session to Session Source Estimate Transfer	8
3.1	Infomax ICA for Session to Session Transfer of Source Estimates	8
3.2	LCMV Beamforming for Session to Session Transfer of Source Estimates	9
4	Experimental Setup and Study Conduction	10
4.1	Hardware and ties to the Cybathlon Challenge	10
4.2	Data Acquisition and Preprocessing	10
5	Empirical Results and Evaluation	11
5.1	Statistical Test Outcome	12
6	Conclusion and Outlook	14
	References	15
A	Appendix	A19
A.1	Whitening Transformation	A19

Abstract

Modern-day Brain-Computer Interfaces (BCIs) are a valuable tool for people suffering from disease of neural origin used to control external applications. However, BCIs are hampered by artefactual source interference with neural sources used to control the BCI. While there are spatial filter methods to decompose the recorded signal into its source estimates, the non-stationary nature of the recorded data does not permit transfer of computed filters for new sessions. Thus, source estimates that are used to make sense of the recorded data are not consistent over sessions. Yet, nowadays state of the art spatial filter methods do not take non-stationary assumptions into consideration. Hence, due to the lack of the spatial filters ability to transfer over sessions the preprocessing of such BCIs are bound to time-costly daily setup procedure. Due to the inability of session transfer for state of the art methods, such as information maximisation independent component analysis (Infomax ICA), we investigate linear constraint minimum variance (LCMV) beamforming to obtain brain related spatial filters for session transfer. Given the beamformers robustness due to second-order statistic estimation, we aim to test for their viability to yield spatial filters that particularly transfer over session. The beamformer was initialised with the spatial patterns of the Infomax ICA and compared against the corresponding spatial filters. The herewith obtained spatial filters were evaluated on real EEG data recorded with a high density EEG using the motor imagery paradigm. Grand mean accuracy was chosen as criterion in order to test for the spatial filters viability for online BCI applications. However, the actual evaluations were performed offline using stratified 10-fold cross-validation. Finally, while the beamformer spatial filters showed no clear trend for session transfer, pairwise permutation tests were performed to show the validity of the empirical results.

Zusammenfassung

Heutige Gehirn-Computer Schnittstellen (GCS) sind ein wertvolles Werkzeug für Menschen die an Krankheiten neuronalen Ursprungs leiden. Allerdings werden GCS durch Interferenz von Artefaktquellen mit neuronalen Quellen behindert, wobei lediglich letztere zur Steuerung der GCS verwendet werden. Zwar gibt es örtliche Filtermethoden um die aufgenommenen Signale wieder in deren ursprüngliche Komponenten zu zerlegen, jedoch erlaubt die nichtstationäre Natur der aufgenommenen Daten keinen Transfer von berechneten örtlichen Filtern auf neue Session. Dennoch nehmen moderne örtliche Filtermethoden keine Nicht-Stationaritätsannahme in Betracht. Daher sind die Vorverarbeitungsmethoden solcher GCS an zeitaufwändigen Kalibrationsphasen auf täglicher Basis gebunden, da die örtlichen Filter nicht die Fähigkeit besitzen auf neue Sitzungen übertragen zu werden. Aufgrund dieser Nicht-Fähigkeit solcher modernen Methoden, zum Beispiel der Informationen Maximierung *Independent Component Analysis* (Infomax ICA), untersuchen wir *linear constraint minimum variance (LCMV) beamforming* um örtliche Filter zu erhalten. Da der *beamformer* durch Schätzung von Statistiken zweiter Ordnung robust ist, wollen wir dessen Einsatz überprüfen solche örtliche Filter zu berechnen, die sich über Sitzungen transferieren lassen. Der *beamformer* wurde mit örtlichen Mustern resultierend aus der *Infomax ICA* initialisiert und mit deren korrespondierenden örtlichen Filtern verglichen. Die hierbei erhaltenen örtlichen Filter wurden auf echten EEG Daten, welche mit einem EEG hoher Elektrodendichte aufgenommen wurden, getestet unter Verwendung des *motor imagery* Paradigmas. Als Evaluationskriterium für den Einsatz in realistischen online GCS Anwendung wurde die mittlere Richtigkeit gewählt. Allerdings wurde die tatsächliche Auswertung offline mittels 10-facher Kreuzvalidierung durchgeführt. Schließlich wurden paarweise Permutationstests durchgeführt um die empirischen Ergebnisse zu validieren, während die untersuchten *beamformer* örtlichen Filter keinen klaren Trend für den Session Transfer zeigten.

Index and Notation

<i>Symbols</i>	<i>Description</i>
$\mathbb{E}[\cdot]$	expectation operator for a given matrix or vector
$\text{cov}[\cdot]$	covariance operator for a given matrix or vector
$\log(\cdot)$	natural logarithm, that is $\log_e(\cdot)$
$\det(\cdot)$	matrix determinant
$H(\cdot)$	Shannon entropy, defined by $H(\cdot) = \sum_{i=1}^{\infty} p(\cdot_i) \log\{p(\cdot_i)\}$

1 Background

The very first results with regard to electrical activity originating from within the brain were first published in 1875 by Richard Caton, in which he observed particular patterns of electrical activity in living animals [12]. These results were measured using needle electrodes and a galvanometer [28]. Fifteen years later, the Polish physiologist Alex Beck published his findings about electrical activity measured with electrodes placed directly on the surface of the brain. These findings about electrical activity in the brain, such as spontaneous fluctuations, evoked potentials, and desynchronisation of brain waves also manifested the believe about the existence of so called brain rhythms [14]. Thus expressing the recorded signals as oscillations is more convenient instead of looking at the raw time series. The term brain rhythms is used to account for distinct brain patterns, that we think is a fundamental part in understanding the different functions of the brain [11, 33]. Ultimately, the German physiologist and psychiatrist Hans Berger was the first person to record electrical activity on the human scalp in 1924 and termed the to date measurements *electroencephalogram* which is commonly abbreviated EEG [6, 24, 33].

Back then, Berger build upon the findings of Carton and emphasized the importance of brain rhythms in clinical diagnosis [19, 45] and set a cornerstone for the next decades of neurological related research. Henceforth, the method of EEG became a basic tool of diagnosing or assisting in diagnosing abnormalities within the brain of living beings [1, 18]. EEG is being used for detection of epileptic seizures [18, 33, 34], supplementary test of brain death [49, 50], monitoring the depth of anesthesia [42], differentiation between psychological disorder and encephalopathy [44], and many more. Nowadays EEG devices usually consist of electrodes, to measure signals from the brain, that is connected via cable to an external amplifier. The amplifier is non-negotiable for the EEG because the electrical activity on the scalp is in the range of microvolts and hence hardly measurable [34, 45, 46] and must be amplified consequently. The electrodes are plugged into an so called EEG cap that is put on the head of the subject to measure the emerging electrical activity from cortical areas. An EEG cap is used to ensure a structured positioning of the electrodes on the scalp allowing for fair interpretability of results across subjects and different instances in time. This setup ensures consistent EEG recordings and thus making the EEG a reliable measurement method.

The extent of EEG usage in research areas and hospitals accelerated in the mid 1900s. Starting from around 1990, EEG scalp recordings were found to be useful for non-invasive Brain-Computer Interfacing (BCI) [31, 37, 52]. The overall goal of a BCI was to aid people with physical disabilities, including paralysis, to perform communication tasks and to control external devices in real life [8, 29]. These communication tasks and control of external devices help in establishing sorts of communication crucial to autonomously perform everyday tasks, thus restoring quality of life for the subject using the BCI [17]. Example applications are shown in the Figure 1 at the bottom.

The EEG characteristics of high mobility and high temporal resolution are beneficial to support the feasibility of those tasks for non-laboratory environments. Brain rhythms, which were found to be a crucial feature in brain functions, were used in order to define the brain to computer communication interface. BCIs are hence bound to these *thought to condition mappings*, commonly called paradigms, employed for the subject to BCI communication. Thus, paradigms render the feasibility of the subject communicating with the BCI, unless the subject is able to control and operate the paradigm [2, 23]. One of the more popular paradigms in EEG based BCIs makes use of the sensorimotor rhythm (SMR) of the brain that relies on the motor commands, for example grasping using either the right or the left hand. Apart from SMR based paradigms, there are other paradigms that showed similar good empirical performance [51, 39]. Any paradigm based on EEG recordings, however, is hampered by the often mentioned low spatial resolution and environmental noise due to the nature of EEG scalp recordings [33, 46, 52]. The wording spatial resolution in the context of this thesis refers to the resolution of brain patterns recorded at the EEG scalp electrodes at each time instant. Fundamentally, there are multiple distinct causes to low spatial resolution which are elaborated on in the following.

In order to analyse the resulting low spatial resolution, we briefly review the process of neural electrical activity propagating from within the brain to the electrodes. Single neuron activity emerges as action potential (AP) in the pyramidal

Figure 1: Brain-Computer Interfaces (BCIs) may function as an universal solution for non-abled or disabled persons to control external applications as depicted in the future. As soon as BCIs are more robust while operating, they can be used to control wheelchairs or even robotic exoskeletons [26].



cell body which cannot be measured by EEG, which is located on the scalp, directly. Nor can the EEG electrodes measure activity elicited by neuron populations directly. Nevertheless, the EEG electrodes are able to record the activity from multiple neuron populations aligned perpendicular to the cranium. EEG measurements are made possible in particular because populations of neurons emit local field potentials when electrical activity is elicited. Now, whenever multiple local field potentials fully or nearly synchronise they mix linearly as described in [34] because of the passive conductance properties of living tissue on a macroscopic scale. Consequently, the linearly mixed, that is to say superposed, electrical activity originating from multiple neuron populations is projected farther and thus such activity is summarized as far field potentials at the electrodes [27, 16]. Now, when there are opposing local field potentials they will rather cancel each other out instead of mixing themselves, yielding potentials of insufficient strength to be measured by scalp electrodes.

Furthermore, smearing of the signals occurs at different stages because of the passive conductance properties of living tissue, cerebrospinal fluid (CSF), and the cranium. Due to the generation processes of the measurable signals, activity of distinct neural sources is likely to be correlated across the brain source-related populations. And to the best of the authors knowledge, the correlation of sources mixed into a single signal also does not allow to determine the actual number of sources as well as the influence of each single source generating the signal. Apart from neural sources, EEG recordings are also prone to contamination by muscular activity occurring close to the EEG cap and ocular activity that is generated by eye movement or blink. The latter, that is to say muscular and ocular activity, is often referred to as artefact. Moreover, there might be portions of noise present in the signal from disparate sources that are due to environmental inference or noise in the EEG system itself.

1.1 Foundations of electroencephalogram characteristics

To start with, we will briefly review physiological characteristics of electroencephalogram (EEG) recorded data, that is the previous mentioned brain rhythms. This shall give the reader a sufficient understanding of the feature type being used in the Brain-Computer Interface (BCI) pipeline for the decoding of actions. To declutter the reading flow in this chapter, whenever the term data is used it refers to EEG data recorded on the human scalp.

First, we need a understanding of the characteristics of the data, hereby, determining physical key aspects underlying the data. The data has units μV (microvolts) normally distributed around $0\mu V$ and varies up to $100\mu V$ that may be exceeded in presence of artefacts or distortion of the recorded signal yielding the data [46]. The data is usually represented in terms of frequency instead of its raw form because of the importance of oscillatory phenomena and thus the data is Fourier transformed [4]. The units of the Fourier transformed data changes from *time in seconds* to *frequency in Hertz* with a few particular frequency bands reflecting different functional patterns that may be observed. The complete frequency band ranges from around $0Hz$ to $2^{-1}f_s Hz$ with f_s the sampling frequency and disperses into

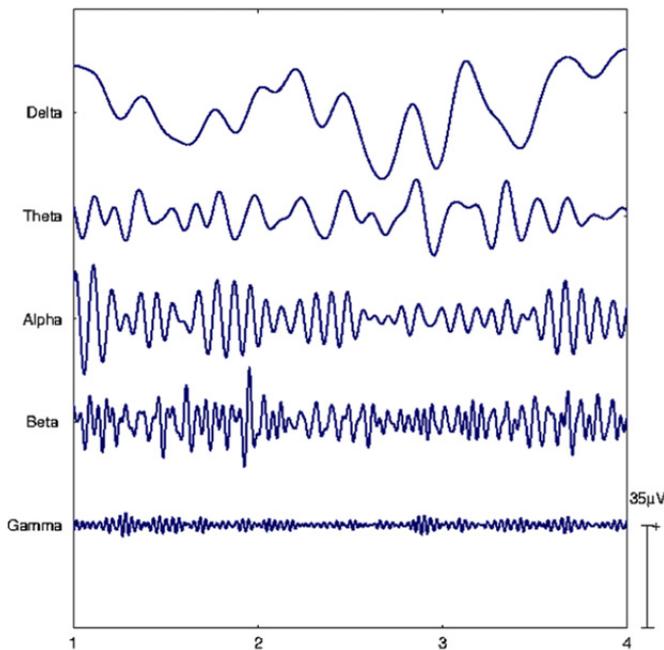


Figure 2: The five major brain rhythms delta, theta, alpha, beta, and gamma related to functions in the brain and patterns in the electroencephalogram. Taken from [46].

- **Delta Band** ($\sim 0.5 - 4$ Hz) is commonly associated with cognitive processes tied to consciousness and sometimes information retrieval [4]. Moreover, relations to deep sleep but also confusion with muscle artefacts have been found in the range of the delta band [43].
- **Theta Band** ($4 - 7$ Hz) which was at first thought to originate from the thalamus, was termed by [48] and hints at a basic role in cognitive processing and in the cortico-hippocampal interaction [4, 43].
- **Alpha Band** ($7 - 14$ Hz) oscillation mainly appear in the posterior half, which targets the parietal and occipital lobe, of the cortex [43]. Alpha rhythms are mostly related to visual tasks.
- **Mu Band** ($8 - 12$ Hz) is the brain rhythm of focus within this work because it is related to the well known motor-imagery paradigm [38].
- **Beta Band** ($15 - 30$ Hz) is a brain rhythm believed to be mostly related to awareness and attention [43].
- **Gamma Band** (above 30 Hz) mainly reflects activity for high cognitive and mental tasks and is the frequency band most prone to muscular artefacts.

These brain rhythms find application in clinical use but are also used for Brain-Computer Interfaces. The brain rhythms are beside others common features used in decoding commands for a paradigm [30, 46]. For the specific choice of sensorimotor (SMR) paradigms, such as the *motor-imagery* (MI), the spectral band called mu band found within the alpha band is often used [30]. However, it is not just the amplitude of the frequency being used but its bandpower [36]. However, a caveat to the MI paradigm is that it is rather not or only partially usable by subjects suffering from motor neuron disease.

1.2 Artefactual signal morphology in electroencephalogram

As previously discussed, common artefacts found in electroencephalogram (EEG) recordings are (A1) muscle artefacts, (A2) ocular artefacts, and (A3) cardiac artefacts. These three types of artefacts are the most often encountered contaminates in the EEG literature and focus of artefact reduction. The morphology of the three most occurring artefactual sources are depicted in Figure 3. Ocular artefacts (EOG) may be caused by horizontal or vertical eye movements but can also be caused by eye blinks. EOG artefacts particularly affect electrodes in the frontal lobe and fronto-central areas. Artefactual interference by eye blinks are more abrupt Muscular artefacts (EMG) arise whenever the subject moves the jaw or swallows and since these activities cannot be suppressed and hence they need to be filtered from the signal.

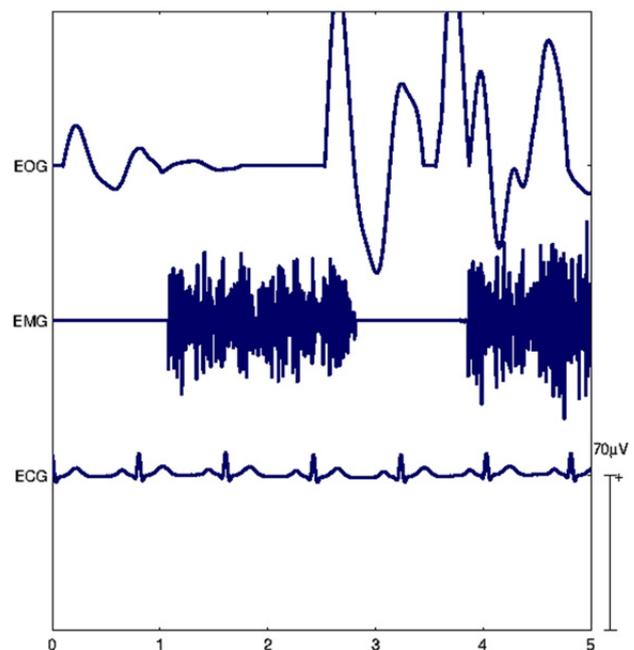


Figure 3: The most prominent artefactual source morphologies in the time domain found in EEG. While these artefactual source are visually identifiable in the time domain, they are present in many of the frequency bands presented in Figure 2.

1.3 Non-Invasive Brain-Computer Interface Pipeline

Throughout this thesis, we will follow the Brain-Computer Interface (BCI) pipeline scheme depicted in Figure 4. This pipeline provides insight to the most important steps to take to build an BCI in which we focus on the field of artefact reduction. To start with, we will elaborate the single steps of the BCI pipeline before we go ahead into the particular sub-field of artefact reduction.

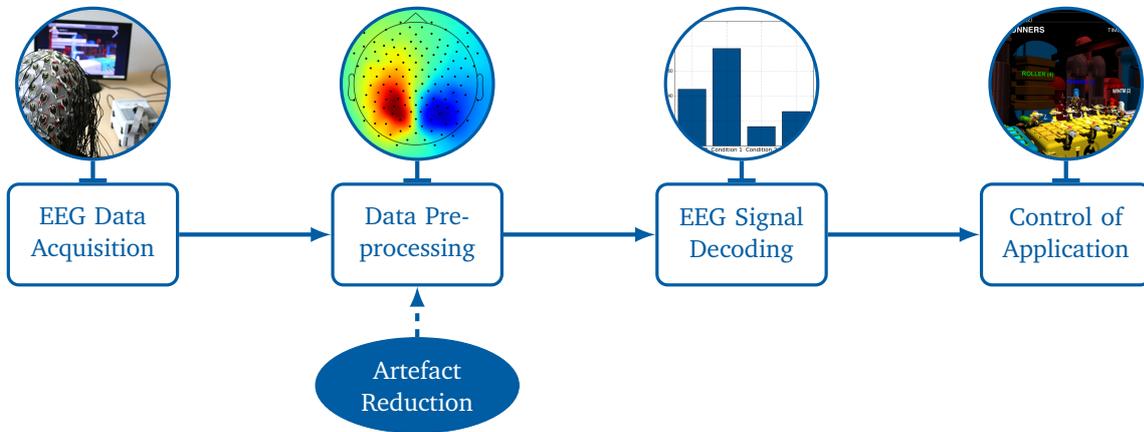


Figure 4: Brain-Computer Interface online pipeline built from non-invasive electroencephalogram. This BCI online pipeline shall motivate and outline the necessity of artefact reduction in the context of BCI. Moreover, this pipeline sketches the basic steps to build a full-fledged BCI.

For non-invasive BCIs in the context of this work, the data is recorded using electroencephalogram from the human scalp to observe brain patterns from populations of neurons. *Data acquisition usually* is the first step to undertake in the pipeline to calibrate the BCI for the associated session, thus the same session the data was obtained. However, before this data can be made use of, one must perform the crucial step of *data preprocessing* to purge the data from artefacts as best as possible. The preprocessing of the data is crucial for the BCI pipeline to function sufficiently well because the pipeline is built on data driven methods. Subsequently, features may be created or extracted from the purged data. These features are then passed onto the *signal decoding* step in which they yield fundamental support in decoding the signals into predefined conditions to control an external *application*. Yet just witnessed the architecture of the rather simple BCI pipeline assumed in this work, we move one level deeper into the pipeline module of artefact reduction. Artefact reduction is a crucial part that helps to construct plausible BCIs and hence particularly important to non-invasive BCIs. The reason for its importance originates from the low spatial resolution that goes hand in hand with the electroencephalogram because we are merely able to record the activity of only a few neuron populations. Hence, the sub-field of artefact reduction also comprises methods for spatial filtering that yields support in noise and artefact attenuation while locating the signal of interest. In the following sections, methods and their material are presented that were used to undertake the spatial filtering part in artefact reduction of the BCI pipeline from Figure 4.

2 Motivation and related work

Non-invasive Brain-Computer Interfaces (BCI) based on EEG measurements offer an affordable and qualitative way of measuring neural signals from the human brain while it does not constitute a threat for the subject. Unfortunately, this type of measurement is prone to artefactual sources, namely non-neural signals, and is moreover hampered by low spatial resolution. While these different factors certainly negatively impact signal quality, there are methods to compromise the issue of low spatial resolution and also artefactual sources. These methods are summarised within the category of spatial filtering and are either of supervised or unsupervised nature. Unsupervised in this context means that the data must not be labelled. On the one hand, there are supervised methods that rely on labelled data to find spatial patterns in the data to discriminate between distinct brain conditions contained in the labels. The most prominent method concerning supervised spatial filtering is the common spatial patterns (CSP) [41]. On the other hand, often used unsupervised methods in this category of spatial filtering are variants of independent component analysis (ICA), principal component analysis (PCA), and second-order blind identification (SOBI) [25, 35, 40, 46]. Those unsupervised methods decompose the signal into its original components without requiring knowledge of the specific brain condition, consequently resulting in suitably applicable spatial filters. These methods produce viable results [10] and allow for applications such as single source examination [15], temporal analysis, and source-localisation [53]. Unfortunately, these methods only function well within sessions so far because they do not consider the non-stationary nature of the decomposed sources underlying the brain patterns.

As a result, BCIs based on such spatial filter methods are left with costly preparation time for setup on a daily basis. Hence, these methods in their current form are prone to be more impractical for real life BCIs. However, there is work done in the field of spatial filtering applicable for session transfer. [32] evaluated common methods in spatial filtering with regards to session transfer in order to simulate an online BCI application. They found the CSP method to perform overall better than ICA variants and Laplacian spatial filters. Yet, it is to notice that CSP requires prior information in the form of labels which is generally not present in online BCI applications. In contrast, unsupervised spatial filters do not require any prior information and hence their usage are more appropriate for online BCI application scenarios. [20] extended unsupervised spatial filtering methods that solely estimates sources regarding at most second-order statistics based on non-stationary assumptions, yet also incorporating prior information due to the usage of CSP. Subsequently, they successfully evaluated their method within and across sessions for an online BCI application scenario, yet requiring additional knowledge about brain condition labels. In contrast, the goal of this thesis builds upon the idea proposed in [21] and aims to robustly estimate sources based on second-order statistics under incorporation of an appropriate forward solution of the brain patterns. The choice of framework to estimate sources is the linear constrained minimum variance (LCMV) beamforming because it yields more robust solutions than existing approaches such as CSP [22]. Similar to the related work, the results are evaluated for an online BCI application while existing results from [21] are to be verified. Finally, results of the Infomax ICA method are compared to the results of the LCMV beamformer. Moreover, the authors of this work propose an extension to the existing LCMV beamformer for session to session transfer that accounts for the non-stationary of the second-order statistical estimates yielding more robust and consistent solutions.

2.1 Thesis Outline

This section agrees on the structure of this thesis and briefly outlines its contents. Chapter 1 contains background reading to firmly introduce the BCI inexperienced reader into the topic. The overall goal of spatial filtering concerned with session-to-session transfer is motivated in chapter 2. Chapter 3 introduces the spatial filtering methods investigated for session-to-session transfer. Moreover, the methods are derived, explained and motivated in the light of session transfer of source estimates. Chapter 4 deals with the details of the study preparations and conduction. An overview is given over the subjects that undergone EEG data acquisition. Subsequently, data preprocessing methods applied before the spatial filter evaluation are explained and their justification is given. Following the data preprocessing, chapter 5 deals with the evaluation. That is justifying the choice of evaluation metric for the spatial filtering session-to-session transfer. Chapter 6 concludes with the empirical performance of the methods and statements of the author regarding the resulting performance for spatial filtering in terms of session-to-session transfer and statistical test outcome. Finally, an outlook for disadvantages of the proposed method and extensions to overcome those disadvantages are given.

3 Spatial Filter Methods used for Session to Session Source Estimate Transfer

This chapter will give an explanation and derivation of the methods for the purpose of recovering temporal source estimates used for session to session transfer. That is the re-usage of *Information maximisation Independent Component Analysis* (Infomax ICA) spatial filters in Section 3.1 and *linear constrained minimum variance (LCMV) beamforming* for temporal subsequent sessions in Section 3.2. Spatial filtering methods are usually used to decompose the mixture of recorded signals contained in the EEG data. Thus, brain patterns meaningful to control a BCI can be obtained and used within the decoding process. While artefactual sources may contaminate the data, they are more likely to be contained in the mixture of signals. Spatial filtering methods try to dissect those artefactual sources from the meaningful brain patterns. Brain patterns can be identified by estimation of temporal source estimates that are of non-stationary nature and hence, as broached in the motivation, can only be used reliably within short time frames. In what follows, unsupervised spatial filtering approaches for within session and session to session transfer of source estimates are presented.

3.1 Infomax ICA for Session to Session Transfer of Source Estimates

Independent Component Analysis (ICA) is a generative model thus is used to model the process of observed data being generated. One common example to illustrate the usage of ICA is the cocktail party problem. In this problem, there are two speakers that emit dissimilar speech signals which mix linearly and are being recorded by two microphones. While the first speech signal mixture contains a greater portion of the first speaker, the other speech signal mixture contains a greater portion of the other speaker. Given the two signal mixtures, ICA attempts to de-mix these signals to obtain the original sources prior to mixing. In technical terms, the sources objective to estimation are defined as $\mathbf{s} = \mathbf{W} \cdot \mathbf{x}$ with $\mathbf{W} = \mathbf{A}^{-1}$ and $\mathbf{W} \in \mathbb{R}^{N \times N}$ for the complete model ICA [25]. The matrix \mathbf{A} hereby is the mixing matrix that arbitrarily mix all the signals and \mathbf{W} is the unmixing matrix to de-mix the signals in its original source estimates. Looking at $\mathbf{W} = \mathbf{A}^{-1}$, we need to assume a linear mixture model, statistical independence of the sources, an identical distribution throughout time, and at last a square equation system [25]. The objective of the infomax ICA, as the name prefix already states, maximises the information, that is equivalent to minimising the mutual information $I(Y; X) = H(Y) - H(X|Y) = H(Y) + H(X, Y) - H(X)$ which is given by

$$\mathbf{W}^* = \arg \min_{\mathbf{W}} \left\{ \sum_{i=0}^N H(\mathbf{W} \cdot \mathbf{x}_i) - H(\mathbf{W} \cdot \mathbf{x}) \right\} \quad (3.1.1)$$

in which \mathbf{x}_i denotes the i -th row element of the vector \mathbf{x} . And with \mathbf{W} being an invertible transformation, we obtain

$$\mathbf{W}^* = \arg \min_{\mathbf{W}} \left\{ \sum_{i=0}^N H(\mathbf{W} \cdot \mathbf{x}_i) - \log(\det(\mathbf{W})) \cdot H(\mathbf{x}) \right\} \quad (3.1.2)$$

Additionally, the $H(\cdot)$ operator denotes the Shannon Entropy which is given by $H(Y) = \sum_{i=1}^{\infty} p(Y_i) \log\{p(Y_i)\}$. With \mathbf{W} , an orthogonal transformation, and $H(\mathbf{x})$, independent of \mathbf{W} , the objective to recover the source estimate takes the form

$$\mathbf{W}^* = \arg \min_{\mathbf{W}} \left\{ \sum_{i=0}^N H(\mathbf{W} \cdot \mathbf{x}_i) \right\}. \quad (3.1.3)$$

The resulting update equation is hence given by

$$\Delta \mathbf{W} \propto \frac{\partial \left[\sum_{i=0}^N H(\mathbf{W} \cdot \mathbf{x}_i) \right]}{\partial \mathbf{W}} = \frac{\partial \sum_{i=0}^N \mathbb{E}[-\log(p(\mathbf{W} \cdot \mathbf{x}_i))]}{\partial \mathbf{W}} \quad (3.1.4)$$

that can be solved with a gradient descent algorithm. An appropriate choice for the output distribution, namely the source distribution $p(\hat{\mathbf{s}}_i) = p(\mathbf{W} \cdot \mathbf{x}_i)$, is given in [5], and remains a parameter to be tuned.

Algorithm 1: Infomax ICA

Data: Raw data $\mathbf{X} \in \mathbb{R}^{N \times T}$ with N sensors and T time points, L the number of sources to estimate

Result: Matrices \mathbf{W} , $\mathbf{A} = \mathbf{W}^{-1}$, \mathbf{S}

Initialize $\mathbf{W} = \mathbf{W}_0$;

Centre \mathbf{X} ;

Whiten \mathbf{X} using PCA [9] with L components;

while \mathbf{W} *not converged* **do**

 | Solve the objective stated in Equation (3.1.4) via gradient descent methods;

A naïve spatial filtering approach forward session to session transfer of source estimates is the re-usage of spatial filters \mathbf{W} yield from previous sessions.

3.2 LCMV Beamforming for Session to Session Transfer of Source Estimates

To start with, we shed light on the motivation of the linear constrained minimum variance (LCMV) beamforming. The motive of LCMV beamforming is to locate source activity from a particular region of interest while attenuating the bandpower of all other regions so they will not infer the source one is interested in. It does so by finding a weight vector that obeys the constraint of keeping sources of interest while minimising the bandpower in all other regions as can be seen from Equation 3.2.1. Van Veen et al. [47] introduced such a method in which they relate the dipole mappings of the brain to surface recordings by exploiting the estimated spatial covariance structure. The objective from the original paper of the LCMV beamforming is given by the constrained optimisation

$$\begin{aligned} \mathbf{w}^* &= \arg \min_{\mathbf{w}} \{ \mathbf{w}^\top \hat{\Sigma} \mathbf{w} \} \\ \text{s.t. } & \mathbf{w}^\top \mathbf{a} = 1 \end{aligned} \quad (3.2.1)$$

with $\hat{\Sigma} \in \mathbb{R}^{N \times N}$ a covariance estimate that solves for the optimal spatial filter \mathbf{w}^* . The derivations for this constrained optimisation is omitted here but can be verified using Lagrange multipliers. The resulting equation yields the optimal filter

$$\mathbf{w}^* = \left(\mathbf{a}^\top \hat{\Sigma}^{-1} \mathbf{a} \right)^{-1} \mathbf{a}^\top \hat{\Sigma}^{-1} \quad (3.2.2)$$

that locates sources of interest encoded by \mathbf{a} . Practical issues may arise for the inversion of the estimated spatial covariance matrix due to an insufficient inner matrix rank. To account for this issue, the authors resort to the methods of diagonal loading, that is adding a constant term to the diagonal of the covariance matrix, and more advanced methods of shrinkage estimators [13]. In order to use the LCMV beamforming for session transfer, the author proposes the reuse of spatial patterns from prior sessions $t - \tau$, $\tau = 1, \dots, t - 1$ while solving for Equation 3.2.2 with $\hat{\Sigma} = \text{cov}(\mathbf{X}) = \mathbb{E}[X_t^2] - \mathbb{E}[X_t]^2$. This method was initially introduced in [21] and, there, it was used to reconstruct temporal activity of non-Gaussian sources for the overcomplete source mixture model. From here on, we refer to this method as Linear Constrained Minimum Variance Spatial Pattern Beamforming (LCMV SPBF).

Algorithm 2: LCMV SPBF

Data: Well conditioned baseline trial spatial covariance estimate $\hat{\Sigma}$, raw data $\mathbf{X} \in \mathbb{R}^{N \times T}$ with N sensors and T time points

Result: Spatial Filter \mathbf{F}

Initialize $\mathbf{W} = \mathbf{W}_0$;

Centre \mathbf{X} ;

Whiten \mathbf{X} using PCA [9];

while \mathbf{W} *not converged* **do**

 └ Solve the objective stated in Equation (3.1.4) via gradient descent methods;

 Compute $\mathbf{A} = \mathbf{W}^{-1}$;

 Visual inspection of $K < L$ paradigm-related sources denoted \mathbf{A}_K ;

$\mathbf{F}^* = \left(\mathbf{A}_K^\top \hat{\Sigma}^{-1} \mathbf{A}_K \right)^{-1} \mathbf{A}_K^\top \hat{\Sigma}^{-1}$;

For a proper reconstruction of the source estimates, the forward solutions in \mathbf{a} need to be uncorrelated which shall hold given the minimisation of the mutual information during the infomax ICA procedure in Equation 3.1.4. In the case of Gaussian distributed forward solutions \mathbf{a} , the reconstruction of the corresponding source will fail because of arbitrary sources that were mixed into one source and due to only considering up to second-order statistics. An exhaustive explanation and analysis for correlated sources in the forward model are given by [5, 25].

4 Experimental Setup and Study Conduction

The study was carried out at the TU Darmstadt with five male subjects aged 26 ± 3.89 years old. Everyone of those subjects were in a physically and mentally healthy state during the study conduction. While data has been recorded, the subject was seated in a comfortable chair and was able to stop the session at any time. The subjects were seated 40cm in front of a computer screen with a frame rate of 60Hz. At the start of the first session, each subject was handed a study guide explaining the data acquisition procedure and the sensorimotor (SMR) paradigm used. Alongside the study guide, the subject was handed a form to fill out anonymised personal information which were allowed to be used within the thesis by the subject explicitly. Those collected information can be found in Table 2. Each subject noted the explicit

Subject	Age	Gender	EEG experience	motor skills		number of sessions
				playing a music instrument and frequency	sportiness and frequency	
S1	30	male	yes	stringed instrument, every day	3 days/week	3
S2	29	male	no	-	1-3 days/week	3
S3	25	male	no	stringed instrument, 1 day/week	2-3 days/week	4
S4	27	male	yes	-	-	4
S5	19	male	no	-	2-3 days/week	3

Table 2: General overview about the subjects, that is their identifier, age, gender, previous EEG experience, their motor skills, and the number of sessions being recorded. The motor skills are listed here because the subjects needed to operate a paradigm related to motor imagery. Hence, their motor related skills are listed to justify possibly negative performance with the motor imagery paradigm.

thought for each condition of the paradigm. Before every ongoing session, the subject was able to see there notes made in the first session to maintain consistency for the chosen thoughts to operate the paradigm. The subject was able to abort the experiment at any time before and during the recording. The individual sessions were scheduled on a weekly basis for each subject and no feedback on performance was given.

4.1 Hardware and ties to the Cybathlon Challenge

The contents of this thesis are, also, currently motivated for and embedded in the Cybathlon Challenge and hence receives support by the collaboration between IAS, Department of Computer Science, Technical University Darmstadt and Brain-Computer Interfaces Group, Department Empirical Inference, Max-Planck Institute for Intelligent Systems.

The hardware used for data acquisition in this thesis was sponsored by Brain Products GmbH¹ for the Cybathlon-Team Athena-Minerva² which I, Thomas Hesse, am part of. The Cybathlon is the first official *Olympics* for bionic athletes with strong ties to research and “*was forged with three aims in mind; to facilitate conversation between academia and industry, to facilitate discussion between technology developers and people with disabilities and to promote the use of robotic assistive aids to the general public*”.³

4.2 Data Acquisition and Preprocessing

The data was recorded using the aforementioned hardware. This setup consisted of the EEG cap in the 10-20 system, 129 electrodes (including the reference and ground electrode), and an amplifier for the EEG data. The electrodes used were gelled to gap and improve conductance properties between the electrodes and the scalp. It was made sure that the impedances of the electrodes were in range 0-20kOhm to actually capture neural data. Afterwards, in an offline step, the data was preprocessed using common average reference (CAR) over all electrodes which is suitable because of the high density of electrodes [7], a 3Hz high-pass Butterworth filter of third order, and zero baseline shift for each electrode before being used for further (pre-)processing. Before recording the paradigm related data, a baseline trial was recorded and the subject was asked to fixate a white cross on top of a black background on a computer screen. Furthermore, the subject was instructed to just fixate the cross on the screen not performing any paradigm related commands.

¹ Brain Products GmbH - Solutions for neurophysiological research

² Cybathlon-Team Athena-Minerva, TU Darmstadt

³ Source: Cybathlon Quote

5 Empirical Results and Evaluation

For the evaluation part, reconsider the pipeline model from Figure 4 introduced earlier. The initial part of the EEG data acquisition was carefully described in Section 4.2. There, we also outlined the preprocessing that was applied after the data was recorded. This preprocessing was consistently applied for the remainder of the pipeline model.

Subsequently, the method presented in Section 3.2 was evaluated against the Infomax ICA derived in Section 3.1 in which the latter was regarded as the baseline performance. The evaluation includes cross-validated within session evaluation and, more important for the application for online BCI applications, across session evaluations. In order to measure the performance of these methods, a logistic regression classifier with binomial log-likelihood was trained on a train set and evaluated on a test set. While there was at most up to forty data points for each class produced by a single session, bootstrapping methods were used in order to artificially upsample the data. To do so, the data of two classes were first subsampled, then stratified, and at last k -fold cross validated with $k = 10$. For the accuracies to converge for each subject respectively, the aforementioned procedure was repeated sufficiently often until the standard deviation of the results converged to zero. Obviously, the aforementioned evaluation procedure was implemented offline for time-locked stimuli. Each motor imagery condition was evaluated which is found in the mu band against the resting phase found in the alpha band (cf. Section 1.1). The results of each subject were averaged afterwards and are shown subsequently in Figure 5.

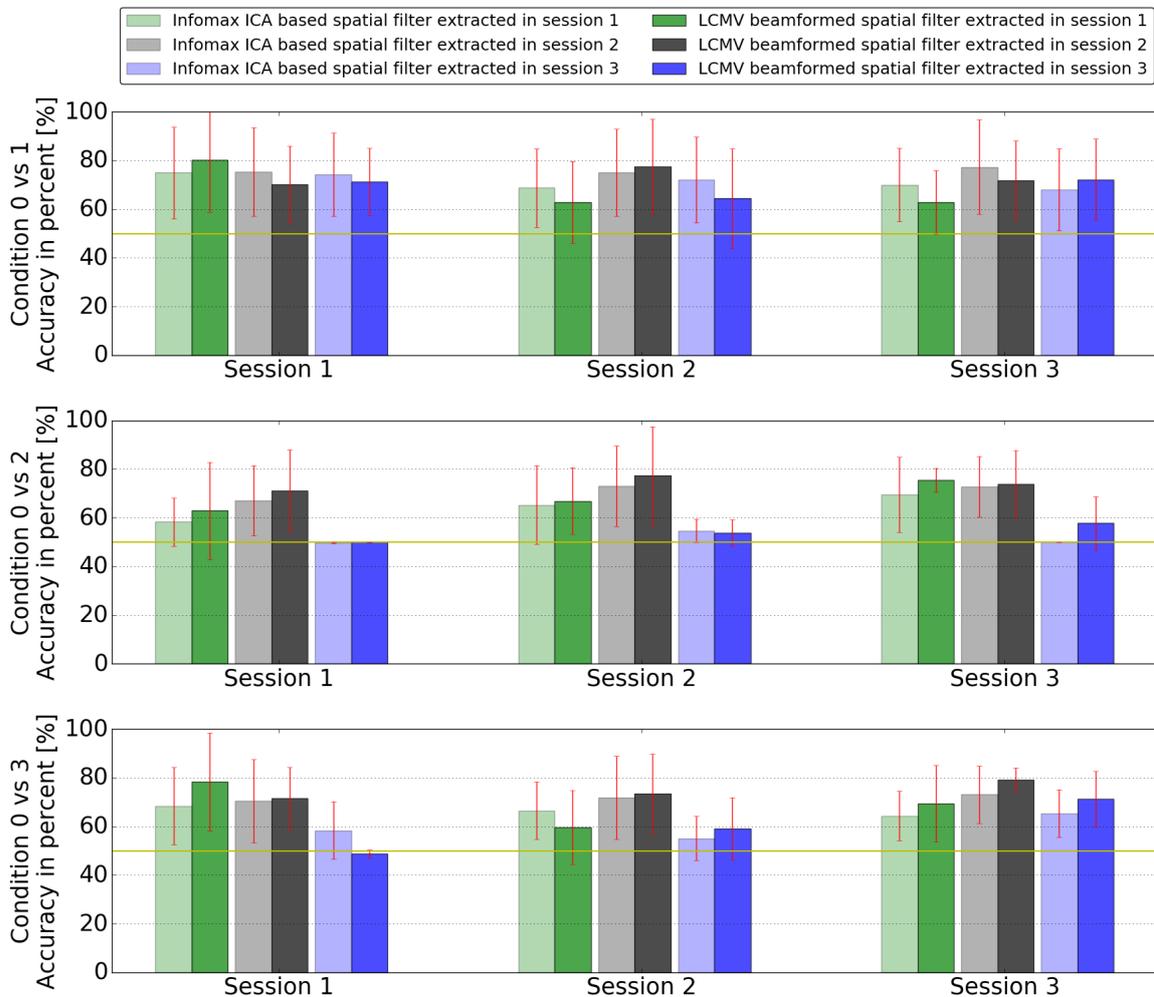


Figure 5: Resulting decoding accuracy of the logistic regression binomial log-likelihood classifier. The colour encodes the session the data was recorded in. The mapping is green for session one, black for session two, blue for session three, and red for session four. The transparent bars show the performance of the Infomax ICA and the dull bars show the performance of the LCMV SPBF method presented in Section 3.2. The red error bars shows the performance variation across the subjects. The yellow line indicates the chance level performance of 50 percent.

The features for the classifier of each subject were three spatial patterns from each session in the mean mu band. These spatial patterns were selected from a set of 64 recovered sources using Infomax ICA. Hence the input features for the bootstrapping applied prior the classification procedure was of the form $F^T X = \Phi \in \mathbb{R}^{3 \times D}$ with D data points representing time-locked trials.

5.1 Statistical Test Outcome

In the previous Section 5, the averaged accuracies with two times their standard deviation were compared to each other. While these results are produced from five subjects, non-parametric statistical testing was applied to account for the validity of the found results. More precisely, pairwise permutation tests were used to examine the true performance of these results. Therefore, we distinguish the model based on its two distinct preprocessing yielding different feature sets and henceforth two models subject to this statistical test. We denote permutation sets $\mathcal{P}_{a,condition} = \{a_{i,condition}\}_{i=1,\dots,5}$ for within session Infomax ICA results and $\mathcal{P}_{b,condition} = \{b_{i,condition}\}_{i=1,\dots,5}$ for within session LCMV SPBF results. Similarly for the session transfer results, we write $\mathcal{P}_{\Delta a,condition} = \{\Delta a_{i,condition}\}_{i=1,\dots,5}$ and $\mathcal{P}_{\Delta b,condition} = \{\Delta b_{i,condition}\}_{i=1,\dots,5}$.

Two null hypotheses were tested, in which the true performance $\mu_a, \mu_b, \mu_{\Delta a}, \mu_{\Delta b}$ is drawn from $\mathcal{P}_{a,condition}, \mathcal{P}_{b,condition}, \mathcal{P}_{\Delta a,condition}, \mathcal{P}_{\Delta b,condition}$ respectively. That is first testing for true equal averaged accuracy within sessions $H_0^* : \mu_a = \mu_b$ and across session $H_0^{**} : \mu_{\Delta a} = \mu_{\Delta b}$. The pairwise permutation test was run with a total of $N = 1.000.000$ runs for both null hypotheses for each binary classifier. Each binary classifier is indexed respective the condition that is subject to testing. The results are depicted in the subsequent in Figure 6 for within session performance and session to session transfer performance.

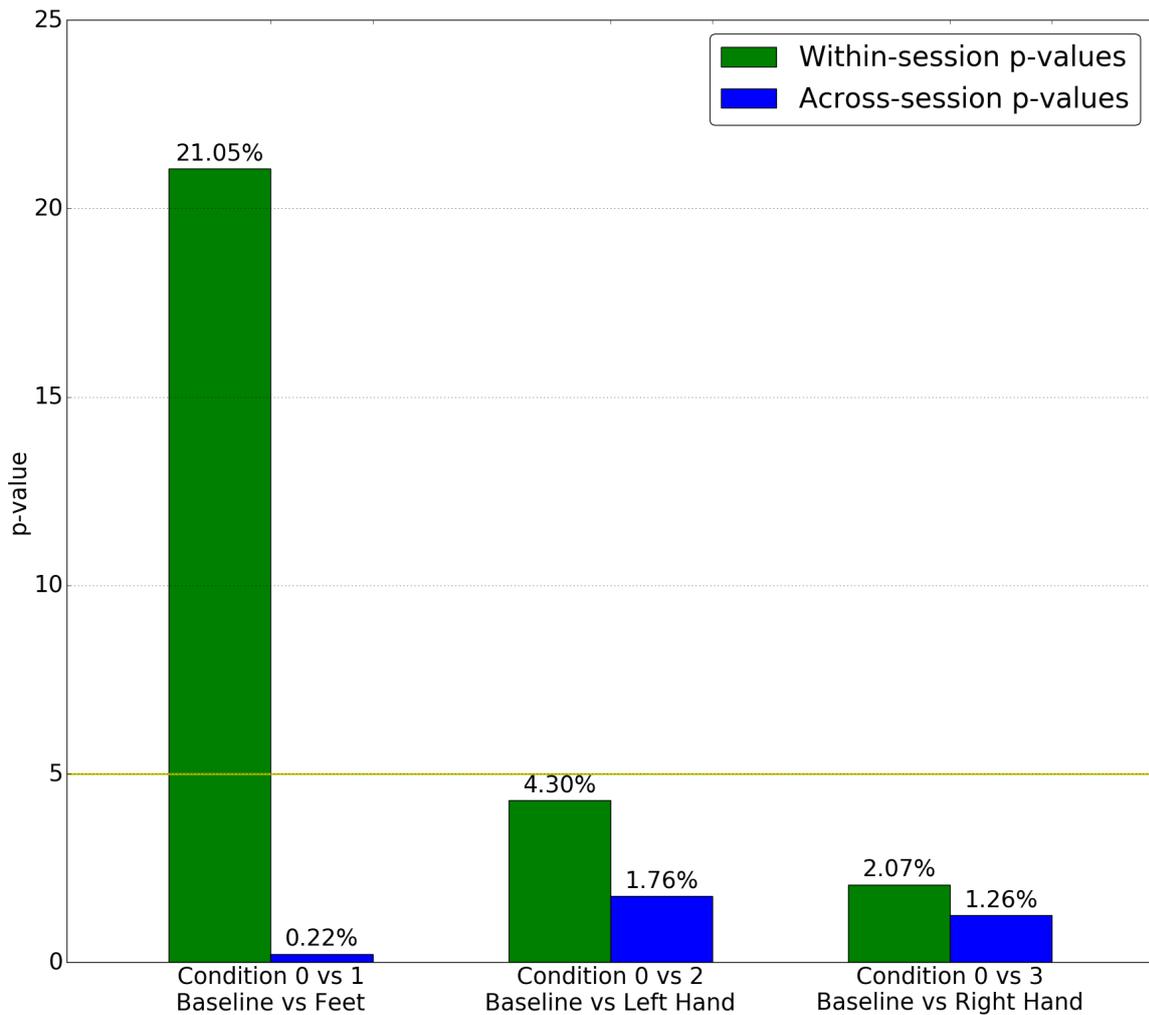


Figure 6: Two-sided pairwise permutation test for the H_0^* in green and H_0^{**} in blue for true equal performance of the compared methods. The plot shows the significance level of 5% and the p-values of the results.

The test statistic for resulting p-values is given by $\mathcal{T}(\mathcal{P}_x, \mathcal{P}_y) = \text{mean}(\mathcal{P}_x) - \text{mean}(\mathcal{P}_y)$ and the associated p-value is hence computed as

$$p = N^{-1} \sum_{i=1}^N \left[\mathcal{T}(\mathcal{P}_x^{(i)}, \mathcal{P}_y^{(i)}) \geq \mathcal{T}(\mathcal{P}_x, \mathcal{P}_y) \right] \quad (5.1.1)$$

with sequential Monte-Carlo roll-outs. The superscripted index $[\cdot]^{(i)}$ signals pseudo-random generated pairwise samples which are obtained by permutation of the pairs $(\mathcal{P}_x, \mathcal{P}_y)$.

The results of the pairwise permutation test show significance for the session transfer results given the significance level of 5%, that is H_0^{**} is accepted. However, for the within session results we fail to either accept or reject H_0^* for the brain condition *feet*.

6 Conclusion and Outlook

We investigated spatial filtering methods for the usage of session transfer for the purpose of an online BCI application. In order to do so, we used the presented BCI pipeline model from Figure 4 and evaluated the methods in terms of accuracy. While the methods were tested with five subjects, non-parametric statistical tests were used to validate the outcome of the empirical results. Yet, LCMV beamforming was solely used by [21, 47] to obtain source estimates within sessions. The method estimates sources up to second-order statistics and hence yields robust results. In this work, we investigated the performance of LCMV beamforming for the use of session to session transfer. The forward solutions yield by methods solving for blind source separation were input to the LCMV beamformer as sources to keep. The forward solutions here were estimated and visually selected for motor-imagery related conditions. Then, the associated spatial filters were compared to the LCMV beamformed aforementioned spatial patterns in terms of accuracy. For the choice of a decoding model we used the log binomial likelihood for a binary classification problem. The feature space used for decoding had dimensions equal to the number of spatial filters for both, LCMV beamforming and Infomax ICA. Overall, the LCMV beamforming of the spatial patterns yield by the Infomax ICA showed to be superior for the within session evaluation as can be seen in Figure 5. While we failed to show significance of the result for the motor-imagery condition of *feet* for a significance level of 5%, the results for the other conditions support the findings in [21].

In a second step, the main focus of this work was to show that these types of filters are applicable for session transfer of the spatial filters. Yet, while the findings show no clear trend to be superior to the spatial filters of the Infomax ICA used for session transfer, they are shown to be significant for this particular use-case. One plausible explanation for this undetermined trend shown in Figure 5 might be the non-stationary nature of the estimated baseline trials. These baseline trials were used to initialise the LCMV beamformer for the associated sessions. In order to account for non-stationarity, the authors suggest methods that transfer structure from the spatial covariance matrices over sessions. Possible choices for spatial covariance structure session transfer are the Riemannian Potato introduced by [3]. Furthermore and finally, this work aims at real life BCIs and hence it is insomuch important to test these results online without time-locked stimuli. This way, the results are more meaningful and valuable to the concept of real life BCIs.

List of Figures

1	External devices that may be controlled with robust Brain-Computer Interfaces in the future.	3
2	Overview of the five major brain rhythms related to functional patterns in the electroencephalogram.	4
3	Morphology of artefactual sources contaminating trials in the time domain.	5
4	Non-Invasive Brain-Computer Interface (BCI) online pipeline assumed throughout this work.	6
5	Resulting decoding accuracy for each session averaged over all subjects.	11
6	Two-sided pairwise permutation test to test for the validity of the empirical results.	12

List of Tables

2	General overview about the subjects, that is their identifier, age, gender, previous EEG experience, their motor skills, and the number of sessions being recorded.	10
---	---	----

List of Algorithms

1	Infomax ICA	8
2	LCMV SPBF	9

References

- [1] R. D. Adams, M. Victor, A. H. Ropper, and R. B. Daroff. Principles of neurology. *Cognitive and Behavioral Neurology*, 10(3):220, 1997.
- [2] B. Z. Allison and C. Neuper. Could anyone use a bci? In *Brain-computer interfaces*, pages 35–54. Springer, 2010.
- [3] A. Barachant, A. Andreev, and M. Congedo. The riemannian potato: an automatic and adaptive artifact detection method for online experiments using riemannian geometry. In *TOBI Workshop IV*, pages 19–20, 2013.
- [4] E. Başar, C. Başar-Eroglu, S. Karakaş, and M. Schürmann. Gamma, alpha, delta, and theta oscillations govern cognitive processes. *International Journal of Psychophysiology*, 39(2–3):241 – 248, 2001.
- [5] A. J. Bell and T. J. Sejnowski. An information-maximization approach to blind separation and blind deconvolution. *Neural computation*, 7(6):1129–1159, 1995.
- [6] H. Berger. Über das elektrenkephalogramm des menschen. *European Archives of Psychiatry and Clinical Neuroscience*, 87(1):527–570, 1929.
- [7] O. Bertrand, F. Perrin, and J. Pernier. A theoretical justification of the average reference in topographic evoked potential studies. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, 62(6):462–464, 1985.
- [8] N. Birbaumer, A. R. Murguialday, M. Wildgruber, and L. G. Cohen. Brain-computer-interface (bci) in paralysis. In *The European Image of God and Man*, pages 483–492. Brill, 2010.
- [9] C. M. Bishop. Pattern recognition. *Machine Learning*, 128, 2006.
- [10] C. Brunner, M. Naeem, R. Leeb, B. Graimann, and G. Pfurtscheller. Spatial filtering and selection of optimized components in four class motor imagery eeg data using independent components analysis. *Pattern Recognition Letters*, 28(8):957–964, 2007.
- [11] G. Buzsaki. *Rhythms of the Brain*. Oxford University Press, 2006.
- [12] R. Caton. Electrical currents of the brain. *The Journal of Nervous and Mental Disease*, 2(4):610, 1875.
- [13] Y. Chen, A. Wiesel, Y. C. Eldar, and A. O. Hero. Shrinkage algorithms for mmse covariance estimation. *IEEE Transactions on Signal Processing*, 58(10):5016–5029, 2010.
- [14] A. Coenen, E. Fine, and O. Zayachkivska. Adolf beck: A forgotten pioneer in electroencephalography. *Journal of the History of the Neurosciences*, 23(3):276–286, 2014.
- [15] P. Comon. Independent component analysis, a new concept? *Signal processing*, 36(3):287–314, 1994.
- [16] R. Q. Cracco and J. B. Cracco. Somatosensory evoked potential in man: far field potentials. *Electroencephalography and clinical neurophysiology*, 41(5):460–466, 1976.
- [17] J. J. Daly and J. R. Wolpaw. Brain–computer interfaces in neurological rehabilitation. *The Lancet Neurology*, 7(11):1032–1043, 2008.
- [18] J. French, P. Williamson, V. Thadani, T. Darcey, R. Mattson, S. Spencer, and D. Spencer. Characteristics of medial temporal lobe epilepsy: I. results of history and physical examination. *Annals of neurology*, 34(6):774–780, 1993.
- [19] P. Gloor. Hans berger on electroencephalography. *American Journal of EEG Technology*, 9(1):1–8, 1969.
- [20] C. Gouy-Pailler, M. Congedo, C. Brunner, C. Jutten, and G. Pfurtscheller. Nonstationary brain source separation for multiclass motor imagery. *IEEE transactions on Biomedical Engineering*, 57(2):469–478, 2010.
- [21] M. Grosse-Wentrup and M. Buss. Overcomplete independent component analysis via linearly constrained minimum variance spatial filtering. *The Journal of VLSI Signal Processing Systems for Signal, Image, and Video Technology*, 48(1-2):161–171, 2007.
- [22] M. Grosse-Wentrup, C. Liefhold, K. Gramann, and M. Buss. Beamforming in noninvasive brain–computer interfaces. *IEEE Transactions on Biomedical Engineering*, 56(4):1209–1219, 2009.

-
- [23] M. Grosse-Wentrup and B. Schölkopf. A review of performance variations in smr-based brain- computer interfaces (bcis). In *Brain-Computer Interface Research*, pages 39–51. Springer, 2013.
- [24] L. F. Haas. Hans berger (1873–1941), richard caton (1842–1926), and electroencephalography. *Journal of Neurology, Neurosurgery & Psychiatry*, 74(1):9–9, 2003.
- [25] A. Hyvärinen and E. Oja. Independent component analysis: algorithms and applications. *Neural networks*, 13(4):411–430, 2000.
- [26] International Business Times (IBTimes). Exoskeletons v wheelchairs: Disability advocates clash with futurists over 'offensive' solution, 2015. [Online; accessed September 20, 2016].
- [27] D. Jewett and J. S. Willston. Auditory-evoked far fields averaged from the scalp of humans. *Brain*, 94(4):681–696, 1971.
- [28] J. F. Keithley. *The story of electrical and magnetic measurements: from 500 BC to the 1940s*. John Wiley & Sons, 1999.
- [29] A. Kübler and N. Birbaumer. Brain–computer interfaces and communication in paralysis: extinction of goal directed thinking in completely paralysed patients? *Clinical neurophysiology*, 119(11):2658–2666, 2008.
- [30] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi. A review of classification algorithms for eeg-based brain–computer interfaces. *Journal of neural engineering*, 4(2):R1, 2007.
- [31] D. J. McFarland, L. M. McCane, S. V. David, and J. R. Wolpaw. Spatial filter selection for EEG-based communication. *Electroencephalography and clinical Neurophysiology*, 103(3):386–394, 1997.
- [32] M. Naeem, C. Brunner, R. Leeb, B. Graimann, and G. Pfurtscheller. Seperability of four-class motor imagery data using independent components analysis. *Journal of neural engineering*, 3(3):208, 2006.
- [33] E. Niedermeyer and F. L. da Silva. *Electroencephalography: basic principles, clinical applications, and related fields*. Lippincott Williams & Wilkins, 2005.
- [34] P. L. Nunez and R. Srinivasan. *Electric fields of the brain: the neurophysics of EEG*. Oxford University Press, USA, 2006.
- [35] J. Onton, M. Westerfield, J. Townsend, and S. Makeig. Imaging human EEG dynamics using independent component analysis. *Neuroscience & Biobehavioral Reviews*, 30(6):808–822, 2006.
- [36] G. Pfurtscheller and F. L. Da Silva. Event-related eeg/meg synchronization and desynchronization: basic principles. *Clinical neurophysiology*, 110(11):1842–1857, 1999.
- [37] G. Pfurtscheller, D. Flotzinger, W. Mohl, and M. Peltoranta. Prediction of the side of hand movements from single-trial multi-channel eeg data using neural networks. *Electroencephalography and clinical Neurophysiology*, 82(4):313–315, 1992.
- [38] G. Pfurtscheller and C. Neuper. Motor imagery and direct brain-computer communication. *Proceedings of the IEEE*, 89(7):1123–1134, 2001.
- [39] G. Pfurtscheller, C. Neuper, and N. Birbaumer. Human brain-computer interface. *Motor cortex in voluntary movements*, pages 367–401, 2005.
- [40] G. Pires, U. Nunes, and M. Castelo-Branco. Statistical spatial filtering for a p300-based bci: tests in able-bodied, and patients with cerebral palsy and amyotrophic lateral sclerosis. *Journal of neuroscience methods*, 195(2):270–281, 2011.
- [41] H. Ramoser, J. Muller-Gerking, and G. Pfurtscheller. Optimal spatial filtering of single trial eeg during imagined hand movement. *IEEE transactions on rehabilitation engineering*, 8(4):441–446, 2000.
- [42] I. J. Rampil. A primer for eeg signal processing in anesthesia. *The Journal of the American Society of Anesthesiologists*, 89(4):980–1002, 1998.
- [43] S. Sanei and J. A. Chambers. *EEG signal processing*. John Wiley & Sons, 2013.
- [44] V. Synek. Prognostically important eeg coma patterns in diffuse anoxic and traumatic encephalopathies in adults. *Journal of Clinical Neurophysiology*, 5(2):161–174, 1988.

-
- [45] M. Teplan. Fundamentals of eeg measurement. *Measurement science review*, 2(2):1–11, 2002.
- [46] J. A. Urigüen and B. Garcia-Zapirain. EEG artifact removal-state-of-the-art and guidelines. *J Neural Eng*, 12(3):031001, June 2015.
- [47] B. D. Van Veen, W. Van Drongelen, M. Yuchtman, and A. Suzuki. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *Biomedical Engineering, IEEE Transactions on*, 44(9):867–880, 1997.
- [48] W. G. Walter and V. Dovey. Electro-encephalography in cases of sub-cortical tumour. *Journal of Neurology, Neurosurgery & Psychiatry*, 7(3-4):57–65, 1944.
- [49] E. F. Wijdicks. The case against confirmatory tests for determining brain death in adults. *Neurology*, 75(1):77–83, 2010.
- [50] E. F. Wijdicks, P. N. Varelas, G. S. Gronseth, and D. M. Greer. Evidence-based guideline update: Determining brain death in adults report of the quality standards subcommittee of the american academy of neurology. *Neurology*, 74(23):1911–1918, 2010.
- [51] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan. Brain–computer interfaces for communication and control. *Clinical neurophysiology*, 113(6):767–791, 2002.
- [52] J. R. Wolpaw and D. J. McFarland. Multichannel eeg-based brain-computer communication. *Electroencephalography and clinical Neurophysiology*, 90(6):444–449, 1994.
- [53] L. Zhukov, D. Weinstein, and C. Johnson. Independent component analysis for eeg source localization. *IEEE Engineering in Medicine and Biology Magazine*, 19(3):87–96, 2000.

A Appendix

Subsequently, auxiliary mathematical concepts that are used within this thesis are derived and explained.

A.1 Whitening Transformation

The whitening transformation is often used to transform an arbitrary set of variables with known covariance onto a new set of uncorrelated variables whose covariance equals the identity matrix. Let $\mathbf{X} \in \mathbb{R}^{m \times n}$ be n -many arbitrary variables of length m with non-singular covariance matrix $\mathbf{C} \in \mathbb{R}^{m \times m}$ and zero mean. The whitened matrix is denoted as

$$\mathbf{Y} = \mathbf{C}^{-\frac{1}{2}} \mathbf{X} \quad (\text{A.1.1})$$

and with $\mathbb{E}[\mathbf{X}\mathbf{X}^\top] = \mathbf{C} = \mathbf{C}^{\frac{1}{2}} (\mathbf{C}^{\frac{1}{2}})^\top$ as a covariance matrix (symmetric and positive definite), it holds that

$$\text{cov}[\mathbf{Y}\mathbf{Y}^\top] \stackrel{\text{A.1.1}}{=} (\mathbf{C}^{-\frac{1}{2}}) \mathbb{E}[\mathbf{X}\mathbf{X}^\top] (\mathbf{C}^{-\frac{1}{2}})^\top = (\mathbf{C}^{-\frac{1}{2}}) \mathbf{C} (\mathbf{C}^{-\frac{1}{2}})^\top = \mathbf{I}_{m \times m} \quad (\text{A.1.2})$$

and hence is a white random vector. ■