Jan Zerfowski*, Martin Wimpff, Paul Anders, Niels Peekhaus, Vincent Jonany, and Surjo R. Soekadar

Towards Multi-Class Optically Pumped Magnetometer-Based Brain-Computer Interfaces

https://doi.org/10.1515/cdbme-2025-0218

Abstract: Magnetoencephalography (MEG) using optically pumped magnetometers (OPMs) offers several advantages over electroencephalography (EEG) and superconducting quantum interference device (SQUID)-based MEG, including improved spatial resolution and versatility. In this study, we investigate the potential of OPM-based MEG for multi-class Brain-Computer Interface (BCI) applications.

We classify neuronal activity related to five mental tasks – hand motor imagery (MI), feet MI, mental rotation, mental subtraction, and word association – using data from 12 participants. The goal is to develop a strategy to select the most discriminative task subsets for each user. The data were processed using the filter-bank common spatial patterns (FBCSP) algorithm and linear discriminant analysis (LDA), with classification accuracy evaluated using block-wise cross-validation. The results show that, in all but one participant, classification accuracies exceeded the commonly accepted 70% threshold for reliable BCI control, particularly with three- and sometimes with four-class combinations. Hand MI and mental rotation emerged as the most distinct tasks, followed by word association and feet MI being included in fewer optimal task combinations. The study demonstrates that user-specific task selection is crucial to maximize BCI performance.

The presented results are promising, but further development is needed, including the implementation of real-time feedback systems, addressing technical limitations, and refining classification techniques. The findings highlight the potential of OPM-based MEG for future clinical applications, particularly in rehabilitative settings, where its high spatial and temporal resolution could significantly enhance BCI systems.

Keywords: Optically pumped magnetometers, Magnetoencephalography, Brain-Computer Interface

***Corresponding author: Jan Zerfowski,** Charité – Universitätsmedizin Berlin, Charitéplatz 1, 10117 Berlin, Germany; PTB Berlin, e-mail: jan.zerfowski@charite.de

Martin Wimpff, Universität Stuttgart, Stuttgart, Germany

Niels Peekhaus, Vincent Jonany, Surjo R. Soekadar, Charité – Universitätsmedizin Berlin

Paul Anders, Physikalisch-Technische Bundesanstalt, Berlin, Germany

1 Introduction

Magnetoencephalography (MEG) using optically pumped magnetometers (OPM) has become an important tool for the investigation of neural functions. Recent studies highlight the potential of OPMs to surpass electroencephalography (EEG) in spatial resolution, and superconducting SQUID-MEG in maintenance cost and versatility [1]. OPMs are a promising technology to advance diagnostics and treatment in many clinical applications like epilepsy [2] and translational neuroscience [3].

Brain-computer interfaces (BCIs) translate neural activity into control signals of external devices, like prostheses or robots that provide direct sensory feedback. They also enable severely paralyzed users with locked-in syndrome (LIS) to communicate [4]. Current BCIs applied in rehabilitation are often based on EEG to analyze and classify voluntary neural activity. Often, event-related (de)synchronization (ERD/ERS) in the alpha and beta frequency bands (8 Hz to 16 Hz and 16 Hz to 32 Hz, respectively) is used as a correlate of mental activity. However, the volume conduction of tissue and skull between the signal source and the electrodes impairs the spatial resolution of EEG and therefore limits the number of decodable states or degrees of freedom.

MEG allows for better localization of discrete neuronal sources, as the magnetic fields pass through tissue with little to no distortion [5]. The classic technique to measure MEG is using superconducting quantum interference devices (SQUIDs), which require cryogenic cooling and are static devices with one-size-fits-all sensor arrays. OPMs, on the other hand, do not require cooling and can move in space during the measurement [5]. This enables bespoke sensor helmets for better sensor fits and studies on children and vulnerable populations unable to sit still for longer periods of time.

While progress in EEG-based BCIs has recently plateaued, OPM-based BCIs are still in the early stage of exploration. At this point, most studies have only developed proof-of-concept for real-time data analysis [6], tested binary or passive paradigms [7–9]. Appropriate best practices and methods for screening users and online processing of data are still lacking.

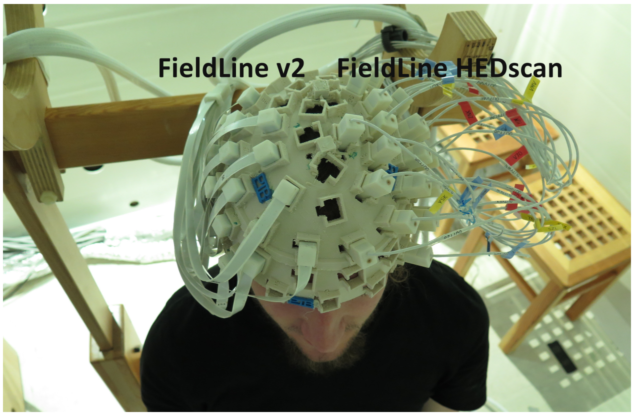

Fig. 1: A participant wearing the 3D-printed helmet within BMSR-2.1. The hemispheres with different sensors are annotated

In this study, we classified neuronal activity based on five different mental tasks and developed a strategy to pick the best performing subset of those tasks for further investigation. We recorded OPM-MEG data of 12 healthy participants performing hand motor imagery (MI), feet MI, mental rotation, mental subtraction and word association. Using methods established in BCIs to process and classify the recorded data, we devised a method to screen each subject and select the most discriminative conditions that should be used for a user-specific BCI. With the development of such a strategy, we present here a proof-of-concept towards multi-class OPM-based BCIs optimized for user-specific performance.

2 Methods

2.1 Sensor System and environment

For this study, a total of 40 OPMs were used (16 v2 and 24 v3/HEDscan OPMs from FieldLine Inc., Colorado, USA). The sensors were inserted into an 3D-printed helmet based on an average head shape [10]. The v2 sensors were distributed on the right hemisphere, the HEDscan sensors on the left hemisphere (see Figure 1)

The experiment was conducted at the Physikalisch-Technische Bundesanstalt (PTB) in Berlin, Germany, in the Berlin Magnetically Shielded Room (BMSR-2.1). It has a very high shielding factor of 10^8 above 6 Hz, static remnant fields of less than $1\,\mathrm{nT}$ and field gradients of under $1\,\mathrm{pT/m}$.

2.2 Experimental Setup and Participants

Five mental tasks were performed based on previous EEG-based BCI studies (see also Figure 2):

Fig. 2: Examples for the presented stimuli (left to right): Hand MI, Feet MI, Mental Rotation; Mental Subtraction, Word Association

- **Hand Motor Imagery**: self-paced, imagined right hand opening and closing motions (approx. 1/s)
- **Foot Motor Imagery**: self-paced, imagined rotations of the feet
- **Mental Subtraction**: Repeated subtraction of two numbers (e.g., 117-12, 105-12, 93-12, ...)
- **Mental Rotation**: Rotation of a simple object before the inner eye
- **Word Association**: Associating words with a given starting letter

Stimuli were presented for 6 seconds at a time, during which participants were asked to continuously and repetitively execute the respective tasks. After each trial, an inter-trial interval of 6 seconds gave participants time to rest before the next trial started. A total of 40 trials per condition were recorded in 5 blocks, between which participants could take a break and continue with the study at their own pace.

BMSR-2.1 features a bidirectional communication system in the room, through which participants were supervised during the experiment. The stimuli were presented using a projector enclosed in a portable magnetic shielding [11].

In total, 12 healthy participants (4 female, 8 male) took part in the study, 9 of which had never participated in BCI experiments before (BCI-naive). All participants received an introduction to the paradigm, the safety mechanisms of BMSR-2.1 and were made aware that they can stop or interrupt the study at any time.

2.3 Data Acquisition and Preprocessing

The 8-layer shielding of BMSR-2.1 effectively eliminates transient magnetic field fluctuations, and thus very little data pre-processing was required. Trials which showed artifacts exceeding 50 pT were eliminated from the data, since those trials most likely indicate that the participant was moving and not focused on the task. For three participants, less than 50% of trials in at least one condition remained after pre-processing. These participants were excluded from further analyses, as insufficient training data can lead to unreliable results.

The data were recorded with the proprietary FieldLine recording software. MNE for python was used to load, label and process the data.

2.4 Data Analysis

Data classification was performed using the filter-bank common spatial patterns (FBCSP, [12]) algorithm combined with linear discriminant analysis (LDA). FBCSP and LDA are established methods for BCI applications that allow for rapid training and processing of the data, while maintaining interpretability of the extracted features.

Since the spatial distribution of alpha and beta oscillatory brain activity depends on the mental task, the FBCSP-based processing is designed to effectively identify those spatial patterns and extract the corresponding signals.

FBCSP was implemented using MNE's minimum-phase filters from $4\,\mathrm{Hz}$ to $40\,\mathrm{Hz}$ at a width of $4\,\mathrm{Hz}$ each, no overlap, and transition bandwidths of $4\,\mathrm{Hz}$. As a result, 9 band-passed CSPs were trained, and for each CSP two filters were selected based on their eigenvalues. Finally, the best six of these 18 spatial filters were selected, based on their mutual information. The band powers of the resulting six channels in FBCSP-space were then used as features for the next analysis step. With the transformed data as input, the scikit-learn LDA implementation was used to retrieve the final class estimation. To determine the accuracy of our classifications, we used a five-fold block-wise cross-validation with an 80%/20% train/test split. The accuracies provided here always refer to the testing accuracy unless stated otherwise.

The percent theoretical chance level for an n-class paradigm is $100/n\%$. However, this only holds true if infinitely many samples are available. For a finite number of samples, the classification errors follow a binomial cumulative distribution [13]. According to the formula provided in [13], the correct chance level thresholds at $p < 10^{-4}$ for a 3, 4 and 5-class classification problem with 40 trials per condition, are 50%, 38.125% and 31.0%, respectively.

3 Results

The current study explores the distinction of five different BCI-related mental tasks when measured with OPMs. The goal is to establish a strategy to select the best user-specific subset of tasks, yielding the highest discriminability and allowing for reliable BCI control. In the BCI community, a threshold of 70% accuracy is often assumed to be the minimum to gain reliable control through operant conditioning. Below 70% users often do not gain control and report that the BCI acts arbitrary and unpredictable. Although the theoretical chance level (computed as $1/\text{number of classes}$) is often used to compare accuracies, we used here the correct empirical chance level based on the number of available trials in accordance with [13]. Figure 3 indicates both theoretical and empirical chance level at $p < 0.0001$. All participants performed significantly above that chance level.

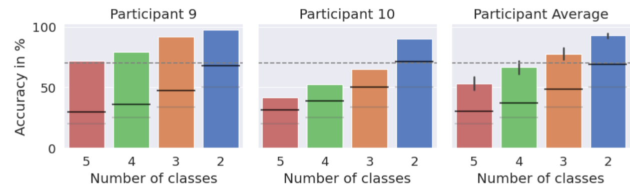

Fig. 3: Classification accuracy for the respective best class combinations. Left: The best participant is shown with 5, 4, 3 and 2 class combinations above the 70% threshold. Center: The worst participant, whose accuracy exceeds 70% only for 2 classes. Right: Average across all 9 analyzed participants (right). The theoretical chance level is indicated with a grey line, the computed chance level for $p < 0.0001$ is shown in black

All except one participant exceeded a classification accuracy of 70% for the best combination of three classes. Three participants performed better than that threshold in the four-class combination, and a single participant performed even better than that on the combination of all five classes.

Further analyses show that not all classes are equally likely to be included in the best-performing subset. However, the best combination of classes is specific to the participant and could not be linked to a regular pattern. Based on the number of occurrences in the best-performing subset, the classes Hand MI and mental rotation are most distinct, followed by word association and feet MI. Mental rotation is included in the fewest sub-selections.

4 Discussion

This study demonstrates the feasibility of classifying multiple mental tasks using OPMs and presents a strategy for selecting the most discriminative tasks for user-specific brain-computer interface (BCI) applications. Our results show that, with appropriate data preprocessing and classification techniques, it is possible to achieve reliable classification accuracies above the commonly accepted 70% threshold in most participants. The optimal combination of tasks varied among individuals, emphasizing the importance of personalized task selection for BCI performance. The presented results indicate that, while reliable five-class discrimination remains challenging, user-specific subsets of three or four classes achieved classification accuracies above 70% in all except one participant.

For this experiment, two independent data acquisition systems (FieldLine v2 and HEDScan/v3) were combined to maximize the head coverage and number of available channels. Since no synchronization between the systems took place, narrow-band oscillations around 5 and 24 Hz, presumably caused by the OPM-internal modulation coils of the opposite system, were visible in the raw data. Although the artifacts are independent of task and time, they might have led to a lowered signal-to-noise ratio, and thus lower accuracy. Furthermore, the mutual disturbances paired with participant movement could have contributed to the artifacts visible in the data of three participants that were excluded from the analyses. For future studies, we recommend using a single system with synchronized modulation and embedded crosstalk suppression.

The results presented here are based on offline processing and classification of the recorded data. The applied filtering and classification algorithms were picked intentionally, as they can be transferred for use in an online pipeline. With the screening strategy established in this work, improved algorithms for artifact rejection and data classification, and the addition of online feedback, the performance of our system could likely be improved significantly.

Future studies should prioritize translating these advancements into practical, real-time applications for clinical and rehabilitation settings. OPM-based BCIs hold significant promise for clinical use, offering a quick and easy setup across diverse populations. Their superior spatial resolution compared to EEG, combined with easier maintenance than SQUID-MEG and high temporal resolution, makes them a compelling focus for further scientific exploration.

Author Statement

Research funding: The presented research was funded through QSens – QHMI1 (03ZU1110DD, 03ZU1110DC), QSens – QHMI2 (03ZU2110FC), NeuroQ (13N16486) and DFG (460785001), European Research Council (ERC) under the project NGBMI (759370), TIMS (101081905) and BNCI2 (101088715), Einstein Stiftung Berlin (ESB); the Federal Ministry of Research and Education (BMBF) under the project NeuroQ (13N16486) **Conflict of interest:** The authors state no conflict of interest. **Informed consent:** Informed consent has been obtained from all individuals included in this study. **Ethical approval:** The research related to human use complies with all the relevant national regulations, institutional policies and was performed in accordance with the tenets of the Helsinki Declaration, and has been approved by the authors' institutional review board or equivalent committee (EA1/338/20).

References

[1] E. Boto et al., "On the Potential of a New Generation of Magnetometers for MEG: A Beamformer Simulation Study," PLOS ONE, vol. 11, no. 8, p. e0157655, Aug. 2016
[2] M. Pedersen, D. F. Abbott, and G. D. Jackson, "Wearable OPM-MEG: A changing landscape for epilepsy," Epilepsia, vol. 63, no. 11, pp. 2745–2753, Nov. 2022
[3] M. Brickwedde et al., "Applications of OPM-MEG for translational neuroscience: a perspective," Transl Psychiatry, vol. 14, no. 1, pp. 1–12, Aug. 2024
[4] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain–computer interfaces for communication and control," Clinical Neurophysiology, vol. 113, no. 6, pp. 767–791, Jun. 2002
[5] E. Boto et al., "Wearable neuroimaging: Combining and contrasting magnetoencephalography and electroencephalography," Neuroimage, vol. 201, p. 116099, Nov. 2019
[6] J. Zerfowski, T. H. Sander, M. Tangermann, S. R. Soekadar, and T. Middelmann, "Real-time data processing for brain-computer interfacing using optically pumped magnetometers," Int. J. Bioelectromagn., vol. 23, no. 2, p. 14/1–6, Aug. 2021.
[7] B. Wittevrongel et al., "Practical real-time MEG-based neural interfacing with optically pumped magnetometers," BMC Biology, vol. 19, no. 1, p. 158, Aug. 2021
[8] A. Y. Paek, A. Kilicarslan, B. Korenko, V. Gerginov, S. Knappe, and J. L. Contreras-Vidal, "Towards a Portable Magnetoencephalography Based Brain Computer Interface with Optically-Pumped Magnetometers," in 2020 42nd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC), Jul. 2020, pp. 3420–3423
[9] G. Tan et al., "A Brain-Computer Interface System based on Magnetoencephalography with Optically Pumped Magnetometers," IEEE Transactions on Instrumentation and Measurement, pp. 1–1, 2025
[10] U. Marhl, A. Jodko-Władzińska, R. Brühl, T. H. Sander, and V. Jazbinšek, "Transforming and comparing data between standard SQUID and OPM-MEG systems," PLOS ONE, vol. 17, no. 1, p. e0262669, Jan. 2022
[11] P. Anders et al., "Magnetically shielded high-resolution visual stimulation for OPM-MEG applications," Biomed Phys Eng Express, vol. 11, no. 2, Feb. 2025
[12] K. Ang, Z. Chin, H. Zhang, and C. Guan, "Filter Bank Common Spatial Pattern (FBCSP) in brain-computer interface," presented at the Proceedings of the International Joint Conference on Neural Networks, Jul. 2008, pp. 2390–2397
[13] E. Combrisson and K. Jerbi, "Exceeding chance level by chance: The caveat of theoretical chance levels in brain signal classification and statistical assessment of decoding accuracy," Journal of Neuroscience Methods, vol. 250, pp. 126–136, Jul. 2015