

Optically pumped magnetometers for a brain-computer interface based on event-related desynchronization

Jan Zerfowski^{1,2,3}

✉ research@janzerfowski.de

Supervisors: Surjo R. Soekadar², Michael Tangermann¹

¹*Donders Institute for Brain, Cognition and Behaviour, Nijmegen, The Netherlands*

²*Charité – Universitätsmedizin Berlin, Berlin, Germany*

³*Physikalisch-Technische Bundesanstalt, Berlin, Germany*

Background Stroke is one of the leading causes of disability worldwide and often responsible for impairments of hand motor function. Rehabilitation and restoration of motor functions can be significantly improved using devices controlled by brain signals, so called brain-computer interfaces (BCIs). Most current BCI systems are based on electroencephalography (EEG), which provides only limited spatial resolution and thus limited versatility of control commands. Compared to EEG, optically pumped magnetometers (OPMs) measure cortical magnetic fields without contact to the scalp and provide a higher spatial resolution and bandwidth. In contrast to superconducting quantum interference device (SQUID)-based magnetoencephalography (MEG), OPMs have low maintenance cost and allow movement in the scanner, making them more applicable in clinical contexts.

Methods We quantify the signal characteristics of a commercially available OPM system (FieldLine Inc., USA) in terms of noise floor, dynamic range and bandwidth to verify its suitability for cortical measurements. We then develop an experiment contrasting resting and right hand grasping imagery to measure modulations of the sensorimotor rhythm (SMR) with 17 OPMs over the left motor cortex of 18 healthy participants. The BCI capabilities of the OPM acquisition system are evaluated with a modular near real-time classification pipeline, which provides visual feedback to the user.

Results The sensor characterization revealed a system noise floor of about $27 \text{ fT}/\sqrt{\text{Hz}}$ at 10 Hz, a bandwidth of 400 Hz and a dynamic range of $\pm 15 \text{ nT}$, fulfilling the minimum requirements for cortical measurements. In 10 of 16 eligible participants, a difference in SMR power between resting and grasping condition could be identified. We show that OPMs are suitable to measure SMR modulations in near real-time and that the classification performance of our pipeline significantly exceeds chance level.

Discussion OPMs allow for the online quantification of voluntary modulations of the sensorimotor rhythm on single-trial basis, a central requirement for many BCI systems used in the rehabilitation of stroke survivors. With their higher spatial resolution compared to EEG, OPMs could be used for more complex classification paradigms and ultimately facilitate a development towards more versatile BCI applications. The increasing availability and sensitivity of commercialized OPM systems allows for the exploration of MEG in new research areas. OPMs are projected to become an important tool in the field of cognitive neuroscience within the next few years.

Keywords: brain-computer interface, optically pumped magnetometer, event-related desynchronization, sensorimotor rhythm



Acknowledgments

I wish to acknowledge everyone who supported me during my Master's studies and throughout this project, academically, financially and morally:

First and foremost, I want to thank Prof. Surjo Soekadar for giving me the opportunity to join his lab and allowing me to pursue this project without any prior knowledge of BCIs or OPMs. He achieved a perfect balance between endorsing my sense to explore the field and guiding me in finding answers to all my questions. Surjo also brought me in touch with Michael Tangermann and I am very grateful for having him as my second supervisor. His great support and feedback throughout the whole project helped me immensely and certainly increased the scientific rigor of this project to a level I could not have achieved otherwise.

A warm thank you goes out to all members of the Clinical Neurotechnology Lab at the Charité – Universitätsmedizin Berlin, in particular Niels Peekhaus and Vincent Jonany who relentlessly discussed OPM-related questions with me, as well as the entire BCI Team who always provided good advice on all my ERD-related questions.

This project would have never been possible without the help from the Optical Magnetometry group at the Physikalisch-Technische Bundesanstalt. They trusted me with using their facilities from day one and made me feel welcome from the start. Most notably, Tilmann Sander-Thömmes and Thomas Middelmann supported me in submitting my first conference entry and provided great insight into the use of OPMs.

A big thank you also goes out to our collaborators in the US: Having weekly exchange meetings with the MEG Core Facility group at the National Institute of Health, namely Stephen Robinson, Allison Nugent and Amaia Benitez, massively improved my understanding of the physics of OPMs and helped prevent avoidable mistakes. The colleagues at FieldLine Inc. also provided deep insight into the sensor technology and always served us with very quick help despite the time difference and unexpected problems we experienced.

Lastly, I am very grateful to Benedikt Ehinger, who managed to re-spark my interest in statistics and showed me how much I don't know yet.

I could have never done this research project without the ideational and financial scholarship from ev. Studienwerk Villigst, which supported me throughout my entire Cognitive Science Bachelor's and Master's degree. I have met great people along the way and am excited to stay in the alumni network and provide academic guidance to young scholarship holders.

Finally, but perhaps most importantly, I want to express the gratefulness I feel towards my family and friends. Social contacts fell short during the COVID-19 pandemic, and only this helped me realize how important it is to value those humans close to my heart. My parents, Olaf and Monika, always encouraged me to go my own way and never ceased to trust in my intuition. They also supported me financially throughout the course of my studies and I cannot put into words how much they have helped me by doing so. I am also incredibly thankful for the time I spent with Niklas Kitzmann, Anastasia Roberts, Chiara Caporuscio and Maximilian Mühlbauer, who welcomed me so warmly when I moved to Berlin in 2020 and continue to make me feel truly at home.

Moving to Berlin to pursue the research with the technology and people I admire has been incredibly rewarding. I am excited to strengthen the bonds with all the people I have met and worked with throughout this research project and during my time in Berlin.

Contents

1. Introduction	4
1.1. Literature Review	5
2. Methods	11
2.1. OPM-MEG System	11
2.2. Sensor Characterization	13
2.3. Participants	14
2.4. Experimental Setup	14
2.5. Data Acquisition	17
2.6. Offline Analysis	17
2.7. Online Analysis	22
3. Results	27
3.1. Sensor Characterization	27
3.2. Offline Analysis	31
3.3. Online Analysis	34
3.4. General Findings	37
4. Discussion	39
5. Conclusion	47
Attributions and Funding	49
Glossary	50
Bibliography	51
Appendix A. Availability of the Source Code and Recorded Data	59
A.1. Python Modules in the Public Domain	60
A.2. BeamBCI	61
Appendix B. Scientific Contributions	62
Appendix C. Methods	63
C.1. Setup for the Characterization Measurements	63
C.2. Data Acquisition Software	64
C.3. Data Analysis	64
Appendix D. Results	70
D.1. Results of the Characterization Measurements	70
D.2. Offline Analysis	73
D.3. Online Analysis	92
Appendix E. Miscellaneous	94
E.1. Motor Imagery Training	94

1. Introduction

Stroke and other neurological disorders are one of the leading causes of disability and death worldwide (World Health Organization & World Bank, 2011; Feigin et al., 2017). The resulting impairments, in particular ones of hand and arm motor function, drastically reduce the quality of life of affected individuals. Restoration and rehabilitation of motor function in stroke survivors has been shown to improve when brain signals are used to control external devices like exoskeleton hands or wheelchairs (Birbaumer & Cohen, 2007; McFarland et al., 2017; Papadopoulos et al., 2022). Such systems, called brain-computer interfaces (BCI), translate measurements of cortical signals into control commands and often use electroencephalography (EEG) to measure the brain activity.

However, the recent development of spin-exchange relaxation free (SERF) optically pumped magnetometers (OPM) shifts the measurement of magnetic signals induced by brain activity into the focus of BCI research. The high accuracy and increased spatial resolution of OPMs compared with EEG and superconducting quantum interference device (SQUID)-based magnetoencephalography (MEG) (Boto et al., 2016), could enable the classification of more complex signals and greatly advance the development of BCI systems.

The body of literature examining OPMs for BCI applications is currently very sparse (see Literature Review). Therefore, thorough studies of basic paradigms and comparisons with established methods are needed before the field can turn to more complex problems.

This study investigated if a commercially available OPM system can measure voluntary modulations of oscillatory activity over the motor cortex and classify the data in near real-time. These modulations are also known as event-related desynchronization/synchronization (ERD/ERS) of the sensorimotor rhythm (SMR) and have been used to control external devices in the past. Even though it has been shown that cortical signals can be measured with OPMs and analyzed in near real-time (Wittevrongel et al., 2021; Zerfowski et al., 2021), so far no rigorous feasibility study of OPMs for an ERD-based BCI has been published at the time of writing.

To develop an ERD-based BCI, multiple research questions were considered separately in this study to examine if OPMs fulfill the basic requirements for such a system.

1. What are the properties and limitations of the OPM system in the available magnetically shielded environment in terms of noise floor, bandwidth and dynamic range?
2. Are the OPMs capable of measuring motor imagery event-related desynchronization in healthy participants and does the spatial distribution of the measured signals align with physiological assumptions?
3. Can the acquired data be analyzed and classified in near real-time such that visual feedback can be given to the user in the shielded environment? What is the performance of such a classification system?
4. What insights were gained while conducting the study? How could these inform the planning of future measurements?

To characterize the commercial OPMs available for this study, external magnetic fields with varying amplitudes and frequencies were applied to the sensors to determine their response characteristics. The hypotheses were that the OPM system has a noise floor of around $15 \text{ fT}/\sqrt{\text{Hz}}$ at 10 Hz with a 3 dB bandwidth of up to 300 Hz. Further, the sensors should retain an approximately linear response for signals of up to $\pm 15 \text{ nT}$. These assumptions were based on personal

1. Introduction

communication with the manufacturer (FieldLine Inc., Boulder, CO, USA) and pertain to the so called “closed loop” mode of operation.

To address the second and third research questions, an experiment involving the presentation of visual cues and feedback to healthy participants was developed: Participants were cued to *relax* or imagine a continuous right hand grasping motion (*close*) in five blocks, with 15 trials per condition — 9 without and 6 with visual feedback. The experiment was conducted on 18 healthy participants seated in a magnetically shielded room using 17 optically pumped magnetometers placed around the EEG electrode location C3.

The data recorded for no-feedback trials were then used to answer the second research question by analyzing the oscillatory amplitude differences in the alpha band (8 Hz to 16 Hz) between *relax* and *close* trials. Based on findings with EEG-BCIs, it was hypothesized that OPMs are capable of measuring modulations of the sensorimotor rhythm (SMR) in the alpha band which occur in the motor cortex during the imagination of movement. Another hypothesis was that the ERD source lies in a small area close to electrode location C3, leading to a dipolar magnetic field pattern around the source with interindividual differences in location, extent and direction.

To address the third research objective, a processing pipeline to analyze and classify the acquired OPM data in near real-time was developed. It was hypothesized that continuous data from the OPM system can be acquired and processed on an experiment computer such that feedback can be provided to the participant within the temporal dynamics of ERD. The goal was then to demonstrate that such a processing and classification pipeline performs above chance level and can be individually trained for multiple subjects. Since the pipeline was retrained on previous data in each block, an improvement of the classification performance over time was expected, which was examined last.

The fourth research question concerned the general findings that were obtained while conducting the study. Since the most published work on OPM-MEG is based on the data of only a few participants, the objective was here to evaluate the functionality and efficiency of the used OPM setup. The insights gained through such evaluation could then inform the planning and conduction of experiments in the future with potentially even more participants or different feedback modalities.

This thesis consists of five chapters whose contents are outlined in the following:
Upcoming is a literature review and subsumption of the current state of BCI and OPM research to provide context for this work and motivate what gaps it aims to fill in the field. Chapter 2 – Methods describes the hardware and the setup used for sensor characterization. It further explains the experimental paradigm as well as the data acquisition and analysis steps. Chapter 3 – Results presents results of the sensor characterization in terms of noise floor, bandwidth and dynamic range. It then focuses on the offline analysis before describing the results of the online analysis and general findings. In chapter 4 – Discussion the presented results are discussed and recommendations for future studies are formulated before this work is concluded with chapter 5.

1.1. Literature Review

Stroke and BCI The World Health Organization’s World report on disability (2011) lists stroke as one of the main contributors for years lived with disability (YLD). Projections indicate that non-communicable disease-related YLDs will increase in the future, explained by general population aging and changing lifestyles (World Health Organization & World Bank, 2011; Feigin et al., 2017).

The most common impairment after stroke is motor impairment which usually affects muscle control on one side of the body (Lawrence et al., 2001), posing a heavy burden for affected

1. Introduction

individuals, their families, and friends. Early interventions are the most effective treatment option for motor impairments but they are often time-consuming and tiring, leading to low compliance (Langhorne et al., 2009). Many established treatment strategies require some remaining voluntary movement ability, rendering them inapplicable for 30 % to 50 % of stroke survivors (Ushiba & Soekadar, 2016). It is therefore essential to develop effective treatment options for severe motor impairments after stroke and other neurological diseases.

In the past years, many studies showed that neural recovery is promoted and the restoration and rehabilitation of motor function is improved using feedback and control of external devices through cortical signals (Pfurtscheller et al., 2000; Soekadar et al., 2011; Ramos-Murguialday et al., 2013; Ushiba & Soekadar, 2016; McFarland et al., 2017; Papadopoulos et al., 2022). Cortical activity can be recorded through different means, the most accurate being invasive technologies like electrocorticography (ECoG). However, non-invasive means are preferable for rehabilitation to reduce the risk of infections and complications, in particular in patient groups with preexisting underlying conditions. Electroencephalography (EEG), as the most prominent non-invasive technology, has been used extensively in the past for the recording and translation of brain activity into control signals. However, the low conductivity of the skull in between cortical source and electrodes induces a current spread and leads to a relatively low spatial resolution in the cm range. This effect is also known as “volume conduction” and can be partly resolved by increasing the number of electrodes, which results in tedious preparation procedures exceeding 1 h for systems with 128 electrodes.

In contrast to EEG, magnetoencephalography (MEG) systems measure the *magnetic fields* induced by cortical activity and are less subject to volume conduction effects. The currently most common technology for measuring MEG are superconducting quantum interference devices, short SQUIDs. The major disadvantage of SQUIDs is the requirement for cryogenic cooling inside a dewar, resulting in high maintenance cost and limited patient populations due to rigid and static sensor configurations. Other technologies like functional magnetic resonance imaging (fMRI) or functional near-infrared spectroscopy (fNIRS) rely on the measurement of oxygenated hemoglobin and thus lack the needed temporal resolution to measure oscillatory brain activity.

Systems to translate cortical signal into feedback or control signals of external devices are also called brain-computer interfaces (BCI, or brain-machine interface, BMI). The most common class of BCIs for motor rehabilitation is based on event-related synchronizations/desynchronizations (ERS/ERD) (Ushiba & Soekadar, 2016). ERS/ERD can be measured as amplitude modulations of the sensorimotor rhythms (SMR) over the motor cortex related to the processing of sensory input or preparation and execution of motor output (Pfurtscheller & Aranibar, 1979). The so called mu rhythm is an SMR which occurs along the rolandic fissure in the alpha frequency band (8 Hz to 16 Hz) and has been used in many BCI applications since it can be voluntarily modulated by most people. A contralateral desynchronization (amplitude downregulation) of the mu rhythm can be induced by motor execution (Pfurtscheller & Aranibar, 1979), but also motor imagery (MI) in healthy users (Pfurtscheller, Neuper, Flotzinger, et al., 1997) and intended movement in paralyzed patients (Kübler et al., 2005).

However, not all humans present with an equally strong SMR modulation during MI, which has led to the controversial concept of “BCI illiteracy”. Some (EEG-) BCI studies attribute the inability to control BCI systems to about 15 % to 30 % of all users (Dickhaus et al., 2009; Allison & Neuper, 2010), but both the term and the estimates have been discussed critically (Thompson, 2019). In a large-scale screening study, Sannelli et al. (2019) developed a categorization of BCI users, by which about 40 % of the users belong to Cat. II or Cat. III. Users in Cat. II (17%) showed poor performance in feedback trials (< 70 %) due to their unstable and sometimes short (< 2 s) SMR modulations. Further, the SMR modulations of Cat. II users sometimes occurred more parietally than over C3/C4 electrode positions. Users in Cat. III (22.5 %) displayed only

1. Introduction

weak SMR with unstable frequencies or modulation in unexpected locations (e.g., ipsilateral). Sannelli et al. (2019) suggested the term “BCI inefficiency” instead of “BCI illiteracy” to avoid laying the responsibility on the user, rather than on the system.

For the perception of fluent control, the delay between user intention and reaction time of the BCI is an important parameter to assess the usability of such system. A 3-second window of temporal integration in sensorimotor control was proposed by Mates et al. (1994) and has been used to argue that the time to exert feedback control should lie below 3 s (Crea et al., 2018). The effect onset of ERD usually is on the order of 1 s (Graimann et al., 2002), which should be possible to verify by evaluating the timecourse of the ERD in the current experiment.

Most publications using ERD-based BCI systems implement paradigms using different combinations of hand, feet and tongue motor execution or imagery. However, the spatial resolution and signal-to-noise ratio usually limit EEG-based systems to one or two degrees of freedom (Abiri et al., 2018). The current work aimed to provide a proof of principle for an OPM-based BCI, but did not aim to maximize classification accuracy. Therefore, the use of source localization or adaptive deep learning methods, which successfully improved the performance of EEG-BCI recently, shall not be attempted here. This is also because different publications often choose diverging performance metrics highly dependent on the task at hand and because it is often difficult to verify the particular features used for classification by complex machine learning algorithms. To compare the BCI developed for this work, the 2014 report by Witkowski et al. has been selected for its similar paradigm and focus on reliability and safety (see section 2.7).

Lal et al. (2005) were the first to demonstrate an MEG-based BCI. They employed a motor imagery paradigm and showed that even with only 7 sensors a good classification performance can be achieved. MEG has since then been used occasionally to realize BCI applications (Lal et al., 2005; Mellinger et al., 2007; Soekadar et al., 2011; Lin et al., 2013; Foldes et al., 2015), but could not overtake EEG due its versatility and wide availability. In a single-trial discrimination study of individual finger movements, Quandt et al. (2012) have shown that MEG is superior over EEG in terms of decoding accuracy. Since the distance of cortical finger representations is on the order of a few mm, they speculated that the variation in magnetic flux patterns could be due to the strong curvature of the cortical sheet. The magnetic field pattern that is expected to arise from an ERD confined to a small area on the cortical sheet is dipolar, which has been confirmed in multiple previous articles (N. J. Hill et al., 2006; Mellinger et al., 2007; Schaworonkow & Nikulin, 2022)

Optically Pumped Magnetometers The principle of using optically pumped atoms to measure weak magnetic fields has been first described in the 1950s by Dehmelt (1957) and developed further by Bell and Bloom (1957). A quantum property known as atomic spin of an alkali metal vapor can be manipulated through the absorption and emission of energy from a light source with a precisely selected wavelength and polarization, inducing a magnetically sensitive state in the vapor. This magnetically sensitive state can be read out and used to measure weak magnetic fields (Bell & Bloom, 1957; Bloom, 1962). Since these first publications, the miniaturization of optical components and lasers, as well as the use of modulation fields, magnetic shielding and high-density vapors has facilitated the development of so called spin-exchange relaxation free (SERF) optically pumped magnetometers (OPM). There is a variety of other schemes to realize OPMs with different sensitivity and dynamic range characteristics, but in this work the term OPM always refers to SERF-OPM technology. For a comprehensible review of the physical and technical principles of SERF-OPMs, the reader is referred to the publication by Tierney et al. (2019).

Xia et al. (2006) were the first to demonstrate a SERF-OPM that can record magnetoencephalography with a noise level of about $3.5 \text{ fT}/\sqrt{\text{Hz}}$ at 10 Hz, comparable to commercial SQUID-MEGs

1. Introduction

(see also Knappe et al., 2010). This finding has sparked the development of commercial OPM systems and facilitated the publication of many other studies investigating sensor miniaturization (Sander et al., 2012), source localization (Boto et al., 2016), wearability (Boto et al., 2018), speech decoding (Dash et al., 2020) and functional connectivity (Boto et al., 2021).

Each magnetometer is an individual device which can be distributed flexibly around the measurement volume and can be placed much closer to the skull, reducing the effective source-sensor distance to the mm-range (Sander et al., 2012). Participants can move during the measurement given low static magnetic fields and gradients in the shielded environment, and OPMs suffer much less from muscle-related movement artifacts (Boto et al., 2018). In contrast to EEG, magnetic fields penetrate both skull and air almost unhindered and therefore strip the need for direct skin contact, the application of conductive gels, and tedious preparation procedures.

One major drawback of the first commercial (SERF-)OPMs is their low intrinsic linear range ($\sim 1\%$ deviation from linear response at $\pm 1\text{ nT}$) (Tierney et al., 2019). For comparison, the earth's magnetic field has a magnitude of approximately $50\text{ }\mu\text{T}$ (in Europe and USA, Alken et al., 2021), which is prohibitively large for measurements without strong magnetic shielding. Different methods to achieve sufficient magnetic shielding have been proposed, using passive (He et al., 2019) and active (Bork et al., 2001; He et al., 2019; Holmes et al., 2019; Iivanainen et al., 2019; Robinson et al., 2022) components. The current commercially available OPMs also feature triaxial field coils which are calibrated before each measurement to counterbalance remnant magnetic fields. The devices by the two most prominent manufacturers can compensate up to $\pm 50\text{ nT}$ (QuSpin Inc., Shah et al., 2018) and $\pm 100\text{ nT}$ (FieldLine Inc., personal communication) of static external fields.

However, the static field offsets can not compensate transient field changes induced by movement or rotation in remnant magnetic fields, power-line noise and other external noise sources like trains passing nearby (Holmes et al., 2019), elevators or mechanical vibrations (Okada et al., 2016). To allow for the measurement of larger fields along the sensitive axis, FieldLine Inc. implemented the so called “closed loop” mode in which the compensation coil currents are dynamically adjusted (see section 2.1 for details). Using this mode, linearity for fields of up to $\pm 15\text{ nT}$ can be achieved, according to the manufacturer (cp. subsection 3.1.2).

The closed loop mode is currently only implemented for axial fields; if transverse fields (perpendicular to the measurement axis) grow too large, they can heavily influence the signal through an effect coined cross-axis projection error (sensitive axis rotation of $3.33^\circ/\text{nT}$, Borna et al., 2022). These effects can be accounted for by using stronger and active shielding, in software (Tierney et al., 2021) or by using motion tracking systems (Seymour et al., 2022). The most recent publication by Robinson et al. (2022) implements a compensation in hardware based on a projection of measured ambient fields and could potentially be adopted by manufacturers.

The accurate localization and reconstruction of source activity is an issue that is currently under heavy investigation by OPM-MEG researchers. While there are long-established approaches for EEG and SQUID-MEG co-registration, the non-standardized acquisition systems and flexibility in OPM placement complicate the localization of sensors relative to each other and the skull/brain. For accurate source reconstruction, a precise localization of the sensors' sensitive volumes and axes in the sub-mm range is required and most labs resort to implementing individual solutions: Rigid, additively manufactured helmets or sensor fixtures are most common in generic (Zetter et al., 2019; R. M. Hill et al., 2020; Nardelli et al., 2020) and individualized (Sander et al., 2020; Marhl et al., 2022) variants and can in principle achieve accurate source reconstruction using beamformers (Boto et al., 2016; Brookes et al., 2021). OPM manufacturers are now beginning to address this issue by developing systems to localize the sensors in a number of different helmet models to achieve flexible and reliable solutions.

Another disadvantage of OPM systems compared to established SQUID-MEG systems is the

1. Introduction

current lack of approvals as medical device for commercially available systems. To date, there are no OPM solutions that are approved by either the European Medical Agency (EMA, Europe) or the US Food & Drug Administration (FDA, USA) for clinical use. Manufacturers are addressing this issue by applying for approval, but these processes are usually tedious and time-consuming. Until solved, OPM systems can thus only be used as research devices with an appropriate approval by responsible ethics committees, but are not ready for extended clinical use at this point.

The OPM-MEG community is currently working together with manufacturers to standardize data formats and establish metadata specifications for sharing and analyzing the data. Open source toolboxes like MNE and FieldTrip are making progress in implementing OPM-specific tutorials and analysis methods. Since the commercially available systems are mostly consisting of individually sold units or bundles with short product cycles, no best practices have been globally adopted yet. First review papers attempt to change this by suggesting data analysis schemes and interference suppression techniques (Seymour et al., 2022).

OPMs and BCI Currently, only few publications have explored the use of OPM technology in the field of BCI research: BCI systems require immediate access to the sensor data for processing, classification and providing feedback to the user in near real-time (also called *online*). The two most prominent manufacturers of OPMs for MEG are currently QuSpin Inc. (Louisville, CO, USA) and FieldLine Inc. (Boulder, CO, USA). Sensor systems by QuSpin Inc. internally digitize the data for access with a proprietary recording software and, in a second step, provide an output port for the user to record the signals. A digital data acquisition system (DAQ) can then be used for real-time analysis of the data. This system has been successfully used by Wittevrongel et al. (2021) with 48 sensors, achieving a 93 % decoding accuracy in a steady state visually evoked potential (SSVEP)-based mind-spelling paradigm. The sensor system by FieldLine Inc. provides a python-based application programming interface (API) through which the sensors can be controlled and the data read out in real-time. In the course of this study a paper was published in which this acquisition method was proven suitable for the classification of occipital alpha (eyes closed vs. eyes open) (Zerfowski et al., 2021). Two publications investigating ERD for BCI were published, but Paek et al. (2020) did not actually perform online classification and only reported offline results. A short conference paper by Fedosov et al. (2021) with comparable methodology to this work presented accuracies of around 0.72 (grasping MI) and 0.64 (relaxation) for online classification but included only a single trained participant.

The current body of literature suggests that OPMs are a new powerful tool to study the human brain and cognitive processes with greater precision and versatility than currently available technologies. It is reasonable to believe that OPM-based BCI systems could surpass the performance of state-of-the-art EEG-based BCI (Papadopoulos et al., 2022) and BCIs based on conventional SQUID-MEG (Dash et al., 2020; Marhl et al., 2021). Many directions to improve the signal quality and robustness of OPM-based measurements are currently being explored, including hardware solutions to reduce remnant magnetic fields, methods for dynamic field compensation, noise cancellation and advanced data processing to identify and eliminate artifacts. To date, there is a lack of standardized approaches and specifications for data recording and analysis, but the rapidly growing OPM community seems to be inclined towards open source tools and major toolboxes are adapting to their users' needs.

This work aims to contribute to the currently small body of literature in the field of OPM-based BCI technology. Despite the first promising publications substantiating the usefulness of OPMs in neurotechnology, more fundamental studies are needed to create a strong base which complex paradigms can build upon. To investigate the central aspects required for an OPM-based BCI,

1. Introduction

event-related desynchronization over the motor cortex induced by hand motor imagery was used in the current study. Designing an experiment with a MI task stripped the need to measure other modalities in parallel, e.g., electromyography (EMG) to measure muscle activity, and mostly eliminated participant movement. The present study reproduces a well-established EEG paradigm with a new sensor technology and thereby represents an important foundational work for the establishment of OPMs in neuroscience.

2. Methods

2.1. OPM-MEG System

To measure the magnetic fields induced by neural activity in the sensorimotor cortex, a commercially available system with 17 single-axis SERF-OPM sensors (FieldLine Inc., Boulder, CO, USA, see Figure 2.1) was used. Synchronized calibration and sample times were achieved by daisy-chaining two separate sensor chassis with 12 magnetometers owned by Charité – Universitätsmedizin Berlin (Germany) and 5 owned by the Physikalisch-Technische Bundesanstalt (PTB, Berlin, Germany). The sampling rate was 1000 Hz with an estimated noise floor of $15 \text{ fT}/\sqrt{\text{Hz}}$ at 10 Hz and a 3 dB bandwidth of 300 Hz (personal communication). Since no official account of the technical specifications is available at the time of writing, these statements were verified in the first stage of this work (see section 3.1). According to the manufacturer, the built-in static field coils can compensate up to $\pm 100 \text{ nT}$ in all 3 axes determined at calibration time, far exceeding the expected remnant fields in the available magnetically shielded environment ($< 10 \text{ nT}$).

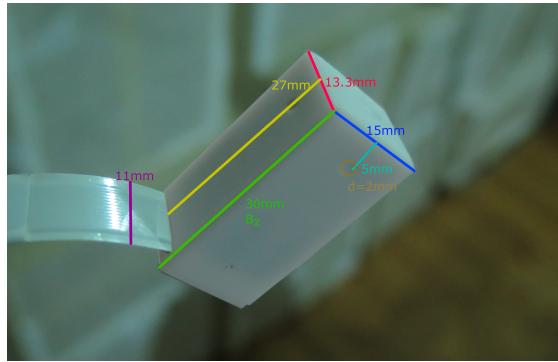


Figure 2.1. The OPMs sold by FieldLine Inc. have a footprint of $13.3 \times 15 \text{ mm}$ and a height of 30 mm

The sensors were operated in *closed loop* mode, a feature that increases the linear measurement range of SERF-OPMs from $\pm 1.5 \text{ nT}$ to up to $\pm 15 \text{ nT}$. To achieve this increased range, a closed loop feedback is implemented in hardware which dynamically adjusts the field compensation coil currents along the sensitive axis to bring it back into the zero-field regime. The output is then the field strength corresponding to the necessary compensation coil current (Nardelli et al., 2019). It is important to add that the compensation is only applied on the sensitive z -axis and therefore transversal fields can still degrade the measurement quality (termed cross-axis projection error, Borna et al., 2022). The current experiment did not involve voluntary movements and the healthy subjects were expected to sit still in the very low remnant fields of the shielded environment. Therefore, no degradation of signal quality due to this effect was expected. The closed loop mode also increases the gain for higher frequencies, supposedly increasing the open loop 3 dB bandwidth of the sensors from $\sim 120 \text{ Hz}$ to up to $\sim 300 \text{ Hz}$ in closed loop mode. Even though this can introduce a higher noise floor for high frequencies (see section 3.1) the manufacturer generally recommends the closed loop mode, in particular since high frequency components were irrelevant in this study.

2. Methods

2.1.1. Magnetic Shielding

The earth's magnetic field ($\sim 50 \mu\text{T}$) and other static and dynamic external sources of magnetic noise (elevators, moving doors, mechanical vibrations) are prohibitively large for the operation of SERF-OPM sensors without magnetic shielding. All measurements were conducted in the Ak3b magnetically shielded room (now VACOSHIELD Advanced, VACUUMSCHMELZE GmbH & Co. KG, Hanau, Germany) at PTB Berlin which features three layers of shielding (2 x Mu-metal + 1x aluminum). It has a footprint of 4 m by 3 m (height: 2.4 m) and achieves a shielding factor of at least 8000 at 10 Hz in the center. An additional active shielding circuit further reduces the remaining fields to below 10 nT and slow fluctuations ($< 3 \text{ Hz}$) usually remain below a few hundred pT.

For the characterization of bandwidth and dynamic range, the tube-shaped Twinleaf MS-2 (Twinleaf LLC, Plainsboro, NJ, USA) was used as an additional level of shielding inside of the MSR (see appendix C.1). It is too small for measurements on the human head (length \times diameter: $360 \times 180 \text{ mm}$) but it can hold multiple sensors for comparative measurements. Its shielding factor of 10^6 is afforded through 4 layers of Mu-metal and it includes a 3-axis coil system which can be used to apply homogeneous magnetic fields of known amplitude and frequency using a function generator and a power source.

The coil system was driven along the longitudinal (z) axis using a low-noise bipolar current supply (CSBA-10, Twinleaf LLC, Plainsboro, NJ, USA) controlled with an arbitrary waveform generator (Agilent 33500B series, Santa Clara, CA, USA). To generate the desired magnetic field amplitudes, the coil constant (105600 nT A^{-1}) and the conversion factor of the power source ($200 \mu\text{A V}^{-1}$) were used to derive the according settings for the waveform generator. The correct magnitude of the applied fields was verified using a commercially available Fluxgate magnetometer with a resolution of 0.1 nT (Fluxmaster, Stefan Mayer Instruments GmbH & Co. KG, Dinslaken, Germany).

2.1.2. Magnetometer setup

The sensor characterization required a mechanically and magnetically stable setup to avoid disturbances that could interfere with the data recordings. The tubular magnetic shield was screwed to a base plate on a cart which was locked in place in the center of the Ak3b. Previous experiments showed that connecting a grounding cable to the shield was needed to avoid the induction of unwanted magnetic fields into the system.

Since the sensors are sensitive to minute differences in magnetic field, only two sensors were placed in the center volume of the tubular shield where best field homogeneity is achieved. A custom sensor fixture to align the sensors along their measurement axis was built using LEGO[®] bricks (Lego A/S, Billund, Denmark) and could be slid into the tube and interlocked in place (see appendix C.1).

The characterization of the sensor noise floor was designed to resemble real participant measurement conditions and was therefore conducted in the Ak3b without further shielding. Due to maintenance operations on the chassis by the manufacturer at the time of measurement, only a set of 10 OPMs was available and placed in the head fixture described below.

To record motor-cortical activity, the sensors had to be located close to the participant's scalp around the C3 electrode location according to the 10-20 system. This was achieved by additively manufacturing a sensor fixture from Soft PLA based on a single head shape but flexible enough to fit on most adult skull sizes. The locations and orientations of the sensor slots in the grid are known with relatively high precision even though the fitting the grid to the participant's head

2. Methods

can introduce small perturbations. As the performed analyses were conducted per-participant and only in sensor space, these perturbations do not impair the interpretability of the results of this study. The OPMs were placed in the grid manually and their positions marked down for the drawing of topological plots.

The fixture placement was done manually over the EEG electrode location C3 (see 2.4.1), as the acquisition of structural MRI images for each participant would have exceeded the scope of this thesis. It has to be mentioned that, due to the grid's flexibility and the lack of co-registration, all 2D sensor-head projections are approximate and can only be used for a qualitative interpretation of the magnetic field patterns.

2.2. Sensor Characterization

Multiple setups were used to confirm the sensor specification unofficially provided by FieldLine Inc., and to quantify the system's signal characteristics. The recommended closed loop mode (Nardelli et al., 2019) attempts a trade-off between sensitivity and bandwidth (Tierney et al., 2021) while simultaneously increasing the dynamic range.

The sensor sensitivity is defined by the frequency dependent noise floor, i.e., the spectrum of the recorded data when no signal is applied to the sensors. These recordings were made with 10 sensors (due to maintenance procedures at the time of recording) in the Ak3b shielded room at PTB. Since the room exhibits slow drifts and attenuated external noise influences, the noise floor obtained here is specific to the *system* rather than each particular *sensor*. Data were recorded consecutively in open and closed loop mode for 4.5 min each. The spectrum analysis was then performed by applying Welch's method with a Hann-tapered sliding window of length 2.048 s with 1.024 s overlap. In a subsequent analysis using the same parameters, a 5 min recording of data in the closed loop mode was used to evaluate if there is a difference in system noise floor between the two available acquisition units.

To assess the dynamic range of the sensors, two different cases of strong field influences were considered and experiments designed for measurement inside the combination of Ak3b and Twinleaf MS-2: The first mode resembles a scenario where strong alternating fields occur along the sensor's measurement axis. Using the field coils in the MS-2 tube, an external sinusoidal field at 12 Hz with increasing amplitude was applied and compared to the measured signal. Applied peak amplitudes in the closed loop mode reached as high as 20 nT (step size 200 pT) while measurements in open loop mode were stopped at 5 nT (step size 50 pT).

The second scenario has even more practical relevance as it quantifies the attenuation of small fields under the influence of large static or slowly drifting offset fields. Here, external offset fields between -7 nT and 7 nT (step size 100 pT) were combined with a 100 pT sinusoid at 12 Hz.

The bandwidth of a sensor is commonly given as the 3 dB point, the frequency at which the amplitude of the measured signal is only $\sim 70.8\%$ of the applied amplitude (i.e., half its power). A sinusoidal field with a fixed amplitude of 50 pT was applied with increasing frequencies from 1 Hz to 490 Hz (Nyquist frequency at 500 Hz; 50 Hz was skipped to avoid interference with power line noise).

The dynamic range and bandwidth measurements were performed with only two randomly selected OPMs placed in the center of the MS-2 shielded tube along its most homogeneous axis (longitudinal z -axis) The measured amplitudes were determined by fitting a sine curve with 12 Hz to the signal with the Levenberg-Marquardt algorithm using scipy's `curve_fit()` function.

2. Methods

2.3. Participants

Participants for this study were recruited through Charité and PTB as well as through a public announcement. Participants were required to be right-handed as assessed through a reduced version of the Edinburgh Handedness Inventory (score > 60) (Oldfield, 1971; Veale, 2014) and have good command of either German or English. The eligible age-range for inclusion in this study was 18 to 60 years. All participants had received at least 2 doses of an approved vaccine against COVID-19 and provided a negative test result before entering the facilities as required by Charité – Universitätsmedizin Berlin and PTB. Ineligible for participation were individuals with previous neurological or psychological diseases or regular medication intake (excluding oral contraceptives). Participants wearing irremovable metal parts which induced large artifacts upon movement were also excluded.

Eighteen healthy, right-handed participants that matched the inclusion criteria were recruited through public announcements and internally through Charité and PTB communication channels. The appointments were scheduled through the open-source calendar tool (terminplaner4.dfn.de) hosted by the Deutsches Forschungsnetz (DFN) and volunteers had to provide their name and email address for communication. Appointment confirmations and invitations were then sent out via email individually with reminders of exclusion criteria and location information. A compensation of €12.5/h was offered to participants not associated with Charité. Three more individuals also undertook the measurement under the same protocol but did not match the inclusion criteria for this study (older than 60 years) and were therefore excluded from further analyses.

2.4. Experimental Setup

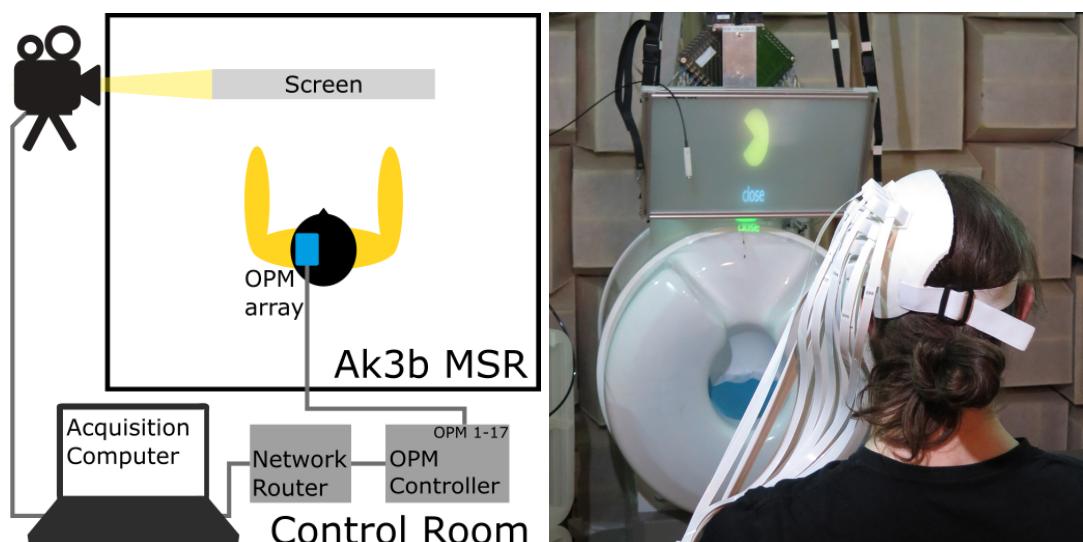


Figure 2.2. Experimental Setup. The user is facing a screen which the cues are projected on. The array of sensors is connected to the control chassis. The control chassis are sending data to the computer via LAN.

The main challenge of this study was the implementation of an online analysis pipeline for

2. Methods

the use with FieldLine's OPM system and feedback presentation in the magnetically shielded chamber.

As every electronic device induces magnetic field noise, potentially rendering the measurement unusable, the extensive experience of the PTB was very helpful in selecting components and designing the environment. A rear-projection of the visual stimuli onto a screen was realized by placing a projector outside the shielded room and redirecting the image through an aperture and two mirrors attached to the ceiling of the MSR. The bidirectional communication system consisted of a small microphone next to the screen and a piezoelectric speaker far away from the sensors which was disconnected while not in use. The OPM data acquisition devices were placed outside of the shielded chamber and the sensor's flat band cables pulled through holes in the wall. The participant was seated in a plastic chair right in the center of the shielded room where magnetic fields and gradients are lowest, with a distance to the screen of approximately 1.2 m.

2.4.1. Protocol

Participants were invited to PTB at a time of their choice and reminded to avoid clothing with irremovable metal parts. Individuals with corrected vision were asked to wear contact lenses since glasses often come with a frame containing metallic parts. After arrival, vaccination cards and tests results were checked and they were guided to the preparation room. Medical-grade face masks were worn at all times except during the experiment when the participant was alone in the MSR to avoid artifacts due to metallic nose pieces in the mask. Participants were provided with a background information sheet, given time to read and ask questions and informed that they could abort the experiment at any time without negative consequences. Compliance with inclusion and exclusion criteria was checked and participants were asked about previous experience with BCI experiments. The documents and forms provided to the participants have been made available online (see appendix A).

To provide an introduction to motor imagery (MI) and acquaint participants with the experiment, a short training established in the Clinical Neurotechnology lab was executed with them (appendix E.1) and the stimuli occurring during the experiment were demonstrated. The MI training gives instructions to execute a continuous grasping motion and gradually cease to move while keeping the imagination as vivid as possible. Participants were asked to not only visualize but also include haptic imagery of how the movement would feel (kinesthetic imagery) to generate stronger sensorimotor modulations (Roc et al., 2020). Then, participants were asked to follow the *relax* and *close* cues from the experiment for a few trials to get used to the timing of the experiment and try out the motor imagery.

When the participant felt acquainted with the task, they were guided into the Ak3b shielded room and seated comfortably in front of the projector screen with their arms resting on a padding to the sides of the body. The sensor fixture was placed centered over electrode location C3 according to the 10-20 system. If necessary, the sensors were taped to the fixture to reduce distance to the skull, since thick or frizzy hair would sometimes push the sensors out of their brackets. Swinging sensor cables upon head movements were avoided by fixing them to the chair or clothes of the participant. The participants were then asked if they feel fit for the experiment of ca. 30 min and advised again to communicate any discomfort during the experiment.

The experiment consisted of five blocks with 30 trials each. Nine *close* and nine *relax* cues in random order were presented without feedback during the first part of each block. Cues were presented for 5 s each, during which right hand grasping motor imagery or no motor imagery should be performed for the full period. The inter-trial interval (ITI) was randomized between 3 s and 5 s to prevent anticipation effects and leave sufficient time for baselining. The recorded data were automatically processed after the first part of each block and presented to the experimenter

2. Methods

for selection of the central ERD frequency (see subsection 2.6.4). If a distinct peak frequency could be found, it was fixed also for future blocks, otherwise a default frequency of 12 Hz was assumed until the next block. After the target SMR frequency was selected by the experimenter, the pipeline was trained on all previous trials and the second part of the block was started after verbal confirmation from the participant. The second part consisted of six trials per condition with visual feedback, again presented in randomized order. A break after each block allowed participants to readjust their seating, take a deep breath and decide when to continue with the next block.

The second part of each block contained 6 *relax* and 6 *close* trials in which visual feedback was provided to the participant. During *close* trials, the feedback consisted of a Pac-Man-shaped yellow circle with an initially missing sector of 60° (mouth) (see Figure 2.3). When a sample was correctly classified as *close*, the mouth closed with a speed of 20°/s (i.e., 3 s for a full close). Signals classified as *relax* led to a stop of the closing motion (rather than reversal) to avoid stress in the participant.

relax feedback was provided through a faded circle with a blurred border, initially starting in red. When a sample was correctly classified as *relax*, the circle's color slowly transitioned to green color within 3 s but halted the transition in case of wrong classifications. For both conditions the feedback was chosen to be simple and distinct, but not distracting.

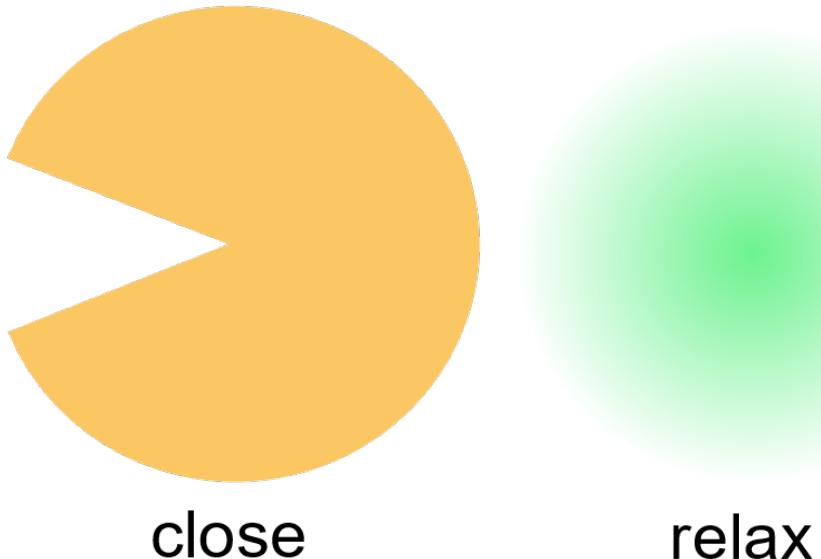


Figure 2.3. Visual Feedback. **Left panel:** Feedback during *close* trials, here open with $\sim 40^\circ$ after ~ 1 s of closing motion. **Right panel:** Feedback during *relax* trials, after full transition (3 s).

The protocol of this study on healthy human subjects was covered by the ethics approval EA1/338/20 given by the ethics committee at Charité – Universitätsmedizin Berlin as part of the Einstein Forschungsvorhaben 2019-558.

2. Methods

2.5. Data Acquisition

For the online classification of OPM sensor data, it was necessary to implement a pipeline which enables the acquisition and analysis in near real-time. In the course of this project and with similar requirements, such a system was built and presented in Zerfowski et al. (2021). In this work, the terms *online* or *near real-time* are preferred, since *real-time* implies a guaranteed maximum response time (often defined by CPU clock cycles). Technically, the requirement for an upper bound of the response time is also fulfilled here (samples are consistently processed faster than new samples arrive), but, since the software runs on Microsoft Windows, the processing time can vary on the order of ms due to the system's scheduler.

FieldLine Inc. provides two methods to acquire and record the sensor data which can be used in parallel. One solution is the FieldLine Recorder, a graphical user interface (GUI) to control the sensors, plot the incoming data and create recordings in the standardized `.fif`-format. Near real-time access to the samples is facilitated through an API written for python 3.8.

Since it is desirable to have the OPM data stream available immediately in a versatile format, LabStreamingLayer (LSL) was chosen to propagate samples between stages of the processing pipeline (and appendix C.2.1). To this end, a module to interface the FieldLine API with LSL was implemented (see appendix A.1.1 and Zerfowski et al., 2021). The module extends the functionality of the proprietary software as it allows to retrieve metadata like sensor serial numbers and calibration values. An application called BeamBCI was originally developed in the Clinical Neurotechnology Lab to facilitate the setting up of EEG-BCI experiments (see also section 2.7 and appendix A.2). Four new modules were written for BeamBCI and slight adaptations made to the codebase to enable the exact purpose of this project:

- Feed the acquired samples into the processing pipeline
- Visualize the data as it propagates through the pipeline
- Control the experiment and start pipeline training
- Present the cues and feedback to the user

For later analyses, the data from all modules were recorded simultaneously and could therefore be used to align trial onsets with the data, both online and offline. This was done using the LabRecorder (see subsection C.2.1) which saves data streams together with metadata in the `.xdf` file format.

2.6. Offline Analysis

2.6.1. Data Preprocessing and Cleaning

External noise sources, sensor errors and movement artifacts can obscure real cortical signal in the recorded data. The most reliable way for finding such artifacts is manual visual inspection, since automated methods currently mostly target movement artifacts and require more sensors.

Different classes of artifacts have been observed in the recorded data and can be pooled into four categories:

Movement artifacts Movement artifacts arise when the OPMs are rotated (moved) in a remnant (gradient) field and were by far the most common artifact in the recorded data. Artifact size would correspond to the range of gradients or remnant fields, in this case they ranged from 20 pT to 100 pT. Not only do movement artifacts conceal real signal if they are in a similar

2. Methods

frequency range, but they can also induce nonlinearities in the sensor or indicate inattention of the participant. If many sensors are available (which was not the case here), movement artifacts can be repaired by modeling interferences as homogeneous homogeneous fields (Tierney et al., 2021), using motion capturing systems (Seymour et al., 2021) or reference sensors for artificial gradiometry (Seymour et al., 2022, also for a review). Here, data containing movement artifacts were just marked as such and ignored in further analyses.

External noise Strong, nonstationary magnetic fields outside of the shielded room, particularly induced by banging doors or moving elevators cannot be completely blocked and will therefore be measured as slow or fast changes in magnetic field. These fields can be as large as several hundred pT in < 1 s and must be removed from the signal because their spectral signature can mask the cortical signal or induce nonlinearities. Most commonly “line noise” (at 50 Hz in the European power grid) is clearly visible in the power spectrum of all sensors. Slow field drifts occur in the MSR due to its low shielding factor at low frequencies (usually < 500 pT at < 0.3 Hz). Both effects are small and thus do not induce nonlinearities in the data, such that they were filtered out using a bandpass filter.

Sensor noise Several factors influence the noise floor of the sensors, including stability of the laser wavelength and power, photodiode noise, quantization effects, etc. Sensor noise is frequency dependent (see also appendix 3.1) and expected to be around $15 \text{ fT}/\sqrt{\text{Hz}}$ at ~ 10 Hz. To achieve the increased bandwidth in the closed loop mode, the sensor gain is increased for higher frequency bands. For some sensors this leads to a much higher noise floor for frequencies above ca. 60 Hz (see section 3.1). Since the performance of the OPMs in the alpha frequency was relatively good, no sensors needed to be excluded from the analysis.

Occasionally, single sensors were exhibiting strong (0.1 nT to 0.5 nT) and irregular (0.2 Hz to 3 Hz), single-sample spikes. The most likely explanation for this is due to cable grounding problems (personal communication with the manufacturer). Pilot data analyses indicated that the spectral signature of the spikes does not impair the data in the frequency of interest.

Muscle artifacts Due to the cubic decay of magnetic fields with distance to the source and missing effects of volume conduction, little to no muscle artifacts are to be expected in OPM data. Only small cardiac artifacts (periodic heart beats, < 10 pT) were visible for some participants. In principle it should be possible to eliminate these artifacts by means of Independent Component Analysis (ICA). In the current study no correction for heart beat artifacts was performed because of their faint nature and to reduce complexity of the preprocessing. Other muscle artifacts (blinking, saccades, swallowing, ...) could not be identified but should be investigated in future studies by means of electrooculography (EOG), electrocardiography, electromyography, etc. (see also Seymour et al., 2022).

To mark artifacts and periods of increased noise in the recorded data, the following steps were performed manually and separately on every dataset.

Loading data The recorded .xdf file was loaded into memory, the relevant LSL streams (OPM data and experimental marker) extracted and an mne.RawArray object containing the markers as mne.Annotations was created from the data. The code used to transform .xdf-files into MNE-Python compatible objects was made available under an open source license, see appendix A.1.3.

Automatic signal rejection Periods in which the recorded data exceeded a threshold of ± 2 nT were marked as bad to identify and exclude periods of potentially nonlinear data due to rapid field fluctuations or large movements.

2. Methods

Manual inspection The entire dataset was manually inspected with MNE’s RawArray.plot() in a window of 10 s. Transient jumps of $> 30 \text{ pT}$ (induced by external noise and movement artifacts) on all channels were marked as bad.

Filtering A copy of the data was created and filtered with a two-pass, zero-phase (non-causal) high-pass Butterworth (IIR) filter with an effective filter order of 10 at 4 Hz (cutoff at 4 Hz: -6 dB). Applying the filter to a copy avoids unwanted signal distortions and non-causal shifts on the original data but still allows to eliminate artifacts (Vanrullen, 2011; Widmann et al., 2015).

Manual inspection The filtered dataset was again manually inspected with a scaling of 10 pT . The filtering made artifacts visible which had previously been obscured by slow and large field drifts. Periods exhibiting such signal distortions or noisy data were additionally marked as bad in the original data.

Saving data The mne.RawArray object was then saved into the original directory with the annotations including experimental markers and periods of bad data.

In a subsequent step, the annotated raw data were segmented into trials (also called epochs in MNE) by extracting the data from -3.5 s to 8.5 s around each *close* or *relax* cue onset. Trials which coincided with data marked as bad were rejected and therefore excluded from further analyses.

Note that the potential overlap between trials ($\pm 3.5 \text{ s}$ around the cue, but ITI between 3 s and 5 s) was intentional to avoid edge artifacts due to filtering and was cropped away in subsequent analyses.

2.6.2. Spectral Filtering

The recorded signals from the motor cortex contain information in a participant-specific frequency band of interest. So called bandpass filters are temporal filters able to extract this information by attenuating other frequency bands in the signal. The filter choice is not trivial since a trade-off between temporal, spectral resolution and other parameters is always necessary and can heavily influence the expected results (Widmann et al., 2015).

The current application requires the online processing of data, meaning that only samples from previous time points are available to the filter. To resemble this fact in the offline analysis, it is therefore essential to use causal filters, since non-causal filters take future data points into account and can introduce systematic underestimation of delays (Widmann et al., 2015). The spectral peak of the sensorimotor rhythm is usually relatively broad, therefore shorter filters with a larger passband are preferable.

A minimum-phase (causal, non-linear phase) finite impulse response (FIR) bandpass filter was used for all offline analyses. Determining the participant-specific central frequency of the ERD was performed by using a passband of $12 \pm 4 \text{ Hz}$. Subsequent analyses were then conducted using a passband bandwidth of $\pm 1 \text{ Hz}$ around the participant’s individual ERD frequency. In both cases, left and right transition bandwidths of 4 Hz were chosen, resulting in a filter of length 825 ms. More details and the filter’s response characteristics are visualized in appendix C.3.1.

2.6.3. Spatial Filtering

The ERD in the motor cortex during hand MI is likely to be strongest in a confined region around electrode location C3 (Pfurtscheller, Neuper, Andrew, et al., 1997). To reduce the unwanted superimposition of signals from distant sources due to volume conduction, it is common to use

2. Methods

a Laplacian filter in EEG applications (McFarland et al., 1997). In contrast to EEG, directed cortical sources appear as structures in MEG with opposing polarities perpendicular to the source direction (Schaworonkow & Nikulin, 2022). Using a Laplacian filter would likely destructively combine these contributions rather than increase signal-to-noise ratio (SNR), whereas other linear recombinations of sensor-level signals can achieve better SNR. Manual selection, as used by Soekadar et al. (2011) or Soekadar, Witkowski, Birbaumer, et al. (2015), was not feasible here as only one session with minimal experimenter intervention was recorded.

The common spatial patterns (CSP) algorithm is a supervised machine learning method to distinguish the spatial correlation of signals in two classes (Fukunaga, 1990; Koles et al., 1990; Lemm et al., 2005). CSP yields spatial filters (linear combinations, also called components) which maximize the variance in one class while minimizing the variance in the other class when applied to the original data. For a signal with n channels, the algorithm yields n components w_i , $i = 1 \dots n$, each associated with an eigenvalue d_i between 0 and 1. The sign of $d_i - 0.5$ indicates the class of which the variance is maximized and the absolute value indicates the component's discriminative power. The components can be interpreted as projectors (also called backward model) into an abstract source space (also called surrogate sensor space). A mathematical derivation of the CSP algorithm and the computation of the corresponding spatial patterns can be found in subsection C.3.2.

The bandpass filter introduced in subsection 2.6.2 was used to extract the power modulations in the alpha band during the motor imagery task with a passband of 8 Hz to 16 Hz. Since the variance of such a filtered signal reflects its band power, the signal variance during event-related desynchronization should be reduced in *close* trials compared to *relax* trials. Based on this physiological assumption, class labels were chosen such that eigenvalues $d_i < 0.5$ reflect a variance maximization in the *relax* condition while variance in *close* trials was minimized. To extract the relevant ERD activity and reduce the signal dimensionality from 17 channels to the abstract source space, only the single one component corresponding to the smallest eigenvalue < 0.5 was thus used in the offline analysis.

The *spatial filters* usually lack interpretability as they are used to filter out non-discriminative information from both conditions. *Spatial patterns*, on the other hand, correspond to the influence of each source/component (in the abstract source space) on the measured data (sensor space) and can therefore indicate the spatial origin of a signal. Plotting the spatial pattern of the selected CSP component thus gives a visual impression of the location, direction and extent of the power attenuation in sensor space for *close* trials in comparison to *relax* trials (see subsection 2.6.6).

As discussed in subsection 2.1.2, the OPM locations are known only approximately (due to the flexibility of the sensor fixture) and not co-registered to the participant's anatomy. The sensor data shown in topological plots is therefore merely an approximate visualization but should not be confused with precise source reconstruction techniques. However, the application of CSP helps to reduce signal dimensionality and reliably extracts spatial patterns to verify physiological assumptions and improve classification results.

2.6.4. Determining the individual sensorimotor rhythm frequency

An important step to fit the parameters of an SMR-based BCI systems to the user, is the determination of the central frequency of the ERD. The first method proposed by Pfurtscheller and Lopes da Silva (1999, section 4.1) involves comparing the power spectra of an activity period and a reference period and has been used extensively in BCI applications.

To find each participant's central frequency of the ERD in the alpha frequency band, the data from the first no-feedback block of the experiment were bandpass filtered between 8 Hz and 16 Hz and cropped between 1 s to 5 s after cue onset to extract the period of the data with the strongest

2. Methods

effect (Pfurtscheller et al., 1998).

The CSP was computed on this data and the first component that maximized the variance in *relax* trials while minimizing variance in the *close* condition (see subsection 2.6.3) was determined and saved to disk. Using this component, the data were projected into the abstract source space. The power spectra of the projected data were then computed using Welch's method with a 2.048 s long Hann window for tapering with 1.536 s of overlap. A visualization of the average spectra for both conditions was displayed with 95 % confidence intervals computed through bootstrapping ($n = 1000$) (Davison & Hinkley, 1997; Graimann & Pfurtscheller, 2006). The SMR frequency was determined as the frequency at which the largest significant difference between *relax* and *close* trials was visible. This frequency was saved for each participant and used to set the spectral filter parameters in subsequent analyses. Participants for which no difference between *relax* and *close* trials could be found, were excluded from later analyses.

2.6.5. ERD timecourse

Since the current BCI system is based on voluntary modulations of the SMR, it is vital to investigate how the ERD evolves after cue onset and if the *relax* and *close* conditions can be distinguished based on that.

To visualize the ERD timecourse, the data from blocks 2 to 5 were filtered with a passband of ± 1 Hz around the individually determined frequency (causal FIR filter, see subsection 2.6.2). Using the CSP component computed on the first block (subsection 2.6.4), the data were projected into the abstract ERD source space. To compute signal power, the variance was then computed in (causal) sliding windows of 400 ms along the filtered data. This choice of window length achieves a balance between instantaneity and smoothness while the data could still be interpreted online and on a single-trial basis, unlike other methods which use the across-trials average over the filtered and squared data.

The resulting data were then downsampled to 100 Hz and potential edge artifacts cut off by cropping the data to -2.5 s to 7.5 s around cue onset. The percent-change in the signal compared to a baseline (-2.0 s to 0 s) was computed as $ERD\% = (A - R)/R \times 100$, where A is data from the *close* trials and R from the baseline period (see Pfurtscheller and Lopes da Silva, 1999).

To compare the SMR modulation in the two conditions, a non-parametric cluster level analysis was performed on the baselined ERD time series (Maris & Oostenveld, 2007). This procedure was chosen to avoid the multiple comparison problem (950 time points) and obtain a good understanding of the consistency and stability of the modulations. The cluster-based permutation test was computed with $n = 1024$ permutations.

Estimating the timecourse of the SMR is vital for the current BCI application. A precise analysis of effect onset timing, however, exceeds the scope of this work and cluster-level analysis tests would not be a valid method for this purpose (Sassenhagen & Draschkow, 2019). Therefore, the presented results identify temporal clusters during which the condition has a significant effect on oscillatory power, but cannot identify the precise timing of effect onset.

2.6.6. Focality

To obtain a qualitative understanding of the distribution of discriminative sensor locations over the motor cortex, a visualization of the CSP spatial patterns was drawn for each participant. Since the sensor patch did not cover the entire head, the determined patterns provide only limited information about the distribution of fields in the brain. However, they do provide insight into the field distribution induced by ERD beneath the sensor patch. Since magnetic fields from a tangential source in the motor cortex around C3 should appear as a dipolar field pattern

2. Methods

(Mellinger et al., 2007; Schaworonkow & Nikulin, 2022), this was also assumed to be visible in the visualized spatial patterns.

To compute a topographic maps, the CSP algorithm was again trained on the data of blocks 2-5 after bandpass-filtering around each participant's individual ERD frequency. The most discriminative spatial pattern (i.e., for the spatial filter with the lowest eigenvalue, see subsection 2.6.3) was then determined. Since the sensor locations were known relative to each other and relative to C3, the spatial pattern could be mapped to the sensor locations on the skull. Values in the spatial patterns were normalized to lie between -1 and 1 and a bilinear interpolation was applied to compute a color coding for the entire surface of the sensor patch.

2.7. Online Analysis

2.7.1. Online processing

Participants of the experiment received visual feedback during motor imagery during 6 *relax* and 6 *close* trials in randomized order in the second part of each block. To provide this feedback during the BCI experiment, a pipeline fast enough to process and classify the incoming OPM data in near real-time was required. A framework developed in the Clinical Neurotechnology lab called *BeamBCI* (Peekhaus et al., publication in preparation) was used as the foundation to implement such a pipeline. The BeamBCI framework streamlines data processing by dividing it up into 6 stages: *Acquisition*, *Preprocessing*, *Classification*, *Task*, *Stimulus* and *Recording*. A visualization of the processing steps is depicted in Figure 2.4. A very similar architecture of the processing pipeline was used by Zerfowski et al. (2021), but for the current study all dependencies on the commercial software NeuroPype (Syntrogi Inc., La Jolla, California, USA) were eliminated.

Acquisition Before further processing steps could be executed, the data had to be acquired and fed into the processing framework. This was realized through a python module which interfaces the FieldLine API, providing immediate access to the sensor data, with the LabStreamingLayer (LSL) protocol. The FieldLine LSL module (see appendix A.1.1) was written as a part of this thesis to allow the data processing in BCI applications and also used by Zerfowski et al. (2021).

Data were streamed with the acquisition system's sampling rate of 1000 Hz and metadata concerning the chassis, sensors and acquisition device parameters were provided in the LSL stream info fields. Network overhead is reduced by the manufacturer through sending data in chunks of 10 samples, effectively reducing the pipeline propagation frequency of incoming samples to 100 Hz. The resulting delay of 10 ms for the oldest sample in the chunk is unproblematic for the current application, since the dynamics of the classification pipeline (windowing over 400 samples) and the brain (reaction time, ERD dynamics, ...) are much slower.

Preprocessing In the second stage of the pipeline, the LSL stream of the acquired sensor data was picked up and fed through a bandpass filter. This step was needed to extract the ERD content from the individual frequency of interest of each participant. The filtering node used a Hamming-windowed FIR-filter generated with `scipy.signal.firwin()` with a length of 499 ms and a passband which was adjusted to ± 1.5 Hz around the participants ERD frequency (see Pipeline Training below and appendix C.3.1). Since only previous time points were available during the online analysis, the data were filtered in a one-pass scheme using `scipy.signal.lfilter()`. An internal buffer with updated initial conditions was maintained in the node such that each chunk could be filtered individually, but as if a sliding window was used.

2. Methods

Classification The *classification* step was afforded through a custom pipeline framework. The pipeline consisted of independent “nodes”, which were initialized before the experiment, retrained individually and could process data by taking in and returning numpy arrays with flexible dimensions ($\#channels \times \#samples$).

After bandpass-filtering around the ERD frequency in the preprocessing step, the first operational node in the classification pipeline applied CSP filters to the data. Only the two most discriminating components were automatically selected for the spatial filtering (see Pipeline Training below). Analogous to the offline analysis, this step amplified the variance difference between close and relax trials and reduced the signal dimensionality down to two channels.

To compute the variance as a means of power estimation, data were then aggregated in a first-in-first-out buffer node with a length of 400 ms. The buffer used an overlap of 100 samples between two subsequent propagations, effectively reducing the pipeline update frequency to 10 Hz.

The chunks of data were then reduced along the temporal axis by computing the variance, reflecting a power estimate in the frequency band of interest in the abstract ERD source space for the last 400 ms. By computing the logarithm of the variance (log-variance), the distribution of the data was brought closer to a Gaussian distribution, which is required for LDA. This operation resulted in samples of dimensions 2×1 (CSP channels \times “time”).

The class prediction was then performed using a linear discriminant analysis (LDA) classifier (see Pipeline Training below), resulting in a single floating point value indicating the predicted class ($relax \leq 0 < close$).

The last processing step involved low-pass-filtering of the class estimate, to smooth out the erratic class switches that were sometimes seen during pilot experiments. To this end, a single-pole filter was implemented. This simple implementation of an infinite-impulse-response (IIR) filter only relies on the previous sample, is defined by a single decay factor and effectively implements a low-pass filter with exponential decay and a 3 dB point at 0.3 Hz (see appendix C.3.1 for more details).

The classification module received control commands from the task module and stored them aligned with the preprocessed data. During the training step, the commands indicating the condition start and stop time points were used for epoching and labeling the data.

Task The *task* module was responsible for controlling the flow of the experiment. In the task module, the experiment could be started, continued after breaks and experimental parameters (repetitions and interval duration, etc.) could be set as well as experimenter notes sent through a separate LSL channel. For potential future uses (e.g., performance-dependent experiment control), the classification values were relayed through the task module but not used in the current setup. The task module also fed the control commands back into the classification module (see above).

Stimulus The *stimulus* module was the last step in the online processing framework and managed the entire visual output of the experiment. Using the control commands from the task module, presentation of the cues was started and stopped accordingly and feedback was displayed (hidden) during feedback (no-feedback) trials. The classification values received from the classification pipeline were used to control the movement and transitions of the *close* and *relax* feedback graphic elements (see subsection 2.4.1).

Recording The final *recording* stage received data from all stages for saving the raw and preprocessed signals, classification values and control commands to a hard drive. All data were saved as .xdf-files using the LabRecorder, the standard tool for saving multiple parallel LSL streams in synchrony.

2. Methods

Due to the buffering with 100 ms overlap, classification values were generated with a frequency of 10 Hz and incorporated the previous ~ 400 ms of data (approximate lower bound, due to filtering). A higher update rate would not have been advantageous due to the slow temporal dynamics of the SMR, a slower update rate leads to slower reaction times of the system and could potentially be perceived by the user as inconsistent control (Crea et al., 2018). To allow monitoring of the signals and identify potential artifacts during the experiment, the bandpass filtered sensor data as well as the CSP-filtered and the classification value were continuously visualized for the experimenter.

The applied methods and parameters were carefully selected after piloting experiments and consist of commonly used methods in EEG-based BCI paradigms. However, this particular pipeline implementation has not been validated and compared with other pipelines used on EEG or MEG data. A systematic evaluation of different processing setups on the recorded data, as well as a comparison of this pipeline with other approaches on available public data should follow to identify possible targets for improvement.

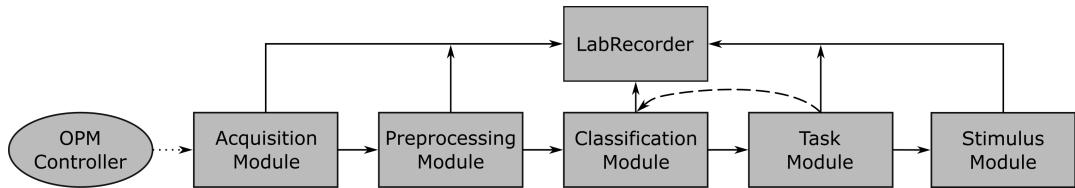


Figure 2.4. Online processing pipeline. The data are transferred from the OPM controller to the acquisition module via Ethernet (dotted line). The data are preprocessed and the classification module predicts a class value, which is then translated into visual feedback. The stimulus module displays the cues and feedback and feeds the current experimental state back into the classifier. All modules forward their output data to the LabRecorder. Straight and dashed lines indicate data transfer via LSL.

Pipeline Training

Since individual differences in central ERD frequency, source location, etc., were expected, the preprocessing and classification pipeline had to be adjusted on a per-participant basis. This so called “training” step was performed during the experiment at the transition between the no-feedback and the feedback part of each block.

Determining each participant’s individual central frequency of the ERD was an important component of the training to allow precise filtering and ensure the success of subsequent processing stages. Since there are no established automatic artifact rejection methods for OPM data, a more robust analysis than in the offline analysis had to be employed here: To select the ERD frequency after the no-feedback phase of a block, all previous trials were bandpass filtered between 8 Hz and 18 Hz and cropped from 1 s to 5 s after cue onset to extract the period of the strongest ERD. The CSP algorithm was computed with the two conditions and the 6 most separating components were applied to the signal to reduce the dimensionality of the data and amplify the modulatory effects of motor imagery. The Fourier transform was then executed on the data after applying a Hann taper to smooth out discontinuities, and the average power spectra during the two conditions were displayed. This allowed the experimenter to identify the frequency with the most dominant difference in power between *relax* and *close* trials as the ERD frequency. If no such frequency could be found, 12 Hz was assumed as the default for the bandpass filter for the following feedback trials until the procedure was repeated in the next block. The selected

2. Methods

frequency of interest was only updated in later trials if a strong alternative peak became visible later, possibly because of artifacts that occurred during the first trials.

After setting the central ERD frequency at the transition between the no-feedback and feedback-part of a block, a retraining of the pipeline was executed. If necessary, first the preprocessing bandpass filter was updated with the new center frequency f and a 3 dB passband of $f \pm 1.5$ Hz, its state reinitialized and the data filtered with the new parameters. The classification pipeline was then retrained with the data from all previous trials automatically and invisible to the participant. Similar to the step-by-step fashion of the processing, the pipeline architecture allowed an independent retraining of all nodes in the order of processing.

The CSP training was performed as explained in the offline analysis, but here the *two* components with best discriminability were saved as filters for subsequent transformations of the data. To bring the training data into the same shape as it would be during processing, the buffer node was cleared and the new data set was piped into the buffer node to acquire buffered 400 ms chunks of the labeled data previously processed by the CSP. The log-variance computation was then simply applied to the data as it did not require any training or parameter adjustments.

Linear discriminant analysis (LDA) was used to reduce the 2 remaining data channels of the power estimate after the log-variance computation to a single class indicator. LDA is an established method in the BCI field and essentially determines the vector projection which maximizes the discriminability between two classes (Fisher, 1936). To compute the linear discriminants, the singular value decomposition method without shrinkage was performed on the labeled data, based on the implementation from scikit-learn. The LDA returned a single floating point value with values above 0 indicating *close* predictions and values ≤ 0 being interpreted as *relax* predictions.

The single-pole low-pass filter had no trainable parameters and was simply applied to the data to acquire the final training predictions.

Combined, these steps took less than 1 s and were conveniently executed right before continuation of the experiment. After the training procedure, all internal states and buffers were reset and reinitialized before the next samples for processing were fed into the pipeline. Due to the inter-trial interval of at least 3 s, all filters and buffers were properly initialized after this clearing before the first classification value was used for visual feedback.

2.7.2. Classification Performance

The visual feedback presented to the participant in the experiment acts as a substitute for proprioceptive feedback through an actual exoskeleton hand. It is therefore relevant to investigate if the control signals generated from the neural data could serve the user in exercises of daily living. For that, the feedback control should be reliable, that means that *close* MI is classified correctly (corresponding to an exoskeleton hand closing when desired). It must also be safe, meaning that *relax* imagery is also classified correctly (i.e., an exoskeleton should only close upon the user's intention).

Witkowski et al. (2014) established that most objects of daily life can be grasped with an average of 50 % of the user's full hand span. A trial in the *close* condition was therefore considered successful when more than 2.5 s of a trial were classified as *close*. *Close* classification samples from the pipeline were interpreted here as "positives", thus a successful trial corresponds to a true positive rate (TPR) $> 50\%$ of the samples between 0 s to 5 s.

For *relax* trials, no grasping motion should be executed. Most common objects of daily life are smaller than 25 % of the full hand span (Witkowski et al., 2014), which corresponds to a false positive rate (FPR) of $< 25\%$. Therefore, *relax* trials with less than 1.25 s of *close* classifications were considered successful.

To evaluate if the performance of the classifier exceeds chance level, the expected success rates

2. Methods

of a chance-level classifier were computed using a cumulative binomial distribution. A chance-level classifier assigns the *close* or *relax* label to a sample with equal probability ($p = 0.5$). Since 50 classification values are generated during a trial of 5 s, more than 25 samples would correspond to the $> 50\%$ reliability threshold and 12 or fewer samples correspond to the $< 25\%$ safety threshold. Accordingly, the expected success rate under a chance level classifier would be 44.39 % for *close* trials and 0.015 % for *relax* trials (see appendix C.3.3 for the derivation). These values were then compared to the obtained success rates using the Wilcoxon signed-rank test to evaluate if the classification pipeline performed better than the chance level classifier.

Since the equal probability assumption in the chance level classifier could be violated if the classifier were biased towards either condition, the empirical chance levels were reported for all analyzed participant subsets. The empirical chance level evaluates the probability of either classification outcome when conducting the experiment and was computed as the average probability for a *close* classification over all participants and conditions. All statistical tests were repeated against the empirical chance level classifier and the results reported in the text.

Block-wise Success Rate To evaluate if the retraining between the no-feedback and the feedback-phase of each block improved the pipeline performance over time, an analysis of the block-dependent success rate was conducted. For this, three different cases were taken into account: The success rate in *close* trials, in *relax* trials, as well as the combined success rate. For all three cases, a linear mixed model was fit to the data with the respective success rate as the dependent variable and the block as independent variable (fixed effect). Each participant was considered an independent group (random effect) to account for individual differences in average success rate. The slope of the linear fit then indicated if there was an increase or decrease of the success rate over time and the corresponding p-value showed if the rate of change was significantly different from 0. For this analysis, only participants with a distinct ERD frequency were considered.

3. Results

3.1. Sensor Characterization

Before the available OPMs could be used in a neuroscientific application, a basic evaluation of the sensors' signal characteristics was performed to answer the first research question. First, the system's noise properties were quantified inside the magnetically shielded environment before dynamic range and bandwidth were measured.

3.1.1. Noise Floor

Measuring the noise floor of the system was the first step of characterizing the signal properties of the used OPMs. Since the sensors are operated in an imperfectly shielded environment during the experiment, external noise influences are as important as the sensor's noise floor. Further, the quantification of the true sensor noise floor is a complex endeavor since sensor and external noise contributions cannot be accounted for easily. The values measured here therefore do not represent the *sensor* noise floor but rather the *system*'s noise floor and will be interpreted as such.

To compute the noise spectra, 10 sensors were positioned in the Ak3b MSR and a recording of 4.5 min was taken in both open and closed loop mode of operation. The noise spectra were then computed using Welch's method and visualized in Figure 3.1.

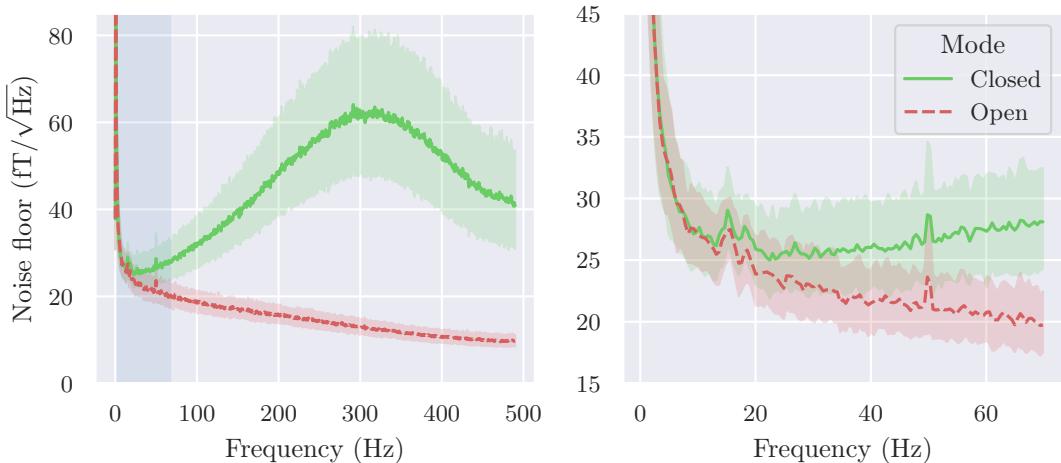


Figure 3.1. Noise spectra measured by 10 FieldLine OPMs as measured in the Ak3b MSR. Both plots show the spectra in open (red) and closed loop (green) mode with 95 % confidence intervals. **Left panel:** The full noise spectrum measured at a sampling frequency of 1000 Hz. **Right panel:** Detail view of the noise spectrum up to 70 Hz. A peak at 50 Hz due to power line noise is clearly visible.

Both spectra indicate large noise contributions for frequencies below 5 Hz with values above $30 \text{ fT}/\sqrt{\text{Hz}}$. The open loop noise spectrum (red) drops with an approximate $1/f$ shape from $53 \text{ fT}/\sqrt{\text{Hz}}$ at 2 Hz, approaching a value of $10 \text{ fT}/\sqrt{\text{Hz}}$ at 450 Hz.

3. Results

The close loop spectrum closely follows the drop of the open loop spectrum up to ~ 30 Hz and then increases again to reach a local maximum of $\sim 62 \text{ fT}/\sqrt{\text{Hz}}$ at around 300 Hz. At 450 Hz it approaches a noise floor of $44 \text{ fT}/\sqrt{\text{Hz}}$.

The 95 % confidence interval widths for the across-sensor average are similar for both modes below 20 Hz ($\sim 6 \text{ fT}/\sqrt{\text{Hz}}$) but rapidly increases for higher frequencies in closed loop mode to about $30 \text{ fT}/\sqrt{\text{Hz}}$ at 300 Hz while decreasing down to $3 \text{ fT}/\sqrt{\text{Hz}}$ for the open loop mode.

Chassis-wise noise floor In practice, when two separate and independent acquisition units are used, possible differences in noise floor between the units are often disregarded. To determine if there is a systematic between-chassis difference in noise floor, the spectrum of a 5 min recording was computed analogous to the analysis above. According to a communication with FieldLine Inc., there should be no difference between the chassis since data digitization is performed on individual data acquisition cards which can even be swapped between chassis and sensors.

The computed spectra are visualized in appendix D.1. One sensor in chassis 1 has an abnormally strong noise profile in higher frequencies, as determined by a z-scoring of the noise floor at 300 Hz. According to the manufacturer, this is due to an increased gain error which could be alleviated by a recalibration (personal communication). After excluding this outlier from the computations, the variability in noise floor is explained better by the sensor-wise gain errors than by between-chassis differences. No significant difference between the noise floors of the two chassis can be seen. Therefore, the noise floor does not seem to be dependent on the chassis but rather on the individual sensor.

3.1.2. Dynamic Range

Amplitude Modulation

Strong and rapid magnetic field changes can occur during a measurement due to participant movement with insufficient shielding or magnetic field drifts. It is important to assess the capability of the sensors to measure those field changes accurately and without distortions.

This aspect of the dynamic range was characterized by applying sinusoidal fields at 12 Hz with a peak amplitude of up to 20 nT in closed loop mode and 5 nT in open loop mode along the sensitive axis of two randomly picked sensors. The measured amplitude was then estimated by using a fixed-frequency sine fit on the data.

Figure 3.2 visualizes the measured/applied amplitude ratio (A_m/A_0) vs. the applied amplitude (A_0). An accurate measurement of the applied fields would correspond to a constant value of $A_m/A_0 = 1.0$.

The open loop mode shows a strong dependence of the measured amplitude on the applied signal's amplitude. While ~ 98 % of the applied sinusoidal amplitude is measured at $A_0 = 100 \text{ pT}$, the amplitude ratio sharply drops down to 80.5 % and 85.5 % at $A_0 = 5 \text{ nT}$. Increasing the amplitude further at this point would not yield useful results.

In closed loop mode, 98 % of A_0 is measured by the OPMs for low amplitudes. However, the ratio is stable at 0.98 up to $A_0 = 14.5 \text{ nT}$. For applied fields larger than 15 nT, a sharp increase in applied amplitude is estimated by the sine fit. Data for fields up to $A_0 = 20 \text{ nT}$ are visualized in the subsubsection D.1.2.

3. Results

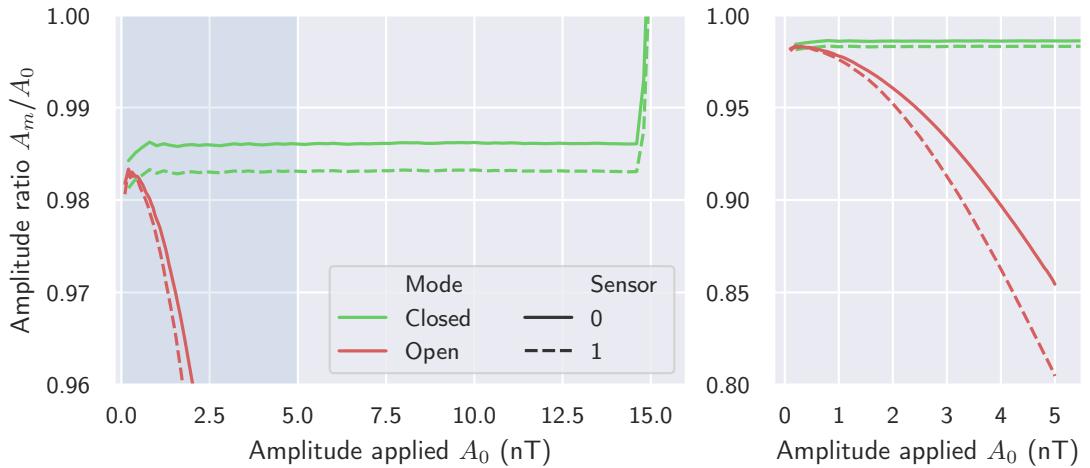


Figure 3.2. Amplitude ratio (A_m/A_0) vs. amplitude applied (A_0) for two OPM sensors. **Left panel:** Applied amplitudes up to 15 nT. The closed loop ratio remains stable around 0.98 up to ~ 14.5 nT. The shaded area indicates the amplitude range covered in the right panel. **Right panel:** Applied amplitudes up to 5 nT to emphasize the sharp fall-off in amplitude ratio for the open loop mode.

Offset Modulation

The second considered aspect of the dynamic range was the ability to measure magnetic field changes under the influence of large offset fields. Such offsets can occur for example through slow field drifts in the shielded room or when the sensors are rotated in a static field due to participant movement.

To assess this type of dynamic range, the sensors were exposed to a sinusoidal field at 12 Hz with a peak-to-peak amplitude of 100 pT. Additional offset fields from -7 nT to 7 nT were applied in steps of 100 pT along the sensitive axis of the sensors. Again, a sine curve fit of fixed frequency was applied to the data to determine how much of the applied sinusoidal amplitude is attenuated depending on the offset field.

Figure 3.3 uses a dB scale on the y-axis to visualize the attenuation induced by the applied offset field. If the sensors measured the full amplitude of the sinusoid independent of the applied offset, the attenuation should be constant around 0 dB.

In open loop mode, the attenuation increases parabolically for positive and negative offsets. The curves exhibit a rapid signal attenuation which crosses the 3 dB line (50 % of applied signal power) at ± 2.8 nT and ± 3.8 nT for the two sensors.

The measured amplitude in the closed loop mode is very stable and does not exhibit attenuation for external fields applied along the sensitive axis up to ± 7 nT.

In accordance with the results presented for the amplitude modulation-related dynamic range, the magnitude of the applied offsets was measured accurately in closed loop mode but shows considerable attenuation in open loop mode. Plots to visualize this dependency can be found in appendix D.1.2.

3. Results

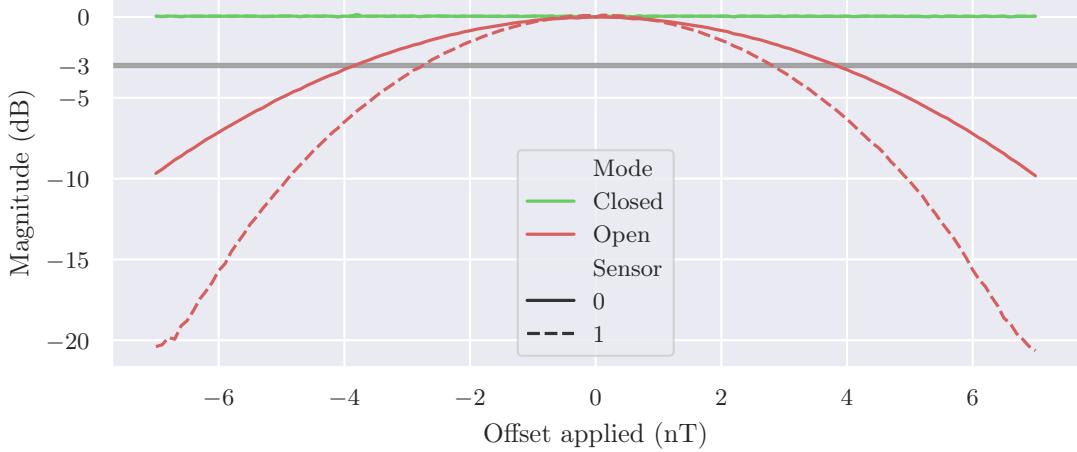


Figure 3.3. Measured amplitude of a sine wave along the measurement axis with constant peak-to-peak amplitude of 100 pT vs. offset applied to the sine. Open loop measurements show a strong attenuation while the closed loop mode maintains a linear relationship.

3.1.3. Bandwidth

Quantifying the frequency response characteristics of a sensor system is particularly important to determine which cortical bands can be measured accurately. Due to their working principle, SERF-OPMs exhibit an exponential signal attenuation in the frequency response. The sensor bandwidth is commonly given as the frequency at which an attenuation of 3 dB (corresponding to 50 % of the applied signal power) is measured.

To determine the bandwidth, a sine wave with peak-to-peak amplitude of 100 pT was frequency modulated between 0 Hz and 490 Hz. A Bode magnitude plot was computed and is shown in Figure 3.4.

In the open loop mode, an almost linear fall-off occurs for frequencies above 25 Hz, reaching an attenuation of 3 dB at around 47 Hz and 80 Hz for the two sensors. Only about 10 % of the applied power (i.e., 10 dB attenuation) were measured at 146 Hz and 235 Hz.

In the closed loop mode, the measured signal maintains a magnitude at the 0 dB level up to around 150 Hz before an amplification of the signal to up to +1.1 dB at 300 Hz occurs for sensor 1 while sensor 0 accurately measures the applied amplitude. The almost linear fall-off for frequencies > 300 Hz reaches the 3 dB point at around 400 ± 10 Hz.

3. Results

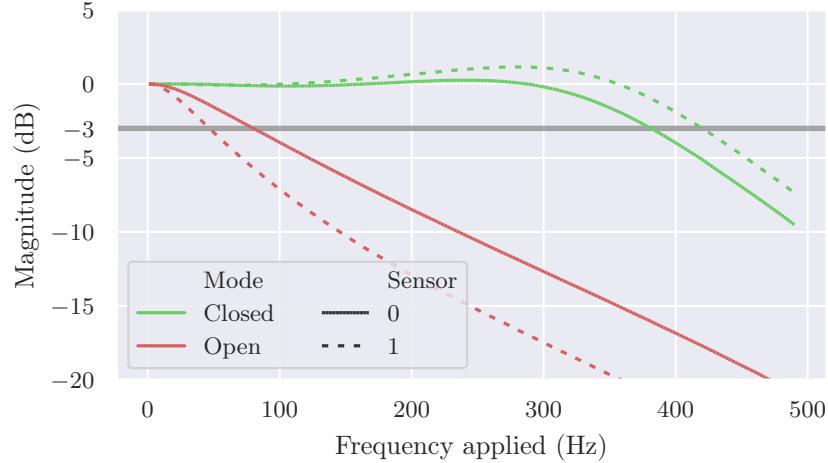


Figure 3.4. Bode magnitude plot (linear x-axis for better visibility) indicating the signal attenuation in dependence on the signal's frequency for two randomly picked OPMs in open and closed loop mode. The gray line indicates the 3 dB point.

3.2. Offline Analysis

3.2.1. ERD Frequency

BCI systems are optimized by incorporating user-specific parameters to detect the feature(s) the system uses. In this case, such a feature is the frequency of the mu rhythm in the alpha band at which the ERD occurs during motor imagery. Previous studies have found that for about 40 % of the population little or no effect of motor imagery might be visible with EEG or MEG (see section 1.1). Determining visibility and frequency of the ERD are important steps to adjust the individual filter parameters for each participant and thereby improve classification results. This analysis could also provide insight as to how many people are eligible for OPM-based BCI.

To determine the central frequency of the ERD, the cleaned data from the no-feedback phase of the first block from each participant were bandpass-filtered between 8 Hz and 16 Hz and then spatially filtered using CSP to obtain a single data channel. Figure 3.5 shows the spectra for two participants, one with a significant difference between *close* and *relax* condition at 12.2 Hz (left panel) and one with no visible difference (right panel).

The same analysis was performed on all 16 participants to identify if an event-related desynchronization was visible, and, if so, record the central frequency the modulation. For 10 participants, a distinct frequency could be identified (see Table 3.1), 6 participants did not show sufficient evidence of an ERD-related difference between the spectra.

Participant	002	003	004	007	009	012	013	014	018	019	Avg.
Frequency (Hz)	12.9	12.2	11.4	11.6	8.0	12.2	12.0	13.2	13.0	13.8	12.0±1.5

Table 3.1. Individual ERD frequency for all participants for which an ERD frequency could be found. The last column *Avg.* shows mean and standard deviation across participants.

Participants without a clear ERD frequency were excluded from subsequent offline analyses.

3. Results

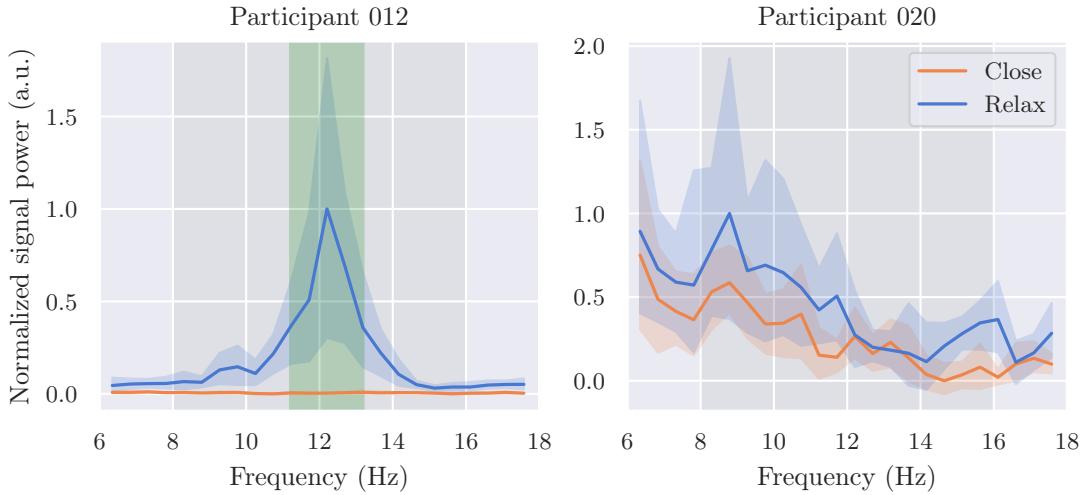


Figure 3.5. Comparison of the power spectra in *relax* and *close* condition for two exemplary participants. The y-axis is in arbitrary units due to the previous application of CSP. The gray area indicates the alpha band from 8 Hz to 16 Hz and the shaded areas around the curves are the 95 % confidence intervals. **Left panel:** A clear and distinct peak in the spectrum of the Relax condition is visible at 12.2Hz. The green area marks the passband of ± 1.5 Hz around the central ERD frequency. **Right panel:** The participant does not exhibit a distinct ERD peak.

Plots for all participants can be found in appendix D.2.1.

3.2.2. ERD Timecourse

Since BCI systems are time-critical applications requiring immediate response, it is important to gain an understanding of the temporal of the mu rhythm modulations as the underlying control signal as measured with OPMs. Therefore, the next step in the offline analysis was an investigation of the ERD timecourse after stimulus onset. From EEG studies it is known that the time dynamics of ERD are relatively slow, meaning that significant changes of the SMR power would be expected at around 1 s after cue onset and offset (Graimann et al., 2002).

Quantification of the ERD over time was achieved by filtering the data of block 2 to 5 around each participant's individual mu rhythm frequency and applying the spatial CSP filter to obtain a single data channel. The variance was computed in sliding 400 ms time windows. All procedures were carefully selected to be causal operations to rule out a systematic underestimation of the effect onset timing. The ERD was then computed by taking the percent-difference against the baseline from -2 s to 0 s before stimulus onset for each trial.

Figure 3.6 visualizes the ERD timecourse averaged over all participants. The timecourses for both *close* and *relax* range around 0 % during the baseline period and show a brief period of non-significant ERS around 0.3 s after stimulus onset. For *relax* trials, the timecourse slowly rises up to ~ 100 % until 1 s after stimulus offset where it stagnates. Trials in the *close* condition show a strong drop of the mu rhythm amplitude at around 0.8 s, remaining stable at -50 % until $t = 4$ s. The 0 %-threshold is crossed again between 4.5 s and 6 s. Both timecourses surprisingly remain above 0 % after the cue disappeared. The cluster-based permutation test found significant

3. Results

clusters in the timecourses from 0.9 s to 4.5 s ($p < 0.001$), indicating a significant effect of the condition on the timecourse.

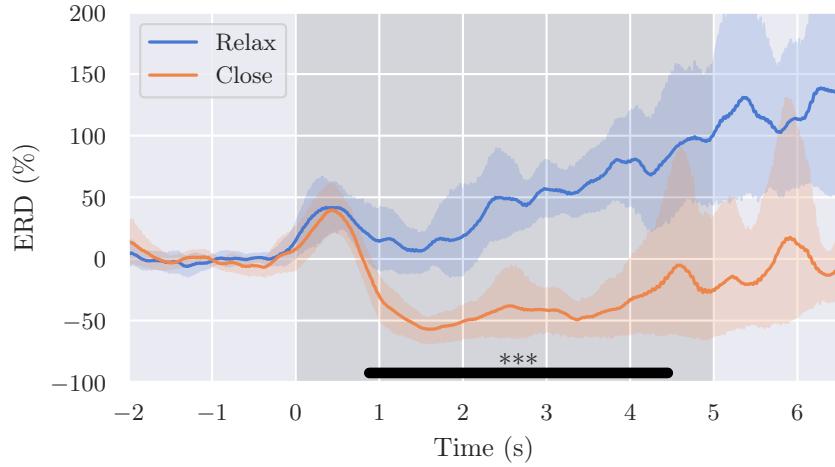


Figure 3.6. Timecourse of the ERD against baseline (−2 s to 0 s before stimulus onset), the shaded margins are the 95 % confidence intervals. The black bar in the bottom marks the cluster with significant difference (***: $p < 0.001$) between *close* and *relax* trials according to a cluster-based permutation test. The gray area indicates the trial duration from 0 s to 5 s.

An individual plot for each participant can be found in appendix D.2.2. It is observed that some participants do not only show strong ERD in the *close* condition, but also ERS during *relax* trials (e.g., p003, p007). However, little to no difference can be found in other participants (p004, p013).

3.2.3. Focality

The last component of the second research question was a qualitative investigation of the physiological plausibility of the spatial filter found for classification. This is important to avoid the classification of noise correlated with the current condition (e.g., looking at the hand in *close* trials) as the predicting feature. Physiologically, the mu rhythm is modulated most strongly in the cortical area associated with hand control approximately beneath the C3 EEG electrode position. Owing to the field structure of magnetic dipoles, a spatial filter for MEG should therefore find a positively and a negatively contributing area on opposing sides of the assumed source. The exact source location of the SMR varies by participant and, due to the irregular cortical surface, small differences in source orientation and location can be amplified in terms of observable magnetic field patterns.

By visualizing the CSP patterns, the contributing sources can be made visible and then judged qualitatively for the visibility of dipolar structures. Since the sensor patch did not cover the whole surface area of the skull, no prediction concerning the field distribution in other parts of the brain can be made. However, the current analysis could already yield insights as to whether the sensor place placement was efficient (i.e., located over the entire extent of the dipole). For this analysis, the CSP algorithm was trained on the bandpass-filtered data of blocks 2–5 and the spatial pattern corresponding to the lowest eigenvalue was selected for visualization (see 2.6.3).

3. Results

Figure 3.7 shows the spatial patterns determined by CSP for two participants, appendix D.2.3 contains the figures for all remaining participants. Dipolar sources can be characterized by having a strongly positive and a strongly negative peak close to each other with an approximately linear slope between them. All participants show a clear dipolar pattern centered between C3 and C1 on the lateral-central axis and spread along the entire frontal-parietal range of the sensor patch. It has to be mentioned again that sensor positioning was performed manually over C3 without subsequent co-registration, such that individual differences of head shape and ERD source location can contribute to the spatial spread of apparent ERD sources.

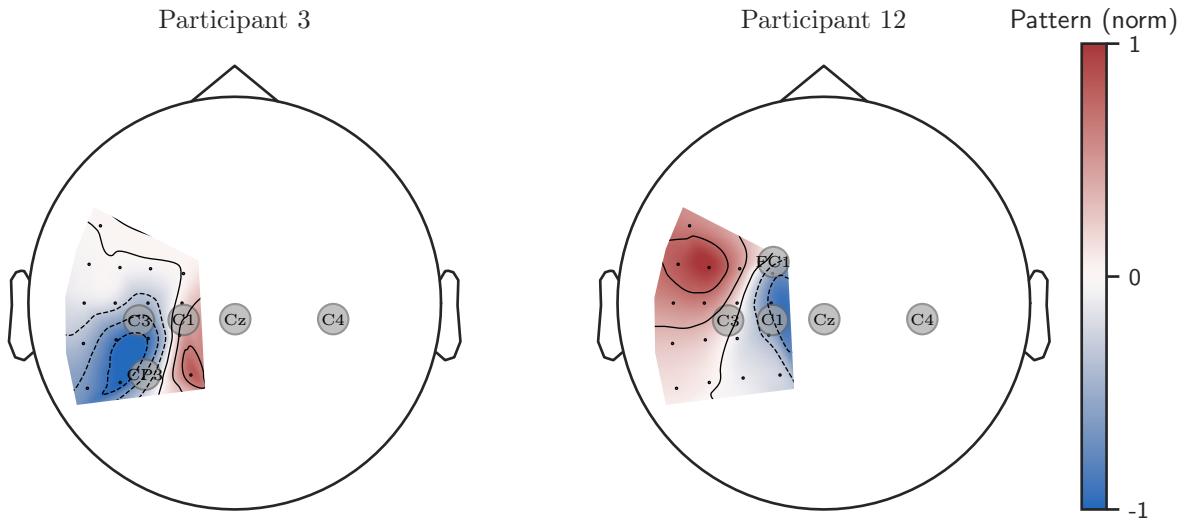


Figure 3.7. Focality of the ERD for two selected participants. Visualization of the CSP pattern with the best discriminability between *relax* and *close* trials. **Left panel:** Participant 3 shows a clear dipolar pattern with a source between CP3 and C1. **Right panel:** A dipolar CSP pattern with a source along the FC1–C3 axis can be seen.

3.3. Online Analysis

The third research question asks for the feasibility and performance of a near real-time data acquisition and classification pipeline, which can provide visual feedback to the user. Online data classification is an important part of any BCI system and was therefore evaluated by implementing a pipeline that could be retrained flexibly and with only minimal experimenter intervention. The publication by Witkowski et al. (2014) was used here as a reference to compare the reliability and safety of the implemented pipeline.

The pipeline was trained after the first block of no-feedback trials based on the individual ERD frequency determined by the experimenter. After the training, which took less than 1 s, the experiment continued immediately into the feedback phase such that idle times for the participant were avoided. However, for some participants no ERD could be found even in the offline analysis after cleaning artifactual and noisy trials. The analyses reported here were therefore performed on all participants ($n = 16$) as well as the subset of participants exhibiting a clear ERD frequency ($n = 10$, see subsection 3.2.1).

3. Results

The classification pipeline ran at an update frequency of 10 Hz, leading to 50 classification values per trial. For *close* trials, the true positive rate (TPR) indicates how many of the 50 updates yielded a correct *close* classification. Since the central feature was the ability to classify *close* MI here, the false positive rate (FPR) indicates the ratio of incorrect classifications during *relax* trials.

In *close* trials the true positive rate (TPR) evaluated to 62.9 % on average for all participants (see also Table 3.2). The corresponding success rate indicates that in 74.2 % of the trials the TPR was above 50 %. A TPR above 50 % was associated with a closing of an exoskeleton hand by over half the hand span, indicating that objects of daily life could be grasped, according to Witkowski et al. (2014). For *relax* trials, the false positive rate is 36.3 %. Trials where the TPR was lower than 25 % were considered successful, leading to a success rate of 40.0 %. An overview of the distributions of TPR and FPR is visualized in Figure 3.8.

To compare the performance of the classification pipeline to a classifier running with chance level classification ($\Pr(\text{close}) = \Pr(\text{relax}) = 0.5$), the expected chance level success rates were computed as 44.39 % for the *close* condition and 0.015 % for the *relax* condition (see appendix C.3.3). A comparison using the Wilcoxon signed-rank test indicated that the measured success rates in both conditions were significantly better than chance level with $p < 0.01$.

Filtering out the participants without clear ERD improves the results, leading to an average TPR of 69.3 % and a success rate of 83.7 % ($p < 0.01$) for *close* trials. The FPR decreases to 30.6 % and the success rate increases to 51.0 % ($p < 0.01$) in *relax* trials.

The empirical chance levels were extremely close to 50 %, with 49.6 % for all participants and 49.9 % for the participants with ERD, hence they were not added to Figure 3.8 to avoid visual clutter. All statistical tests were repeated with the corresponding empirical chance levels and resulted in the same significance estimates.

Even though the success rate of the classifier is significantly above chance level and the average TPR of 62.9 % (69.3 % for participants with ERD) indicate reliable control, safety concerns of the developed system must be addressed. As in Witkowski et al. (2014), the safety criterion was defined as < 25 % FPR in *relax* trials. This criterion was violated frequently: In $61.85 \pm 20.77\%$ of cases for all participants and in $46.31 \pm 17.11\%$ of cases for participants with clear ERD.

Condition	All participants		Participants with ERD		Witkowski et al. Avg. Close
	Avg. Close	Success rate	Avg. Close	Success rate	
Close	62.9 %	74.2 % ($p < 0.01$)	69.3 %	83.7 % ($p < 0.01$)	63.59 %
Relax	36.3 %	40.0 % ($p < 0.01$)	30.6 %	51.0 % ($p < 0.01$)	36.11 %
Combined	57.1 %		67.3 %		

Table 3.2. Average *close* rates and success rates by condition. Compared here are the statistics for all participants, participants with a clear ERD frequency and the results reported by Witkowski et al. (2014). The p -values reported in the success rate columns are as compared against a chance-level classifier using the Wilcoxon signed-rank test.

3. Results

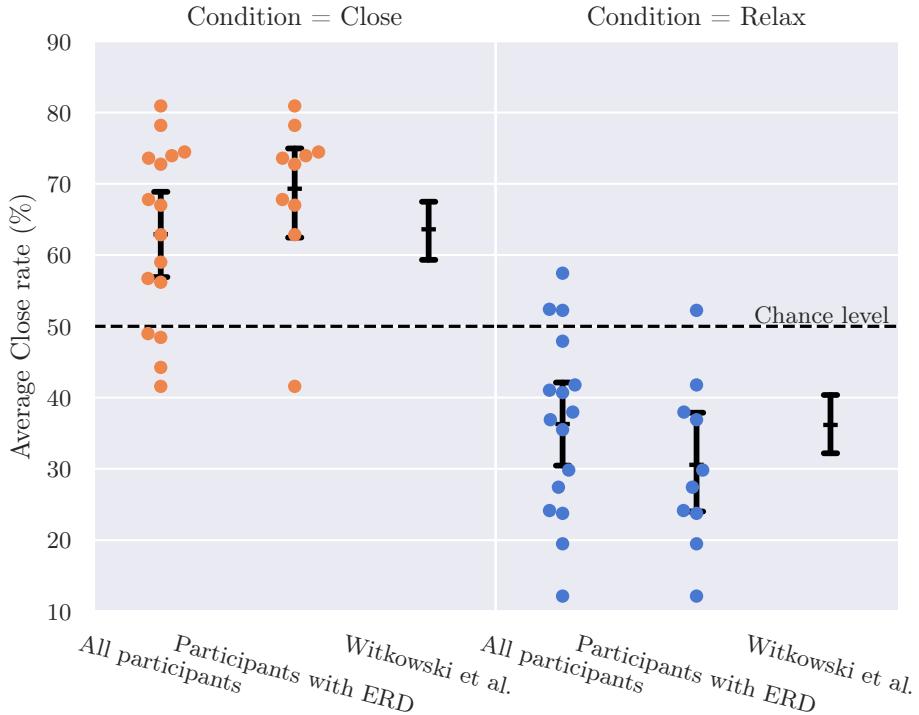


Figure 3.8. Average rate of *close* classifications during feedback trials. Each dot corresponds to one participant, the error bars indicate the mean and the 95 % confidence intervals. The leftmost column of each panel contains all 16 participants, the center column only participants with visible ERD. The rightmost column shows the mean and confidence intervals as reported by Witkowski et al. (2014) for comparison. **Left panel:** *Close* trials, **Right panel:** *Relax* trials.

3.3.1. Block-wise Success Rate

Lastly, a per-block analysis was conducted to examine if the classification performance in participants with ERD improved over time, due to the repeated training on previous data. To study this, the success rates were computed per condition and as the combined success rate and grouped by block 1 to 5. For each case, a mixed linear model was fit to the data with the block as fixed effect and the success rate as dependent variable. Participants were used for random effect grouping to control for interindividual differences.

Figure 3.9 displays the results of the analysis. For the *close* condition, the model indicates a non-significant ($p = 0.267$) increase in success rate of 1.167 %/block. The model fit on the *relax* finds a slightly steeper, non-significant ($p = 0.245$) negative slope of the success rate (-2.833 %/block). The last model was fit on the combined success rate, which is simply the average of the two previous success rates, since they were exactly balanced by design (5 \times 6 of *relax* and *close* trials). Accordingly, the model found a non-significant ($p = 0.491$) negative slope of -0.833 %/block.

3. Results

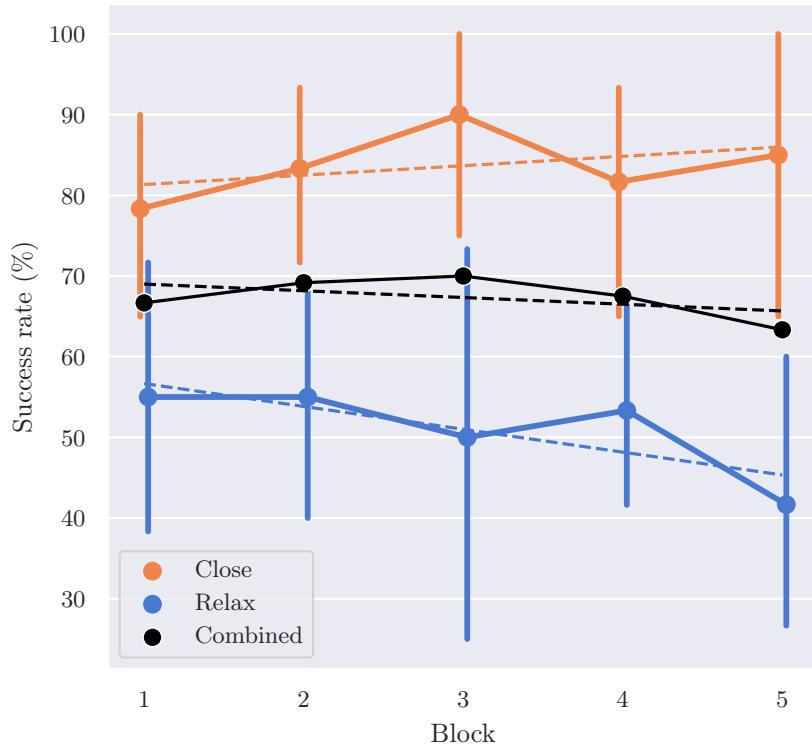


Figure 3.9. Success rate vs. experiment block. The blue dots indicate the success rate in *relax* trials, orange is used for the *close* trials, black shows the combined success rate. The error bars visualize the 95 % confidence intervals. The dashed lines are the mixed linear model fits.

3.4. General Findings

The fourth research question asks for the insights gained through the conduction of the study and how these could inform future measurements. To date, there is no “standard” way of conducting neuroscientific experiments with OPM technology and there are no established analysis pipelines. The knowledge obtained during the organization of the participant measurements and the data acquisition with healthy subjects shall therefore be documented here briefly. Even though Seymour et al. (2022) present some recommendations for OPM experiments (i.e., for interference suppression), the knowledge acquired here pertains to the specific case of an OPM-BCI and 18 instead of only 2 participants.

The participant recruitment through Charité – Universitätsmedizin Berlin and PTB communication channels yielded 2/3 of all participants, other participants were acquired through public announcements. Scheduling of the appointments and acquisition of name and email address of prospective participants was performed through terminplaner4 hosted by the Deutsches Forschungsnetz (DFN) rather than the more common commercial website doodle.com. Despite its imperfect user interface and occasional confirmation emails being sent out in the wrong language, the tool worked well for the organization and scheduling of the experiment and complies with the European General Data Protection Regulation (GDPR), which was the decisive advantage here.

3. Results

Communication with the participants to coordinate the arrival at the measurement facilities, COVID-mitigation requirements, etc., before and after the experiment was done via email and no issues occurred. All participants arrived on time and the rescheduling of an appointment was only necessary in a single case. Since MEG is prone to artifacts due to metallic items, reminders for the participants to avoid wearing clothing with metal parts were sent out in advance. They reportedly prevented some cases where participants would have worn clothes with metal parts otherwise (bra, belt, ...) and helped prepare clean lab clothing in cases where participants could not comply.

Regulations to avoid COVID-19 infections were set by PTB and Charité – Universitätsmedizin Berlin, namely requiring completed vaccination status, self-testing before arrival and mask-wearing. No case of COVID-19 was reported after the experiment despite the average 7-days incidence rate of ~ 320 cases/100 000 inhabitants in Berlin during that time. Eighteen participants were measured on 10 working days in a period of two weeks, the duration of each measurement from pickup at the gate until leaving the facilities never exceeded 1.5 h. The maximum number of participants recorded on a single day was 5, proving that an efficient scheduling with multiple participants per day for an OPM study is doable and safe.

Of the 18 participants complying with the inclusion criteria, two had to be excluded from further analysis after the recording: One retrospectively reported a retainer, which should have been covered by the data acquisition form asking for “metal parts” and even providing “amalgam fillings” as an example. An updated version of the form could possibly contain more examples, e.g., retainers or ear rings, to avoid the repetition of this error.

Another participant did not follow the experimental instructions correctly, as was discovered during a debriefing. The debriefing after the experiment was not formally a part of the protocol, but every participant was asked how they felt during the experiment. Such a conversation after the experiment could reveal important details to improve reproducibility of an experiment, such as distractions in the MSR or inconveniences that did not occur during the pilot measurements. A more formalized debriefing should therefore be an integral part of future experimental protocols.

The 17 available OPMs operated stably throughout all measurements and there were no interruptions during the experiments. Positioning the OPM patch usually took less than 5 min and only in a few cases tape on the sensor patch was required to apply pressure and keep the sensors close to the scalp. Since the OPMs’ tips can be in contact with the scalp and warm up during operation, participants were asked if the temperature felt uncomfortable at any point during the measurement. No complaints were made about the warmth coming from the sensor patch. In fact, most participants indicated that the quiescence in the room, the repetitive experimental paradigm and the slight warming on the head had a soothing effect, but no cases of falling asleep during the measurement were reported.

The trials within the two phases of each block were presented in a randomized fashion and with a random ITI of 3 s to 5 s. Some participants stated that, especially during later trials, switching back into the *relax* state after *close* trials became increasingly hard, possibly explaining the decreasing performance in the *relax* in comparison to the increasing performance in the *close* condition (subsection 3.3.1). A slightly longer ITI, e.g., 4 s to 6 s, could help with the return to baseline and might avoid overlapping trial effects (see also subsection 3.2.2).

Participants had the chance to take a break after each block and could advise the experimenter to continue when they were ready. On average, participants continued after a break of ~ 48 s with the longest break taken after the third block (~ 53 s). All participants finished the experiment within less than 31 min.

4. Discussion

This work evaluated the suitability of commercially available SERF optically pumped magnetometers to implement a brain-computer interface based on modulations of the sensorimotor rhythm over the human motor cortex. In particular to aid the restoration of motor skills in stroke survivors, developing non-invasive means to realize external device control through the near real-time analysis of cortical activity is major goal in the field of clinical neuroscience. Compared to SQUID-MEG, OPM-MEG achieves an increased spatial resolution (Boto et al., 2016) and much lower maintenance cost while allowing movement in the scanner (Boto et al., 2018). As opposed to EEG, MEG does not suffer from conductive inhomogeneities in the skull and does not require direct contact with the scalp, making it less invasive and quicker to set up than EEG. These advantages justify the hypothesis that OPM-based BCI systems could translate brain activity into control signals with higher spatial precision and therefore allow for more complex control of external devices.

The current body of literature on OPM technology for BCI is sparse and requires the investigation of basic BCI paradigms before targeting complex device control with multiple degrees of freedom. Further, using OPMs in standard experiments will ultimately allow benchmarking and comparisons to EEG and SQUID-MEG and help manufacturers make improvements to receive the currently lacking clinical approval for their devices.

The current study constitutes a first step towards establishing OPM technology in the field of clinical neuroscience. Measuring the OPM system's signal characteristics provides evidence that it fulfills the requirements for measuring cortical signals in a shielded environment. In an analysis of the no-feedback trials involving right hand motor imagery, a physiologically realistic spatiotemporal difference of the SMR amplitude between the *relax* and *close* condition was found in a majority of the participants. The OPM system also features an API allowing near real-time access to the data, which was used for the implementation of an online classification pipeline providing visual feedback. Retraining the pipeline during the experiment with the participant's data led to a classification performance significantly above chance level, showing that OPMs are suitable to be used in BCI settings.

Sensor Characterization The first research question asked for the properties and limitations of the OPM system in the available magnetically shielded environment in terms of noise, bandwidth and dynamic range. The OPM manufacturer, FieldLine Inc., provides only unofficial specifications of these signal characteristics, which made a quantification necessary to determine if their OPM system is principally suited for the recording of neuronal signals from the human motor cortex. FieldLine Inc. recommends using the closed loop mode of the sensors, as it massively enhances the dynamic range and adds only little noise in the frequency range of interest for this study.

The empty room recordings revealed a system noise floor of $\sim 27.3 \text{ fT}/\sqrt{\text{Hz}}$ at 10 Hz inside the Ak3b MSR at the PTB in Berlin, which is higher than the unofficial sensor noise floor specification of $15 \text{ fT}/\sqrt{\text{Hz}}$. Assessing the true sensor noise floor as given by the manufacturer is complex because external noise influences would have to be disentangled from the internal sensor noise, exceeding the scope of this thesis. However, for the applied setting of this study, the system noise floor was the more relevant quantity compared to the sensor noise floor, since external noise influences could not be completely shielded off.

By comparing the system noise floor to the power of the SMRs over the motor cortex, an estimate of the achievable signal-to-noise ratio could be computed. Unfortunately, there is no literature examining the amplitude of the SMR with on-scalp measurements to date, which is why

4. Discussion

SQUID-MEG studies are the only reliable source to gauge expected fields. Mellinger et al. (2007) reported amplitudes on the order of magnitude of $100 \text{ fT}/\sqrt{\text{Hz}}$ measured with SQUID-MEG, providing a lower bound of the expected signal power because of the much larger distance between sensor and source compared to on-scalp OPMs. Dividing the lower bound of the SMR power by the OPM system's noise floor then results in a theoretical SNR of at least ~ 4 .

With the presented results, the manufacturer's noise floor specification can neither be confirmed nor rejected, but it is realistic that the discrepancy between measurement and specification is due to external magnetic fields. The shielding factor of ≥ 8000 (78 dB) at 10 Hz in the MSR (VACUUMSCHMELZE GmbH & Co. KG, Hanau, 2019) is high, but still allows for some transient magnetic fields to penetrate through the walls. However, the peak of $60 \text{ fT}/\sqrt{\text{Hz}}$ at 300 Hz only in the closed loop mode can not be explained by external noise. This effect is due to the frequency-dependent gain applied within the acquisition system to increase the closed loop bandwidth to up to 300 Hz. The gain parameters are configured for each sensor individually, but the quality of the configuration can degrade as the sensors age.

The effect of degrading laser stability and gain parameters also explains the considerable variability between the noise floors in individual sensors in the chassis-wise analysis (paragraph 3.1.1 and appendix D.1.1). The sensor noise floor variability is much larger between the sensors than between the chassis, indicating that the sensors are individual units independent of the acquisition device. This assumption has also been confirmed by the manufacturer, who also stated that the gain errors observed in some of the sensors could be eliminated by a recalibration of the sensors (personal communication).

To assess the dynamic range, sinusoidal magnetic fields of varying amplitudes were applied to two randomly selected OPMs. The dynamic range of the sensors was measured to be around $\pm 14.5 \text{ nT}$ in closed loop mode and both tested magnetometers underestimated the applied fields by about 2 %. While the strong open loop decay is in line with the functional principle of SERF-OPMs (Tierney et al., 2019), it is eliminated entirely through the dynamic compensation in closed loop mode. The sharp increase of the measured amplitude when the applied amplitude is increased above 15 nT has been communicated with the manufacturer. This type of railing occurs due to the limited power output of the current source supplying the compensation coils, but will be improved upon in future firmware and hardware versions.

The capability to measure strong field modulations of up to almost 15 nT clearly shows that the magnetometers are suited for measurements in the shielded environment. Cross-axis projection errors have not been quantified here but are negligible in studies not involving large head movements due to the low remnant fields in the Ak3b MSR ($< 10 \text{ nT}$) (Borna et al., 2022). Signals could be measured accurately on top of offset fields of up to 7 nT, showing that even the largest ($< 500 \text{ pT}$) slow ($< 3 \text{ Hz}$) field fluctuations occurring in the MSR do not result in non-linearities in the measured fields.

There was a discrepancy of 2 % between the applied and measured fields, which could be explained by multiple contributing factors: Firstly, the accuracy of the OPM calibration depends on the stability of laser power and wavelength which are inaccessible for the user. It is possible that a linear offset is caused by inaccurate calibration values, but only a recalibration and subsequent reevaluation by the manufacturer could indicate the size of this effect. The magnitude of the applied magnetic fields have been validated with a Fluxgate magnetometer, but its resolution in the nT-range restricts the accuracy of these validation measurements to a few %. A last source of error could be a small deviation in the current-to-field constant of the Twinleaf MS-2 shield in combination with the used current supply unit. Tracing the exact origin of the measured deviations would require a sophisticated setup with precisely calibrated devices, which exceeds the scope of this work. Overall, the linear error between applied and measured sinusoidal amplitude is not expected to impair the quality of the cortical measurements, as only relative measures of

4. Discussion

SMR were used here.

Lastly, the closed loop 3 dB bandwidth was measured to be approximately 400 Hz, which by far exceeds the frequency range of interest in the alpha band (up to 16 Hz). As implied by physical limitations of SERF-OPMs (Tierney et al., 2019), the frequency-dependent attenuation in the open loop mode is very strong and drops below the 3 dB point at around 60 Hz. In the closed loop mode, this attenuation is compensated through a frequency-dependent gain parameter calibrated by the manufacturer to achieve a linear response in frequency space.

The two randomly selected magnetometers show that there is an inter-sensor variability in the quality and accuracy of measurements. The stronger dependency of sensor 1 on external fields and frequency in the open loop mode shows that fluctuations of the laser stability or other sensor components can impair the sensors' accuracy. A recalibration of the sensor gains in the closed loop mode could alleviate these issues (personal communication with the manufacturer) and restore the precision of the sensors without replacement.

The quantified signal characteristics presented here clearly indicate that the used magnetometers should be capable of measuring cortical signals robustly and with high accuracy. Since the ERD is a relative quantity, the linear offset of the sensors does not pose a problem for this study as no exact units were required. The variability between the sensors appeared relatively small compared to the measured signals, and the noise floor, dynamic range and bandwidth surpass the minimum requirements to measure brain activity by far.

Offline Analysis The second research question required the offline analysis of recorded data to determine if significant SMR modulations in the alpha band could be measured, and to investigate the physiology of the recorded signals. The data from the no-feedback phases of the experiment involved grasping motor imagery and relaxation, which were compared by means of spectral and spatial filtering to evaluate the modulatory effect of the two conditions.

A frequency with significant mu rhythm modulations could be identified for 10 of the 16 participants after 9 trials of hand motor imagery and relaxation each. This rate of approximately 38 % of participants for which no SMR modulations were found is consistent with the categorizations developed by Sannelli et al. (2019), based on a large-scale EEG study. They found that for around 40 % of study participants the source of SMR modulations was either not located over the motor cortex, unstable in frequency or not detectable at all. In the current study, a flexible patch with only 17 sensors centered over the C3 electrode was used, not covering the ipsilateral C4 or parietal regions of the skull. The 3D printed patch fit on all participants' heads and sensors were pushed into the slots and fixed in place using tape before starting the experiment. Even though participants did not report any incidents, slight shifting down of the sensor patch or hair pushing the sensors away from the skull cannot be ruled out entirely. Considering the results by Sannelli et al. (2019), the lack of whole head coverage, possible sensor movements and that MEG has lower sensitivity for radial sources compared to EEG (Ahlfors et al., 2010), it is a sign in favor of the suitability of OPM that an ERD frequency could be found for over 60 % of the participants in this study.

After determining the individual ERD frequency, computing the timecourse of the mu rhythm power revealed clear event-related desynchronization during *close* trials and synchronization in some participants during *relax* trials. These results are in line with many previous studies based on EEG and SQUID-MEG and confirm the suitability of OPM technology for measuring modulations of the SMR. To avoid an underestimation of effect timing, causal filters were used in all analyses and cluster-based permutation testing revealed a cluster of significant difference ($p < 0.001$) in the time interval from 0.9 s to 4.5 s. The *close* ERD timecourse indicates that, even though the grand average ERD remained below 0 % until $t > 5$ s, the effect became less stable

4. Discussion

towards the end of the trials. This instability could be due the short-lasting SMR modulations in some individuals (Pfurtscheller et al., 1998; Sannelli et al., 2019) and anticipatory effects towards the end of the trial, as the trial duration was always exactly 5 s. Sannelli et al. (2019) also reported that some participants had difficulties maintaining the ERD in no-feedback trials.

For some participants, a clear ERS could be observed during *relax* trials. Why this ERS was maintained after the end of the trial remains unclear, although it could be an artifact due to delayed relaxation in some users when *relax* trials followed on *close* trials after a short random ITI.

The temporal dynamics on the order of 1 s after cue onset are in line with Graimann et al. (2002) and show that the SMR modulations measured using OPMs are perfectly within the 3 s time frame of sensorimotor integration (Mates et al., 1994)

Visualizing the individual spatial patterns computed using CSP indicated clear dipolar structures in the spatial distribution of channels to discriminate *relax* and *close* trials. The sensor patch was centered over C3 and did not provide full head coverage, hence no claims concerning other source locations can be made and the observed patterns could be part of larger patterns covering the whole head. However, the observed patterns are in line with physiological expectations of the characteristics of magnetoencephalographic cortical sources, namely being dipolar and with individual differences in location and direction (Mellinger et al., 2007; Schaworokow & Nikulin, 2022). The visible patterns, in combination with the significant differences in *relax* and *close* ERD, show that the chosen magnetometer setup covers a sufficient area of the scalp, both in terms of density and placement, to pick up SMR activity in most participants.

Online Analysis One prerequisite for the clinical use of OPMs in feedback-based rehabilitation paradigms is the continuous analysis and classification of the acquired data. The third component of this thesis therefore investigated the performance of a computation pipeline that could perform such online data processing. A preliminary version of the pipeline developed for this project was first presented in Zerfowski et al. (2021), being capable of classifying eyes open/closed from occipital alpha signals.

The modular pipeline developed for this experiment received the OPM data with 17 channels and a sampling frequency of 1000 Hz in chunks of 10 samples. With an effective update frequency of 100 Hz, bandpass filtering and the CSP algorithm were applied before a sliding window buffer further downsampled the data to 10 Hz. Despite running on a non-real-time system, the pipeline processed incoming samples fast enough to achieve a stable classification rate of 10 Hz without running into buffer overflows or discarding data. The modularity of the pipeline architecture allowed automated training during the experiment with minimal experimenter intervention. The training took less than 1 s and could be performed without introducing waiting times for the participant in the scanner. The output of the pipeline, consisting of floating point values whose sign indicates the class label, were integrated into visual feedback during feedback trials. Feedback was provided in the form of a closing Pac-Man shape (*close*) and a color-changing circle (*relax*) and projected onto a screen in the MSR from a projector placed outside of it. The setup worked flawlessly throughout the series of days and participants and no issues with the near real-time acquisition setup could be found.

An analysis of the classification data revealed that the implemented classifier was significantly better than a chance level classifier ($p < 0.01$). This held true for both the theoretical ($p = 50\%$) and the empirical ($p = 49.6\% - 49.9\%$) chance level classifier, for both all participants and only the participants with significant ERD. When filtering for participants with significant ERD, the online classification setup achieved a success rate of 83.7 % in *close* and 51.0 % in *relax* trials. These results are comparable with the EEG-based BCI presented by Witkowski et al. (2014).

However, the combined success rate of 67.3 % is below the value of 70 % which is often assumed

4. Discussion

to be the minimum threshold for efficient BCI control (Kübler et al., 2001; Sannelli et al., 2019). The low combined success rate can be mostly attributed to the relatively low success rate in *relax* trials, due to the frequent violations of the safety criterion defined as < 25 % FPR within a trial (i.e., *close* classification rate < 25 % in a *relax* trial). For BCI systems suffering from this problem, it has become common to enhance them with EOG “veto” control to stop a closing motion and implement asynchronous control (Witkowski et al., 2014; Soekadar, Witkowski, Vitiello, et al., 2015; Crea et al., 2018). Assuming that the EOG-enhancement of the EEG-BCI by Witkowski et al. (2014) could yield similar improvements in the current study, adding EOG to the paradigm would lift the average success rate far above 70 %. It is also important to mention here that the pipeline training was performed without any automated cleaning of the data and on only 9 trials per condition. Thus, any large artifact during a trial in the first block could mask real effects, induce large distortions in the trained weights and decrease the pipeline performance unquantifiably.

Lastly, an analysis of the block-wise success rate was conducted to examine if the repeated training had an effect on the performance of the pipeline. Three mixed linear models were fit to different success rate metrics (*close*, *relax* and combined success rate), but none found a significant influence of the block index on the success rate. This is surprising, since it indicates that the classification performance after the first no-feedback block (i.e., after only 9 trials per condition) already reached a level at which not much improvement could be made by including more trials in subsequent steps of retraining. Usually, EEG-BCIs require much more training data to achieve efficient classification, since they are prone to overfitting due to the large number of dimensions in the sensor-space (Blankertz, Tomioka, et al., 2008). In agreement with Sannelli et al. (2011), it is reasonable to assume that the only 17 channels used here avoided such overfitting and allowed to find a good fit after only very few trials. The stagnation of within-session classification performance is also in line with results reported by Blankertz, Losch, et al. (2008). It could alternatively occur due to balancing effects between the onset of fatigue after the third or fourth experimental block and training effects in the participants and the classification pipeline. A dedicated offline analysis involving the block-wise and cross-validated training of the pipeline on the cleaned datasets could help to identify the source of this effect and should be performed to inform the development of future OPM-based BCI systems.

General findings As the fourth research objective, the conduction of the study was reflected upon to summarize insights which could guide the experimental design of future studies. As of today, there is no established “standard” way of conducting OPM-MEG research and most studies involve only few participants.

Communication with the 18 participants, which were recruited through lab-internal and public communication channels, was done via email and worked flawlessly. Most participants could avoid clothing with metal parts and only a single participant had to be excluded after recording due to potentially artifactual data. Study participants were measured at the facilities of the PTB over the course of two weeks and the self-testing, vaccination and masking requirements were effective at fending off the spread of COVID-19 during the ongoing pandemic.

During the briefing for the experiment, participants were informed about the experimental protocol, could ask questions, and tried the motor imagery paradigm before going into the MSR. In a single case, the motor imagery training failed and the participant did not follow the instructions correctly, leading to their exclusion from further analyses. This was discovered during an informal debriefing, a structured version of which should be made part of future studies to discover potential flaws and issues early on.

The duration of the experiment itself seemed appropriate, even though some participants reported fatigue after the fourth block because of the soothingly warm sensor tips, the quiet

4. Discussion

environment, and the repetitive task. The duration of the breaks during the experiment could be chosen by the participant, which helped them to bring back their focus. All participant appointments were finished within under 1.5 h (including admission to the facilities and debriefing), which is in part also due to the quick setup of the sensor patch (usually < 5 min), much less than would be expected for a common EEG experiment.

Implications The present study demonstrated the suitability of commercially available optically pumped magnetometers for the implementation of a motor imagery-based brain-computer interface. Despite being a “standard” paradigm in the development of MI-BCI, no publications have shown the efficiency of such an OPM-based system for as many participants. The conference paper by Fedosov et al. (2021) involved only a single participant and lacked in-depth descriptions of the used methods, while Paek et al. (2020) did not achieve online real-time control.

To demonstrate the online classification of MI-related SMR synchronization/desynchronization with OPMs, multiple steps were performed: Firstly, a quantitative analysis of the signal characteristics of the available OPMs was conducted to verify the unofficially provided manufacturer specifications. Even though multiple studies with commercially available OPM sensors have been published, only few used the sensors from FieldLine Inc.. Additionally, most studies list only the *sensor* instead of the *system* noise floor, which is less useful to evaluate the sensors in realistic setups. The presented results show that the OPM system by FieldLine Inc., in combination with the available magnetically shielded environment, should achieve the necessary sensitivity for cortical measurements and that linearity of the sensors should be maintained in a no-movement paradigm.

In the second step, data of 16 participants were analyzed to determine if significant event-related desynchronizations could be distinguished between a *relax* and a *close* condition. For 10 participants, a frequency of significant power decrease in the alpha range in *close* trials could be determined, being in line with results on EEG-BCI efficiency (Sannelli et al., 2019). The timecourse of the ERD in eligible participants was also in agreement with other studies, instabilities towards the end of trials were most probably due to the experimental protocol and the common difficulty of maintaining strong MI-ERD for longer periods of time (Sannelli et al., 2019). To find and filter for the ERD source in a stream of 17 channels, the CSP algorithm was used and the spatial patterns visualized in a subsequent step. For all eligible participants, a dipolar pattern in the vicinity of C3 was found, consistent with Mellinger et al. (2007) and Schaworonkow and Nikulin (2022). This confirms again that OPMs have capabilities comparable with SQUID-MEG, but they could be used in more versatile settings due to their flexibility in placement and improved spatial resolution (Boto et al., 2016).

The third, central, research objective was the implementation of a near real-time pipeline to classify the continuous OPM data stream and provide visual feedback. A success rate of 67.3 % far above chance level was computed from the online classification data, showing that BCI control is possible in principle with OPMs. However, the achieved efficiency is below the common 70 % mark for reliable control, presumably due to the relatively low success rate in *relax* trials. This could potentially be resolved through the addition of EOG, as is often done in EEG-BCI systems used for rehabilitation. One particular feature of the system was the low number of trials needed to achieve such a classification performance, the reason for which could be the comparatively low number of sensors (avoiding overfitting) or the high SNR of OPMs.

As the fourth component, a summary of the general findings in terms of organization and conduction of the study was provided. It was shown that an OPM experiment with 18 participants can be conducted safely and without disruptions and that only a few details should be adjusted for future attempts to study an OPM-based BCI.

4. Discussion

Limitations The presented BCI system is the first iteration of its kind and naturally allows for improvements and enhancements, as it was only designed as a proof of principle with validated assumptions and methods selected from previous EEG and SQUID-MEG research.

Most importantly, the selected paradigm only attempted to distinguish between *relax* and *close* MI in a synchronous classification paradigm. This binary control realizes only a single degree of freedom which could be used to open or close an exoskeleton hand, but it is unclear how well the classification would generalize to a two-class paradigm, e.g., left vs. right hand. However, the distinction between the ERD induced by motor execution or motor imagery on both sides would most likely require a larger sensor setup (e.g., 16 sensors over both C3 and C4) to cover both source locations (compare Mellinger et al., 2007). The system further does not allow continuous control by the user, due to the trial-based protocol of the experiment. This shortcoming could be eliminated easily after a brief training phase involving only 9 trials, as shown here. After such training, the classification pipeline is running in near real-time and provides a continuous classification feasible for feedback control.

Since there are no established methods for the automated cleaning of OPM data at the moment, the online pipeline used all available data from previous trials for the training. Even though many EEG-BCI systems follow the same approach, any artifact occurring during the comparatively low number of training trials (9 per condition in the first training; 39 per condition in the last training) could lead to disproportionate distortions of the trained parameters and degrade the classification performance to an unknown degree. A distinction of noise and artifacts which were commonly observed during the manual preprocessing steps of the OPM data was presented in subsection 2.6.1. Informed by this distinction, which lists artifacts not common with EEG or SQUID-MEG, an automated trial rejection could improve the performance of the BCI system.

All 16 eligible participants of this study were healthy, right-handed individuals. Therefore, the number of participants is insufficient to generalize the results for a larger population. Further, no claims concerning clinical populations can be made, as the study would have to be repeated with an “intended movement” paradigm for stroke patients.

Recommendations Currently, no OPM manufacturer has undergone the required safety evaluations to obtain a medical device approval by either the European Medical Agency (EMA, Europe) or the US Food & Drug Administration (FDA, USA). This is one of the most important steps that has to be taken by the manufacturers of such systems for researchers to acquire ethical approval and allow larger studies on clinical populations to be conducted.

In general, other established data acquisition and analysis methods should be explored in the context of OPMs, to improve the signal quality and establish best-practices. For data acquisition, the introduction of capable compensation methods for external magnetic fields both online (dynamic field compensation, Robinson et al., 2022), and offline (homogeneous field correction, Tierney et al., 2021) could considerably improve the SNR of OPM data. Further, a thorough characterization of OPM-specific artifacts is still missing, in particular a quantification of the effects of participant movements in remnant fields and muscle artifacts such as blinking, swallowing and heart beats. To determine these characteristics, modalities like 3D sensor tracking, electromyography, electrooculography or electrocardiography should be added to the data acquisition setup to correlate the OPM data with the potential physiological noise sources.

Informed by these other measurements, the artifact characteristics could be translated into sensible parameters for data preprocessing methods (see Seymour et al., 2022, for a review of the current state) and best-practices that could then be adopted by the OPM-MEG community.

These data rejection methods could not only aid the experimenter in the preprocessing, they could also enable an automated data preprocessing and trial rejection for real-time applications and BCI setups.

4. Discussion

No standard procedures to place or localize OPMs relative to the skull/brain have been established to date, but they would be needed to implement reliable source reconstruction and localization procedures. Extraction of relevant signals could be significantly improved by using co-registration methods, as multiple studies have shown (Boto et al., 2016; Xygonakis et al., 2018; Marhl et al., 2019; Zetter et al., 2019; Pfeiffer et al., 2020; Schaworonkow & Nikulin, 2022).

The performance of OPM-based BCI has certainly not reached its peak yet: Minute differences in the cortical source location of, e.g., the finger representations or reach-and-grasp movements, can lead to largely different magnetic dipole fingerprints which could already be distinguished in SQUID-MEG and EEG (Quandt et al., 2012; Schwarz et al., 2017). The increased spatial resolution and advancements in source localization make OPM technology a promising candidate to study the neural correlates of motor processes in the future. To accomplish more complex classification problems, improvements to the processing pipeline like as filter bank CSP (Ang et al., 2008) with a variant of multi-class CSP (Grosse-Wentrup & Buss, 2008; Lindig-León & Bougrain, 2014) or advanced machine learning methods (Lotte et al., 2018) should be considered. Additional signal modalities such as EOG could be introduced to implement asynchronous paradigms and a “veto” function (Witkowski et al., 2014; Crea et al., 2018). This study could only present visual feedback as the use of an MEG-proof exoskeleton would have exceeded the scope of this thesis. Since regular exoskeleton hands used in EEG-BCI applications contain electrically driven finger actuators inducing unknown and potentially large magnetic field artifacts, they could not be used in conjunction with OPMs. However, a non-magnetic exoskeleton hand using Bowden cables from a previous SQUID-MEG study conducted in Tübingen (Soekadar et al., 2011) is available and should be easy to integrate with the existing modular pipeline.

5. Conclusion

Brain-computer interfaces are an important tool for the restoration of movement and communication in disabled individuals and can often lead to drastic improvements in their quality of life. Current non-invasive state-of-the-art BCI systems are based on electroencephalography, limiting the versatility of such systems due its relatively low spatial resolution. In contrast, optically pumped magnetometers measure the magnetic field fluctuations induced by cortical activity and promise to achieve more precise source localization as well as lower maintenance cost than SQUID-based MEG (Boto et al., 2016). OPMs are therefore a promising new technology to gain more knowledge about cognitive processes in the brain and to develop BCI applications that could massively enhance the capabilities of clinical neurotechnology in the near future. The intention of the current study was to provide a proof of principle for a motor imagery-based BCI using 17 commercially available OPM sensors.

To implement such a BCI, the first research objective aimed to assess the suitability of the OPM system's signal characteristics to record cortical data in the available shielded environment. These characteristics were quantified in terms of noise floor, dynamic range and bandwidth: The system's noise floor of $27 \text{ fT}/\sqrt{\text{Hz}}$ at 10 Hz is higher than the sensor noise of commercial SQUIDs-MEG, but the much smaller source-sensor distance balances this effect such that cortical oscillations could be measured with a high signal-to-noise ratio. The OPM system features a dynamic field compensation mode to achieve linear operation of the sensors in longitudinal external fields of up to $\pm 14.5 \text{ nT}$, which is well within the limits of the static fields and slow drifts in the magnetically shielded environment. To study cortical oscillations in the alpha range, a (3 dB) bandwidth of at least 16 Hz was needed, which was by far exceeded by the measured sensor bandwidth of $\sim 400 \text{ Hz}$. From this first analysis, it was concluded that the signal characteristics of the OPM system should be sufficient to measure cortical signals from the motor cortex in the available MSR.

In a second step, a right hand motor imagery experiment involving separated no-feedback and feedback trials was developed. A grid of 17 OPM sensors was placed on the scalp around the C3 electrode location and visual cues were presented on a screen indicating if the participant should *relax* or perform grasping motor imagery (*close*). The experiment consisted of 5 blocks, each of which contained 9 trials per condition without feedback and 6 trials per condition in which visual feedback was presented. The visual feedback was controlled by a modular processing pipeline continuously running in the background and classifying the incoming data. In total, 18 participants were measured, two of which had to be excluded after the recording because they did not comply with the instructions.

The no-feedback data were analyzed offline to determine how well the expected modulations of the sensorimotor rhythm could be measured, how they developed over time and where they were located. In 10 of the 16 eligible participants, a frequency which displayed a significant desynchronization of the SMR in the alpha band during grasping motor imagery, could be determined. A significant difference in the ERD timecourse between *close* and *relax* trials after 1 s was shown, which is in line with previous EEG and MEG studies, and fast enough to warrant suitability of the ERD signal for use in a near real-time BCI. Within the limited region of the skull covered by the sensor patch, individual dipolar field patterns were observed in a qualitative analysis of the most discriminative spatial pattern determined by the CSP algorithm. The individual distribution of the dipolar field patterns is a major advantage of MEG-based BCI as compared to EEG, since small differences in source location on the irregular cortical surface can lead to distinct changes in the resulting field patterns used for classification.

5. Conclusion

The third research objective asked for the performance of the classification pipeline that provided the visual feedback to the user. The pipeline had a simple modular setup, implementing a bandpass filter to extract the signal around the individual ERD frequency, CSP to reduce the data dimensionality, a buffer to downsample the signal to 10 Hz and compute the sliding log-variance over windows of 400 ms, before classifying the signal with the LDA algorithm. Despite the fast rate of 1000 incoming samples per second on a non-real-time system, the processing pipeline ran smoothly without any buffer overflows or having to discard data since only relatively simple filter and matrix operations were used. The regular training of the pipeline took less than a second and happened with minimal experimenter intervention, except the determination of the participant's individual ERD frequency at the start of the experiment. The determined reliability and safety of the feedback control were comparable to the EEG-based system by Witkowski et al. (2014) and the combined classification performance of 67.3 % was significantly better than a chance level classifier. However, it did not reach the common threshold of 70 % for reliable BCI control. Regular violations of the safety threshold indicate that improvements, for instance by implementing EOG-based control, are needed for rehabilitative purposes and could even facilitate the use of the system in asynchronous BCI paradigms.

This is the currently largest published study systematically investigating optically pumped magnetometers for brain-computer interfacing. Best-practices for conducting OPM experiments are only beginning to be established, but the implemented protocol worked well and seemed to have only minor flaws in terms of participant communication and duration of the experiment. However, the presented results are clear evidence that OPMs are suited in principle for the implementation of motor imagery-based BCI systems and are well in line with other publications assessing EEG and SQUID-based BCIs. Not only did this study prove that OPMs can measure ERD over the motor cortex, it also showed that large studies with multiple subjects per day can be conducted safely. Preparation times were considerably shorter and less invasive compared to EEG, since no direct skin contact or conductive gels were needed.

Many labs and companies around the world are currently developing and improving OPM sensors for the measurement of biomagnetic signals. Simultaneously, existing OPM-MEG systems become more affordable and easier to use, and more neuroscience researchers have growing interest in the technology. Through the emergence of easier and cheaper solutions for both passive and active magnetic shielding and field compensation, OPMs are becoming accessible to research groups that had no access MEG before. Standardized procedures for data preprocessing, co-registration and source reconstruction are being developed rapidly by the flourishing OPM-MEG community, and toolboxes are constantly improving their support for the new systems.

With all these developments, it can be expected that optically pumped magnetometers will become an integral part of cognitive neuroscience within the next few years. Particularly the field of clinical neuroscience could profit from the adoption of OPM technology – new paradigms requiring precise source localization and movement in diverse clinical patient groups are projected to be possible and could induce a step change in the development of rehabilitation approaches for disabled patients.

Attributions and Funding

This research study was conducted and this thesis was written by me, Jan Zerfowski. In the following I point out in which parts of this work I received help from others and attribute their contributions.

Source code and data availability The majority of the source code used for the data acquisition and analysis has been made available under an open source license. Please refer to appendix A for an overview of where the source code can be found, a reference to the use of external python packages and an explanation of how to obtain the recorded OPM data in an anonymized format.

Sensor Characterization The paradigm to characterize the sensors was developed by me, together with Thomas Middelmann, Tilmann Sander-Thömmes and Tido Willms (all PTB). I constructed the used sensor fixtures and recorded all data myself. The data analyses were developed by me and I wrote the entire code to execute the analyses.

Experimental Paradigm The experimental paradigm was developed by me, with useful hints in terms of trial durations, repetitions, etc. provided by Niels Peekhaus, Mareike Vermehren, Annalisa Colucci, Alessia Cavallo and Elisa Braufse (all Clinical Neurotechnology Lab). The motor imagery training sheet (appendix E.1) was developed by Mareike Vermehren and Annalisa Colucci. The flexible 3D sensor patch with brackets for the FieldLine sensors was developed and printed by Tilmann Sander-Thömmes (PTB). The projector and communication system in the MSR were provided by PTB and maintained by Tilmann Sander-Thömmes. Participant recruitment and scheduling was performed entirely by me, as well as the briefing and measurement of the participants. A single participant recording was performed by Niels Peekhaus under my supervision.

Processing Pipeline I developed the general architecture of the processing pipeline and wrote the code for the processing nodes myself. Only the bandpass filter node, the buffer node and the single pole filter node were implemented with the help of Niels Peekhaus. Niels Peekhaus wrote a large portion of the pre-existing BeamBCI framework and modules, including the basic code for the stimulus presentation module, which was edited by me. During the writing of this thesis, large contributions to the code of BeamBCI were made by me.

Data analysis All data analyses were developed and the code to perform the analyses was written by me. A small fraction of the code that was used to insert the relative coordinates of the flexible sensor patch into an mne.Info object was written by Urban Mahrl in megtools (see also appendix A).

Funding Funding for this research project was provided by

Einstein Stiftung Berlin Einstein Forschungsvorhaben 2019-558

European Research Council ERC StG NGBMI 759370

Bundesministerium für Bildung und Forschung Zukunftscluster QSENS (QHMI, 03ZU1110DD)

Glossary

- API** Application Programming Interface. 9, 17, 22, 39
- BCI** Brain Computer Interface. 4–7, 9, 10, 15, 21, 22, 24, 25, 31, 32, 34, 37, 39, 43–48
- CSP** Common Spatial Pattern. 20–25, 32–34, 42, 44, 47, 48, 59
- DAQ** Data Acquisition System. 9
- EEG** Electroencephalography. 4–10, 13, 20, 24, 31–33, 39, 41, 43–48
- EOG** Electrooculography. 18, 43, 45, 46, 48
- ERD** Event-Related Desynchronization. 4–7, 9, 16, 19–26, 31–36, 41–45, 47, 48, 57, 58, 65–78
- ERS** Event-Related Synchronization. 4, 6, 32, 33, 42
- FIR** finite impulse response (filter). 19, 21, 23
- FPR** False Positive Rate (rate of *close* classifications in *relax* trials). 26, 35, 43
- ITI** Inter-trial Interval. 15, 19, 38, 42
- LDA** Linear Discriminant Analysis. 23, 25, 48
- LSL** LabStreamingLayer, see also appendix C.2.1. 17, 18, 22–24, 56
- MEG** Magnetoencephalography. 4–10, 20, 24, 31, 33, 38–41, 43–48, 54
- MI** Motor Imagery. 6, 10, 15, 20, 25, 35, 44, 45
- MSR** Magnetically Shielded Room. 12, 15, 18, 27, 38–40, 42, 44, 47, 49
- OPM** Optically Pumped Magnetometer. 4, 5, 7–13, 15, 17, 18, 20, 24, 27–32, 37–49, 54
- PTB** Physikalisch-Technische Bundesanstalt. 11–15, 37–39, 43, 52
- SERF** Spin-Exchange Relaxation Free. 4, 8, 11, 30, 39–41
- SMR** sensorimotor rhythm. 4–7, 16, 20, 21, 24, 32, 33, 39–42, 44, 47
- SNR** Signal to Noise-ratio. 20, 40, 45
- SQUID** Superconducting QUantum Interference Device. 4, 6, 8–10, 39–41, 44–48
- TPR** True Positive Rate (rate of *close* classifications in *close* trials). 26, 35

Bibliography

- Abiri, R., Borhani, S., Sellers, E., Jiang, Y., & Zhao, X. (2018). A comprehensive review of EEG-based brain-computer interface paradigms. *Journal of Neural Engineering*, 16. <https://doi.org/10.1088/1741-2552/aaf12e>
- Ahlfors, S. P., Han, J., Belliveau, J. W., & Hämäläinen, M. S. (2010). Sensitivity of MEG and EEG to Source Orientation. *Brain Topography*, 23(3), 227–232. <https://doi.org/10.1007/s10548-010-0154-x>
- Alken, P., Thébault, E., Beggan, C. D., Amit, H., Aubert, J., Baerenzung, J., Bondar, T. N., Brown, W. J., Califf, S., Chambodut, A., Chulliat, A., Cox, G. A., Finlay, C. C., Fournier, A., Gillet, N., Grayver, A., Hammer, M. D., Holschneider, M., Huder, L., ... Zhou, B. (2021). International Geomagnetic Reference Field: The thirteenth generation. *Earth, Planets and Space*, 73(1), 49. <https://doi.org/10.1186/s40623-020-01288-x>
- Allison, B. Z., & Neuper, C. (2010). Could Anyone Use a BCI? In D. S. Tan & A. Nijholt (Eds.), *Brain-Computer Interfaces* (pp. 35–54). Springer London. https://doi.org/10.1007/978-1-84996-272-8_3
- Ang, K., Chin, Z., Zhang, H., & Guan, C. (2008). Filter Bank Common Spatial Pattern (FBCSP) in brain-computer interface, 2390–2397. <https://doi.org/10.1109/IJCNN.2008.4634130>
- Bell, W. E., & Bloom, A. L. (1957). Optical Detection of Magnetic Resonance in Alkali Metal Vapor. *Physical Review*, 107(6), 1559–1565. <https://doi.org/10.1103/PhysRev.107.1559>
- Birbaumer, N., & Cohen, L. G. (2007). Brain-computer interfaces: Communication and restoration of movement in paralysis. *The Journal of Physiology*, 579(3), 621–636. <https://doi.org/10.1113/jphysiol.2006.125633>
- Blankertz, B., Kawanabe, M., Tomioka, R., Hohlefeld, F., Müller, K.-R., & Nikulin, V. (2007). Invariant common spatial patterns: Alleviating nonstationarities in brain-computer interfacing. In J. Platt, D. Koller, Y. Singer, & S. Roweis (Eds.), *Advances in neural information processing systems*. Curran Associates, Inc. <https://proceedings.neurips.cc/paper/2007/file/6aab1270668d8cac7cef2566a1c5f569-Paper.pdf>
- Blankertz, B., Losch, F., Krauledat, M., Dornhege, G., Curio, G., & Muller, K.-R. (2008). The Berlin Brain-Computer Interface: Accurate performance from first-session in BCI-naïve subjects. *IEEE Transactions on Biomedical Engineering*, 55(10), 2452–2462. <https://doi.org/10.1109/TBME.2008.923152>
- Blankertz, B., Tomioka, R., Lemm, S., Kawanabe, M., & Muller, K.-r. (2008). Optimizing Spatial filters for Robust EEG Single-Trial Analysis. *IEEE Signal Processing Magazine*, 25(1), 41–56. <https://doi.org/10.1109/MSP.2008.4408441>
- Bloom, A. L. (1962). Principles of Operation of the Rubidium Vapor Magnetometer. *Applied Optics*, 1(1), 61–68. <https://doi.org/10.1364/AO.1.000061>
- Bork, J., Hahlbohm, H., & Klein, R. (2001). The 8-layered magnetically shielded room of the PTB: Design and construction. https://www.ptb.de/cms/fileadmin/internet/fachabteilungen/abteilung_8/8.2_biosignale/8.21/mssr.pdf
- Borna, A., Iivanainen, J., Carter, T. R., McKay, J., Taulu, S., Stephen, J., & Schwindt, P. D. D. (2022). Cross-Axis projection error in optically pumped magnetometers and its implication for magnetoencephalography systems. *NeuroImage*, 247, 118818. <https://doi.org/10.1016/j.neuroimage.2021.118818>
- Boto, E., Bowtell, R., Krüger, P., Fromhold, T. M., Morris, P. G., Meyer, S. S., Barnes, G. R., & Brookes, M. J. (2016). On the Potential of a New Generation of Magnetometers for

Bibliography

- MEG: A Beamformer Simulation Study. *PLOS ONE*, 11(8), e0157655. <https://doi.org/10.1371/journal.pone.0157655>
- Boto, E., Hill, R. M., Rea, M., Holmes, N., Seedat, Z. A., Leggett, J., Shah, V., Osborne, J., Bowtell, R., & Brookes, M. J. (2021). Measuring functional connectivity with wearable MEG. *NeuroImage*, 230, 117815. <https://doi.org/10.1016/j.neuroimage.2021.117815>
- Boto, E., Holmes, N., Leggett, J., Roberts, G., Shah, V., Meyer, S. S., Muñoz, L. D., Mullinger, K. J., Tierney, T. M., Bestmann, S., Barnes, G. R., Bowtell, R., & Brookes, M. J. (2018). Moving magnetoencephalography towards real-world applications with a wearable system. *Nature*, 555(7698), 657–661. <https://doi.org/10.1038/nature26147>
- Brookes, M. J., Boto, E., Rea, M., Shah, V., Osborne, J., Holmes, N., Hill, R. M., Leggett, J., Rhodes, N., & Bowtell, R. (2021). Theoretical advantages of a triaxial optically pumped magnetometer magnetoencephalography system. *NeuroImage*, 236, 118025. <https://doi.org/10.1016/j.neuroimage.2021.118025>
- Crea, S., Nann, M., Trigili, E., Cordella, F., Baldoni, A., Badesa, F. J., Catalán, J. M., Zollo, L., Vitiello, N., Aracil, N. G., & Soekadar, S. R. (2018). Feasibility and safety of shared EEG/EOG and vision-guided autonomous whole-arm exoskeleton control to perform activities of daily living. *Scientific Reports*, 8(1), 10823. <https://doi.org/10.1038/s41598-018-29091-5>
- Dash, D., Wisler, A., Ferrari, P., Davenport, E. M., Maldjian, J., & Wang, J. (2020). MEG Sensor Selection for Neural Speech Decoding. *IEEE Access*, 8, 182320–182337. <https://doi.org/10.1109/ACCESS.2020.3028831>
- Davison, A. C., & Hinkley, D. V. (1997). *Bootstrap Methods and their Application*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511802843>
- Dehmelt, H. G. (1957). Slow Spin Relaxation of Optically Polarized Sodium Atoms. *Physical Review*, 105(5), 1487–1489. <https://doi.org/10.1103/PhysRev.105.1487>
- Dickhaus, T., Sannelli, C., Müller, K.-R., Curio, G., & Blankertz, B. (2009). Predicting BCI performance to study BCI illiteracy. *BMC Neuroscience*, 10(1), P84. <https://doi.org/10.1186/1471-2202-10-S1-P84>
- Fedorov, N., Shevtsov, O., & Ossadtchi, A. (2021). Motor-Imagery BCI with Low-Count of Optically Pumped Magnetometers. *2021 Third International Conference Neurotechnologies and Neurointerfaces (CNN)*, 16–18. <https://doi.org/10.1109/CNN53494.2021.9580251>
- Feigin, V. L., Abajobir, A. A., Abate, K. H., Abd-Allah, F., Abdulle, A. M., Abera, S. F., Abyu, G. Y., Ahmed, M. B., Aichour, A. N., Aichour, I., Aichour, M. T. E., Akinyemi, R. O., Alabed, S., Al-Raddadi, R., Alvis-Guzman, N., Amare, A. T., Ansari, H., Anwari, P., Ärnlöv, J., ... Vos, T. (2017). Global, regional, and national burden of neurological disorders during 1990–2015: A systematic analysis for the Global Burden of Disease Study 2015. *The Lancet Neurology*, 16(11), 877–897. [https://doi.org/10.1016/S1474-4422\(17\)30299-5](https://doi.org/10.1016/S1474-4422(17)30299-5)
- Fisher, R. A. (1936). THE USE OF MULTIPLE MEASUREMENTS IN TAXONOMIC PROBLEMS. *Annals of Eugenics*, 7(2), 179–188. <https://doi.org/10.1111/j.1469-1809.1936.tb02137.x>
- Foldes, S. T., Weber, D. J., & Collinger, J. L. (2015). MEG-based neurofeedback for hand rehabilitation. *Journal of NeuroEngineering and Rehabilitation*, 12(1), 85. <https://doi.org/10.1186/s12984-015-0076-7>
- Fukunaga, K. (1990). *Introduction to statistical pattern recognition* (2. ed). Acad. Press.
- Graimann, B., Huggins, J. E., Levine, S. P., & Pfurtscheller, G. (2002). Visualization of significant ERD/ERS patterns in multichannel EEG and ECoG data. *Clinical Neurophysiology*, 113(1), 43–47. [https://doi.org/10.1016/S1388-2457\(01\)00697-6](https://doi.org/10.1016/S1388-2457(01)00697-6)

Bibliography

- Graimann, B., & Pfurtscheller, G. (2006, January 1). Quantification and visualization of event-related changes in oscillatory brain activity in the time-frequency domain. In C. Neuper & W. Klimesch (Eds.), *Progress in Brain Research* (pp. 79–97). Elsevier. [https://doi.org/10.1016/S0079-6123\(06\)59006-5](https://doi.org/10.1016/S0079-6123(06)59006-5)
- Grosse-Wentrup, M., & Buss, M. (2008). Multiclass Common Spatial Patterns and Information Theoretic Feature Extraction. *IEEE Transactions on Biomedical Engineering*, 55(8), 1991–2000. <https://doi.org/10.1109/TBME.2008.921154>
- He, K., Wan, S., Sheng, J., Liu, D., Wang, C., Li, D., Qin, L., Luo, S., Qin, J., & Gao, J.-H. (2019). A high-performance compact magnetic shield for optically pumped magnetometer-based magnetoencephalography. *Review of Scientific Instruments*, 90(6), 064102. <https://doi.org/10.1063/1.5066250>
- Hill, N. J., Lal, T. N., Schröder, M., Hinterberger, T., Widman, G., Elger, C. E., Schölkopf, B., & Birbaumer, N. (2006). Classifying Event-Related Desynchronization in EEG, ECoG and MEG Signals. In K. Franke, K.-R. Müller, B. Nickolay, & R. Schäfer (Eds.), *Pattern Recognition* (pp. 404–413). Springer. https://doi.org/10.1007/11861898_41
- Hill, R. M., Boto, E., Rea, M., Holmes, N., Leggett, J., Coles, L. A., Papastavrou, M., Everton, S., Hunt, B. A., Sims, D., Osborne, J., Shah, V., Bowtell, R., & Brookes, M. J. (2020). Multi-Channel Whole-Head OPM-MEG: Helmet Design and a Comparison with a Conventional System. <https://doi.org/10.1101/2020.03.12.989129>
- Holmes, N., Tierney, T. M., Leggett, J., Boto, E., Mellor, S., Roberts, G., Hill, R. M., Shah, V., Barnes, G. R., Brookes, M. J., & Bowtell, R. (2019). Balanced, bi-planar magnetic field and field gradient coils for field compensation in wearable magnetoencephalography. *Scientific Reports*, 9(1), 14196. <https://doi.org/10.1038/s41598-019-50697-w>
- Iivanainen, J., Zetter, R., Grön, M., Hakkarainen, K., & Parkkonen, L. (2019). On-scalp MEG system utilizing an actively shielded array of optically-pumped magnetometers. *NeuroImage*, 194, 244–258. <https://doi.org/10.1016/j.neuroimage.2019.03.022>
- Knappe, S., Sander, T. H., Kosch, O., Wiekhorst, F., Kitching, J., & Trahms, L. (2010). Cross-validation of microfabricated atomic magnetometers with superconducting quantum interference devices for biomagnetic applications. *Applied Physics Letters*, 97(13), 133703. <https://doi.org/10.1063/1.3491548>
- Koles, Z. J., Lazar, M. S., & Zhou, S. Z. (1990). Spatial patterns underlying population differences in the background EEG. *Brain Topography*, 2(4), 275–284. <https://doi.org/10.1007/BF01129656>
- Kübler, A., Nijboer, F., Mellinger, J., Vaughan, T. M., Pawelzik, H., Schalk, G., McFarland, D. J., Birbaumer, N., & Wolpaw, J. R. (2005). Patients with ALS can use sensorimotor rhythms to operate a brain-computer interface. *Neurology*, 64(10), 1775–1777. <https://doi.org/10.1212/01.WNL.0000158616.43002.6D>
- Kübler, A., Neumann, N., Kaiser, J., Kotchoubey, B., Hinterberger, T., & Birbaumer, N. P. (2001). Brain-computer communication: Self-regulation of slow cortical potentials for verbal communication. *Archives of Physical Medicine and Rehabilitation*, 82(11), 1533–1539. <https://doi.org/10.1053/apmr.2001.26621>
- Lal, T. N., Birbaumer, N., Schölkopf, B., Schröder, M., Hill, N. J., Preissl, H., Hinterberger, T., Mellinger, J., Bogdan, M., Rosenstiel, W., & Hofmann, T. (2005). A brain computer interface with online feedback based on magnetoencephalography. *Proceedings of the 22nd International Conference on Machine Learning - ICML '05*, 465–472. <https://doi.org/10.1145/1102351.1102410>
- Langhorne, P., Coupar, F., & Pollock, A. (2009). Motor recovery after stroke: A systematic review. *The Lancet Neurology*, 8(8), 741–754. [https://doi.org/10.1016/S1474-4422\(09\)70150-4](https://doi.org/10.1016/S1474-4422(09)70150-4)

Bibliography

- Lawrence, E. S., Coshall, C., Dundas, R., Stewart, J., Rudd, A. G., Howard, R., & Wolfe, C. D. A. (2001). Estimates of the Prevalence of Acute Stroke Impairments and Disability in a Multiethnic Population. *Stroke*, 32(6), 1279–1284. <https://doi.org/10.1161/01.STR.32.6.1279>
- Lemm, S., Blankertz, B., Curio, G., & Muller, K.-R. (2005). Spatio-Spectral Filters for Improving the Classification of Single Trial EEG. *IEEE Transactions on Biomedical Engineering*, 52(9), 1541–1548. <https://doi.org/10.1109/TBME.2005.851521>
- Lin, P. T., Sharma, K., Holroyd, T., Battapady, H., Fei, D.-Y., & Bai, O. (2013, June 19). A High Performance MEG Based BCI Using Single Trial Detection of Human Movement Intention. In F. Signorelli (Ed.), *Functional Brain Mapping and the Endeavor to Understand the Working Brain*. InTech. <https://doi.org/10.5772/54550>
- Lindig-León, C., & Bougrain, L. (2014). A comparison between different Multiclass Common Spatial Pattern approaches for identification of motor imagery tasks. Retrieved August 18, 2022, from <https://hal.inria.fr/hal-01092272>
- Lotte, F., Bougrain, L., Cichocki, A., Clerc, M., Congedo, M., Rakotomamonjy, A., & Yger, F. (2018). A review of classification algorithms for EEG-based brain-computer interfaces: A 10 year update. *Journal of Neural Engineering*, 15(3), 031005. <https://doi.org/10.1088/1741-2552/aab2f2>
- Lotte, F., & Guan, C. (2011). Regularizing Common Spatial Patterns to Improve BCI Designs: Unified Theory and New Algorithms. *IEEE Transactions on Biomedical Engineering*, 58(2), 355–362. <https://doi.org/10.1109/TBME.2010.2082539>
- Marhl, U., Jodko-Władzinska, A., Brühl, R., Sander, T. H., & Jazbinšek, V. (2019). Application of source localization algorithms in magnetoencephalography: Test on a new generation of magnetometers. *2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, 357–362. <https://doi.org/10.23919/MIPRO.2019.8756998>
- Marhl, U., Jodko-Władzinska, A., Brühl, R., Sander, T. H., & Jazbinšek, V. (2021). Comparison Between Conventional SQUID Based and Novel OPM Based Measuring Systems in MEG. In T. Jarm, A. Cvetkoska, S. Mahnič-Kalamiza, & D. Miklavcic (Eds.), *8th European Medical and Biological Engineering Conference* (pp. 254–261). Springer International Publishing. https://doi.org/10.1007/978-3-030-64610-3_30
- Marhl, U., Jodko-Władzinska, A., Brühl, R., Sander, T. H., & Jazbinšek, V. (2022). Transforming and comparing data between standard SQUID and OPM-MEG systems. *PLOS ONE*, 17(1), e0262669. <https://doi.org/10.1371/journal.pone.0262669>
- Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, 164(1), 177–190. <https://doi.org/10.1016/j.jneumeth.2007.03.024>
- Mates, J., Müller, U., Radil, T., & Pöppel, E. (1994). Temporal integration in sensorimotor synchronization. *Journal of Cognitive Neuroscience*, 6(4), 332–340. <https://doi.org/10.1162/jocn.1994.6.4.332>
- McFarland, D. J., Daly, J., Boulay, C., & Parvaz, M. A. (2017). Therapeutic applications of BCI technologies. *Brain-Computer Interfaces*, 4(1-2), 37–52. <https://doi.org/10.1080/2326263X.2017.1307625>
- McFarland, D. J., McCane, L. M., David, S. V., & Wolpaw, J. R. (1997). Spatial filter selection for EEG-based communication. *Electroencephalography and Clinical Neurophysiology*, 103(3), 386–394. [https://doi.org/10.1016/S0013-4694\(97\)00022-2](https://doi.org/10.1016/S0013-4694(97)00022-2)
- Mellinger, J., Schalk, G., Braun, C., Preissl, H., Rosenstiel, W., Birbaumer, N., & Kübler, A. (2007). An MEG-based Brain-Computer Interface (BCI). *NeuroImage*, 36(3), 581–593. <https://doi.org/10.1016/j.neuroimage.2007.03.019>

Bibliography

- Nardelli, N. V., Krzyzewski, S. P., & Knappe, S. A. (2019). Reducing crosstalk in optically-pumped magnetometer arrays. *Physics in Medicine & Biology*, 64(21), 21NT03. <https://doi.org/10.1088/1361-6560/ab4c06>
- Nardelli, N. V., Perry, A. R., Krzyzewski, S. P., & Knappe, S. A. (2020). A conformal array of microfabricated optically-pumped first-order gradiometers for magnetoencephalography. *EPJ Quantum Technology*, 7(1), 1–11. <https://doi.org/10.1140/epjqt/s40507-020-00086-4>
- Novi, Q., Guan, C., Dat, T. H., & Xue, P. (2007). Sub-band Common Spatial Pattern (SBCSP) for Brain-Computer Interface. *2007 3rd International IEEE/EMBS Conference on Neural Engineering*, 204–207. <https://doi.org/10.1109/CNE.2007.369647>
- Okada, Y., Hämäläinen, M., Pratt, K., Mascarenas, A., Miller, P., Han, M., Robles, J., Cavallini, A., Power, B., Sieng, K., Sun, L., Lew, S., Doshi, C., Ahtam, B., Dinh, C., Esch, L., Grant, E., Nummenmaa, A., & Paulson, D. (2016). BabyMEG: A whole-head pediatric magnetoencephalography system for human brain development research. *Review of Scientific Instruments*, 87(9), 094301. <https://doi.org/10.1063/1.4962020>
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9(1), 97–113. [https://doi.org/10.1016/0028-3932\(71\)90067-4](https://doi.org/10.1016/0028-3932(71)90067-4)
- Paek, A. Y., Kilicarslan, A., Korenko, B., Gerginov, V., Knappe, S., & Contreras-Vidal, J. L. (2020). Towards a Portable Magnetoencephalography Based Brain Computer Interface with Optically-Pumped Magnetometers. *2020 42nd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC)*, 3420–3423. <https://doi.org/10.1109/EMBC44109.2020.9176159>
- Papadopoulos, S., Bonaiuto, J., & Mattout, J. (2022). An Impending Paradigm Shift in Motor Imagery Based Brain-Computer Interfaces. *Frontiers in Neuroscience*, 15. Retrieved March 18, 2022, from <https://www.frontiersin.org/article/10.3389/fnins.2021.824759>
- Pfeiffer, C., Ruffieux, S., Andersen, L. M., Kalabukhov, A., Winkler, D., Oostenveld, R., Lundqvist, D., & Schneiderman, J. F. (2020). On-scalp MEG sensor localization using magnetic dipole-like coils: A method for highly accurate co-registration. *NeuroImage*, 212, 116686. <https://doi.org/10.1016/j.neuroimage.2020.116686>
- Pfurtscheller, G., & Aranibar, A. (1979). Evaluation of event-related desynchronization (ERD) preceding and following voluntary self-paced movement. *Electroencephalography and Clinical Neurophysiology*, 46(2), 138–146. [https://doi.org/10.1016/0013-4694\(79\)90063-4](https://doi.org/10.1016/0013-4694(79)90063-4)
- Pfurtscheller, G., Guger, C., Müller, G., Krausz, G., & Neuper, C. (2000). Brain oscillations control hand orthosis in a tetraplegic. *Neuroscience Letters*, 292(3), 211–214. [https://doi.org/10.1016/S0304-3940\(00\)01471-3](https://doi.org/10.1016/S0304-3940(00)01471-3)
- Pfurtscheller, G., & Lopes da Silva, F. H. (1999). Event-related EEG/MEG synchronization and desynchronization: Basic principles. *Clinical Neurophysiology*, 110(11), 1842–1857. [https://doi.org/10.1016/S1388-2457\(99\)00141-8](https://doi.org/10.1016/S1388-2457(99)00141-8)
- Pfurtscheller, G., Neuper, C., Andrew, C., & Edlinger, G. (1997). Foot and hand area mu rhythms. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, 26(1-3), 121–135. [https://doi.org/10.1016/s0167-8760\(97\)00760-5](https://doi.org/10.1016/s0167-8760(97)00760-5)
- Pfurtscheller, G., Neuper, C., Schlogl, A., & Lugger, K. (1998). Separability of EEG signals recorded during right and left motor imagery using adaptive autoregressive parameters. *IEEE Transactions on Rehabilitation Engineering*, 6(3), 316–325. <https://doi.org/10.1109/86.712230>
- Pfurtscheller, G., Neuper, C., Flotzinger, D., & Pregenzer, M. (1997). EEG-based discrimination between imagination of right and left hand movement. *Electroencephalography and Clinical Neurophysiology*, 103(6), 642–651. [https://doi.org/10.1016/S0013-4694\(97\)00080-1](https://doi.org/10.1016/S0013-4694(97)00080-1)

Bibliography

- Quandt, F., Reichert, C., Hinrichs, H., Heinze, H., Knight, R., & Rieger, J. (2012). Single trial discrimination of individual finger movements on one hand: A combined MEG and EEG study. *NeuroImage*, 59(4), 3316–3324. <https://doi.org/10.1016/j.neuroimage.2011.11.053>
- Ramos-Murguialday, A., Broetz, D., Rea, M., Läer, L., Yilmaz, Ö., Brasil, F. L., Liberati, G., Curado, M. R., Garcia-Cossio, E., Vyziotis, A., Cho, W., Agostini, M., Soares, E., Soekadar, S., Caria, A., Cohen, L. G., & Birbaumer, N. (2013). Brain-machine interface in chronic stroke rehabilitation: A controlled study. *Annals of Neurology*, 74(1), 100–108. <https://doi.org/10.1002/ana.23879>
- Robinson, S. E., Andonegui, A. B., Holroyd, T., Hughes, K. J., Alem, O., Knappe, S., Maydew, T., Griesshammer, A., & Nugent, A. (2022). Cross-Axis Dynamic Field Compensation of Optically Pumped Magnetometer Arrays for MEG. *NeuroImage*, 262, 119559. <https://doi.org/10.1016/j.neuroimage.2022.119559>
- Roc, A., Pillette, L., Mladenovic, J., Benaroch, C., Bernard, N., Jeunet, C., & Lotte, F. (2020). A review of user training methods in brain computer interfaces based on mental tasks. *Journal of Neural Engineering*, 18. <https://doi.org/10.1088/1741-2552/abca17>
- Sander, T. H., Jodko-Władzińska, A., Hartwig, S., Brühl, R., & Middelmann, T. (2020). Optically pumped magnetometers enable a new level of biomagnetic measurements. *Advanced Optical Technologies*, 9(5), 247–251. <https://doi.org/10.1515/aot-2020-0027>
- Sander, T. H., Preusser, J., Mhaskar, R., Kitching, J., Trahms, L., & Knappe, S. (2012). Magnetoencephalography with a chip-scale atomic magnetometer. *Biomedical Optics Express*, 3(5), 981–990. <https://doi.org/10.1364/BOE.3.000981>
- Sannelli, C., Vidaurre, C., Müller, K.-R., & Blankertz, B. (2011). CSP patches: An ensemble of optimized spatial filters. An evaluation study. *Journal of Neural Engineering*, 8(2), 025012. <https://doi.org/10.1088/1741-2560/8/2/025012>
- Sannelli, C., Vidaurre, C., Müller, K.-R., & Blankertz, B. (2019). A large scale screening study with a SMR-based BCI: Categorization of BCI users and differences in their SMR activity. *PLOS ONE*, 14(1), e0207351. <https://doi.org/10.1371/journal.pone.0207351>
- Sassenhagen, J., & Draschkow, D. (2019). Cluster-based permutation tests of MEG/EEG data do not establish significance of effect latency or location. *Psychophysiology*, 56(6), e13335. <https://doi.org/10.1111/psyp.13335>
- Schaworonkow, N., & Nikulin, V. V. (2022). Is sensor space analysis good enough? Spatial patterns as a tool for assessing spatial mixing of EEG/MEG rhythms. *NeuroImage*, 253, 119093. <https://doi.org/10.1016/j.neuroimage.2022.119093>
- Schwarz, A., Ofner, P., Pereira, J., Sburlea, A. I., & Müller-Putz, G. R. (2017). Decoding natural reach-and-grasp actions from human EEG. *Journal of Neural Engineering*, 15(1), 016005. <https://doi.org/10.1088/1741-2552/aa8911>
- Seymour, R. A., Alexander, N., Mellor, S., O'Neill, G. C., Tierney, T. M., Barnes, G. R., & Maguire, E. A. (2021). Using OPMs to measure neural activity in standing, mobile participants. *NeuroImage*, 244, 118604. <https://doi.org/10.1016/j.neuroimage.2021.118604>
- Seymour, R. A., Alexander, N., Mellor, S., O'Neill, G. C., Tierney, T. M., Barnes, G. R., & Maguire, E. A. (2022). Interference suppression techniques for OPM-based MEG: Opportunities and challenges. *NeuroImage*, 247, 118834. <https://doi.org/10.1016/j.neuroimage.2021.118834>
- Shah, V., Osborne, J., Orton, J., & Alem, O. (2018, February 22). Fully integrated, standalone zero field optically pumped magnetometer for biomagnetism. In S. M. Shahriar & J. Scheuer (Eds.), *Steep Dispersion Engineering and Opto-Atomic Precision Metrology XI* (p. 51). SPIE. <https://doi.org/10.1117/12.2299197>
- Soekadar, S. R., Witkowski, M., Birbaumer, N., & Cohen, L. G. (2015). Enhancing Hebbian Learning to Control Brain Oscillatory Activity. *Cerebral Cortex*, 25(9), 2409–2415. <https://doi.org/10.1093/cercor/bhu043>

Bibliography

- Soekadar, S. R., Witkowski, M., Mellinger, J., Ramos, A., Birbaumer, N., & Cohen, L. G. (2011). ERD-Based Online Brain–Machine Interfaces (BMI) in the Context of Neurorehabilitation: Optimizing BMI Learning and Performance. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 19(5), 542–549. <https://doi.org/10.1109/TNSRE.2011.2166809>
- Soekadar, S. R., Witkowski, M., Vitiello, N., & Birbaumer, N. (2015). An EEG/EOG-based hybrid brain-neural computer interaction (BCI) system to control an exoskeleton for the paralyzed hand. *Biomedical Engineering / Biomedizinische Technik*, 60(3), 199–205. <https://doi.org/10.1515/bmt-2014-0126>
- Thompson, M. C. (2019). Critiquing the Concept of BCI Illiteracy. *Science and Engineering Ethics*, 25(4), 1217–1233. <https://doi.org/10.1007/s11948-018-0061-1>
- Tierney, T. M., Alexander, N., Mellor, S., Holmes, N., Seymour, R., O'Neill, G. C., Maguire, E. A., & Barnes, G. R. (2021). Modelling optically pumped magnetometer interference in MEG as a spatially homogeneous magnetic field. *NeuroImage*, 244, 118484. <https://doi.org/10.1016/j.neuroimage.2021.118484>
- Tierney, T. M., Holmes, N., Mellor, S., López, J. D., Roberts, G., Hill, R. M., Boto, E., Leggett, J., Shah, V., Brookes, M. J., Bowtell, R., & Barnes, G. R. (2019). Optically pumped magnetometers: From quantum origins to multi-channel magnetoencephalography. *NeuroImage*, 199, 598–608. <https://doi.org/10.1016/j.neuroimage.2019.05.063>
- Ushiba, J., & Soekadar, S. R. (2016, January 1). Chapter 6 - Brain–machine interfaces for rehabilitation of poststroke hemiplegia. In D. Coyle (Ed.), *Progress in Brain Research* (pp. 163–183). Elsevier. <https://doi.org/10.1016/bs.pbr.2016.04.020>
- VACUUMSCHMELZE GmbH & Co. KG, Hanau. (2019). Magnetically Shielded Rooms. Retrieved March 18, 2022, from [https://www.vacuumschmelze.com/03_Documents/Brochures/Magnetically%20Shielded%20Rooms%20\(MSR\).pdf](https://www.vacuumschmelze.com/03_Documents/Brochures/Magnetically%20Shielded%20Rooms%20(MSR).pdf)
- Vanrullen, R. (2011). Four Common Conceptual Fallacies in Mapping the Time Course of Recognition. *Frontiers in Psychology*, 2. Retrieved March 1, 2022, from <https://www.frontiersin.org/article/10.3389/fpsyg.2011.00365>
- Veale, J. F. (2014). Edinburgh Handedness Inventory – Short Form: A revised version based on confirmatory factor analysis. *L laterality*, 19(2), 164–177. <https://doi.org/10.1080/1357650X.2013.783045>
- Widmann, A., Schröger, E., & Maess, B. (2015). Digital filter design for electrophysiological data – a practical approach. *Journal of Neuroscience Methods*, 250, 34–46. <https://doi.org/10.1016/j.jneumeth.2014.08.002>
- Witkowski, M., Cortese, M., Cempini, M., Mellinger, J., Vitiello, N., & Soekadar, S. R. (2014). Enhancing brain-machine interface (BMI) control of a hand exoskeleton using electrooculography (EOG). *Journal of NeuroEngineering and Rehabilitation*, 11(1), 165. <https://doi.org/10.1186/1743-0003-11-165>
- Wittevrongel, B., Holmes, N., Boto, E., Hill, R. M., Rea, M., Libert, A., Khachatrian, E., Van Hulle, M. M., Bowtell, R., & Brookes, M. J. (2021). Practical real-time MEG-based neural interfacing with optically pumped magnetometers. *BMC Biology*, 19(1), 158. <https://doi.org/10.1186/s12915-021-01073-6>
- World Health Organization & World Bank. (2011). *World report on disability 2011*. World Health Organization. Retrieved March 20, 2022, from <https://apps.who.int/iris/handle/10665/44575>
- Xia, H., Ben-Amar Baranga, A., Hoffman, D., & Romalis, M. V. (2006). Magnetoencephalography with an atomic magnetometer. *Applied Physics Letters*, 89(21), 211104. <https://doi.org/10.1063/1.2392722>

Bibliography

- Xygonakis, I., Athanasiou, A., Pandria, N., Kugiumtzis, D., & Bamidis, P. D. (2018, August 1). *Decoding Motor Imagery through Common Spatial Pattern Filters at the EEG Source Space*. Computational Intelligence and Neuroscience. <https://doi.org/10.1155/2018/7957408>
- Zerfowski, J., Sander, T. H., Tangermann, M., Soekadar, S. R., & Middelmann, T. (2021). Real-time data processing for brain-computer interfacing using optically pumped magnetometers. *International Journal of Bioelectromagnetism*, 23(2), 14/1–6. Retrieved August 11, 2021, from <http://www.ijbem.org/volume23/number2/14.pdf>
- Zetter, R., Iivanainen, J., & Parkkonen, L. (2019). Optical Co-registration of MRI and On-scalp MEG. *Scientific Reports*, 9(1), 5490. <https://doi.org/10.1038/s41598-019-41763-4>

A. Availability of the Source Code and Recorded Data

All code which was used during the acquisition and analysis of data for this study has been made available online, as explained in the following. Some repositories contain code or files that have not been created only by me (also compare chapter 5), which is why the code publication is split up into two parts, a public and a non-public repository.

Public Code The public code repository is available at https://github.com/jzerfowski/msc_thesis (mirror at https://janzerfowski.de/research/msc_thesis/msc_code_public.zip). It contains the following files and folders:

- **readme.md** Readme file detailing the structure of the repository
- **thesis.pdf** The thesis pdf-file
- **requirements.txt** file to reconstruct the virtual environment (venv) that has been used with python 3.8 to run all the source code
- **analysis** contains the code that has been used to perform all analyses and create the figures presented in this thesis
- **create_participant** is a utility program that uses template files to create all necessary data folders and files to for a new participant in the experiment
- **data_organizer** provides an interface to the folders with the recorded data for easy access to raw and epoched data. It has been used extensively in the analysis scripts
- **DataProcessors** contains a selection of the pipeline nodes that were also required in the offline analysis. All nodes are available in the BeamBCI module in the non-public repository.
- **fieldline_lsl** (submodule), containing the used version of the public fieldline_lsl module
- **fif2csv** (submodule), containing the most recent version of the public fif2csv module
- **for_participant** contains the documents and forms participants were provided with before the experiment and which were used to acquire relevant information
- **misc** Miscellaneous files for the organization of the repository
- **Triggerduino** (submodule), containing the most recent version of the public Triggerduino software
- **xdf2mne** (submodule), containing the used version of the public xdf2mne module

To accomplish the data acquisition, analysis and visualization of the results, heavy use of common open source packages for python was made. The major software packages I used shall be listed here for completeness, but a detailed list including the used package version to reconstruct the virtual environment is found in the **requirements.txt**. python 3.8, numpy, scipy, scikit-learn, matplotlib, seaborn, MNE, LabStreamingLayer, pyxdf, statsmodels, pandas

A. Availability of the Source Code and Recorded Data

Non-public Code Access to the code in the private repository can be granted upon request (email to research@janzerfowski.de) after consultation of all involved parties. It contains the following files and folders:

- `positioning_utils` contains code and files to insert relative sensor positions into an `mne.Info` object for sensor localization and plotting. It also contains the models for the sensor patch created by Tilmann Sander-Thömmes, which is why it is not publicly available. The code to perform matrix transformations has been written by Urban Mahrl and is inserted as a git submodule from <https://github.com/UrbanM/megtools>.
- `pythonbci_dev_opm` contains the code for pythonBCI (now called BeamBCI), partly written by Niels Peekhaus and Elisa Braufse, publication in preparation (see appendix A.2)

Data Access to the data (~ 10 GB) can be granted upon request (email to research@janzerfowski.de). It is planned to publish the data in a public domain repository after completion of this thesis.

A.1. Python Modules in the Public Domain

A.1.1. FieldLine LSL

FieldLine Inc. provides an extensive API that includes multiple methods to configure and obtain data from the sensors. However, no generic and versatile way to propagate the data from the API was available. It was therefore decided to build an interface between the FieldLine API and LabStreamingLayer. The public repository can be found at https://github.com/jzerfowski/fieldline_lsl and in the public repository under `fieldline_lsl`.

A.1.2. fif2csv

`fif2csv` is a simple python script to save .`fif`-files as .`csv`-files. It takes several command line arguments to customize the formatting and delimiters. Due to the less efficient encoding, the output files can be up to 8 times larger than the input files. The public repository can be found at <https://github.com/jzerfowski/fif2csv> and in the public repository under `fif2csv`.

A.1.3. xdf2mne

A small tool to transform time series data from a .`xdf`-file into `mne`'s `RawArray` data structure. The public repository can be found at <https://github.com/jzerfowski/xdf2mne> and in the public repository under `xdf2mne`.

A.1.4. Triggerduino – Analog Trigger

For some pilot experiments, multiple data streams (OPM-MEG, EMG, stimulus marker) were to be aggregated in a single device with as little temporal delay as possible. For that, a device with which the temporal shift between incoming data sources could be determined was developed based on the Arduino microcontroller platform.

The public repository can be found at <https://github.com/jzerfowski/Triggerduino> and in the public repository under `Triggerduino`.

In line with a de-facto standard, both FieldLine's sensor chassis as well as the available MEG system in the PTB have an analog input which operates between 0 V and 5 V. FieldLine provides an SMA-input jack, while the MEG device sports a BNC-Socket. These analog inputs can be

A. Availability of the Source Code and Recorded Data

used to receive short trigger pulses that indicate the beginning of a new experimental condition or mark a simultaneous pulse on all devices which can be used to subsequently obtain the temporal delay between the data streams.

Using an Arduino Nano and two 12-bit Digital-to-Analog Converter (DAC) chips based on the MCP4725, we built a device we accordingly named Triggerduino. The number of outputs can be increased relatively simply by adding more DAC chips with unique I2C-addresses. Further, we programmed an API in python which communicates with the Arduino via its serial connection. This makes it easy to send triggers from any python-based experiment framework such as, for instance, PsychoPy.

A.2. BeamBCI

BeamBCI stands for Berlin adaptable modular BCI and is a software written for python 3.6 and higher. The software did justice to its name, since slight adaptions and new modules made it perfectly suited for the OPM-based BCI. BeamBCI implements linear propagation of recorded samples through the first five modules. The modules resemble the processing stages *Acquisition*, *Preprocessing*, *Classification*, *Task* and *Stimulus*. A sixth *Recording* module allows to save the output of all stages as an .xdf file with all data streams synchronized with internal system time precision. It is based on LabRecorder (available at <https://github.com/labstreaminglayer/App-LabRecorder>) The publication and open source licensing of the software is currently still in progress but an analogous overview can be found in Zerfowski et al. (2021).

B. Scientific Contributions

Over the course of this thesis project, multiple contributions to the OPM-MEG community have been published:

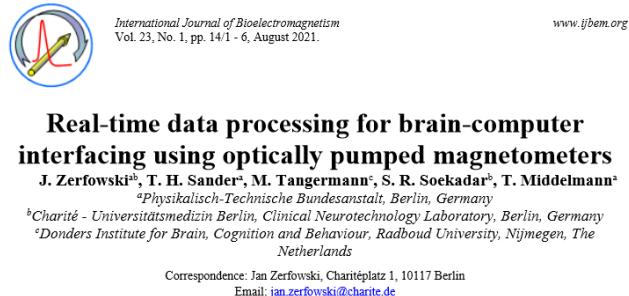


Figure B.1. The conference paper “Real-time data processing for brain-computer interfacing using optically pumped magnetometers” (Zerfowski et al., 2021) can be found here https:////janzerfowski.de/research/paper/icbem2021_paper.pdf



Figure B.2. Left poster: Interactive poster made for the Synapsium symposium at Radboud University. Interactive version: <https:////janzerfowski.de/research/poster/synapsium2021.html>, static pdf version: <https:////janzerfowski.de/research/poster/synapsium2021.pdf>.
Right poster: Poster presented on the BIOMAG 2022 conference in Birmingham https:////janzerfowski.de/research/poster/202208_biomag_birmingham_poster_Zerfowski.pdf

C. Methods

C.1. Setup for the Characterization Measurements

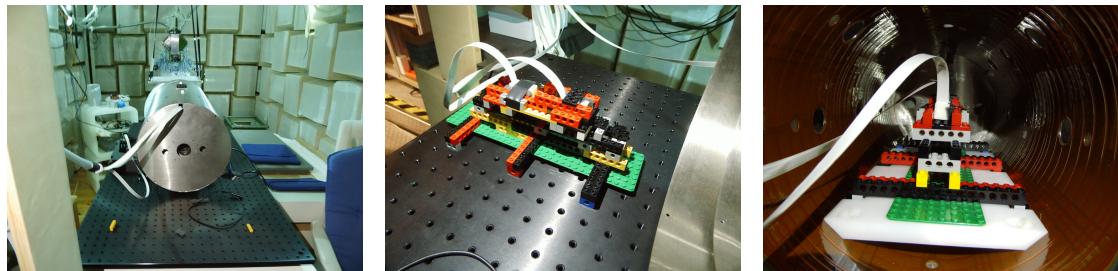
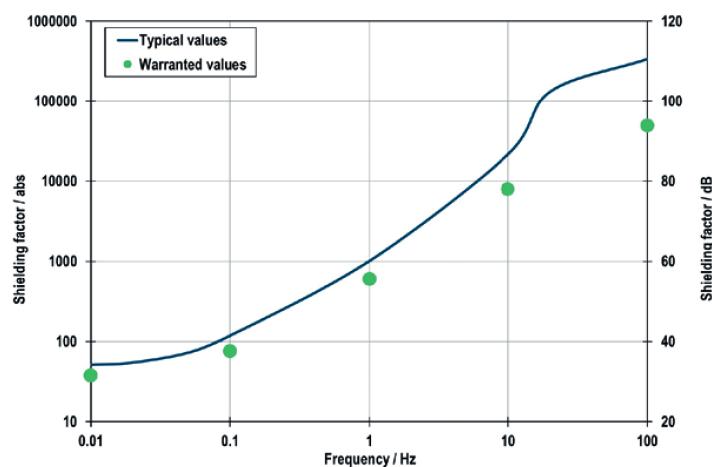


Figure C.1. Setup of the characterization measurements. The magnetically shielded tube was placed inside the Ak3b shielded room. Two sensors were mounted along the Z-axis of the tube. Sinusoidal magnetic fields were applied on the z-axis coils to characterize the sensors (see section 2.2)

TYPICAL SHIELDING FACTORS OF VACOSHIELD ADVANCED 3 X 4 M (13.1 X 9.8 FT)



The typical values have been determined by real data of VAC MSRs.

Figure C.2. Shielding factors of Magnetically Shielded Room Ak3b (now called Vacoshield Advanced) used for the measurements. Plot from VACUUMSCHMELZE GmbH & Co. KG, Hanau (2019), accessed 4.1.2022.

Magnetically Shielded Room

C. Methods

C.2. Data Acquisition Software

C.2.1. LabStreamingLayer

LabStreamingLayer (LSL) is a TCP-based streaming protocol that is widely used in Neuroscientific applications. It allows transferring samples, triggers and control commands between devices and software which can then process the data. LSL is based on liblsl, a library written in C++ that can be compiled on any platform and can therefore operate between operating systems or architectures.

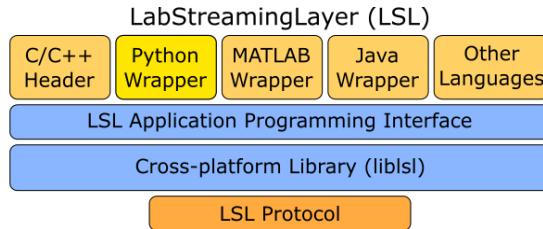


Figure C.3. The protocol structure of LabStreamingLayer

In this project, the LSL protocol was used for all inter-software communication. LSL is available at <https://github.com/sccn/labstreaminglayer>.

Data transmitted through an LSL stream can be recorded using the LabRecorder (<https://github.com/labstreaminglayer/App-LabRecorder>) in the .xdf file format and read using pyxdf (<https://github.com/xdf-modules/pyxdf>).

C.3. Data Analysis

C.3.1. Spectral Filters

Bandpass Filter used in Offline Analysis

The minimum-phase FIR filter used in the offline analysis is shown in Figure C.4. It was generated using MNE's `filter()` method, whose output is reproduced here:

```
Setting up band-pass filter from 9 - 11 Hz
FIR filter parameters
-----
Designing a one-pass, non-linear phase, causal bandpass filter:
- Windowed time-domain design (firwin) method
- Hamming window with 0.0194 passband ripple and 53 dB stopband attenuation
- Lower transition bandwidth: 4.00 Hz
- Upper transition bandwidth: 4.00 Hz
- Filter length: 825 samples (0.825 sec)
```

C. Methods

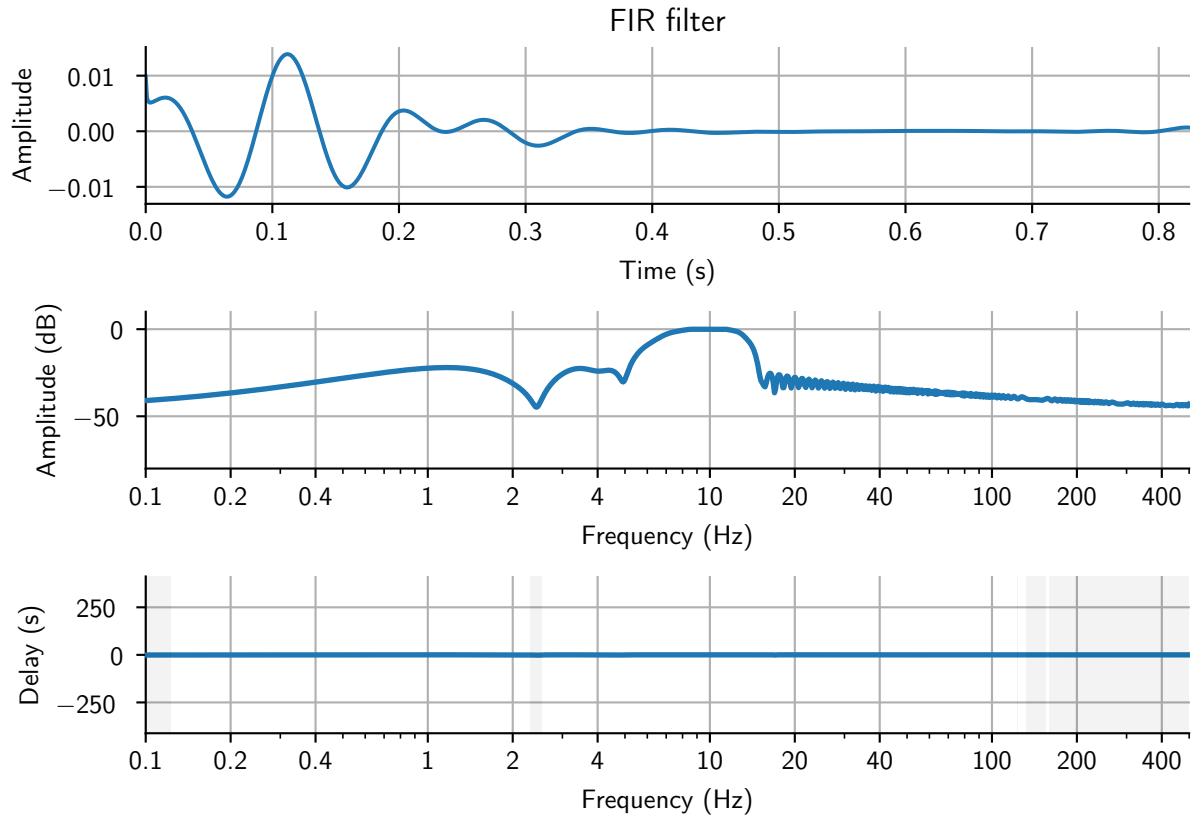


Figure C.4. Minimum-phase FIR filter response as used in the offline analyses for an exemplary subject with an ERD frequency of 10 Hz. Passband 10 ± 1 Hz, left and right transition bandwidth: 4 Hz, using a Hamming window function.

Bandpass Filter used in Online Analysis

The linear-phase FIR filter used for the offline analysis is shown in Figure C.5. Implementation details can be found in the methods section in section 2.7.

C. Methods

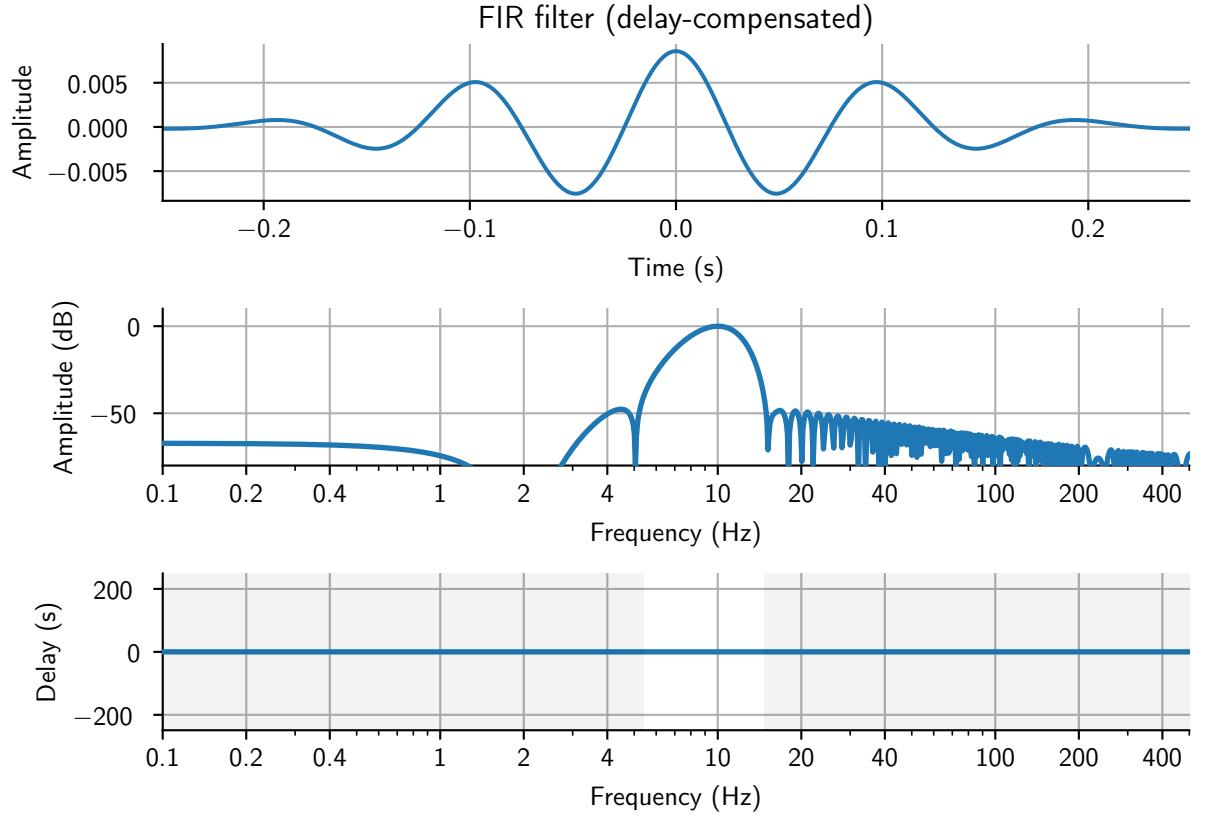


Figure C.5. Linear-phase FIR filter response as used in the online processing for an exemplary subject with an ERD frequency of 10 Hz. Passband 10 ± 1.5 Hz, length 499 samples, using a Hamming window function.

Single-Pole Filter in Online Analysis

To smoothen the classification values, a single-pole IIR filter was used in the online pipeline. It uses only the single previous sample to implement an exponential decay and depends on a single parameter, the time constant τ . Using τ and the sampling frequency $f_s = 10$ Hz a decay factor can be computed $d = 1 - e^{-1/(\tau * f_s)} = 0.81$. The filter output $y[n]$ is then computed as $y[n] = y[n - 1] + d(x[n] - y[n - 1])$, where $x[n]$ is the current sample input.

The filter response is shown in Figure C.6.

C. Methods

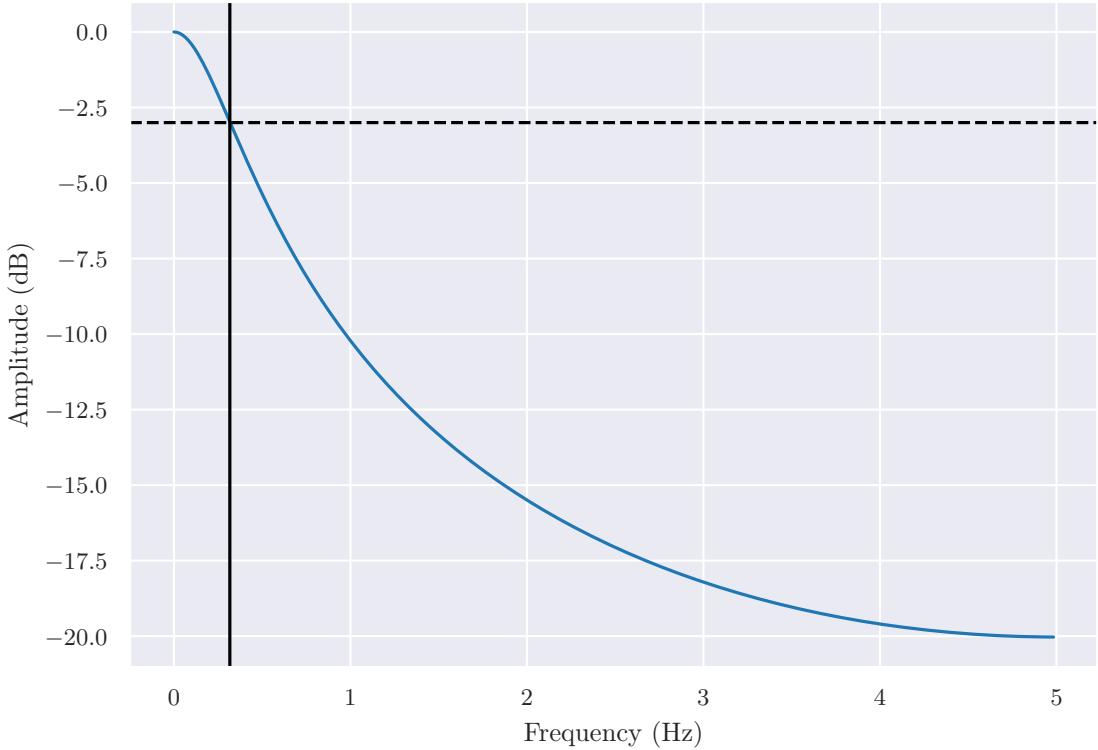


Figure C.6. Frequency response of the single pole filter with a time constant of 0.5 at a sampling frequency of 10 Hz. The 3 dB point is indicated by the black line at 0.3 Hz.

C.3.2. Common Spatial Patterns

The common spatial patterns (CSP) algorithm is a powerful method for BCI systems as it can be used to compute spatial filters which achieve the highest discriminability of two classes in terms of variance. Many variations of CSP have been proposed Blankertz et al., 2007; Novi et al., 2007; Ang et al., 2008; Lotte and Guan, 2011. In this study the most basic formulation of CSP has been used, which will be explained in the following:

To achieve maximum signal discriminability between two classes A, B , compute a set of weight vectors (components) $w_i \in \mathbb{R}^C$ such that w_i maximize the variance ratio between the two classes X_A and X_B

$$\arg \max_{w_i} \frac{\text{var}(w_i X_A)}{\text{var}(w_i X_B)},$$

where $X \in \mathbb{R}^{C \times T}$ is the sensor-space signal with T samples in C channels. $X_A \in \mathbb{R}^{C \times |T_A|}$ then contains the samples recorded in class A . $T_A \in \mathbb{R}$ is the set of indices corresponding to samples in condition A ; analogous for B . It is assumed that the sensor-space signals $X_{A,B}$ are bandpass filtered such that their variances reflect the band power and shifted to have a mean of zero for each channel.

C. Methods

The covariance matrix Σ^c for class $c \in \{A, B\}$ can be computed as

$$\Sigma^c = \frac{1}{|T_c|} \sum_{t \in T_c} X_c(t) X_c(t)^T$$

where $X_c(t)$ indicates the sample at time t in class c .

The vectors w_i are then contained as columns vectors in a matrix $W = [w_1, w_2, \dots, w_C]$ which can be computed using the generalized Eigendecomposition of the covariance matrices of the two classes:

$$W^T \Sigma^A W = \Lambda^A \quad (\text{C.1})$$

$$W^T \Sigma^B W = \Lambda^B, \quad (\text{C.2})$$

with the convention to scale W such that $\Lambda^A + \Lambda^B = I$. Solving equation (C.1) then corresponds to solving the generalized eigenvalue problem

$$\Sigma^A w_i = \lambda_i \Sigma^B w_i. \quad (\text{C.3})$$

where $\lambda_i = \lambda_i^A / \lambda_i^B$, $\lambda_i^c = w_i^T \Sigma^c w_i$ and Λ^c is a diagonal matrix with λ_i^c as the values on its diagonal. Then $\lambda_i^c \geq 0$ is the variance in the corresponding surrogate channel and $\lambda_i^A + \lambda_i^B = 1$.

`d, W = scipy.linalg.eigh(Sigma^A, b=Sigma^A + Sigma^B)` computes the solution for the generalized eigendecomposition problem and returns the eigenvalues (d_i) and a matrix with all eigenvectors (corresponding to the components w_i).

The matrix of *spatial filters* W acts as a backward model, providing information about how measured values in channels correspond to source activity. Such filters lack interpretability as they are used to filter out non-discriminative information from both conditions. For better interpretation usually the *spatial patterns* (A) are provided, which correspond to the influence of each source activity on the measured data. To compute the spatial patterns from spatial filters and the data, a forward model must be computed. This forward model should minimize the error between the original signal and backward-forward projected signal. Luckily, the Roger-Moore pseudoinverse defined as $A^+ = A^T (AA^T)^{-1}$ (`scipy.linalg.pinv()`) has exactly that property

$$\arg \min_A \|X - AS\| = XS^+ \quad (\text{C.4})$$

where $S = W^T X$ is the signal in the surrogate sensor space and A is the forward model.

To derive \hat{A} let, then

$$\begin{aligned} \hat{A} &= \arg \min_A \|X - AS\| \\ &= XS^+ \\ &= XS^T (SS^T)^{-1} \\ &= XX^T W (SS^T)^{-1} \\ &= \Sigma_X W \Sigma_S^{-1} \end{aligned}$$

The spatial patterns a_i corresponding to the spatial filters w_i can be found in the corresponding rows of \hat{A} .

C. Methods

C.3.3. Derivation of chance level classifier success rates

The results from section 3.3 were compared to a chance-level classifier with equal probability for assigning *close* or *relax* as label output from the classification pipeline. Within 5 s, 50 samples were generated. From this, it was derived that a *close* trial is successful when > 25 samples are classified as *close* and a *relax* trial is successful when ≤ 12 samples are classified as *close*.

A cumulative binomial distribution can be used to compute the expected success rates under such a chance-level classifier: Assuming that X follows a binomial distribution with $p = 0.5$ and $n = 50$, the probability for k *close* classifications ($\Pr(X = k)$) is

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}. \quad (\text{C.5})$$

The probability of obtaining less than or equal to k *close* classifications is then, for $p = 0.5$,

$$\Pr(X \leq k) = \sum_{i=0}^k \binom{n}{i} p^i (1-p)^{n-i} = \sum_{i=0}^k \binom{n}{i} 0.5^n. \quad (\text{C.6})$$

The expected success rate for *close* trials is then

$$\Pr(X > 25) = 1 - \Pr(X \leq 25) = 1 - 0.556138 = 0.443862$$

and the expected success rate for *relax* trials computes as

$$\Pr(X \leq 12) = 0.000152932$$

D. Results

D.1. Results of the Characterization Measurements

D.1.1. Noise Floor

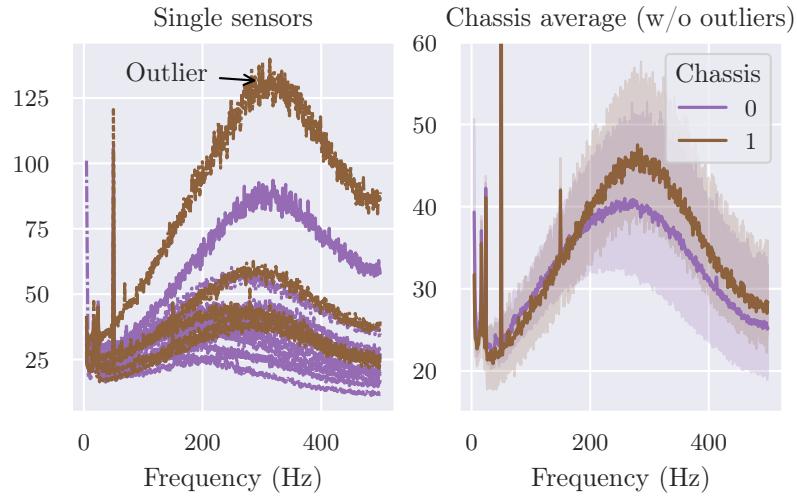


Figure D.1. Comparison of the system noise floors, distinguished by chassis. **Left panel:** Per-sensor system noise floor. A single outlier with abnormal noise profile was identified by computing the z-score values of the noise at 300 Hz and thresholding above 3.0. **Right panel:** Averaged system noise floor with outliers excluded from the computation. The shaded 95 % confidence intervals are overlapping.

D.1.2. Dynamic Range

Amplitude Modulation

D. Results

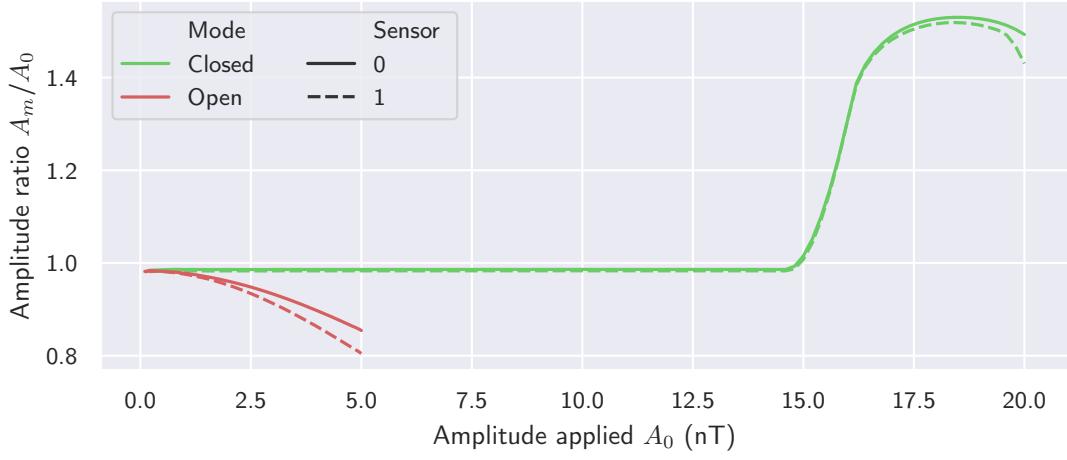


Figure D.2. Full range of Figure 3.2. For applied amplitudes larger than 14.5 pT, the estimated amplitude shoots up to 150% at 17.5 nT. This is an artifact due to a software error that will be fixed in future versions.

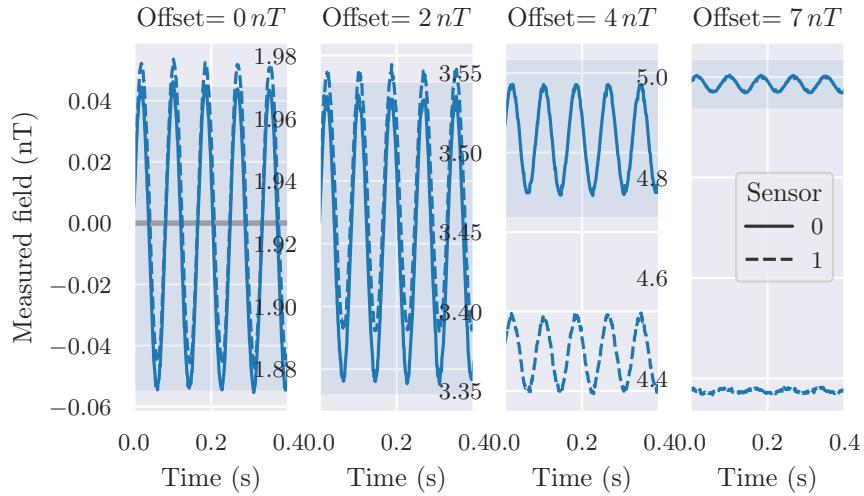


Figure D.3. Recorded signal during the dynamic range amplitude modulation in the sensors' open loop mode. The separate panels show 400 ms of data with amplitudes from $A_0 = 1$ nT up to $A_0 = 4.5$ nT. The gray bars indicate the expected peak-to-peak amplitude according to A_0 .

D. Results

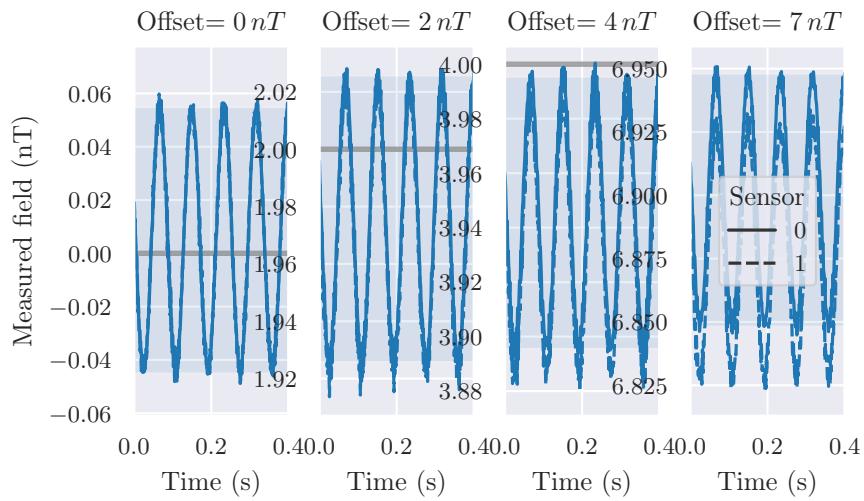


Figure D.4. Recorded signal during the dynamic range amplitude modulation in the sensors' closed loop mode. The separate panels show 400 ms of data with amplitudes from $A_0 = 8 \text{ nT}$ up to $A_0 = 16 \text{ nT}$. The gray bars indicate the expected peak-to-peak amplitude according to A_0 .

D. Results

Offset Modulation

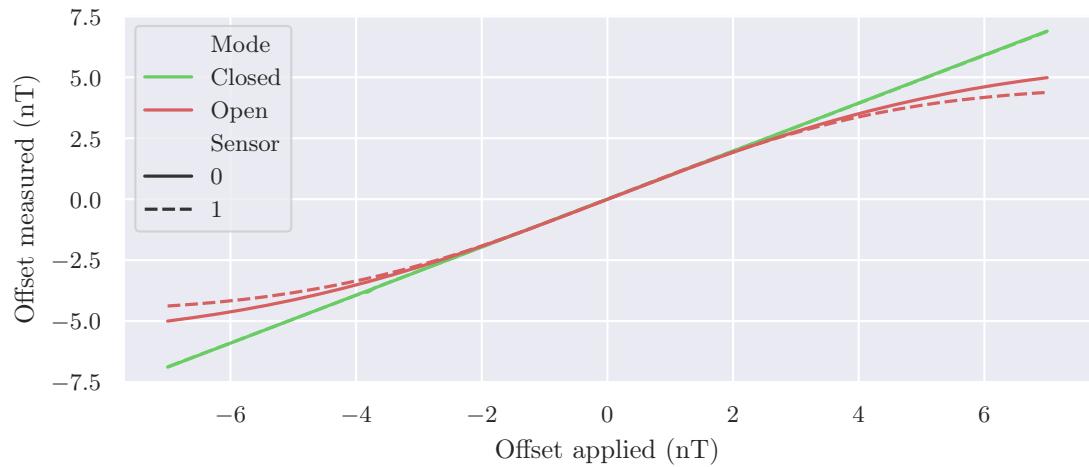


Figure D.5. Measured offset vs. applied offset for the dynamic range measurements described in subsubsection 3.1.2.

D.2. Offline Analysis

D.2.1. Individual ERD spectra

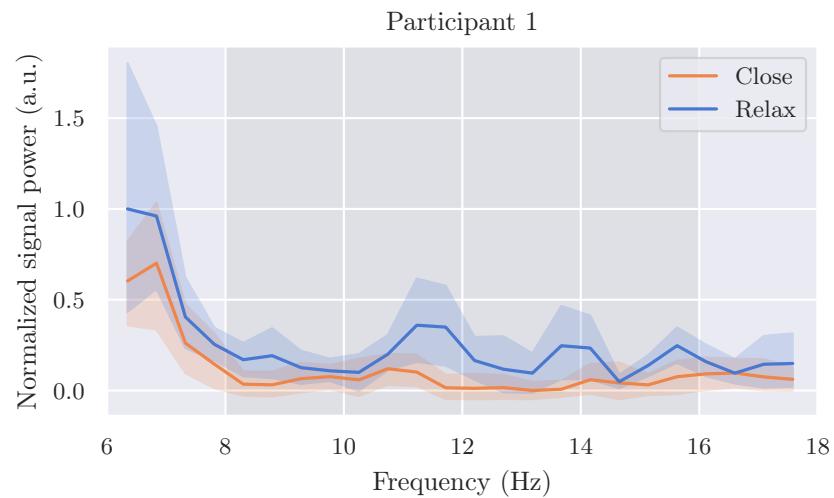


Figure D.6. Participant 001: ERD frequency plot. No ERD frequency could be determined for this participant.

D. Results

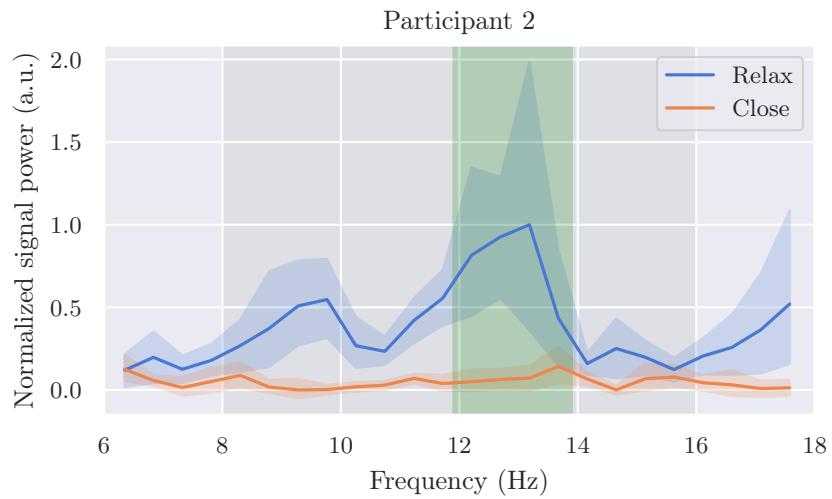


Figure D.7. Participant 002: ERD frequency plot. The ERD frequency was determined to be 12.9 Hz (± 1 Hz)

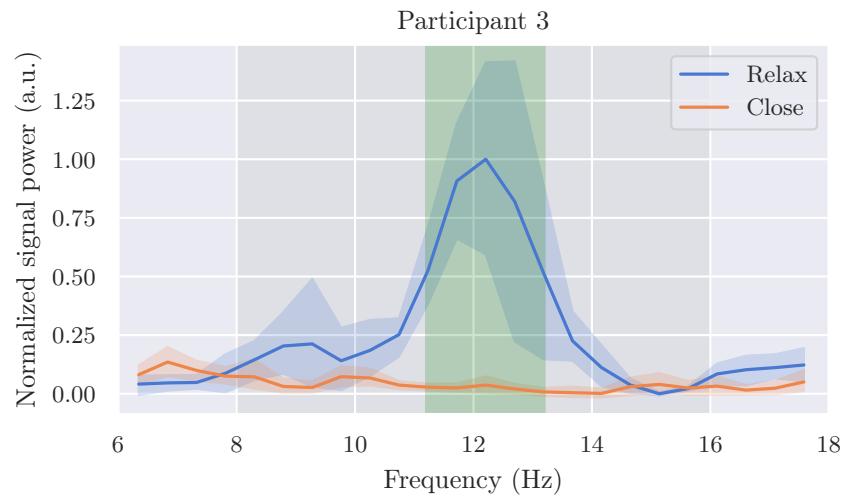


Figure D.8. Participant 003: ERD frequency plot. The ERD frequency was determined to be 12.2 Hz (± 1 Hz)

D. Results

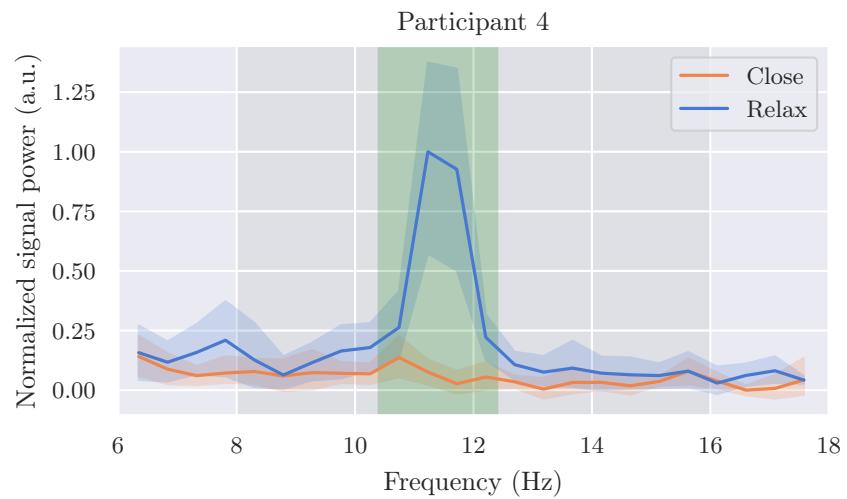


Figure D.9. Participant 004: ERD frequency plot. The ERD frequency was determined to be 11.4 Hz (± 1 Hz)

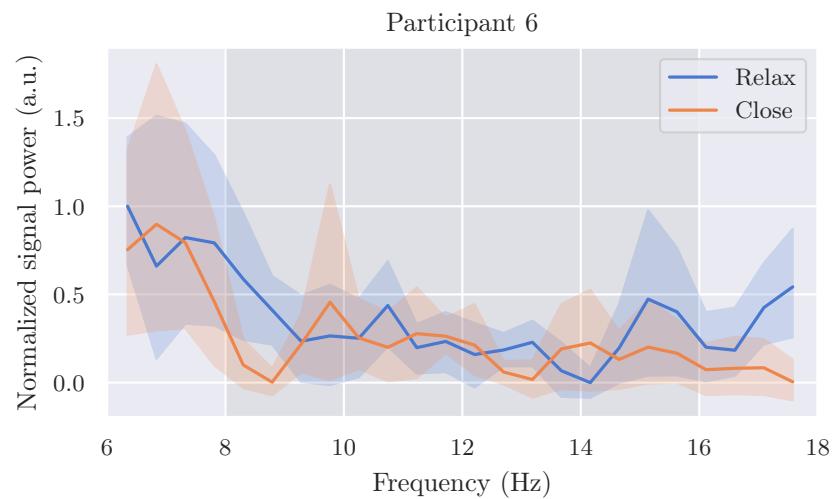


Figure D.10. Participant 006: ERD frequency plot. No ERD frequency could be determined for this participant.

D. Results

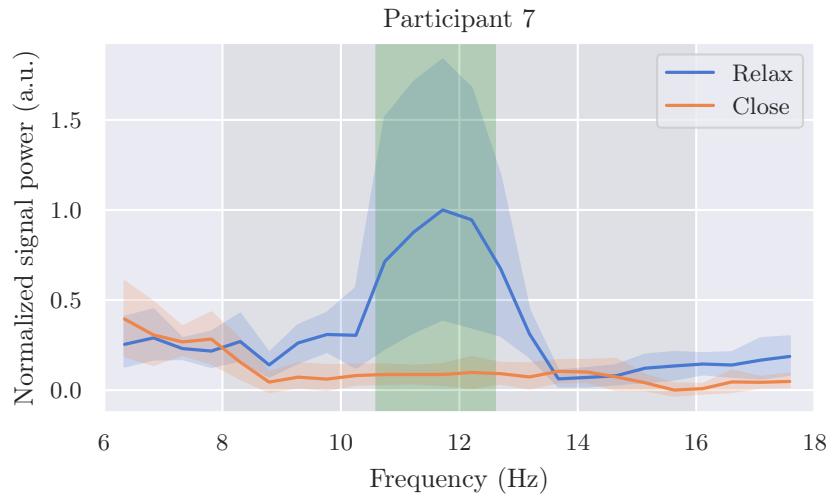


Figure D.11. Participant 007: ERD frequency plot. The ERD frequency was determined to be 11.6 Hz (± 1 Hz)

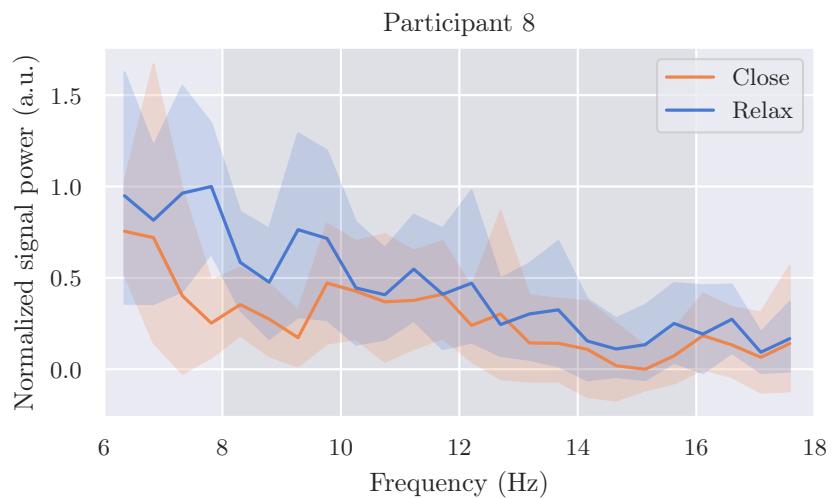


Figure D.12. Participant 008: ERD frequency plot. No ERD frequency could be determined for this participant.

D. Results

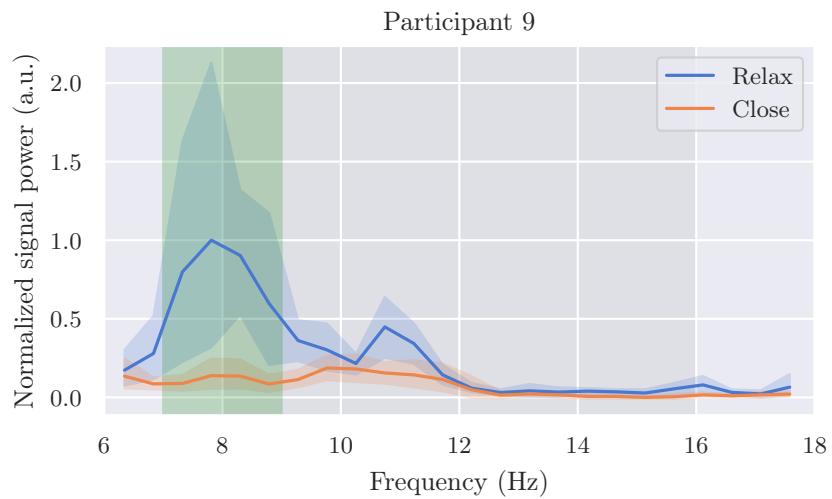


Figure D.13. Participant 009: ERD frequency plot. The ERD frequency was determined to be 8.0 Hz (± 1 Hz)

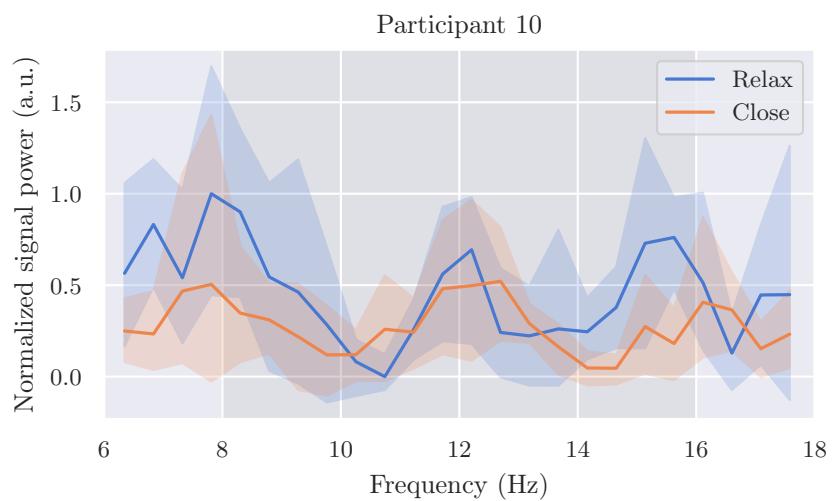


Figure D.14. Participant 010: ERD frequency plot. No ERD frequency could be determined for this participant.

D. Results

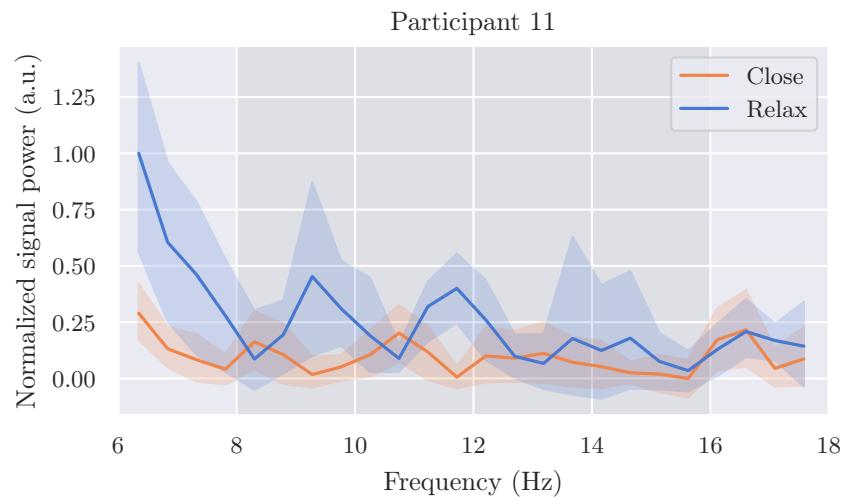


Figure D.15. Participant 011: ERD frequency plot. No ERD frequency could be determined for this participant.

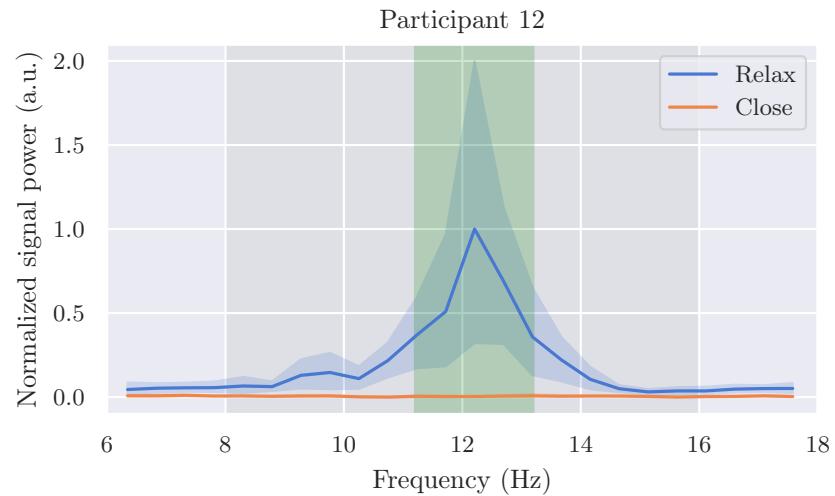


Figure D.16. Participant 012: ERD frequency plot. The ERD frequency was determined to be 12.2 Hz (± 1 Hz)

D. Results

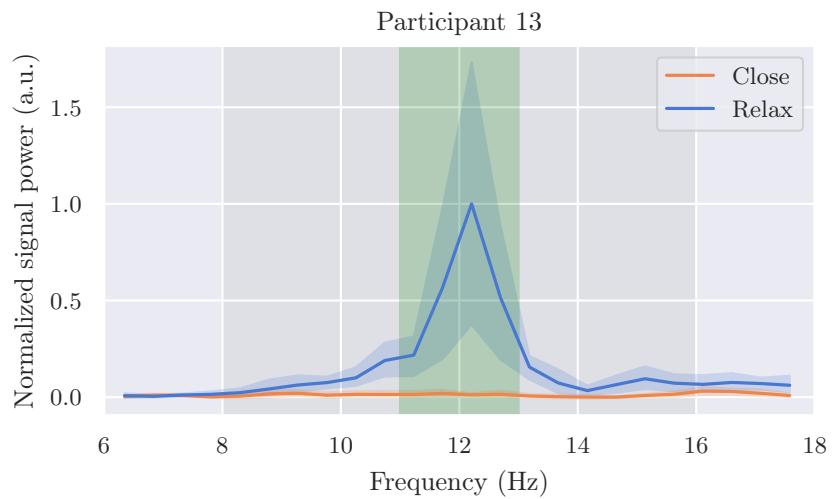


Figure D.17. Participant 013: ERD frequency plot. The ERD frequency was determined to be 12.0 Hz (± 1 Hz)

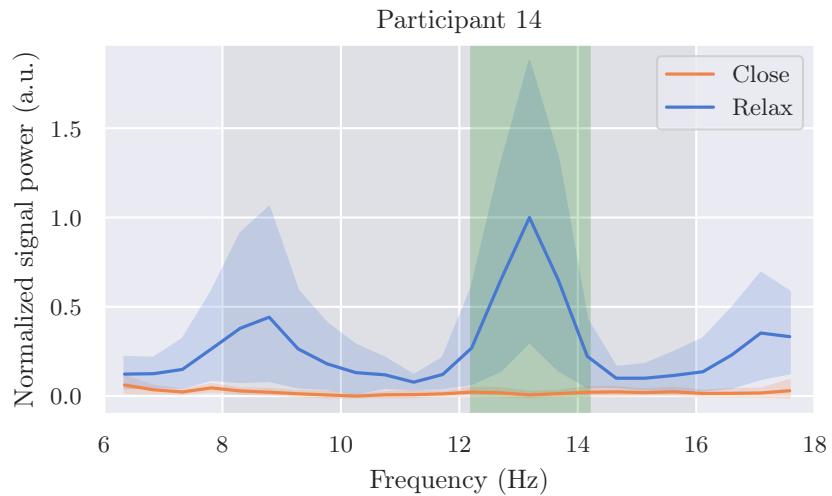


Figure D.18. Participant 014: ERD frequency plot. The ERD frequency was determined to be 13.2 Hz (± 1 Hz)

D. Results

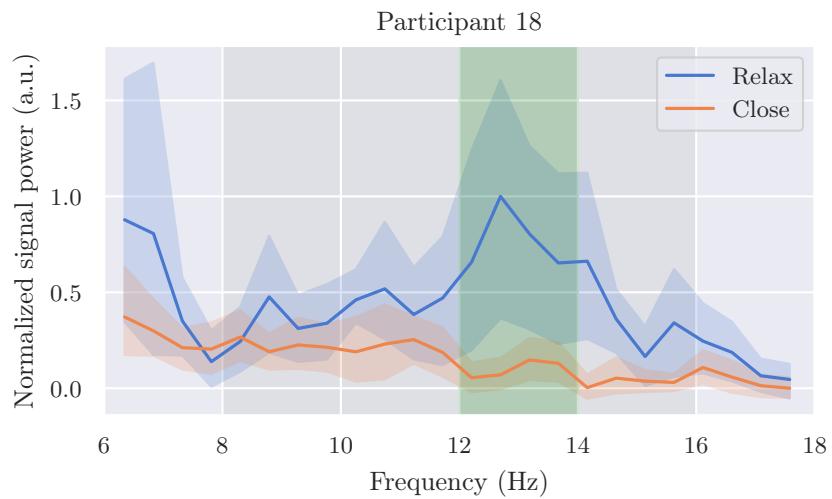


Figure D.19. Participant 018: ERD frequency plot. The ERD frequency was determined to be 13.0 Hz (± 1 Hz)

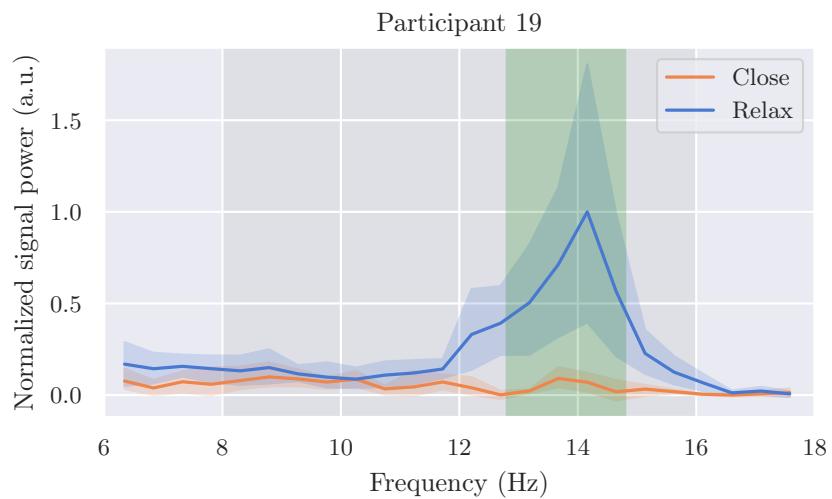


Figure D.20. Participant 019: ERD frequency plot. The ERD frequency was determined to be 13.8 Hz (± 1 Hz)

D. Results

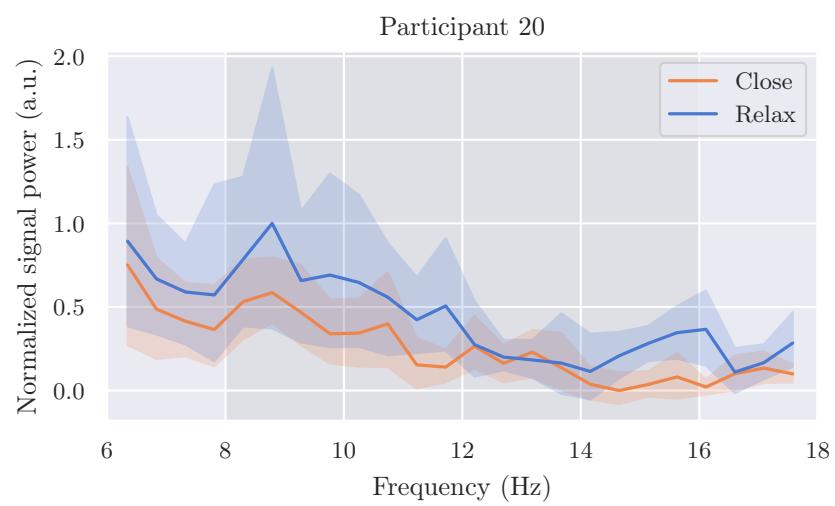


Figure D.21. Participant 020: ERD frequency plot. No ERD frequency could be determined for this participant.

D. Results

D.2.2. Individual ERD timecourses

The timecourses of the ERD against baseline (-2 s to 0 s before stimulus onset) are shown here for all participants for which an ERD frequency could be determined in the subsection 3.2.1. The shaded margins in all plots are the 95 % confidence intervals across trials. The black bars in the bottom mark the clusters with significant difference between *close* and *relax* trials according to a cluster-based permutation test ($n = 1024$ permutations, *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$). The gray area indicates the trial duration from 0 s to 5 s.

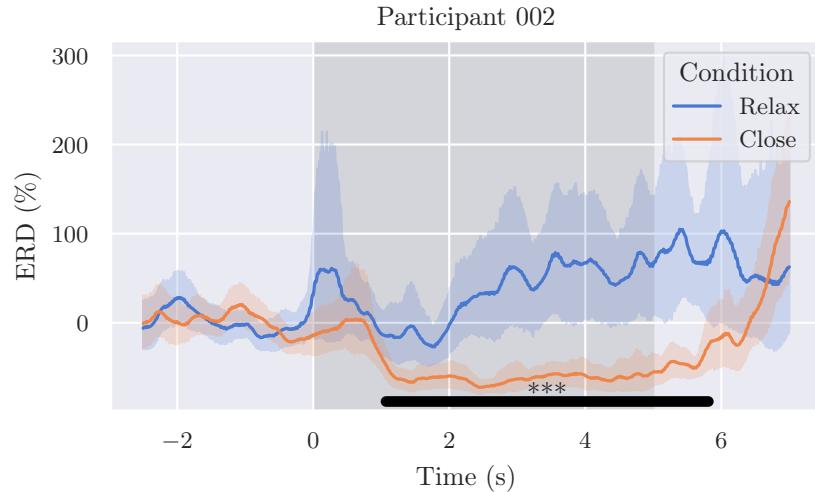


Figure D.22. Timecourse of the ERD for participant 002

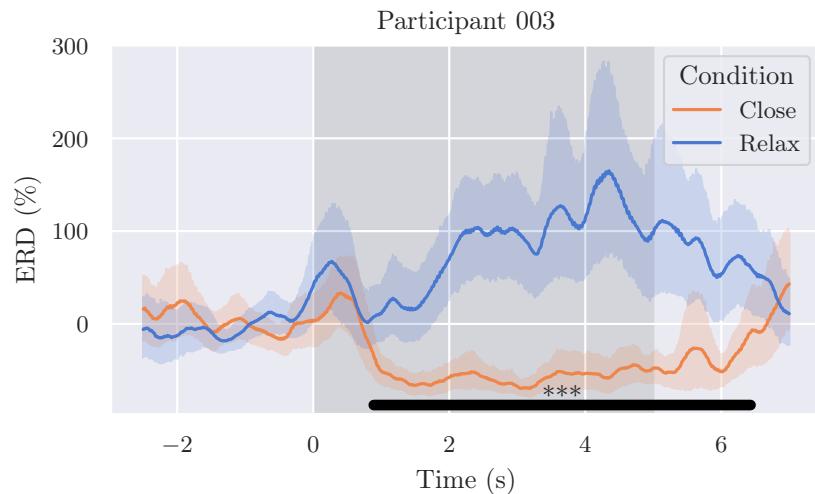


Figure D.23. Timecourse of the ERD for participant 003

D. Results

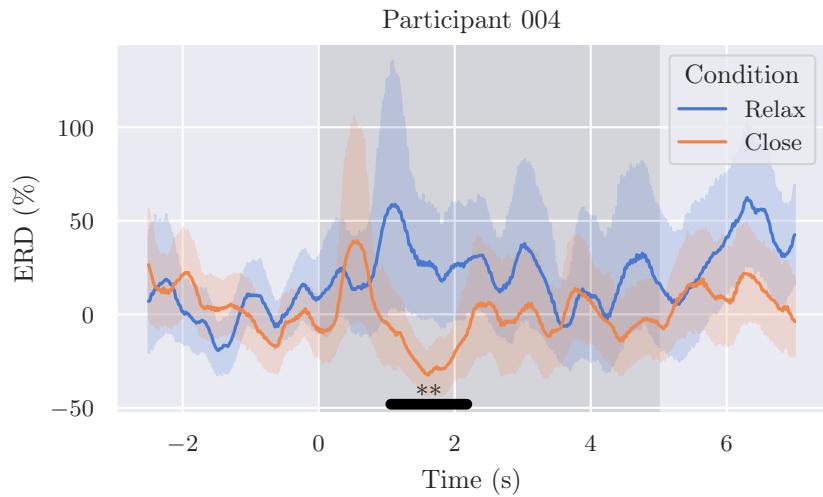


Figure D.24. Timecourse of the ERD for participant 004

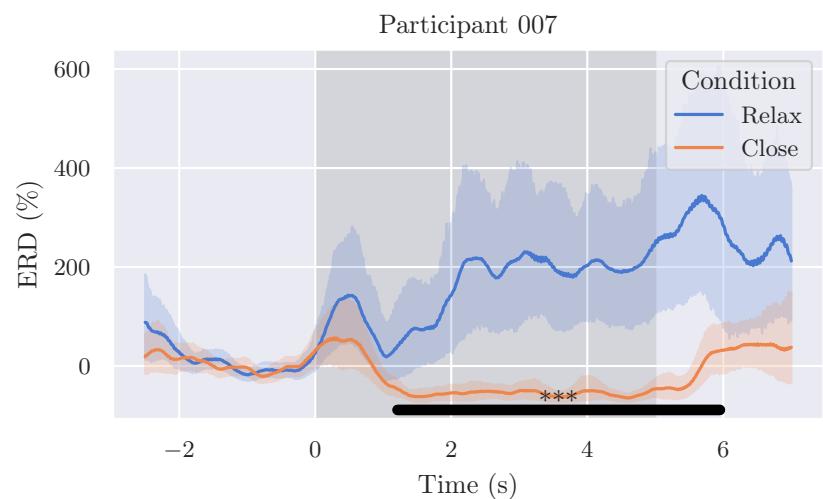


Figure D.25. Timecourse of the ERD for participant 007

D. Results

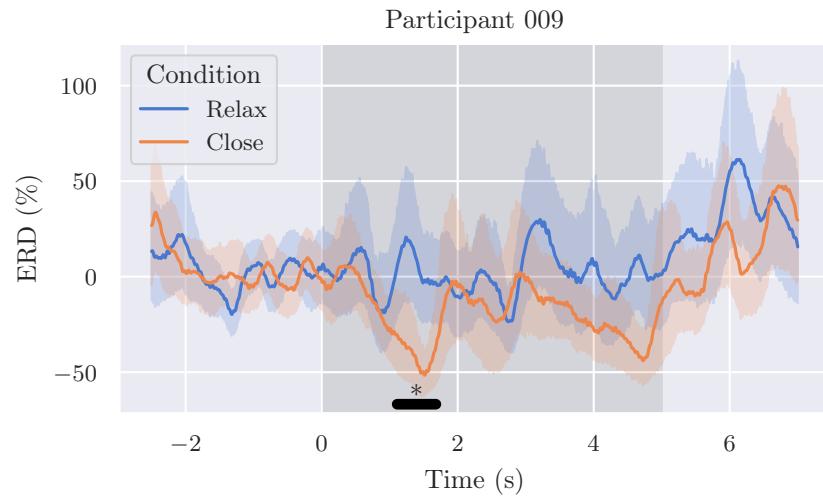


Figure D.26. Timecourse of the ERD for participant 009

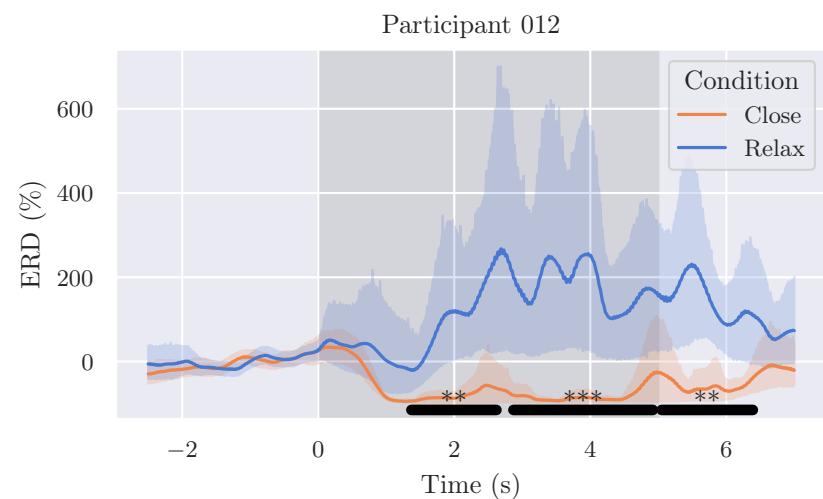


Figure D.27. Timecourse of the ERD for participant 012

D. Results

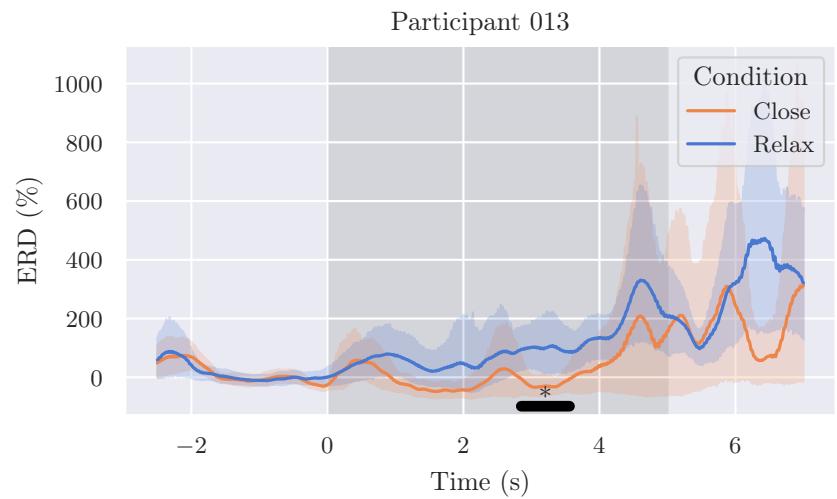


Figure D.28. Timecourse of the ERD for participant 013

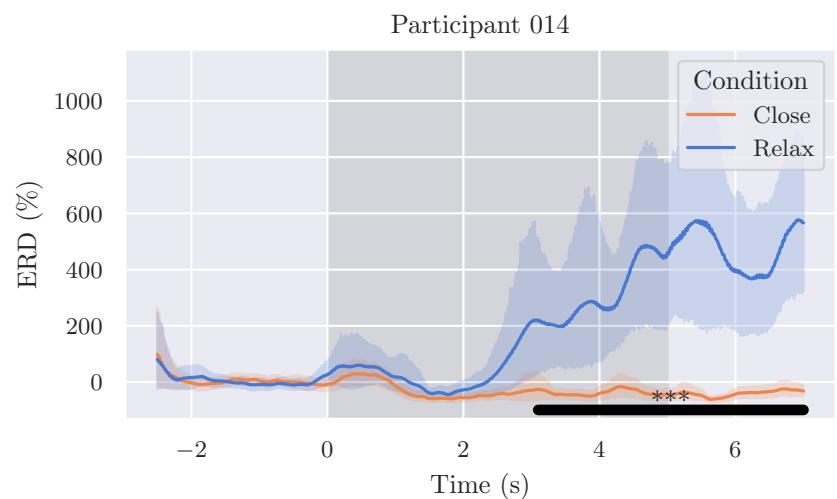


Figure D.29. Timecourse of the ERD for participant 014

D. Results

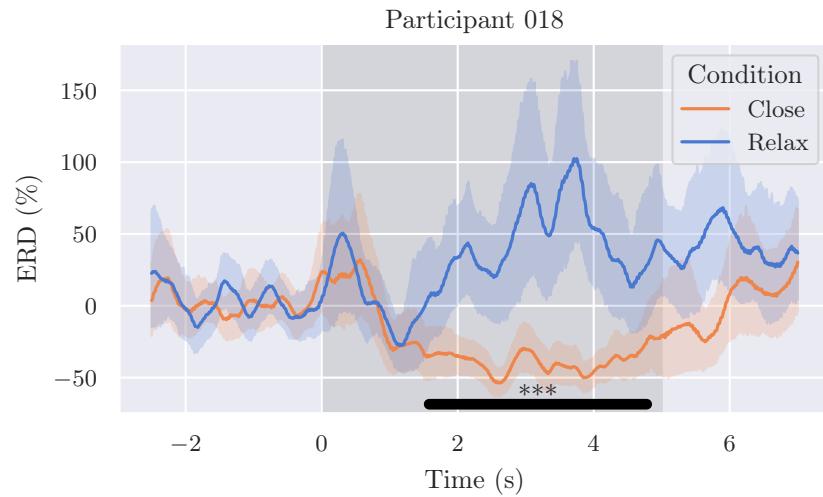


Figure D.30. Timecourse of the ERD for participant 018

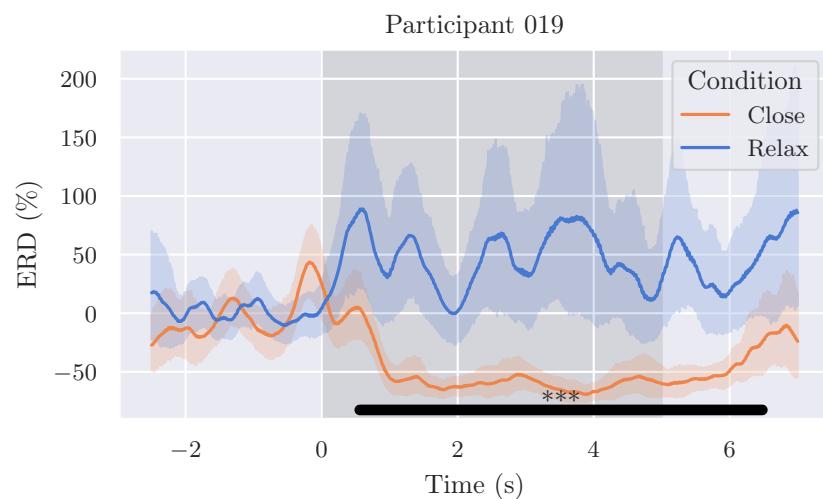


Figure D.31. Timecourse of the ERD for participant 019

D. Results

D.2.3. Focality

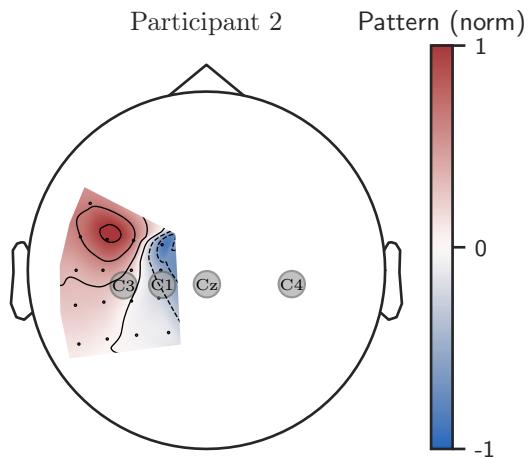


Figure D.32. Participant 002: Visualization of CSP pattern corresponding to the lowest eigenvalue.

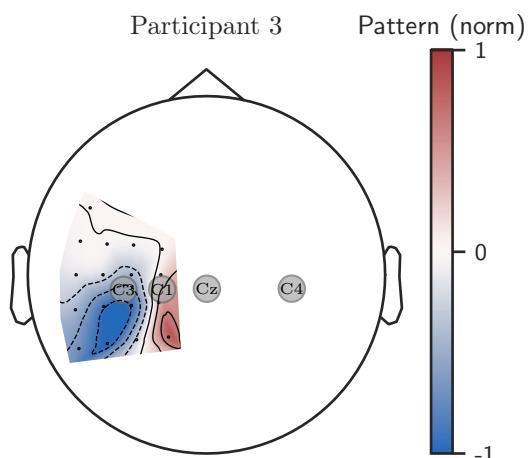


Figure D.33. Participant 003: Visualization of CSP pattern corresponding to the lowest eigenvalue.

D. Results

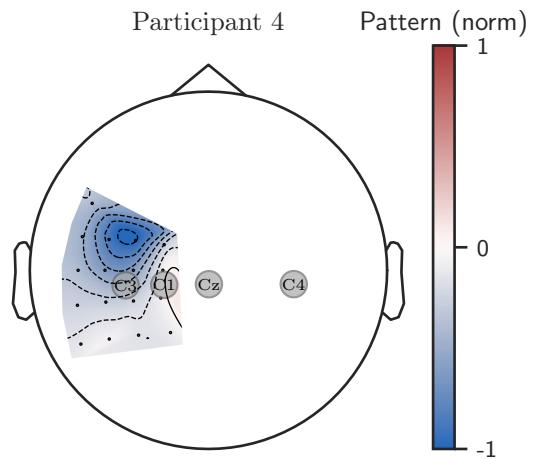


Figure D.34. Participant 004: Visualization of CSP pattern corresponding to the lowest eigenvalue.

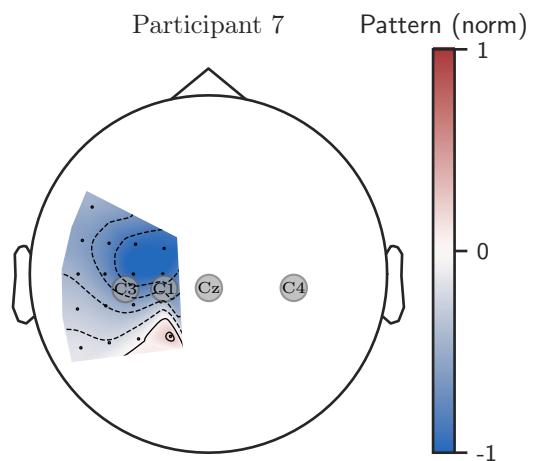


Figure D.35. Participant 007: Visualization of CSP pattern corresponding to the lowest eigenvalue.

D. Results

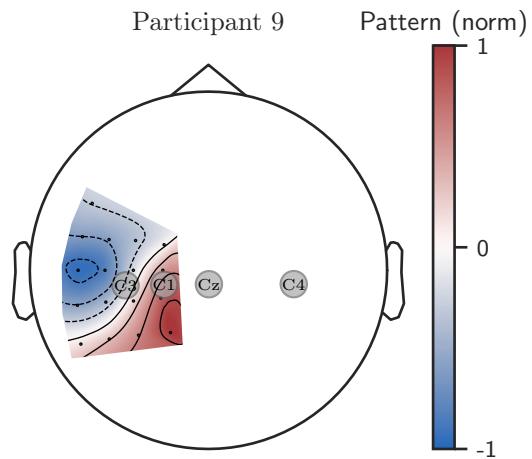


Figure D.36. Participant 009: Visualization of CSP pattern corresponding to the lowest eigenvalue.

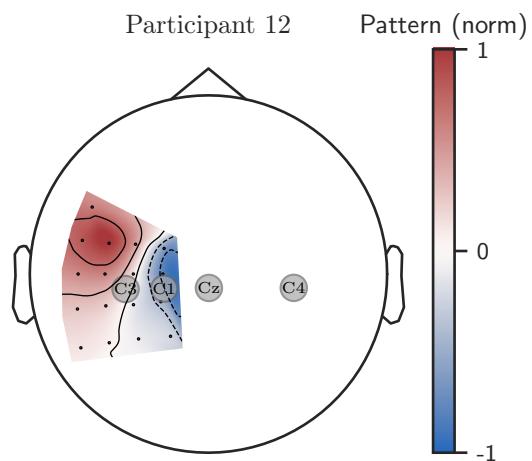


Figure D.37. Participant 012: Visualization of CSP pattern corresponding to the lowest eigenvalue.

D. Results

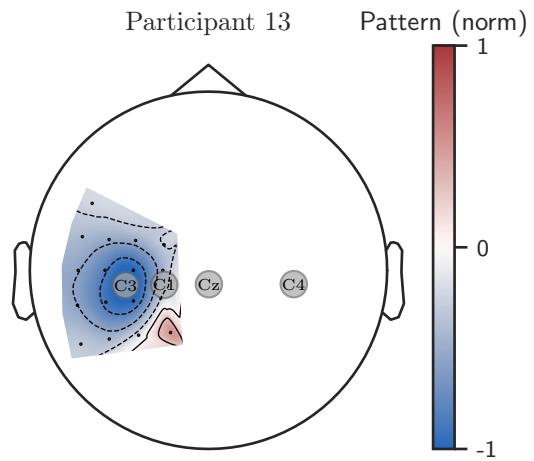


Figure D.38. Participant 013: Visualization of CSP pattern corresponding to the lowest eigenvalue.

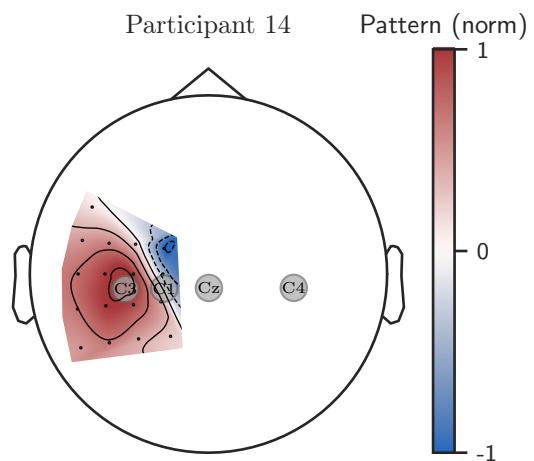


Figure D.39. Participant 014: Visualization of CSP pattern corresponding to the lowest eigenvalue.

D. Results

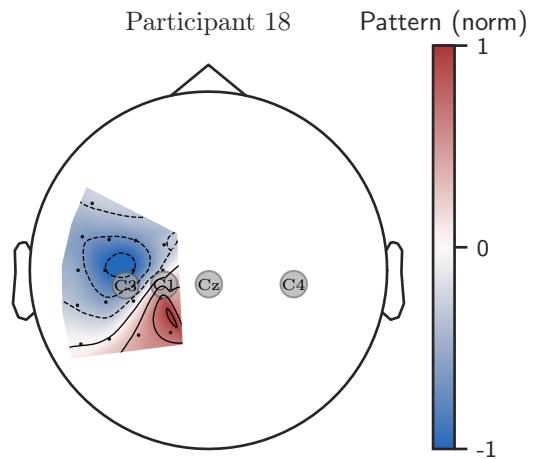


Figure D.40. Participant 018: Visualization of CSP pattern corresponding to the lowest eigenvalue.

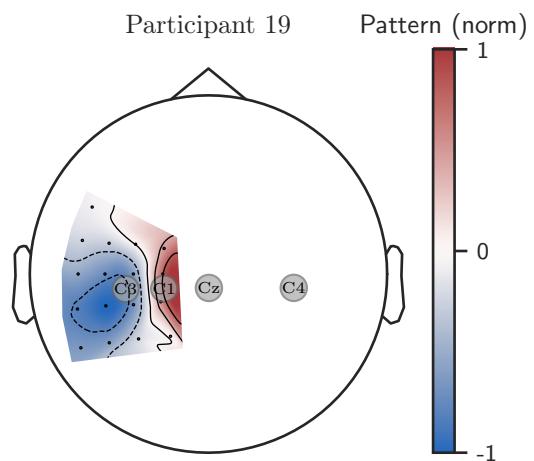


Figure D.41. Participant 019: Visualization of CSP pattern corresponding to the lowest eigenvalue.

D. Results

D.3. Online Analysis

D.3.1. Pipeline evaluation

analyze_csp_eigenvalues.py:

```
For participant 001, both eigenvalues are >0.5 in blocks [0 1]
(participant does NOT have an ERD)
For participant 002, the average eigenvalue is 0.279
For participant 003, the average eigenvalue is 0.151
For participant 004, the average eigenvalue is 0.471
For participant 006, the average eigenvalue is 0.411 (participant does NOT have an ERD)
For participant 007, the average eigenvalue is 0.23
For participant 008, both eigenvalues are >0.5 in blocks [1 2]
(participant does NOT have an ERD)
For participant 009, the average eigenvalue is 0.451
For participant 010, both eigenvalues are >0.5 in blocks [0 1 2 3 4]
(participant does NOT have an ERD)
For participant 011, the average eigenvalue is 0.437 (participant does NOT have an ERD)
For participant 012, the average eigenvalue is 0.0758
For participant 013, the average eigenvalue is 0.16
For participant 014, the average eigenvalue is 0.223
For participant 018, the average eigenvalue is 0.332
For participant 019, the average eigenvalue is 0.266
For participant 020, the average eigenvalue is 0.473 (participant does NOT have an ERD)
```

D.3.2. Block-wise success rate statistics

D. Results

```

Model for Close condition      Mixed Linear Model Regression Results
=====
Model:      MixedLM Dependent Variable: success_exo_percent
No. Observations: 50    Method:      REML
No. Groups:   10    Scale:       110.4706
Min. group size: 5    Log-Likelihood: -198.8492
Max. group size: 5    Converged:    Yes
Mean group size: 5.0

Coef. Std.Err. z P>|z| [0.025 0.975]
-----
Intercept     80.167 7.416 10.811 0.000 65.632 94.701
Block         1.167 1.051 1.110 0.267 -0.893 3.227
participant_id Var 428.391 22.415
=====

Model for Relax condition      Mixed Linear Model Regression Results
=====
Model:      MixedLM Dependent Variable: success_exo_percent
No. Observations: 50    Method:      REML
No. Groups:   10    Scale:       594.8005
Min. group size: 5    Log-Likelihood: -230.5189
Max. group size: 5    Converged:    Yes
Mean group size: 5.0

Coef. Std.Err. z P>|z| [0.025 0.975]
-----
Intercept     59.500 9.400 6.330 0.000 41.076 77.924
Block        -2.833 2.439 -1.162 0.245 -7.613 1.947
participant_id Var 229.312 7.468
=====

Model for combined success rate      Mixed Linear Model Regression Results
=====
Model:      MixedLM Dependent Variable: success_exo_percent
No. Observations: 50    Method:      REML
No. Groups:   10    Scale:       146.3676
Min. group size: 5    Log-Likelihood: -202.7778
Max. group size: 5    Converged:    Yes
Mean group size: 5.0

Coef. Std.Err. z P>|z| [0.025 0.975]
-----
Intercept     69.833 6.711 10.406 0.000 56.680 82.987
Block        -0.833 1.210 -0.689 0.491 -3.205 1.538
participant_id Var 289.368 13.774
=====
```

Figure D.42. Output of the model fitting described in section 2.7.

E. Miscellaneous

E.1. Motor Imagery Training

A sheet with the following text has been handed to participants before the experiment. They could both read the text and receive the instructions from the experimenter. The instructions have been developed by BCI group in the Clinical Neurotechnology lab and are subject to continuous improvement.

E.1.1. Preparation

Ask the participant to sit comfortably, close their eyes and lay the hand supine over the desk or on the arm rest. Ask him to focus the attention on the forearm and on the hand: where are they located and what is their position in space. Ask him to close the hand in a strong fist and hold it close as strong as possible for a few seconds. Make the participant focusing on the hand and the forearm again and notice that the arm should feel now “warmer”. Let the participant focus on the hand and the sensations it evokes for a couple of seconds more.

Ask the participant to start slowly opening and closing his hand and repeat this movement. It is important to ask him to maximize this movement, so to fully open and fully close the hand. While the participant keeps moving the hand, direct his attention to the movement itself and to the physical sensations it elicits, for example on the back of the hand while stretching the fingers or on the palm when the finger nails reach it. After having repeated the movement 4/5 times, ask the participants (while keep moving) to gradually reduce the actual movement, but at the same keep imagining the hand to fully open and fully close, by remembering/recreating the same physical sensation he experienced before. After a couple of seconds, the participant should stop moving the hand, but keep imagining his hand to open and close.

If the participant can imagine/mentally visualize vividly his hand movement, the training has been successfully completed!

E.1.2. Motor imagery task tips

Suggest to participants to look at their hand while imagining moving it, and to repeat the closing/grasp motion for 4/5 times during 5s lasting cues. Every participant needs to find his own strategy, but for some very experienced musicians it may be easier to imagine moving the hand while playing rather than grasping an object (true for any highly experienced movement). For other participants, it helps to imagine squeezing a ball or a bottle during the task. Even when successfully performing motor imagery, participants need to actively relax in between tasks. When providing real-time feedback (especially during the first runs), remind the participant to keep focusing on the motor imagery task despite too good/ too bad feedback. The calibration may not be optimal.

E.1.3. Relax tips

There is no standard training for relaxation. Some suggestions are: focusing on your breathing to calm down and relax, defocusing (staring at one spot on the screen and emptying your mind from any thoughts), thinking at a relaxing place or activity. Absolutely to avoid: thinking at activity related to hand movement or keep thinking at the preceding/following motor imagery task.