EXPLORING MULTIFRACTALITY AND TEMPORAL DYNAMICS IN MOTOR IMAGERY

COMPARING LINEAR AND NON-LINEAR APPROACHES

JOS PRINSEN

THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE IN COGNITIVE SCIENCE & ARTIFICIAL INTELLIGENCE

DEPARTMENT OF
COGNITIVE SCIENCE & ARTIFICIAL INTELLIGENCE
SCHOOL OF HUMANITIES AND DIGITAL SCIENCES
TILBURG UNIVERSITY

STUDENT NUMBER

2040874

COMMITTEE

dr. Travis Wiltshire
dr. Paris Mavromoustakos Blom

LOCATION

Tilburg University
School of Humanities and Digital Sciences
Department of Cognitive Science &
Artificial Intelligence
Tilburg, The Netherlands

DATE

May 21, 2024

WORD COUNT

A lot

ACKNOWLEDGMENTS

Some room for acknowledgements.

EXPLORING MULTIFRACTALITY AND TEMPORAL DYNAMICS IN MOTOR IMAGERY

COMPARING LINEAR AND NON-LINEAR APPROACHES

JOS PRINSEN

Abstract

This thesis investigates the dynamics of brain activity during motor imagery using Multifractal Detrended Fluctuation Analysis to address limitations of traditional linear methods like Event-Related Desynchronization. The study aims to characterize the multifractal properties of EEG signals during MI, compare these properties between high and low-performing groups, and analyze their temporal evolution over multiple training sessions. Utilizing EEG data from 62 participants, results reveal that MI signals exhibit stronger multifractal characteristics compared to rest, particularly in the ipsilateral hemisphere. High performers show significant contralateral ERD, while low performers lack clear lateralization, suggesting that multifractal measures might be useful to address BCI inefficiency. Temporal analysis indicates significant improvement in Motor Imagery accuracy over sessions, with decreasing multifractal complexity, highlighting their potential in capturing learning dynamics. These findings suggest that integrating multifractal analysis with traditional methods could potentially enhance BCI systems, improving communication and independence for individuals with severe motor impairments.

1 DATA SOURCE, ETHICS, CODE, AND TECHNOLOGY STATEMENT

The data used in this thesis was made open source in parallel to the following paper: Mindfulness Improves Brain–Computer Interface Performance by Increasing Control Over Neural Activity in the Alpha Band **?**. The thesis did not involve any data collection. The owner of the data made it publicly available, and gave consent for using it in further publication. The data was acquired through Figshare: Human EEG Dataset for Brain-Computer Interface and Meditation. All figures and images were created specifically for this project. No code from other papers was utilized, however, some parts of the ERD calculation were based on a paper by citeMa2020 . All packages and package versions are listed in the coding environment provided with the thesis on Github: https://github.com/JosPrinsen/MasterThesis.git. ChatGPT was used for generating both Python and R code for the analysis, as well as for rephrasing texts to improve academic language use. Zotero was employed for reference management.

2 INTRODUCTION

Imagine a technology that not only responds to our commands but understands our intentions. Brain-computer interfaces (BCIs) promise such an interface, with Motor Imagery (MI) serving as a cornerstone for these revolutionary systems (Padfield, Zabalza, Zhao, Masero, & Ren, 2019). Yet, despite significant advancements, MI BCIs frequently face inefficiencies, specifically for a small group of BCI inefficient users (Zhang, Li, Zhang, Yao, & Xu, 2020). While some argue that the users are incapable of generating the correct brain activity during the MI task, others argue that the classification systems are incapable of generalizing over everyone (Zhang et al., 2020). Moreover, there are large differences in the patterns generated across trials and over different sessions/days (Zhang et al., 2020). Traditional linear analysis methods, while foundational, might not capture all components of the brain's intricate communication patterns (Aftanas & Golocheikine, 2002).

This paper proposes to explore a non-linear analysis method, namely Multifractal Detrended Fluctuation Analysis (MFDFA), to delve deeper into the brain related dynamics that occur during MI. MFDFA is a statistical method used to analyze scale-invariant properties and multifractality in non-stationary time series data (Kelty-Stephen, Lane, Bloomfield, & Mangalam, 2022; Kelty-Stephen, Palatinus, Saltzman, & Dixon, 2013). In other words, MFDFA is a method used to detect and measure varying patterns in small and large time scales/fragments over different statistical moments (Kelty-Stephen et al., 2013). Multifractal analysis has been used to analyze diverse bio-physiological signals, such as detecting epileptic zones in EEG (Sikdar, Roy, & Mahadevappa, 2018), however, there is a lack of research of multifractality relating to MI. While multifractality is promising, given the increase in accuracy when used in machine learning classifiers, there are no papers examining or interpreting multifractal behavior in MI (Brodu, Lotte, & Lécuyer, 2012). The main goal of this thesis is to understand how these nonlinear metrics behave during MI, investigating whether, and in what way, MFDFA metrics could be used to enhance BCI's in the future. Additionally, by investigating the temporal evolution of MI across sessions, this study aims not just to understand the degree of multifractality that occurs during MI, but also how these complex measures evolve over time during MI training. Essentially, this temporal analysis seeks to uncover how individual differences in learning speed and MI improvement manifest in the multifractal properties of EEG signals collected during MI, possibly offering new insights into the cognitive and neural mechanisms underlying MI-BCI training. Moreover, since there are no studies showing the evolution of MFDFA metrics of brain dynamics over time/sessions, it might expand our understanding of the non-linear processes that govern brain activity.

To guide this exploratory analysis, this thesis seeks to address several key research questions that aim to deepen our understanding of the multifractal characteristics of EEG signals during MI. Specifically, this research will investigate the following:

RQ1 *What are the Multifractal characteristics of EEG signals during motor imagery?*

RQ2 *Are MFDFA metrics different based on the performance of participants during MI?*

RQ3 *How do the Multifractal properties of EEG signals evolve with motor imagery training, and how does this compare to traditional linear descriptors?*

These questions are exploratory in nature, and focus on contrasting the complex, non-linear dynamics of brain activity associated with MI with those extracted through traditional linear methods. Ultimately, the findings of this paper might contribute to a greater understanding of the dynamics during MI, which in turn could improve BCI's.

Advancing the understanding and efficiency of BCIs holds significant societal importance. Enhanced BCIs can provide improved communication tools, increased independence, and a higher quality of life for individuals with severe motor impairments. For example, effective BCIs could enable people with disabilities to control prosthetic limbs, operate wheelchairs, or communicate through digital interfaces, all through the power of thought (Tariq, Trivailo, & Simic, 2018). Addressing the inefficiencies in current BCI systems is crucial for making these technologies more reliable and accessible. Additionally, improved BCI technology can reduce healthcare costs and the burden on caregivers by promoting self-sufficiency among users (Geronimo & Simmons, 2020). Ultimately, the findings of this research could be a stepping stone to drive technological progress, where advanced BCI technology is accessible to all who need it.

3 RELATED WORK

3.1 *Introduction to Motor Imagery*

MI entails mentally rehearsing a motor action without physical execution (Padfield et al., 2019). This cognitive simulation leverages neural pathways akin to those involved in actual movement (Sheahan, Ingram, Žalalytė, & Wolpert, 2018). BCI systems are capable of classifying the intent from the brain activity generated from this mental simulation, which presents a viable alternative for individuals with mobility impairments or limb loss, such as those affected by paralysis or stroke (Padfield et al., 2019). By envisioning movements, users can interface with prosthetic devices or engage in neurorehabilitation exercises, fostering improvement in motor control through repeated mental practice (Bovend'Eerdt, Dawes, Sackley, Izadi, & Wade, 2010). In neurorehabilitation patients practice MI and enhance their ability to vividly imagine and simulate motor actions without actual movement, which in turn also positively impacts control over physical motor actions (Padfield et al., 2019). Given these promising applications, MI has been extensively studied using techniques such as electroencephalography (EEG), providing a non-invasive means to capture the brain's electrical activity with high temporal resolution, making it suitable for studying dynamic changes in brain activity associated with MI tasks (Lotze & Halsband, 2006; Padfield et al., 2019)

3.2 *Even-related Desynchronization/Synchronization*

Previous research has demonstrated a modulation of oscillatory activity during MI tasks through triggering event-related neural potentials (ERP's) particularly in the mu (8-13 Hz) and beta (13-30 Hz) frequency bands over sensorimotor regions (SMR) of the brain (Pfurtscheller, 2000). This modulation, known as event-related desynchronization (ERD) followed by an event-related synchronization (ERS) signifies a decrease in power, followed by in increase in power within the mu and beta frequency bands over the sensorimotor areas, corresponding to an increase in neuronal activity related to the imagined or anticipated motor act (Neuper, Wörtz, & Pfurtscheller, 2006; Pfurtscheller, 2000; Pfurtscheller & Da Silva, 1999). The magnitude of ERD/ERS, quantified by the decrease followed by an increase in power, in the mu and beta bands correlates with the degree of engagement of the SMR region (Nakayashiki, Saeki, Takata, Hayashi, & Kondo, 2014; Pfurtscheller & Da Silva, 1999). Studies have shown that as individuals practice motor imagery tasks, the SMR is more engaged, which is evident from an increase in the magnitude of the ERD (Nakayashiki et al., 2014; Prasad, Herman, Coyle, McDonough, & Crosbie, 2010). While the strength of the ERD changes when improving on MI tasks, the mechanics of such ERD patterns are not consistent between participants, and even over trials within participants (Huang et al., 2023; Wriessnegger, Müller-Putz, Brunner, & Sburlea, 2020). For example, contralateralization may vary, with some participants showing more pronounced bilateral or ipsilateral ERD patterns, especially during complex or less familiar tasks (Lee, Lee, Kim, & Kim, 2020).

3.3 *Lateralization in Motor Imagery*

Lateralization in the context of motor imagery refers to the asymmetric activation of brain hemispheres, particularly evident in EEG patterns (Lee et al., 2020). Lateralization as described by the lateralization index (LI) refers to the difference in power between the hemispheres, indicating the degree of engagement of sensorimotor regions during motor imagery tasks (Nam, Jeon, Kim, Lee, & Park, 2011). High LI values suggest stronger engagement of the contralateral hemisphere, which is critical for the accuracy and performance of MI-BCI systems, whereas negative LI values suggest that ipsilateral is dominating the interaction (Lee et al., 2020; Nam et al., 2011). Studies have shown that the ERD is significantly larger in the hemisphere contralateral to the movement (Doyle, Yarrow, & Brown, 2005). This contralateral dominance is essential for motor selection processes, allowing early preparation and shorter reaction times in tasks where the laterality of the movement is predictable (Doyle et al., 2005; Pfurtscheller & Da Silva, 1999). The pattern of lateralization, however, is not consistent across all individuals, reflecting differences in neural architectures and learning strategies (Huang et al., 2023). Some participants may show more pronounced bilateral or ipsilateral ERD patterns, especially during complex or less familiar tasks (Huang et al., 2023; Wriessnegger et al., 2020).

3.4 *Variability in MI descriptors*

The variability observed between subjects has been attributed to factors such as age, gender, and living habits, which influence the brain's topography and electrophysiological properties, and has been named the BCI inefficiency problem (Huang et al., 2023; Zhang et al., 2020). This diversity in neural architecture and functional organization means that a BCI model trained on one individual often performs poorly when applied to another (Huang et al., 2023). Next to inter subject variability, there is also intra subject variability, where the ERD, and lateralization is different between trials, and session within the same participant (Huang et al., 2023; Zhang et al., 2020). It is believed that intra subject variability arises from fluctuations in psychological and physiological states, such as levels of fatigue, relaxation, concentration, and task familiarity (Huang et al., 2023; Perquin, Van Vugt, Hedge, & Bompas, 2023). These changes can significantly impact the consistency of EEG recordings over time, leading to a decline in BCI performance even within the same individual. The presence of both inter- and intra-subject variability undermines the assumption of independent and identically distributed data sets, which is a cornerstone of conventional machine learning frameworks that are used for classifying MI interactions (Huang et al., 2023).

3.5 *Addressing BCI inefficiency*

While some researchers try to address this inter and intra subject variability through means of adaptive classification, for example, through transfer learning. Others are exploring new features to use in the classification of MI (Brodu et al., 2012). For a comprehensive review of traditional features that are employed for MI, please refer to (Bashashati, Fatourechi, Ward, & Birch, 2007). More recently however, the use of non-linear features has been suggested. According to Kelty-Stephen et al. (2013), it is possible that traditional linear analyses do not fully capture the complexity of physiological data, including EEG activity, as they often assume homogeneity in the data. Non-linear analyses such as multifractal analysis might be able to account for this, as these metrics are designed to capture the intricacies of data that exhibit variability across multiple scales (Ihlen & Vereijken, 2010; Kelty-Stephen et al., 2022).

3.6 *Multifractal Detrended Fluctuation Analysis*

An example of such a multifractal analysis is MultiFractal Detrended Fluctuation Analysis (MFDFA), which is a statistical tool used to explore the complex and self-similar properties of time series data across different scales (Ihlen & Vereijken, 2010; Kelty-Stephen et al., 2022). At its core, MFDFA explores how the statistical moments (e.g., variance or the mean) of signals change as one zooms in or out across different time scales, which in turn reveals the degree of interactivity between time scales (Kelty-Stephen et al., 2013). In terms of brain activity, this interactivity could be interpreted as different neuronal substructures or populations interacting or communicating over multiple spatiotemporal scales, causing emergent behavior visible as brain networks or global activity (Ihlen & Vereijken, 2010).

MFDFA plots yield several important measures that provide insights into the complexity and variability of time series data. The Hurst exponent (H(q)), derived from the slope of the log-log plot of the fluctuation function against different statistical moments (q), indicates the degree of long-range dependence in the data for each scale. The multifractal spectrum (D(h)) visually represents the distribution of singularities, with its width reflecting the range of multifractal behaviors; a wider spectrum indicates greater multifractal complexity. Scaling exponents (Tau(q)) describe how (statistical) moments scale with window size. When Tau(q) shows a non-linear curve it suggests the presence of multifractal behavior. The singularity strength (h) and its range (hmax - hmin) highlight the intensity and diversity of irregularities in the signal, while the maximal fractal dimension (D(hmax)) identifies the most prevalent scaling behavior. For a more indepth explanation on MFDFA, please refer to Kelty-Stephen et al. (2013).

MFDFA has widely been used in a variety of different research fields, such as geography, economics and for physiological measures (Gu & Zhou, 2006). MFDFA has also been used to analyze EEG signals. For example, in a study by Gaurav, Anand, and Kumar (2021), MFDFA metrics extracted from EEG were used to classify between 6 levels of cognitive states associated with motor, perception, and attention tasks. The study utilized H(q) values calculated from EEG segments of 10 seconds, reaching an accuracy of around 91%, demonstrating that Hurst exponents can be used to differentiate between tasks which are expected to generate varying levels of complexity in the brain. Moreover, MFDFA features were used for classifying emotions from EEG activity, achieving an average classification accuracy of 84.5% for positive emotions and 82.5% for negative emotions (Paul, Mazumder, Ghosh, Tibarewala, & Vimalarani, 2015).

MFDFA measures aren't the only type of non-linear algorithms that capture multifractal dynamics. While multifractal cummulants capture slightly different dynamics, both aim to capture the multifractal structure/complexity of a timeseries. Interestingly, multifractal cummulants were used to classify MI. Brodu et al. (2012) demonstrates over 13 datasets that multifractal cumulants, can be used for classification of left/right hand MI with an accuracy of 75.8%. Moreover, combining these multifractal features with band powers increased the classification accuracy for nearly all datasets to an average of 80.3%. This shows that multifractal features can be used to capture the dynamics needed for MI classification.

Despite the successful application of multifractal features in MI classification, there is a noticable gap in the understanding of the multifractal behaviour relating to MI dynamics. While both multifractal cummulants and MFDFA features have been successfully used in classifiers for different cognitive tasks, this was always done without any literature explaining how or why the features could help in improving classification accuracy. Instead of incorporating multifractal values directly into a classifier, this thesis aims to take a step back, and try to understand the multifractal dynamics that occur during MI. By analyzing the Hurst exponent, multifractal spectrum, and scaling exponents, the main goal is to uncover the underlying complexities and interactions within MI-related EEG signals, with as ultimate goal to both understand the brain better, but also to increase understanding of the multifractal features that can be used in MI classification.

Considering that MI is suffering from inter- trial, and session variability, it is crucial to not investigate one session in isolation, but to investigate these metrics through a temporal scope. Techniques such as autocorrelation (AC) and detrended fluctuation analysis (DFA) can provide additional layers of understanding specifically in the temporal domain (Perquin et al., 2023). These methods have been utilized effectively to analyze temporal dependencies and reveal patterns between trials during sessions of reaction time series, as demonstrated in the study by Perquin et al. (2023) on sensorimotor variability. Autocorrelation helps in quantifying the correlation of a time series with itself over various delays, which can expose the persistence of dependencies across multiple trials (Xu, De Barbaro, Abney, & Cox, 2020). DFA, on the other hand, can assess the scaling properties of the data, highlighting long-range correlations that are not immediately apparent through simple statistical measures (Hardstone et al., 2012). In other words, these two techniques are used to quantize the degree to which trials within a session are similar to itself, and whether there are long term relationships between the trials. Performing a longitudinal analysis by calculating Autocorrelation and DFA for both linear ERD metrics and MFDFA metrics might provide insights into the intra-trial and intra-session variability in each of these metrics. By examining how multifractal measures evolve across multiple sessions, a temporal analysis might provide insights into how individual differences in learning speed and MI improvement are reflected in the multifractal properties of EEG signals collected during MI. Notably, no studies to date have performed a longitudinal analysis of MFDFA values in EEG metrics, let alone for Motor Imagery. This novel approach aims to reveal how multifractal dynamics change over time, potentially identifying specific multifractal characteristics that correlate with MI practice. It would be especially interesting to compare these multifractal metrics to the standard linear metrics for MI to understand their relative effectiveness and potential complementary insights.

4 METHOD

The following section is dedicated to describing the methods and dataset used to perform a longitudinal study of multifractal dynamics for MI.

4.1 *Dataset Description*

The dataset was comprised of de-identified electroencephalogram data from 62 healthy participants, collected during a BCI Motor Imagery study that explored the impact of mind-body awareness training on BCI performance (Stieger et al., 2020). Participants underwent 7 to 11 BCI training sessions, each aimed at enhancing their ability to control a computer cursor through intended directional thoughts in both one-dimensional (1D) and two-dimensional (2D) tasks. For this study a subset was taken, where only the first 7 sessions with only 1D tasks were considered. The directional thoughts were either Left or Right (L/R) imagined movement.

EEG signals were recorded during the Motor Imagery sessions using a 64-electrode setup, of which 62 were included in the final dataset. The data were collected with a sampling rate of 1000 Hz and stored in MATLAB-readable (.mat) format, structured to include raw EEG signals and associated metadata. Each session consisted of approximately 450 trials, with mixed 1D tasks and 2D tasks. Only 1D tasks were included for this study, totalling ot 225 trials per session. Each trial consisted of a rest phase of 2 seconds, a 2 second presentation phase, during which either the left or right side was highlighted, and a MI phase, during which the participant was asked to move a circle from the middle of the screen to either the left or right side of the screen, based on their L/R intent. The screen provided feedback to the user, by moving the circle based on the classification results of the participant, which was based on an Autoregressive classification algorithm, which is further described in Stieger et al. (2020). The duration of the MI phase varied depending on how fast the participant completes the task. The task is completed once the circle reaches either the L/R side of the screen, which is stored as either successful or unsuccessful based on whether the participant was able to move the circle to the correct side. When the participant was not able to complete the MI phase within 6 seconds, the trial was marked as a timeout, which was also stored as an unsuccessful result. On a timeout, a "forced result" was also stored, which is based on the side the circle was on when the task timed out. For example, if the circle was at the correct side, while not fully touching the side of the screen when timed out, the forced result would still be marked as successful, whereas the result would be marked as unsuccessful. Performance over the entire session is stored as the number of successful trials divided by the total number of trials. A total of 685 trials were removed from the dataset, which were deemed to have artifacts in the EEG data. The clean dataset consisted of 96981 trials, over 434 sessions The sessions of all participants were divided into two groups; a "good" performing group, and a "bad" performing group, based on a k-means clustering algorithm, with k=2. The algorithm split the sessions in two groups based on their performance over the session. In total 160 session were included in the "good" performing group, whereas 274 were classified as "bad".

4.2 *Quantification of Motor Imagery Dynamics*

The data from Stieger et al. (2020), was already pre-processed, with artifacts marked, as well as a bandpass filter between 0.1 and 100 Hz. Data from the C3 and C4 channels were further bandpassed to capture only the Mu (8-13Hz) and Beta (13-30Hz) frequency bands over the left and right sensorimotor regions. The specific frequency bands, as well as channels of interest were chosen due to their known relevance in motor control and cognitive engagement, and are typically the main frequency bands investigated during MI. The raw EEG data of the Rest and MI phase were squared to investigate magnitude of power within the bandpassed frequency range. Bandpass and squaring the data was chosen over an alternative method like a Time-Frequency method for computational efficiency. After the amplitudes were extracted, the power in the Rest phase was averaged, and used as a baseline for the MI phase. The ERD values were extracted from the MI based on a formula from Pfurtscheller and Da Silva (1999), which follows as:

$$\begin{aligned} ERD_{C3} &= \frac{PowerMI_{C3} - PowerREST_{C3}}{PowerREST_{C3}} \\ ERD_{C4} &= \frac{PowerMI_{C4} - PowerREST_{C4}}{PowerREST_{C4}} \end{aligned} \tag{1}$$

These values are averaged over the trial, giving the average decrease in Mu/Beta power between the Baseline Rest phase and the MI phase

4.3 *Parameters for Multifractal Detrended Fluctuation Analysis*

MFDFA metrics are extracted from the baselined amplitudes that were described previously. The MFDFA package as described in X was used for the entire analysis. MFDFA estimations are extremely sensitive, and therefore setting the correct parameters is of utmost importance. In MFDFA, the q-range critically influences the analysis by weighting fluctuations in the time series, with positive q-values emphasizing larger fluctuations and negative values accentuating smaller ones. The chosen range of -5 to 15 is based on Gaurav et al. (2021), which used this range for MFDFA analysis for EEG. This range prevents computational errors and the dominance of extreme fluctuations in smaller timeseries typical for EEG. Furthermore the detrending order is set to 2, which is a sufficient detrending order according to Gaurav et al. (2021) for EEG data. In MFDFA, the range of scales used during analysis influences the estimates that are calculated. The minimal scale primarily impacts the granularity of the analysis, specifically influencing how small-scale fluctuations are observed and quantified. According to Ihlen and Vereijken (2010), the minimal value should be based on the fastest significant frequency component in the data, as well as its sampling frequency. Since both the Rest and MI data were band-passed at 100Hz, the fastest significant frequency component is at 100Hz. Considering a sampling frequency of 1000Hz, we divide the sampling frequency by the fastest frequency component, setting a minimal scale of 10. This ensures that each segment captures at least one full cycle of the fastest dynamics, while ensuring that the analysis is stable and that the estimation of local Root Mean Square (RMS) fluctuations avoids random noise that might appear significant in smaller segments. Conversely, the maximal scale, typically less than 1/10th of the total data points (Ihlen & Vereijken, 2010), ensures there are enough segments for robust statistical analysis, influencing the ability to observe long-range correlations and larger structural patterns within the data. Therefore, the maximum scale of the Rest period is set to be 200, as the Rest period has 2000 samples per trial. The MI period has varying signal lengths, which complicates things slightly. One approach could be to have varying maxima for each trial. This would give the most tailored MFDFA analysis for each trial, however, this can lead to inconsistencies when comparing the results across different trials, sessions or participants. As this is one of the main goals of the study, an alternative approach is used, namely to select a common maximal scale for all trials. In this case the common maximal scale is based on the median of signal lengths, which was 3120 samples. Therefore, the maximum scale for all MI trials was set to 312. The downside is that this may limit the ability to detect scale-dependent features for large time periods, as their analysis will be constrained to the smaller maximal scale. Alternatively, if the signal is very short, the estimate might have a higher variance, as there are too few segments to compare for robust statistical analysis (Ihlen & Vereijken, 2010). The median was chosen, as this ensures that most trials are within the correct range of scales, while keeping the scales consistent over all trials to enable comparison between the estimates.

4.4 *Comparing multifractality between Rest and MI*

The differences between the MFDFA metrics during the Rest period and the MI period are compared, between the two performance groups. This is done by comparing the singularity spectra, the curves of Generalized Hurst exponents over different moments of q, and by investigating Tau(q) over q. Moreover, four metrics are investigated to quantize the degree and strength of multifractality. Namely, multifractal width ($\text{MFDFA}_{\text{Width}}$), as characterized by the width of singularity strengths (h), the range of generalized Hurst exponents (delta H(q)), furthermore, possible truncation is examined. This highlights the increase or decrease of multifractal properties that arise during the imagination of a certain direction, by comparing it to the Resting state brain activity. Lastly, the difference in $\text{MFDFA}_{\text{Width}}$ (delta h) is calculated between MI and Rest periods, giving a baselined multifractal width ($\text{MFDFA}_{\text{Width-Dif}}$), which indicates the increase in complexity due to imagining direction. Analyzing these metrics will give an overall insight into the multifractal dynamics of MI compared to Rest, helping to answer RQ1, as well as providing the differences between the performance groups, providing insights into RQ2.

4.5 *Lateralization*

Next, the phenomenon of lateralization in ERD, $\text{MFDFA}_{\text{Width}}$, and $\text{MFDFA}_{\text{Width-Dif}}$ metrics were explored. First, all three metrics were compared between contralateral and ipsilateral channels. For example, if a participant was tasked to imagine left-side movement, the activity from the C4 channel was contralateral, whereas activity from the C3 channel being ipsilateral. To statistically test for differences between contra- and ipsilateral metrics of ERD, and $\text{MFDFA}_{\text{Width-Dif}}$, and $\text{MFDFA}_{\text{Width}}$, 3 bonferonni corrected Wilcoxon signed rank tests were employed, given the non-normally distributed metrics.

After testing for differences in the metrics between hemispheres, the degree and direction of lateralization were examined using the lateralization index. The lateralization index (LI), as described in Doyle et al. (2005). can be quantized as:

$$LI = \frac{(Power_{\text{C3Left}} - Power_{\text{C4Left}}) + (Power_{\text{C4Right}} - Power_{\text{C3Right}})}{2} \qquad (2)$$

Applying the same formulaic structure to $\text{MFDFA}_{\text{Width-Dif}}$, gives us the degree of lateralization of (multifractal) complexity (formula 3). Note that for clarity, $\text{MFDFA}_{\text{Width-Dif}}$, was abbreviated with MFDFA

$$LI_{\text{MFDFA}} = \frac{(MFDFA_{\text{C3Left}} - MFDFA_{\text{C4Left}}) + (MFDFA_{\text{C4Right}} - MFDFA_{\text{C3Right}})}{2} \tag{3}$$

Outliers were identified and removed using the interquartile range (IQR). Specifically, values below Q1 - 1.5IQR or above Q3 + 1.5IQR were considered outliers. Given that the data deviated from a normal distribution, a Kruskal-Wallis test was employed, to determine if there were significant differences in LI's between the different metrics (LI-ERD and LI-MFDFA) and performance conditions (Low and High). For significant Kruskal-Wallis test results, pairwise comparisons were conducted using a Wilcoxon signed-rank test with Bonferroni correction to identify specific group differences. Differences were visually inspected through boxplots, illustrating the distribution of lateralization indices across the different performance groups and metrics. Moreover, both LI, and LI_{MFDFA} for both performance levels were statistically compared to 0, using Bonferroni corrected Wilcoxon signed-rank test, to verify whether the lateralization differed from a bilateral interaction. A metric statistically different from 0 would indicate the presence of a dominating hemisphere.

4.6 *Temporal effects on ERD and MFDFA*

To investigate the temporal effects of both ERD and MFDFA a mix of DFA, ACF, and linear Mixed-effect models were employed. Specifically inter Session, and inter trial dynamics were investigated, providing insights regarding RQ3.

4.6.1 *Session based analysis*

To investigate changes over sessions, six different linear mixed-effect models were employed, each tailored to a specific metric of interest. The metrics analyzed included task accuracy, as well as, $\text{MFDFA}_{\text{Width}}$-, $\text{MFDFA}_{\text{Width-Dif}}$-, ERD-, LI-, and LI_{MFDFA}-values. A random effect is introduced for Subjects, in order to account for inter subject variability. The following linear mixed-effect models were used:

$$\text{ERD} \sim \text{Session} + (1|\text{Subject}) \tag{4}$$

$$\text{MFDFA}_{\text{Width}} \sim \text{Session} + (1|\text{Subject}) \tag{5}$$

$$\text{MFDFA}_{\text{Width}-\text{Dif}} \sim \text{Session} + (1|\text{Subject}) \tag{6}$$

$$\text{LI} \sim \text{Session} + (1|\text{Subject}) \tag{7}$$

$$\text{LI}_{\text{MFDFA}} \sim \text{Session} + (1|\text{Subject}) \tag{8}$$

$$\text{Accuracy} \sim \text{Session} + (1|\text{Subject}) \tag{9}$$

For each linear mixed-effect model, the standard coefficient and p-values were extracted to assess the significance and magnitude of the changes over sessions. The standard coefficient provided a measure of the effect size, while the p-values were used to determine the statistical significance of the observed changes. These values were carefully inspected to draw conclusions about how each metric changes over the course of MI training. One disadvantage of this method is that it provides a linear estimation of change over session, possibly missing nonlinear patterns or more complex temporal dynamics that may occur during MI training.

4.6.2 *Long term dependency in ERD and MFDFA*

Detrended Fluctuation Analysis was used to investigate the long range temporal correlations within sessions. Each trial has a value for ERD or $\text{MFDFA}_{\text{Width}}$ or $\text{MFDFA}_{\text{Width-Dif}}$, which can be interpreted as a timeseries, where each trial is a time point within the analysis (Perquin et al., 2023). By calculating DFA over trials for each session, we can track the long term dependency within each session, and compare them between sessions. For each session, the DFA exponent was computed for the ERD, $\text{MFDFA}_{\text{Width}}$, $\text{MFDFA}_{\text{Width-Dif}}$ metrics. The scales that were used are on a log scale in the range of 4 to 52, as the time series were short (225 samples). A detrending order of 2 was used, based on visual inspection of consistency of scaling exponents derived from the log transformed root mean squares of the detrending functions. For a detrending order of 2, the scaling exponent derived from the log_rms plot was consistent across scales, suggesting that the detrending effectively normalized the data. The DFA exponent extracted from this analysis indicates the presence of LRTC, with values typically between 0.5 (indicating no correlation/random process) and 1.5 (indicating a highly correlated process). DFA values were visually compared between ERD, $\text{MFDFA}_{\text{Width}}$, and $\text{MFDFA}_{\text{Width-Dif}}$, for each session between both high and low performing clusters.

4.6.3 *Autocorrelation*

ACF analysis was utilized to estimate the degree of autocorrelation within the data across different lags, where each lag represents a subsequent trial.

For instance, a lag of 1 indicates the correlation between a metric's value at a given trial and its value at the next trial, while a lag of 2 represents the correlation between the metric's value at a given trial and its value two trials ahead, and so forth. This analysis helps in understanding the persistence or memory in the data by examining how the values of a metric are dependent on previous trials. Significant autocorrelations at specific lags indicate how past trials influence subsequent ones, possibly revealing patterns such as trends or cycles in the data. ACF plots were generated and compared between session for ERD-, $\text{MFDFA}_{\text{Width}}$-, and $\text{MFDFA}_{\text{Width-Dif}}$-metrics.

4.6.4 *Inter-trial variability*

To investigate inter-trial variability, LMMs were employed to analyze how ERD, $\text{MFDFA}_{\text{Width}}$, $\text{MFDFA}_{\text{Width-Dif}}$ change across trials. Moreover, the $\text{MFDFA}_{\text{Width}}$ during rest was investigated to possibly investigate changes in cognitive states such as fatigue. This analysis aimed to understand the fluctuations and patterns within sessions, providing insights into the consistency and variability of these metrics across individual trials. Similar to the Sessional LMM analysis, difference between subjects are accounted for by introducing a Random effect.

$$\text{ERD} \sim \text{Trial} + (1|\text{Subject}) \tag{10}$$

$$\text{MFDFA}_{\text{Width}} \sim \text{Trial} + (1|\text{Subject}) \tag{11}$$

$$\text{MFDFA}_{\text{Width}-\text{Dif}} \sim \text{Trial} + (1|\text{Subject}) \tag{12}$$

$$\text{MFDFA}_{\text{Width}-\text{Rest}} \sim \text{Trial} + (1|\text{Subject}) \tag{13}$$

5 WARNING

The following 2 sections (results and discussion) will likely change for the second submission. There was a mistake in my pipeline which might have had minor to major impact on nearly my entire analysis. I allude to this in my discussion, but have since confirmed the mistake. I'd still appreciate feedback on the structure, and general interpretation of the findings.

6 RESULTS

This study applied MFDFA to EEG signals recorded during MI training to investigate the non-linear properties of these signals. The first research question aimed to understand the multifractal characteristics of EEG signals during MI and how they relate to the underlying neural mechanisms established in the literature. This was achieved by examining the metrics generated by MFDFA and investigating the degree of multifractality, and lateralization in these metrics. Moreover, in line with RQ2, these metrics were compared between two performance clusters, comparing the multifractal properties of MI related brain activity between high and low-performing groups. The third research question focused on the temporal evolution of multifractal properties of EEG signals with motor imagery training over multiple sessions, addressing BCI inefficiency by examining inter-trial and inter-session variability. This included a comparison with traditional linear descriptors to assess the potential advantages of MFDFA in capturing the dynamic changes associated with MI practice. The following sections present the detailed findings for each research question, highlighting the multifractal characteristics, performance-based differences, and temporal evolution of the EEG signals during motor imagery tasks.

6.1 *Multifractal Analysis*

By comparing the multifractal characteristics of EEG signals during rest and MI, specific changes in brain activity associated with the cognitive process of imagining movement can be identified. This comparison helps establish a baseline for multifractal complexity, allowing the unique aspects of MI to be pinpointed and potentially improving the accuracy and efficacy of BCI applications. The differences between Rest and MI were compared for both performance clusters. The singularity spectrum in 2 shows the singularity dimensions D(h) against the singularity strength h for four conditions: MI-successful, MI-unsuccessful, Rest-successful, and Rest-unsuccessful. For the same conditions, figure 1 shows the Generalized Hurst exponent (H(q)) over different statistical moments (q)

6.1.1 *Generalized Hurst exponents*

Figure 1 shows the H(q) values over statistical moments q. The spectrum $\Delta H(q)$ values are 1.491 for MI-unsuccessful and 1.485 for MI-successful, showing close similarity in multifractal behavior with a marginally taller spectrum for the unsuccessful condition. Considering MI-unsuccessful was higher than MI-successful, this indicates that there are stronger multi-

Figure 1: Generalized Hurst exponents over different statistical moments of q. Simlar to the Singularity spectrum, the dark points (Dark pink, and Dark blue), indicate the aggregate of successful trials, wheareas light points (Light pink and Light blue) indicate the aggregate of unsuccessful trials

fractal characteristics during unsuccessful trials. Interestignly, comparing the multifractality strength between clusters in rest show similar results. The spectrum $\Delta H(q)$ values are 0.846 for Rest-unsuccessful and 0.836 for Rest-successful, indicating a marginally higher multifractal strength in the unsuccessful cluster. Overall, the MI conditions exhibit far stronger multifractal characteristics compared to the resting conditions, reflecting more complex underlying dynamics during motor imagery tasks.

Long-term persistence, as described by Hurst exponent H(2), was also extracted from Figure 1. This metric is identical to values found during regular (monofractal) DFA. For the successful MI condition the H(2) exponent was X, while for unsuccessful MI, it was X. Successful and unsuccessful rest conditions had H(2) exponents of X, and X respectively. These values indicate a large degree of long-range correlation. Specifically, the Hurst exponent values greater than 0.5 suggest persistent behavior, meaning that the future trends are likely to follow past trends. It should be noted however, that regularly one does not expect values above 1, indicating that these values are irregularly high. Therefore, these values will not be interpreted as the degree of long term persistence, but rather a relative change of persistence between conditions. The higher Hurst exponents in MI conditions compared to Rest conditions imply stronger

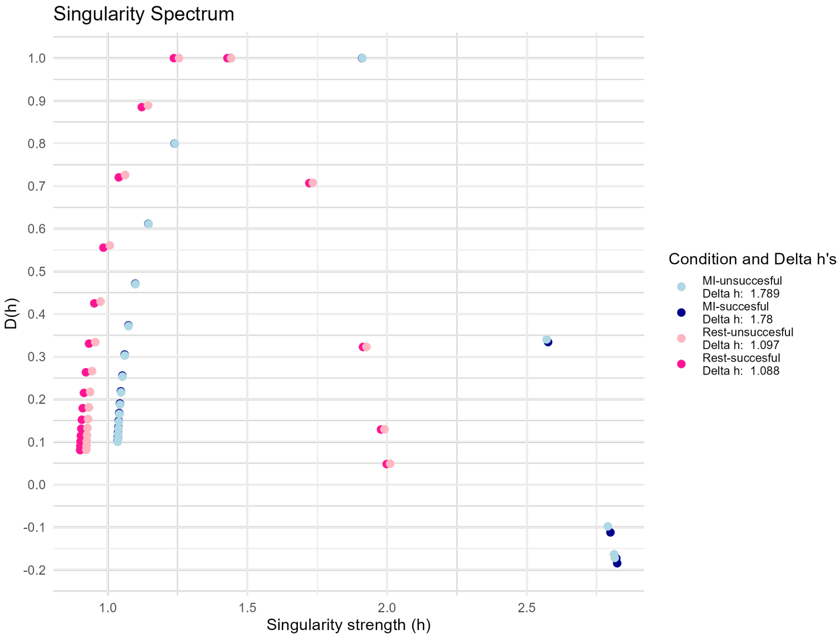

Figure 2: Singularity spectrum of EEG activity during rest (pink) and MI (blue). The dark points (Dark pink, and Dark blue), indicate the aggregate of successful trials, wheareas light points (Light pink and Light blue) indicate the aggregate of unsuccessful trials

long-term persistence, indicating that the MI tasks have a more stable long-term dynamic. Moreover, the successful Rest conditions seem to have slightly higher generalized Hurst exponents, indicating marginally more stable long-term dynamics compared to the unsuccessful rest.

6.1.2 *Singularity Spectrum*

Figure 2 shows the Singularity spectrum. This plot shows the distribution of scaling properties present in the Signal. Higher Signularity strength (h) show more regularity, and big fluctuations, wheareas lower singularity strength indicate the presence of small fluctuations. For both MI conditions, the structure of the singularity spectrum was very similar, indicating a similar evolution law. Furthermore, the singularity spectrum of both MI-successful and MI-unsuccessful exhibit left truncation, indicating that the system is characterized by multifractality with more pronounced smaller fluctuations, and that it is insensitive to local fluctuations with large magnitudes (Ihlen & Vereijken, 2010). This is further evident from the large density of points for relatively low values for h. This indicates that the level of "roughness", or small fluctuations are prominent in the data, while more smooth activity is less prominent (as evident by the low amount of points with high levels of h). Multifractal dimension or D(h) shows for differing levels of h how impactful these fluctuations are for the timeseries. This indicates that while the system seems to be comprised of prominently small fluctuations, slightly larger slightly larger fluctuations seem to have a grander impact on the dyanmics of the system, when these (more rarely) occur. The multifractal width, characterized as Δh is 1.789 for MI-unsuccessful and 1.780 for MI-successful. This indicates that there is a marginally larger variety of scaling exponents for MI-unsuccessful. This indicates that the complexity of MI-unsuccessful is marginally higher. For the resting conditions, the range of Δh is 1.097 for Rest-unsuccessful and 1.088 for Rest-successful, suggesting a less pronounced degree of multifractality compared to the MI conditions, with unsuccessful rest showing a slightly broader range compared to successful rest.

6.2 *Lateralization*

Following section reports on the results of the lateralization analysis, which is a critical aspect of classifying MI. By comparing contralateral and ipsilateral brain activity, as well as examining the LI, the aim is to determine the degree and direction of lateralization. Considering lateralization is such an important aspect of the ERD, investigating $\text{MFDFA}_{\text{Width}}$, and $\text{MFDFA}_{\text{Width-Dif}}$ might provide new insights into the dynamics of MI.

To test whether there were statistical differences in the metrics between contra- and ipsilateral hemispheres 3 bonferonni corrected Wilcoxon signed rank tests were employed. Note that the channel from which the activity was compared varied, as this was based on the direction of the task. For example, the C3 electrode was contralateral for right direction imagination, whereas it was ipsilateral for left direction imagination. The Wilcoxon signed rank test did not show significant differences in MFDFA width values ($W = 4529201518$, p-value = 0.789). Interestingly, $\text{MFDFA}_{\text{Width-Dif}}$ did show a significant effect ($W = 4554746472$, p-value = <0.05). This suggests that when isolating the additional dynamics from MI of that of Rest, there seems evidence for lateralization in complexity.

To further investigate the degree of lateralization, the LI's from Motor Imagery across different performance groups and metrics is illustrated in the two plots presented in Figure **??**. Figure **??** a) shows significant differences between Low and High performance groups in LI-ERD metric ($p < 0.001$, effect size = 0.38). This indicates that high performers exhibit a more negative lateralization index, suggesting pronounced contralateralization, while low performers show values close to zero, indicating bilateral activity. Figure 3 b) shows the LI-MFDFA metric, which encompasses the difference in multifractal width between the hemispheres, which shows

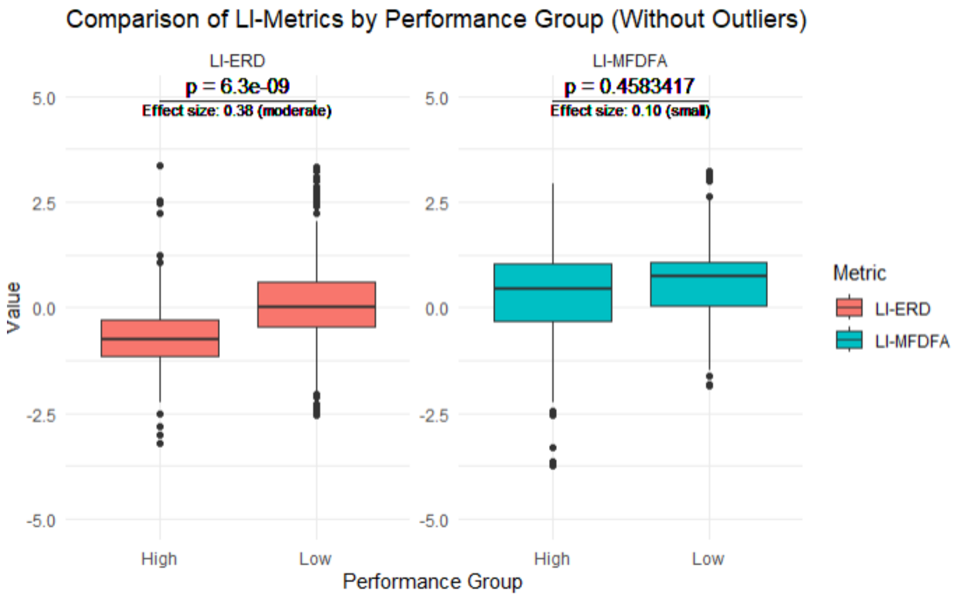

Figure 3: Comparison of high and low clusters between Lateralization Indexes of ERD and $\text{MFDFA}_{\text{Width-Dif}}$, showing contralaterallization happening for high performers for ERD. Moreover, LI-MFDFA shows ipsilateralization for both high and low performers

no significant difference between the performance groups (p = 0.458), suggesting that no evidence for differences in lateralization were found. A Wilcoxon signed rank test confirms that both LI-ERD and LI-MFDFA for high performers, aswell as LI-MFDFA for low performers differ from 0, meaning that they differ from bilateral dominance ($p < 0.001$ and $p < 0.01$, respectively). Contrary to this, LI-ERD for low performers cannot statistically be distinguished from 0, providing no evidence for contra- or ipsilaterally dominated dynamics (p= 0.208).

6.3 *Temporal dynamics*

The sessional analysis, as described in Table 1. revealed a significant effect on the accuracy metric, while other metrics did not show significant changes over sessions. Specifically, the accuracy metric exhibited a significant positive standardized coefficient (0.144) with an estimate of 0.0104 ($p < 0.001$), indicating that participants improved in their ability to generate the right brain activity for MI classification over sessions. In contrast, the other metrics, namely, ERD, $\text{MFDFA}_{\text{Width}}$, $\text{MFDFA}_{\text{Width-Dif}}$, LI, and LI_{MFDFA}, did not show significant (linear) changes over sessions. The standardized coefficients for these metrics were relatively small and their p-values were greater than 0.05, indicating non-significant effects. These findings suggest

Table 1: Summary of Session models

Metric	Estimate	Std. Error	p-value	Standardized Coefficient
ERD	0.0035	0.0032	0.281	0.041
$MFDFA_{Width}$	-0.0005	0.00737	0.483	-0.027
$MFDFA_{Width\text{-}Dif}$	0.0017	0.002159	0.431	0.026
LI	0.0093	0.0340	0.728	0.016
LI_{MFDFA}	0.0267	0.0328	0.417	-0.049
accuracy	0.0104	0.0029	0.0003	0.144

that while participants accuracy improved significantly, metrics describing MI remained relatively stable across training sessions.

6.3.1 *Detrended Fluctuation Analysis*

Figure 4. illustrates the mean DFA scaling exponent values across trials of different sessions for three distinct metrics: ERD, $MFDFA_{Width}$, and $MFDFA_{Width\text{-}Dif}$. The plot shows that the $MFDFA_{Width\text{-}Dif}$ metric (represented in blue) consistently has higher DFA values, indicating a higher degree of long-range temporal correlations in this metric compared to ERD (red) and $MFDFA_{Width}$ (green). The dotted black line indicates the border of alpha = 0.5, indicating no LTC. The ERD metric exhibits the lowest DFA values, even showing anticorrelated dynamics for low performers. This suggests that increases in the ERD early on indicate decreases in the ERD later on in the session or vice versa. Moreover, the ERD seems to be hovering around an DFA exponent of 0.5, indicating little to no temporal structure.

6.3.2 *Autocorrelation*

Figure 5. shows the autocorrelation plots generated over trials. The average autocorrelation value over all participants are plotted over Lags for each session. Lag 0 is not plotted, was not plotted, as this is always 1, since all values correlate fully with itself. The dotted red line indicates the confidence interval to determine whether the autocorrelation is different from 0. The ACF plot showing ERD values shows that at low lags there is a small amount of autocorrelation , which drops under the significance value for most session around a lag of 18. This indicates that there is low correlation up to 18 trials after the recorded trial. Interestingly, at lag 1, the ERD metric is anti correlated with itself. This indicates that when ERD values are high, the next trial likely has a lower ERD value, and vice versa. ACF values from $MFDFA_{Width}$ however, show that most values are indistinguishable from 0 autocorrelation. It does seem like

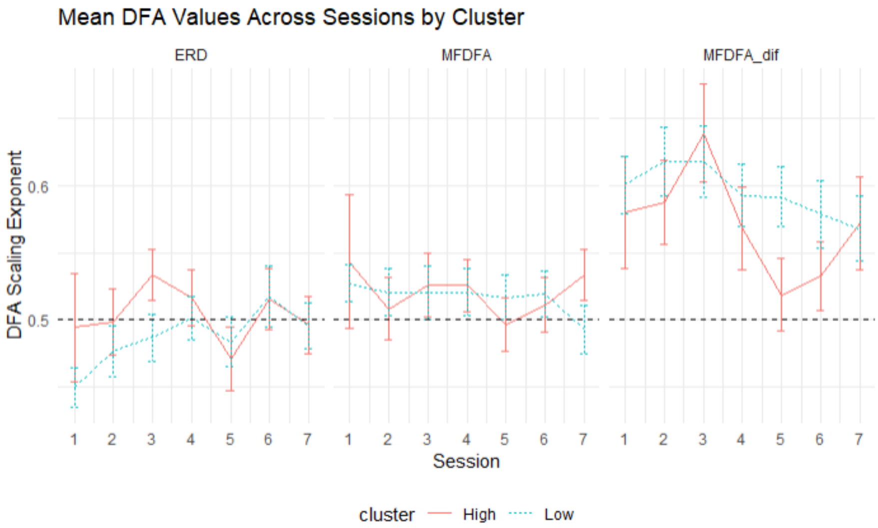

Figure 4: DFA results over sessions averaged over all participants. ERD shows little to no LTC, even indicating anticorrelation for low participants in early sessions. This seems to become more stable for higher sessions. $\text{MFDFA}_{\text{Width}}$ seems to have a low degree of LTC, which is fairly stable over sessions. On the contrary, $\text{MFDFA}_{\text{Width-Dif}}$ seems to start of with a scaling exponent of around 0.6, slowly decreasing over sessions.

there is very little (but significant) autocorrelation values for Session 1, 6, and 7. It should be noted however, that for both ERD and $\text{MFDFA}_{\text{Width}}$ these autocorrelation values are very low, indicating a high degree of between trial variability. $\text{MFDFA}_{\text{Width-Dif}}$ seems to show clearer signs of autocorrelation. Interestingly, for a low number of lags (up to 8), lower session numbers seem to have higher autocorrelation values.

6.3.3 *Inter-trial variability*

The analysis of changes over trials as described in 2 revealed significant effects with $p<0.001$ on all measurements, with notable trends observed in the standardized coefficients. The ERD metric demonstrated a negative standardized coefficient (-0.047), with an estimate of -0.00003, indicating a decrease in ERD values over trials. Similarly, $\text{MFDFA}_{\text{Width}}$, and $\text{MFDFA}_{\text{Width-Dif}}$ exhibited a small negative standardized coefficient (-0.031 and -0.035 respectively), suggesting a slight decrease in multifractal complexity. Conversely, the multifractal complexity during rest periods showed a positive standardized coefficient (0.020), indicating an increase of complexity over trials.

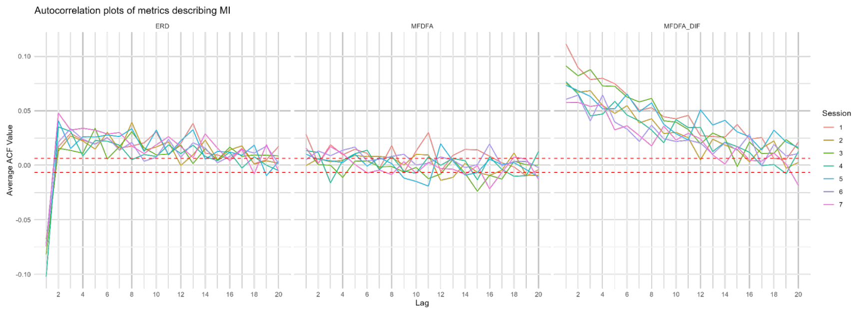

Figure 5: ACF plots for ERD, $\text{MFDFA}_{\text{Width}}$, and $\text{MFDFA}_{\text{Width-Dif}}$ showing the degree of autocorrelation over different lags (trials)

Table 2: Summary of Trial models

Metric	Estimate	Std. Error	p-value	Standardized Coefficient
ERD	-0.000266	0.000018	< 0.001	-0.047273
$\text{MFDFA}_{\text{Width}}$	-0.000099	0.000010	< 0.001	-0.031431
$\text{MFDFA}_{\text{Width-Rest}}$	0.000085	0.000013	< 0.001	0.019600
$\text{MFDFA}_{\text{Width-Dif}}$	-0.000184	0.000017	< 0.001	-0.034878

7 DISCUSSION

A broader multifractal spectrum was observed during MI compared to rest. This indicates that the neural activity involved in motor imagery has an observable broader range of scaling exponents describing the signal. Moreover, the range of Generalized Hurst exponents was also larger, indicating stronger multifractal characteristics. The observable differences between MI and Rest conditions are not very surprising. During MI, participants mentally rehearse movements, which requires the engagement of various neural networks and cognitive processes; these networks likely increase the complexity and variation in fluctuations within the EEG signal. These findings, as well as the significant differences between multifractal widths of Rest and MI indicate that complexity could potentially be used to differentiate between these mental states.

A more interesting observation is the truncation that appears in MI compared to the rest condition. This indicates that the brain activity during MI is more characterized by multifractality with more pronounced smaller fluctuations, compared to a more even split of small and large fluctuations. This is an interesting finding, as this suggests that the small fluctuations dominate the MI interaction. This might be un-intuitive, as we characterize MI as an ERD/ERS, however, these results suggest, it is not the large changes in activity that dominate the interaction. This could suggest that small fluctuations cause the ERD/ERS that follows imagination. Interestingly, the D(h) values of more stable, larger fluctuations, are considerably higher. This suggests that while these specific scaling behaviors are less frequent, they form significant and substantial structures within the data, which is perfectly in line with our understanding of (de)-synchronization.

7.1 *High Hurst exponents*

When investigating the Hurst exponent for H(2), an unexpectedly high Hurst exponent of 1.2 was found, which is surprising given the theoretical expectations. Hurst exponents greater than 0.5 typically indicate persistent behavior, while values exceeding 1 suggest an extreme regularity that seems implausible for the inherently irregular nature of EEG data. This observation aligns with other studies utilizing multifractal analysis, which also report similarly high Hurst exponents. Several hypotheses were considered to investigate potential causes for these strange results. Firstly, the possibility that the high Hurst exponents could be an artifact of the multifractal analysis method itself was explored. However, upon examining regular DFA values for randomly selected participants, values exceeding 1 were found, suggesting that the anomaly is not confined to the multifractal approach. Another hypothesis was that squaring the data might induce more regularity, thus inflating the Hurst exponents. To test this, time-frequency methods using wavelets were employed to estimate power across different frequency bins. Despite this methodological shift, the high Hurst exponents persisted, indicating that the anomaly was not mitigated by this alternative analysis. In reviewing existing literature, no conclusive explanations for these high Hurst exponent values were found within the scope of this investigation. While it is possible that there are underlying reasons for these observations that were not addressed in this study, no definitive cause was identified in the papers examined. When interpreting these findings, caution is advised. Although comparisons between conditions may provide insights, the unusually high Hurst exponents should be carefully considered. The potential for artifacts or methodological influences cannot be ruled out, and these factors may impact the reliability of the interpretations. While the Hurst exponent suggests extreme long term correlation, it might still be possible to compare the degree of LTC between conditions. Comparing Rest and MI states shows that there seems to be more LTC in MI tasks. This is somewhat surprising, considering that we are not expecting changes in mental processes, and thus expect

One hypothesis, which builds upon the findings of the singularity spectrum, could be that during the MI task, there are a lot of differences in small fluctuations, which forms large stable trends (ERD/ERS). In other words, the overall trends become more stable, as a result of small groups of neurons (de)-synchronizing. This highlights the interactions between local fluctuations and global patterns, which might be a critical factor in understanding (successful) motor imagery.Further investigating the effects that small fluctuations have on large fluctuations could potentially provide a biomarker for distinguishing motor imagery more successfully.

7.2 *Lateralization of the ERD*

As expected, the contralateral hemisphere shows greater desynchronization during motor imagery, as indicated by a more negative LI, particularly in high performers. This aligns with the well-established notion that motor imagery of a specific hand induces more significant changes in the EEG power spectrum in the hemisphere opposite to the imagined movement (Neuper & Pfurtscheller, 2001). High performers exhibit this contralateral dominance more distinctly, which likely contributes to their higher classification accuracy using ERD-based classifiers.

Low performers, on the other hand, do not show significant contralateral dominance, which might explain their poorer classification outcomes. Sessions that were deemed low performers were showing an LI-ERD around 0, indicating no clear contra- or ipsilateral dominated power. These findings are congruent with current literature, and shows that classifiers based on ERD/ERS need clear lateralized neural signals for effective MI classification (Nam et al., 2011). Considering that the majority of sessions were classified as a low performer, as described in section X, and failed to show significant contralateralization, it further shows the detriment of BCI inefficiency. This lack of distinct lateralization could reflect less efficient neural engagement or a less clear representation of the imagined movement in the brain (Zhang et al., 2020). It should be noted however, that both accuracy as the LI-ERD are aggregated over the entire trial. This indicates that participants in the low performing group might sometimes be able to generate the correct contralateralization, whereas in other trials, they might not be able to do this.

7.3 *Lateralization of Multifractality and Neural Complexity*

Interestingly, the analysis of multifractal width reveals a different pattern. While ERD is predominantly contralateral, the increase in multifractal width during MI is more ipsilateral. This suggests that the ipsilateral hemisphere experiences a greater increase in signal complexity during MI. This signal complexity, as described by a wider range of scaling exponents suggests more pronounced and varied multifractal behavior in the ipsilateral hemisphere. One possible explanation for this ipsilateral complexity increase could be related to the nature of multifractality, which quantifies how fluctuations at different scales influence the overall signal. During MI, while the contralateral hemisphere shows a clear change in power due to active engagement in the imagined task, the ipsilateral hemisphere might be involved in integrative processes that do not necessarily result in power changes but rather in an increase in signal variability and complexity. This could involve processes such as inhibitory control, sensory integration, and the modulation of neural networks to maintain coordination during the imagined movement (Bassett & Sporns, 2017). An alternative hypothesis could be that during (de-)synchronization of the motor cortex, the complexity goes down on the contralateral hemisphere, making the ipsilateral hemisphere appear dominant according to the lateralization index. Essentially, as neurons in the contralateral hemisphere become more (de-)synchronized in response to motor imagery, the variability and irregularity of the EEG signal decrease, leading to a narrower multifractal spectrum. This hypothesis hinges on the behavior of neural synchronization and desynchronization during motor imagery. MI involves the coordinated firing of large groups of neurons, which is essential for simulating the motor plan. Desynchronization can occur due to the recruitment of neurons that fire in a more coordinated and rhythmic manner, reducing the overall variability of the signal. Future research could investigate, and compare the full multifractal spectrum for contra- and ipsilateral hemispheres. Specifically, looking into the central tendencies of singularities might give insights into whether the contralateral hemisphere is becoming more coordinated, and thus less complex. Moreover, investigating signals at more electrodes over the SMR region might provide nuances that are not available through the limited analysis that was applied. Another difference between the lateralization indexes for ERD's and MFDFA widths is that contrary to the LI-ERD metric, LI-MFDFA did not display significant differences between performance groups (Figure X.b). This might suggest that while participants in low-performing sessions lack a clear contralateralization in terms of alpha/beta power, they may still be modulating their brain activity effectively. In other words, the stable multifractal properties across different performance levels imply that underlying neural dynamics during motor imagery are maintained, even if not distinctly lateralized in power. Therefore, incorporating MFDFA metrics into classifiers could help address inter-trial variability by leveraging the nuanced and stable complexity measures, potentially improving classification accuracy for users with less pronounced ERD patterns. These findings align with the branch of literature that attributes BCI inefficiency to the limitations of the classifier rather than the participants' ability to modulate their brain activity, while still accounting for previous reports indicating that users may not be correctly performing a contralateral ERD/ERS. Furthermore, it suggests that MFDFA widths could potentially be used in classifiers to combat BCI inefficiency.

7.4 *Temporal Analysis*

The sessional analysis revealed a significant improvement in the accuracy metric for MI classification, with participants showing increased ability to generate the correct brain activity over the course of 7 sessions of MI training, as evidenced by a positive standardized coefficient of 0.1438 ($p < 0.001$). This suggests learning took place. In contrast, ERD values did not significantly change over sessions. This is contradictory compared to other literature on MI improvement. It is possible that the statistical analysis wasn't granular enough. The average difference in power between MI and baseline was considered, meaning that the entire MI period was condensed into one average. It is likely that non-ERD/S activity contributed to lowering the average. Moreover, for this specific analysis, the average over all trials was taken, small differences might not be significant due to small sample size.

The sessional analysis revealed a significant improvement in the accuracy metric for MI classification, with participants showing increased ability to generate the correct brain activity over the course of 7 sessions of MI training, as evidenced by a positive standardized coefficient of 0.1438 ($p < 0.001$). This suggests learning took place. In contrast, ERD values did not significantly change over sessions. This is contradictory compared to other literature on MI improvement. It is possible that this is due to the way the ERD was quantified. The average difference in power between MI and baseline was considered, meaning that the entire MI period was condensed into one average, which likely captures non-ERD/S activity. The small positive effect observed in the difference in multifractal width (mfdfa_dif_values) between rest and MI periods over sessions suggest a gradual increase in the scaling properties induced by the MI task. This implies that as participants practice MI tasks, the neural activity during MI might become increasingly distinct from the resting state in terms of its scaling behavior. The stability of overall multifractal width during MI (mfdfa_values) could be interpreted as the overall complexity in the SMR regions not changing. This potentially highlights an important takeaway when working with multifractal width, namely, that this metric captures the overall scaling properties of the entire time series. In other words, it includes interactions between fluctuations that might reflect noise or background activity from the brain. This suggests that if we wish to extract the largest amount of information from this metric, it is useful to first compare it to a baseline resting state, similar to how we treat raw data. When further interpreting this difference between complexity during MI and the difference between MI and Rest, we could interpret this through the lens of a combination of distributed resource theory and Hebbian learning. Distributed resource theory posits that cognitive processes, including neural adaptations, are not isolated but distributed across various neural and external resources. This framework can help explain why the overall multifractal width remains stable; the brain might be reallocating its resources more efficiently without increasing overall complexity. Moreover, it would be in line with the priming hypothesis as described in Section X. Hebbian learning, which is often summarized as "cells that fire together wire together," suggests that repeated MI practice strengthens specific neural pathways, enhancing their efficiency and specificity. These strengthened connections could lead to improved performance (as indicated by increased accuracy) and more distinct scaling properties during MI (reflected in $\text{MFDFA}_{\text{Width-Dif}}$), even if the overall $\text{MFDFA}_{\text{Width}}$ remains unchanged. Future research should aim to confirm these findings of difference in scaling properties during and between MI and Rest, as it could provide deeper insights into the distribution of resources by the brain, and could have implications for cognitive load, as well as learning. The LIs for both power and multifractal width did not show significant changes over the course of the MI training sessions. This stability suggests that the patterns of hemispheric engagement seem to remain consistent throughout the training period. One possible reason for the lack of change in LIs could be that the brain establishes an effective neural strategy involving lateralized activity early in the training, which remains sufficient for the task at hand. This would have rather large implications however, as accuracy increases over sessions. It would indicate that the improvements in MI task performance are due to more efficient or refined neural processing within the established lateralization pattern rather than changes in the degree of lateralization itself. Perhaps a more likely interpretation could be that the dataset was not large enough or sensitive enough to detect subtle changes in lateralization over time. Especially since the LI data is a singular metric per session, which might not be enough data for statistically showing small effects. Future research with larger datasets or more sensitive measures might be needed to detect any finer changes in lateralization that could occur with prolonged MI practice. Most of the findings indicate that improvements in motor imagery MI training occur slowly and incrementally over many sessions. This gradual progression highlights the slow pace of neural adaptations associated with MI, suggesting that extended practice is necessary for significant changes to become evident in measurable metrics, such as lateralization. Consequently, more sessions may be required to observe substantial improvements, underscoring the need for prolonged and consistent training to achieve meaningful advancements in motor imagery skills.

7.5 *Temporal dependency*

The findings from the DFA and ACF analyses indicate that $\text{MFDFA}_{\text{Width-Dif}}$ consistently exhibits higher temporal dependency compared to ERD and $\text{MFDFA}_{\text{Width}}$. This both suggests that trials in close succession for $\text{MFDFA}_{\text{Width-Dif}}$ are more similar, while also reflecting stronger LTC. Conversely, ERD displays some degree of anticorrelation in both ACF and DFA metrics, particularly noticeable at lag 1 for ACF, implying that high ERD values are likely followed by lower values in the next trial/future trials, indicating a dynamic adjustment mechanism where immediate successive trials differ significantly. Interestingly, the decrease in LTC for high performers in later sessions is a noteworthy and somewhat unexpected finding. One possible explanation could be that high performers, having achieved proficiency in the task, are more capable of adjusting their brain activity, modulating complexity according to the task needs. Literature on temporal dependencies suggests that these measures could potentially reflect various aspects of cognitive and behavioral processes, such as mental flexibility, attentional states, and task performance. For instance, higher temporal dependencies might indicate less mental flexibility and a more rigid adherence to a consistent action throughout a task (Perquin et al., 2023). It would be interesting to investigate whether DFA over $\text{MFDFA}_{\text{Width-Dif}}$ could be a neural correlate for this phenomena, and whether sessional improvement has an effect on this. It should be noted however that neither ERD metrics or Multifractal Width seem to corroborate these findings, showing little to no temporal dependency.

7.6 *inter trial variability*

The observed differences in ERD, as described in Section X, are somewhat surprising, indicating that power differences between MI and rest become less pronounced with an increasing number of trials. This could suggest changes in cognitive states such as fatigue or diminishing levels of concentration in participants. This finding, however, contradicts the theory of priming in MI. Priming refers to the enhancement of motor system readiness due to repeated mental rehearsal of movements, which typically leads to more efficient neural processing over trials. This repeated exposure is expected to reduce the effort required to engage relevant neural circuits, typically making it easier to produce larger ERDs. An alternative explanation for the behavior of ERDs could be related to the processing of the ERD/S metric. One methodological choice that was made was to square the signal to extract power from the raw signal. Squaring the signal converts negative voltage in the EEG signal into positive power, allowing for this metric of power to describe both desynchronization (negative) and synchronization (positive). This conversion is only accurate when the signal is centered around zero. This is usually performed by subtracting the baseline from the signal. However, if the baseline fails to center the data around zero, it might lead to deviations, where positive values are more largely increased, and where small values (i.e. values that are normally below zero), are not equally increased. Since both MI and Rest data were squared before this baselining, it is a feasible explanation as to why some participants exhibit a positive ERD while others show negative ERDs. A revision to the pipeline is advised, where ERD/ERS is first calculated for each timepoint, and then squared. Note that the results from the lateralization were only marginally influenced by this, as the same margin of error would have been included for both hemispheres, showing similar degrees of lateralization. This hypothesis of an imperfect ERD/ERS calculation is further supported by the contradicting findings of Multifractal complexity. According to the results in section X, there is evidence suggesting a decrease of multifractal complexity over trials, potentially indicating more efficient use of motor units. Interestingly, the multifractal complexity during Rest seems to increase, indicating that the "Background" activity during rest becomes more complex. It is possible that this is the effect of priming on the background activity of the brain. It is possible that complexity increases as a consequence of the brain anticipating a next MI trial. Sessions

8 CONCLUSION

The findings of this research offer a fresh perspective on the brain dynamics underlying MI, particularly through the lens of MFDFA. By investigating the multifractal characteristics of EEG signals during MI, this study has highlighted the complex and varied nature of neural activity associated with motor imagery. Moreover, it provides new insights that could potentially be used to combat BCI inefficiency.

One of the pivotal findings is the significant ipsilaterally dominated lateralization in MFDFA widths. It's especially promising since the performance did not seem to have an impact on this dynamic, suggesting that MFDFA metrics can be used to combat BCI inefficiency.

Interestingly, the analysis revealed that during MI, smaller fluctuations in neural activity are more prominent than larger ones, however, H(q) values were higher for less prominent fluctuations. This is might suggest that small fluctuations drive ERD/ERS formation, possibly hinting at the possibility of using the relation of small and large fluctuations as new neurocorrelates.

Furthermore, the temporal analysis of MI training sessions demonstrated that participants improve their ability to generate the correct brain activity over time, although very slowly. The stability of multifractal width across sessions indicates that while overall neural complexity remains consistent, the specific scaling properties induced by MI become more distinct with practice. This finding suggests that MFDFA widths could provide a more neural correlate for adaptation and learning during MI training, offering valuable insights for enhancing BCI training protocols.

In conclusion, this study highlights the potential of multifractal analysis as a powerful tool for understanding and enhancing MI-based BCIs. By capturing the complex, non-linear dynamics of brain activity, MFDFA metrics offer a richer and more detailed view of the neural processes underlying motor imagery. These findings not only contribute to the theoretical understanding of MI but also have practical implications for the design of more effective and reliable BCIs. Integrating multifractal analysis with traditional methods could lead to significant improvements to MI classification, ultimately improving the quality of life for individuals with severe motor impairments.

REFERENCES

Aftanas, L., & Golocheikine, S. (2002). Non-linear dynamic complexity of the human EEG during meditation. , *330*(2), 143–146. Retrieved 2024-05-20, from https://linkinghub.elsevier.com/retrieve/pii/S0304394002007450 doi: 10.1016/S0304-3940(02)00745-0

Bashashati, A., Fatourechi, M., Ward, R. K., & Birch, G. E. (2007). A survey of signal processing algorithms in brain–computer interfaces based on electrical brain signals. , *4*(2), R32–R57. Retrieved 2024-05-17, from https://iopscience.iop.org/article/10.1088/1741-2560/4/2/R03 doi: 10.1088/1741-2560/4/2/R03

Bassett, D. S., & Sporns, O. (2017). Network neuroscience. *Nature neuroscience*, *20*(3), 353–364.

Bovend'Eerdt, T. J., Dawes, H., Sackley, C., Izadi, H., & Wade, D. T. (2010).

An integrated motor imagery program to improve functional task performance in neurorehabilitation: a single-blind randomized controlled trial. *Archives of physical medicine and rehabilitation*, *91*(6), 939–946.

Brodu, N., Lotte, F., & Lécuyer, A. (2012). Exploring two novel features for EEG-based brain–computer interfaces: Multifractal cumulants and predictive complexity. , *79*, 87–94. Retrieved 2024-02-26, from https://www.sciencedirect.com/science/article/pii/S0925231211006291 (Publisher: Elsevier)

Doyle, L. M., Yarrow, K., & Brown, P. (2005). Lateralization of event-related beta desynchronization in the EEG during pre-cued reaction time tasks. , *116*(8), 1879–1888. Retrieved from https://linkinghub.elsevier.com/retrieve/pii/S1388245705001458 doi: 10.1016/j.clinph.2005.03.017

Gaurav, G., Anand, R. S., & Kumar, V. (2021). EEG based cognitive task classification using multifractal detrended fluctuation analysis. , *15*(6), 999–1013. Retrieved 2024-02-12, from https://link.springer.com/10.1007/s11571-021-09684-z doi: 10.1007/s11571-021-09684-z

Geronimo, A., & Simmons, Z. (2020). Telebci: remote user training, monitoring, and communication with an evoked-potential brain-computer interface. *Brain-Computer Interfaces*, *7*(3-4), 57–69.

Gu, G.-F., & Zhou, W.-X. (2006). Detrended fluctuation analysis for fractals and multifractals in higher dimensions. *Physical Review E*, *74*(6), 061104.

Hardstone, R., Poil, S.-S., Schiavone, G., Jansen, R., Nikulin, V. V., Mansvelder, H. D., & Linkenkaer-Hansen, K. (2012). Detrended fluctuation analysis: a scale-free view on neuronal oscillations. *Frontiers in physiology*, *3*, 23105.

Huang, G., Zhao, Z., Zhang, S., Hu, Z., Fan, J., Fu, M., ... Dan, G. (2023). Discrepancy between inter-and intra-subject variability in EEG-based motor imagery brain-computer interface: Evidence from multiple perspectives. , *17*, 1122661. Retrieved 2024-02-12, from https://www.frontiersin.org/articles/10.3389/fnins.2023.1122661/full (Publisher: Frontiers)

Ihlen, E. A., & Vereijken, B. (2010). Interaction-dominant dynamics in human cognition: Beyond $1/f$ fluctuation. , *139*(3), 436. Retrieved 2024-02-26, from https://psycnet.apa.org/journals/xge/139/3/436/ (Publisher: American Psychological Association)

Kelty-Stephen, D. G., Lane, E., Bloomfield, L., & Mangalam, M. (2022). Multifractal test for nonlinearity of interactions across scales in time series. , *55*(5), 2249–2282. Retrieved 2024-02-11, from https://link.springer.com/10.3758/s13428-022-01866-9 doi: 10.3758/s13428

-022-01866-9

Kelty-Stephen, D. G., Palatinus, K., Saltzman, E., & Dixon, J. A. (2013). A tutorial on multifractality, cascades, and interactivity for empirical time series in ecological science. , *25*(1), 1–62. Retrieved 2024-02-11, from http://www.tandfonline.com/doi/abs/10.1080/10407413.2013.753804 doi: 10.1080/10407413.2013.753804

Lee, S., Lee, C.-H., Kim, H., & Kim, D.-j. (2020). Lateralization of alpha oscillation under preparation lead to efficiency of motor imagery: Related with performance of classification. In *2020 IEEE international conference on systems, man, and cybernetics (SMC)* (pp. 2502–2505). IEEE. Retrieved from https://ieeexplore.ieee.org/document/9283439/ doi: 10.1109/SMC42975.2020.9283439

Lotze, M., & Halsband, U. (2006). Motor imagery. , *99*(4), 386–395. Retrieved from https://www.sciencedirect.com/science/article/pii/S0928425706000210 (Publisher: Elsevier)

Nakayashiki, K., Saeki, M., Takata, Y., Hayashi, Y., & Kondo, T. (2014). Modulation of event-related desynchronization during kinematic and kinetic hand movements. , *11*(1), 90. Retrieved 2024-02-11, from https://jneuroengrehab.biomedcentral.com/articles/10.1186/1743-0003-11-90 doi: 10.1186/1743-0003-11-90

Nam, C. S., Jeon, Y., Kim, Y.-J., Lee, I., & Park, K. (2011). Movement imagery-related lateralization of event-related (de) synchronization (erd/ers): motor-imagery duration effects. *Clinical Neurophysiology*, *122*(3), 567–577.

Neuper, C., & Pfurtscheller, G. (2001). Event-related dynamics of cortical rhythms: frequency-specific features and functional correlates. *International journal of psychophysiology*, *43*(1), 41–58.

Neuper, C., Wörtz, M., & Pfurtscheller, G. (2006). ERD/ERS patterns reflecting sensorimotor activation and deactivation. , *159*, 211–222. Retrieved 2024-02-11, from https://www.sciencedirect.com/science/article/pii/S0079612306590144 (Publisher: Elsevier)

Padfield, N., Zabalza, J., Zhao, H., Masero, V., & Ren, J. (2019). EEG-based brain-computer interfaces using motor-imagery: Techniques and challenges. , *19*(6), 1423. Retrieved 2024-02-26, from https://www.mdpi.com/1424-8220/19/6/1423 (Publisher: MDPI)

Paul, S., Mazumder, A., Ghosh, P., Tibarewala, D., & Vimalarani, G. (2015). Eeg based emotion recognition system using mfdfa as feature extractor. In *2015 international conference on robotics, automation, control and embedded systems (race)* (pp. 1–5).

Perquin, M. N., Van Vugt, M. K., Hedge, C., & Bompas, A. (2023). Temporal structure in sensorimotor variability: A stable trait, but what for? , *6*(3), 400–437. Retrieved 2024-02-12, from https://link.springer

.com/10.1007/s42113-022-00162-1 doi: 10.1007/s42113-022-00162-1

Pfurtscheller, G. (2000). Spatiotemporal ERD/ERS patterns during voluntary movement and motor imagery. , *53*, 196–198. Retrieved 2024, from https://www.sciencedirect.com/science/article/pii/S1567424X09701576 (Publisher: Elsevier)

Pfurtscheller, G., & Da Silva, F. L. (1999). Event-related eeg/meg synchronization and desynchronization: basic principles. *Clinical neurophysiology*, *110*(11), 1842–1857.

Prasad, G., Herman, P., Coyle, D., McDonough, S., & Crosbie, J. (2010). Applying a brain-computer interface to support motor imagery practice in people with stroke for upper limb recovery: a feasibility study. , *7*(1), 60. Retrieved 2024-02-11, from https://jneuroengrehab.biomedcentral.com/articles/10.1186/1743-0003-7-60 doi: 10.1186/1743-0003-7-60

Sheahan, H. R., Ingram, J. N., Žalalytė, G. M., & Wolpert, D. M. (2018). Imagery of movements immediately following performance allows learning of motor skills that interfere. , *8*(1), 14330. Retrieved 2024-02-26, from https://www.nature.com/articles/s41598-018-32606-9 (Publisher: Nature Publishing Group UK London)

Sikdar, D., Roy, R., & Mahadevappa, M. (2018). Epilepsy and seizure characterisation by multifractal analysis of eeg subbands. *Biomedical Signal Processing and Control*, *41*, 264–270.

Stieger, J. R., Engel, S., Jiang, H., Cline, C. C., Kreitzer, M. J., & He, B. (2020). Mindfulness improves brain–computer interface performance by increasing control over neural activity in the alpha band. , *31*(1), 426–438. Retrieved from https://doi.org/10.1093/cercor/bhaa234 (_eprint: https://academic.oup.com/cercor/article-pdf/31/1/426/34841148/bhaa234.pdf) doi: 10.1093/cercor/bhaa234

Tariq, M., Trivailo, P. M., & Simic, M. (2018). Eeg-based bci control schemes for lower-limb assistive-robots. *Frontiers in human neuroscience*, *12*, 312.

Wriessnegger, S. C., Müller-Putz, G. R., Brunner, C., & Sburlea, A. I. (2020). Inter-and intra-individual variability in brain oscillations during sports motor imagery. , *14*, 576241. Retrieved 2024, from https://www.frontiersin.org/articles/10.3389/fnhum.2020.576241/full (Publisher: Frontiers Media SA)

Xu, T. L., De Barbaro, K., Abney, D. H., & Cox, R. F. (2020). Finding structure in time: Visualizing and analyzing behavioral time series. *Frontiers in Psychology*, *11*, 521451.

Zhang, R., Li, F., Zhang, T., Yao, D., & Xu, P. (2020). Subject inefficiency phenomenon of motor imagery brain-computer interface: Influence factors and potential solutions. , *6*(3), 224–241. Retrieved 2024-02-26, from http://journals.sagepub.com/doi/10.26599/BSA.2020.9050021 doi: 10.26599/BSA.2020.9050021