DECODING MULTI-CHANNEL EEG SIGNALS OF UPPER LIMB MOTOR IMAGERY IN RIEMANNIAN SPACE

by

Mohammed Jamal Abdalwahid Abdalwahid

B.S., in Nuclear Medical Imaging, National Ribat University, 2013

Submitted to the Institute of Biomedical Engineering
in partial fulfillment of the requirements
for the degree of
Master of Science
in
Biomedical Engineering

Boğaziçi University
2025

DECODING MULTI-CHANNEL EEG SIGNALS OF UPPER LIMB MOTOR IMAGERY IN RIEMANNIAN SPACE

APPROVED BY:

Prof. Dr. Ahmet Ademoğlu .....
(Thesis Advisor)

Assist. Prof. Dr. Pınar Senay Özbay .....

Assist. Prof. Dr. İpek Şen .....

DATE OF APPROVAL: 25 August 2025

To my Family

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my supervisor, Prof. Dr. Ahmet Ademoğlu, for his steadfast guidance, constructive feedback, and unwavering support throughout this research. His clarity of thought, critical insight, and high academic standards have continually challenged me to grow and develop as a researcher. Sir, you have been a beacon of knowledge.

I especially value the autonomy I was given to explore ideas independently, balanced with the reassurance that thoughtful advice was always available when needed. This combination of freedom and mentorship greatly enriched my learning experience.

I am also deeply grateful for the time and effort he dedicated to reviewing my work and providing detailed comments, even amidst a demanding schedule. For this and much more, I am truly thankful.

I would also like to extend my sincere thanks to Prof. Dr. Esin Öztürk Işık for providing the computational resources that made this research possible. As she once remarked, our lab is full of enjoyment, Hocam, you made it even more so with your generosity and support.

Finally, I wish to acknowledge the many friends I met during this chapter of my life, especially my coffee-cup companions and lab mates: Hüden Neşe for making the lab an inspiring and enjoyable place to work, and for Belal Tavashi as a good friend, fellow coffee-cup companion, and thesis-writing partner, for their camaraderie and encouragement.

ACADEMIC ETHICS AND INTEGRITY STATEMENT

I Mohammed Jamal Abdalwahid Abdalwahid hereby certify that I am aware of the Academic Ethics and Integrity Policy issued by the Council of Higher Education (YÖK) and I fully acknowledge all the consequences due to its violation by plagiarism or any other way.

Name : Signature:

Date:

ABSTRACT

DECODING MULTI-CHANNEL EEG SIGNALS OF UPPER LIMB MOTOR IMAGERY IN RIEMANNIAN SPACE

Decoding motor imagery (MI) from electroencephalography (EEG) is essential for brain-computer interface (BCI) applications. This study investigates the use of movement-related cortical potentials (MRCPs) enhanced with the xDAWN spatial filter and classified using Riemannian geometry-based methods: Minimum Distance to Mean (MDM) and Tangent Space Mapping combined with Linear Discriminant Analysis (TSM-LDA). Within-subject and cross-subject scenarios were evaluated for both binary (movement versus rest) and multiclass tasks, examining the effects of analysis window length (100 ms–1 s) and frequency bands (0.3–3 Hz and 7–30 Hz). Results indicate that within-subject binary decoding achieved up to 71% accuracy, with peak discriminative activity occurring approximately 1.2–1.3 seconds after cue onset, while cross-subject binary decoding reached around 67% accuracy with peak latency approximately 1.3–1.4 seconds. Multiclass decoding achieved 22–24% accuracy within-subject, dropping to approximately 20% in the cross-subject evaluation. Longer temporal windows consistently improved performance, and xDAWN filtering enhanced the signal-to-noise ratio. Confusion matrix analysis revealed that most misclassifications occurred between similar movement types. These findings demonstrate the potential of MRCP–based Riemannian decoding for neuroprosthetic control and highlight the importance of advanced feature extraction methods to enable robust, calibration-free BCI systems.

Keywords: Motor Imagery (MI), Electroencephalography (EEG), Brain-Computer Interface (BCI), Movement-Related Cortical Potentials (MRCPs), xDAWN, Riemannian, Minimum Distance to Mean (MDM), Tangent Space Mapping (TSM).

ÖZET

ÜST UZUVLARDA MOTOR HAREKET İMGELEMENİN RIEMANN UZAYINDA ÇOK KANALLI EEG SİNYALLERİ İLE İNCELENMESİ

Motor hareket imgelemeyi (MI) elektroensefalografi (EEG) ile çözümlemek, beyin-bilgisayar arayüzü (BCI) uygulamaları için temel bir adımdır. Bu çalışma, xDAWN uzamsal filtresi ile zenginleştirilmiş hareketle ilişkili kortikal potansiyeller (MRCP) kullanılarak, Ortalamaya En Yakın Mesafe (MDM) ve Teğet Uzaya Haritalama ile birleştirilmiş Doğrusal Ayrıştırım İncelemesi (TSM-LDA) olarak bilinen Riemann geometrisine dayalı yöntemlerle sınıflandırılmasını incelemektedir. Hem tek katılımcı (within-subject) hem de katılımcılar arası (cross-subject) senaryolar için ikili (hareket vs. dinlenme) ve çoklu sınıf görevleri için değerlendirilmiş ve analiz pencere uzunluğu (100 ms–1 s) ile frekans bantlarının (0.3–3 Hz ve 7–30 Hz) etkileri incelenmiştir. Sonuçlar, denek-içi ikili sınıflandırmada %71'e kadar doğruluk elde edildiğini gösterirken, en yüksek ayrıcı aktivitenin olarak komut başlangıcından 1.2–1.3 saniye sonra meydana geldiğini, denekler-arası ikili sınıflandırmada ise doğruluğun yaklaşık %67 ve en yüksek gecikme 1.3–1.4 saniye olduğunu ortaya koymaktadır. Çoklu sınıflandırmada denek-içi doğruluk 22–24% iken, denekler-arası değerlendirmede yaklaşık %20 seviyesine düşmektedir. Daha uzun zaman pencereleri performansı tutarlı bir şekilde iyileştirmiş ve xDAWN süzgeci sinyal-gürültü oranını artırmıştır. Karıştırma matrisi analizi, en hatalı sınıflandırmaların benzer hareket tipleri arasında meydana geldiğini göstermiştir. Bu bulgular, MRCP tabanlı Riemann çözümlemesinin nöroprotezlerin denetimi alanında kullanım potansiyelini göstermekte ve sağlam, kalibrasyon gerektirmeyen BCI sistemleri için gelişmiş özellik çıkarım yöntemlerinin önemini vurgulamaktadır.

Anahtar Sözcükler: Motor İmgeleme (MI), Elektroensefalografi (EEG), Beyin-Bilgisayar Arayüzü (BCI), Hareketle İlişkili Kortikal Potansiyeller (MRCP), xDAWN, Riemann, Ortalamaya En Yakın Mesafe (MDM), Teğet Uzaya Haritalama (TSM).

TABLE OF CONTENTS

ACKNOWLEDGMENTS iv
ACADEMIC ETHICS AND INTEGRITY STATEMENT v
ABSTRACT vi
ÖZET vii
LIST OF FIGURES xi
LIST OF TABLES xv
LIST OF SYMBOLS xvi
LIST OF ABBREVIATIONS xviii
1. INTRODUCTION 1
1.1 Decoding 1
1.2 Neural Decoding 2
1.3 Problem Statement 3
1.4 Objective of the Study 5
1.5 Research Questions 5
1.6 Organization of the study 6
2. BACKGROUND 7
2.1 Brain 7
2.2 Relevant Neuroanatomy for EEG 7
2.2.1 Cerebral Cortex 7
2.3 Relevant Neurophysiology for EEG 8
2.4 EEG 10
2.4.1 Generation of EEG Signals 10
2.5 EEG Montages 12
2.6 EEG Artifact 13
2.6.1 Artifact Detection and Correction 13
2.7 Motor Imagery 14
2.7.1 Motor Imagery Overview 14
2.7.2 EEG Signatures of Motor Imagery: ERD/ERS 15
2.7.3 Movement-Related Cortical Potentials (MRCPs) 16

2.8 Brain-Computer Interfaces (BCI) 17
2.8.1 Classification in BCI Problems and Requirements for a Modern BCI 17
2.8.2 BCI mode 20
2.8.3 Paradigm 21
2.8.3.1 Motor Imagery (MI) Paradigm 21
2.8.3.2 P300 or Event-Related Potential (ERP) Paradigm . . 21
2.8.3.3 Steady-State Visual Evoked Potential (SSVEP) Paradigm 21
2.8.3.4 Other Paradigms 21
2.8.4 Riemannian Geometry in BCI Classification: 22
2.9 Riemannian Geometry 22
2.9.1 From Motor Imagery EEG to Covariance Matrices 23
2.9.1.1 Symmetric Positive Definite (SPD) Matrices 24
2.9.2 Riemannian Metrics for SPD Matrices 25
2.9.2.1 Affine-Invariant Metric (AIM) 25
2.9.2.2 Log-Euclidean Metric (LEM) 25
2.9.3 xDAWN Spatial Filter 26
2.9.4 Riemannian Metrics and Classification 27
3. MATERIAL AND METHODS 28
3.1 Dataset 28
3.2 Preprocessing 29
3.3 Classification 32
3.3.1 Classification in Riemannian Space 32
3.3.2 Tangent Space 33
3.3.3 Permutation 37
3.3.4 False Discovery Rate (FDR) 38
3.3.5 Evaluation Metrics 38
3.4 Application 40
3.4.1 Binary Task Construction 40
3.4.2 Movement vs. Movement Classification 40
3.4.3 Evaluation Procedure 40
3.4.4 Feature Ablation 41

4. RESULTS 44
4.1 Binary Classification 44
4.1.1 Within-Subject Analysis 44
4.1.2 Feature Ablation 47
4.1.2.1 Effect of Removing xDAWN Spatial Filter 47
4.1.2.2 Feature Ablation in the 7–30 Hz Band 48
4.1.3 Cross-Subject Binary Case 57
4.2 Multiclass Classification 58
5. DISCUSSION 64
6. CONCLUSION 68
6.1 Scope and Limitations 69
6.2 Significance of the Study 70
REFERENCES 72

LIST OF FIGURES

Figure 1.1 Illustration of the difference between symbolic interpretation and literal perception. The ibis, sacred to the god Thoth in ancient Egypt, represented wisdom and knowledge to those familiar with the symbol, but appeared as an ordinary bird to others. 2

Figure 2.1 Brain structure illustration (adapted from [1]). 8

Figure 2.2 Schematic of cortical neurons generating postsynaptic potentials (PSPs) and the resulting signals recorded by a scalp EEG electrode. 10

Figure 2.3 Typical ERD/ERS pattern observed in a healthy individual (adapted from [2]). 15

Figure 2.4 Movement-related cortical potential (MRCP) (adapted from [3]). 16

Figure 2.5 Basic design of a BCI system (adapted from [4]). 18

Figure 2.6 Illustration of the difference between Euclidean and Riemannian distances. On a flat map, distances between two points (cities) are measured as straight lines (Euclidean), whereas on the Earth's curved surface, the shortest path follows the geodesic along the sphere (Riemannian distance). This analogy helps understand distances between SPD matrices in EEG-based BCI classification. 23

Figure 3.1 Grand average of movement-related cortical potentials (MRCPs) recorded from four motor-related EEG channels (FCz, Cz, C3, C4). The signals are stacked for better visualization of waveform patterns. 30

Figure 3.2 Grand average of MRCPs from the same four channels, displayed without stacking to illustrate their relative amplitudes. 31

Figure 3.3 Grand average of ERD/ERS bands from the same four channels. 31

Figure 3.4 Illustration of the Riemannian Minimum Distance to Mean (RMDM) classifier applied to a two-class problem (adapted from [5]). 32

Figure 3.5 Main steps of MDM pipeline. 33

Figure 3.6	Points are mapped to the tangent space at the reference point G. (adapted from [5]).	35
Figure 3.7	Main steps of TSM + LDA pipeline.	36
Figure 3.8	Pipeline on the SPD manifold (adapted from [6]).	36
Figure 3.9	Illustration of the non-overlapping sliding window approach for EEG feature extraction. Each window segment is independent.	42
Figure 3.10	Illustration of the overlapping sliding window approach for EEG feature extraction. Consecutive windows share part of their data.	43
Figure 4.1	Boxplots showing classification performance across time centers for different window lengths, using (**Top**) MDM and (**Bottom**) TSM + LDA classifiers. Four window widths are compared: 100 ms, 400 ms, 700 ms, and 1 s. Each box represents the distribution of scores, highlighting both central tendency and variability as a function of window duration.	46
Figure 4.2	Proportion of subjects with significant decoding after FDR correction for the TSM+LDA pipeline. The x-axis indicates the center of the time window, and the y-axis shows the window width. Overall, the patterns illustrate changes in the extent of significant decoding. With xDAWN spatial filtering (**Top**), slow cortical potentials (0.3–3 Hz) yield stronger and more consistent decoding. Without xDAWN filtering (**Bottom**), decoding is weaker and less consistent, with smaller regions of significance.	49
Figure 4.3	Proportion of subjects with significant decoding after FDR correction for the MDM pipeline. The x-axis indicates the center of the time window, and the y-axis shows the window width. Overall, the patterns illustrate changes in the extent of significant decoding. With xDAWN spatial filtering (**Top**), slow cortical potentials (0.3–3 Hz) yield stronger and more consistent decoding. Without xDAWN filtering (**Bottom**), decoding is weaker and less consistent, with smaller regions of significance.	50

Figure 4.4 Median FDR-corrected p-values at each time center for MDM (**Top**) and TSM + LDA (**Bottom**) classification, pooling results from all four window lengths (100, 400, 700 ms, 1 s), with and without the xDAWN spatial filter (0.3–3 Hz). Markers indicate time points that reached statistical significance after FDR correction ($\alpha = 0.05$). The red dashed line indicates the FDR significance threshold. 51

Figure 4.5 Median FDR-corrected p-values at each time center for MDM (**Top**) and TSM + LDA (**Bottom**) classification, pooled across all four window lengths (100, 400, 700 ms, 1 s) and two frequency bands (0.3–3 Hz and 7–30 Hz). Markers indicate time points reaching statistical significance after FDR correction ($\alpha = 0.05$). Red dashed line shows the FDR significance threshold. 52

Figure 4.6 Proportion of subjects with significant decoding after FDR correction for the MDM (**Top**) and TSM + LDA (**Bottom**) pipelines in the oscillatory activity band (7–30 Hz). Compared to Figure 4.2 (**Top**) and Figure 4.3 (**Top**), this heatmap shows a temporal shift in the peak significance from around 1.2 s to approximately 1.9 s. 53

Figure 4.7 Histogram of significant decoding times across subjects. (**Top**): MDM; (**Bottom**): TSM + LDA. xDAWN peaks occur earlier (∼1.2–1.4 s) than oscillatory features (∼1.6–1.8 s). 54

Figure 4.8 Decoding accuracy over time using MDM (**Top**) and TSM + LDA (**Bottom**) for four different window widths (100 ms, 400 ms, 700 ms, and 1 s). The x-axis represents the center of the analysis window in seconds, and the y-axis shows mean classification accuracy across subjects. The dashed line indicates chance level (50%). Both classifiers show that decoding accuracy increases to a maximum around 1.2 s for all window widths, then gradually decreases. Longer windows (700 ms and 1 s) achieve higher performance, whereas the 100 ms window shows the lowest accuracy. 59

Figure 4.9 Normalized confusion matrices (in %) for the multiclass within-subject study using TSM + LDA (**Top**) and MDM (**Bottom**) classifiers with a 1 s analysis window. The matrices show the average classification performance across subjects for each movement class, where the diagonal represents correctly classified trials and off-diagonal entries indicate misclassifications. 61

Figure 4.10 Normalized confusion matrices (in %) for the multiclass cross-subject study using TSM + LDA (**Top**) and MDM (**Bottom**) classifiers with a 1 s analysis window. The matrices show the average classification performance across subjects for each movement class, where the diagonal represents correctly classified trials and off-diagonal entries indicate misclassifications. 63

LIST OF TABLES

Table 2.1	Different EEG frequency bands, each defined by a particular frequency range and associated mental states (adapted from [7]).	11
Table 2.2	Advantages and disadvantages of EEG as a neuroimaging technique (adapted from [8, 9, 10]).	11
Table 2.3	Types of EEG artifacts.	14
Table 4.1	Summary statistics for each window using MDM and TSM + LDA classifiers.	45
Table 4.2	Summary statistics for each window using MDM and TSM + LDA classifiers after replacing the xDAWN spatial filter with simple covariance matrices (SCM).	47
Table 4.3	Summary statistics for each window using MDM and TSM + LDA classifiers after removing xDAWN in the 7–30 Hz frequency band.	48
Table 4.4	Summary statistics for each window using MDM and TSM + LDA classifiers in a cross-subject study.	57
Table 4.5	Summary statistics for two windows using MDM and TSM + LDA classifiers in a within-subject multiclass study.	60
Table 4.6	Summary statistics for each window using MDM and TSM + LDA classifiers in the cross-subject multiclass study.	62
Table 5.1	Mapping of Problems, Objectives, and Research Questions to Results and Discussion Sections.	67

LIST OF SYMBOLS

$+$	Combination operator (e.g., TSM + LDA)
α	Alpha band (8–12 Hz) / significance level
β	Beta band (12–30 Hz)
γ	Gamma band (>30 Hz)
δ	Delta band (0.5–4 Hz)
θ	Theta band (4–8 Hz)
μ	Mu rhythm (8–13 Hz, sensorimotor cortex)
$X \in \mathbf{R}^{C \times N}$	EEG trial matrix with C channels and N samples
$\mathbf{v} \in \mathbf{R}^{n}$	Generic vector for SPD positivity check
$\delta_R(P_1, P_2)$	Riemannian (geodesic) distance between P_1 and P_2
$\delta_L(P_1, P_2)$	Log-Euclidean distance between P_1 and P_2
G	Riemannian mean of a set of SPD matrices (reference point for tangent space)
S_i	Matrix in tangent space corresponding to P_i
s_i	Vectorized tangent-space feature of trial i
$I(k)$	Index set of trials belonging to class k
$\|I(k)\|$	Number of trials in class k
C	Number of EEG channels
L	Number of SPD matrices (trials)
TP	True Positive count
TN	True Negative count
FP	False Positive count
FN	False Negative count
M_{ij}	Number of samples with true class i predicted as j
D	Original dataset
D'	Permuted dataset
$\hat{D}$	Set of k permuted datasets
k	Number of permutations

f	Classifier being evaluated
$e(f, D)$	Performance of classifier f on dataset D
p_i	Uncorrected p-value from permutation test
m	Number of comparisons/tests
$\mathbf{w}$	Projection vector in LDA
$\mathbf{w}^*$	Optimal projection vector that maximizes class separability
$J(\mathbf{w})$	Objective function (ratio of between-class to within-class variance)
S_B	Between-class scatter matrix
S_W	Within-class scatter matrix
$\boldsymbol{\mu}_i$	Mean vector of class i
C_i	Set of samples belonging to class i
$\mathbf{x}$	Feature/sample vector
$\delta_i(\mathbf{x})$	Linear discriminant function for class i
Σ	Shared covariance matrix (assumed equal across classes)
π_i	Prior probability of class i

LIST OF ABBREVIATIONS

ACNS	American Clinical Neurophysiology Society
AIM	Affine-Invariant Metric
BCI	Brain-Computer Interface
BP	Bereitschaftspotential (Readiness Potential)
EEG	Electroencephalography / Electroencephalogram
ERP	Event-Related Potential
ERD	Event-Related Desynchronization
ERS	Event-Related Synchronization
FDR	False Discovery Rate
FIR	Finite Impulse Response
ICA	Independent Component Analysis
LDA	Linear Discriminant Analysis
LEM	Log-Euclidean Metric
LOSO	Leave-One-Subject-Out
LW	Ledoit-Wolf (covariance regularization)
MDM	Minimum Distance to Mean (classifier)
MDRM	Minimum Distance to Riemannian Mean
MI	Motor Imagery
MRCP	Movement-Related Cortical Potential
P300	Positive EEG deflection 300 ms after stimulus onset
PCA	Principal Component Analysis
SCM	Sample Covariance Matrix
SCI	Spinal Cord Injury
SCP	Slow Cortical Potential
SCPS	Slow Cortical Potential Shifts
SPD	Symmetric Positive Definite (matrix)
SSVEP	Steady-State Visual Evoked Potential
SSVER	Steady-State Visual Evoked Response

SVM	Support Vector Machine
TSM	Tangent Space Mapping
xDAWN	Supervised Spatial Filtering Algorithm for ERPs

1. INTRODUCTION

1.1 Decoding

The human desire to communicate is not new. From very early times, people tried to put their thoughts into forms that others could understand-sometimes by speaking, sometimes with symbols, and sometimes in more hidden ways. Long before computers or modern neuroscience existed, ancient civilizations developed systems to transform information into special forms and then to interpret them. They did this to preserve knowledge, send secret messages, or go beyond the limits of normal speech.

From the Cuneiform writing of the Sumerians in Mesopotamia to the hieroglyphs of ancient Egypt, these first writing systems represented a major step in externalizing human thought. Such symbolic languages were themselves like codes that only trained scribes or scholars could accurately interpret them. The ability to decode an information depends on knowledge and context, just as the hieroglyph of an ibis can convey deep symbolic meaning to an ancient Egyptian scribe while appearing as "just a bird" to the untrained eye (Figure 1.1).

Similarly, early military and diplomatic ciphers, like the Spartan scytale or later Morse code, are examples of invented methods to keep messages hidden from unintended recipients. These historical practices, even if simple by today's standards, illustrate a fundamental principle that persists in modern systems: converting internal meaning into external signals and interpreting them correctly. This principle is central to modern brain research, where scientists attempt to decode neural signals to understand actions, thoughts, or perceptions. Just as ancient messages need translation to reveal their meaning, brain signals must be decoded to reveal intentions or planned movements, forming the basis of technologies such as brain-computer interfaces (BCIs).

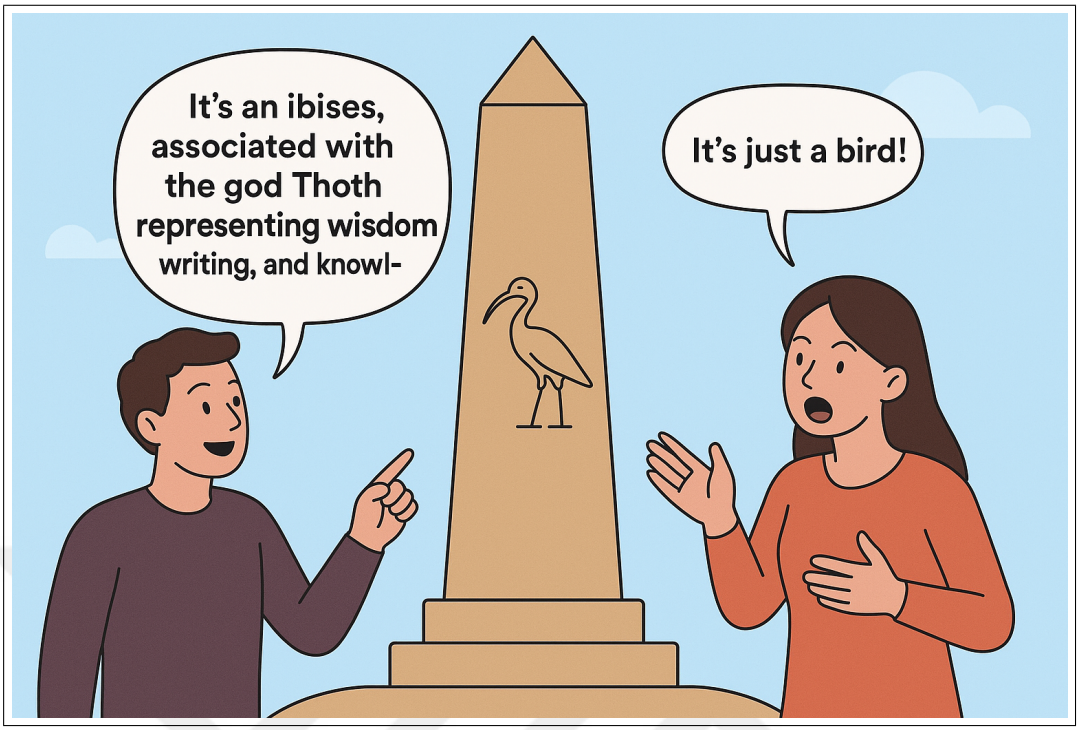

Figure 1.1 Illustration of the difference between symbolic interpretation and literal perception. The ibis, sacred to the god Thoth in ancient Egypt, represented wisdom and knowledge to those familiar with the symbol, but appeared as an ordinary bird to others.

1.2 Neural Decoding

Decoding the brain's activity, much like interpreting ancient scripts, demands a detailed understanding of both the structure of the signals and their contextual origins. In the way that the Rosetta Stone provided scholars with a breakthrough in understanding Egyptian hieroglyphs, contemporary approaches-ranging from advanced decoding algorithms to machine learning models offer powerful means to unravel the brain's intricate activity patterns.

Neural decoding is creating a mathematical mapping from brain activity to external events or actions. In the sensory domain, this can involve visual, auditory, or other sensory information received by the individual. In the motor domain, it refers to the state or movement of the skeletomuscular system [11].

Neural decoding can also be seen as a form of pattern recognition. A set of neuron spike times forms a pattern, and the decoder aims to determine which stimulus or movement is associated with that pattern [11].

In the world of BCIs, this principle reaches its most practical and transformative form. The challenge is no longer encrypting messages for secrecy but decoding neural activity to access unspoken thoughts, intentions, or commands. Neural signals in the brain are, in a way, analogous to coded messages of ancient times-rich with meaning, yet not immediately interpretable without a decoding framework.

Understanding how the brain encodes and sends movement signals is not only theoretically interesting but also essential for practical applications. In particular, for people with spinal cord injuries, decoding these signals enables control of external devices such as motor neuroprostheses, providing the potential to restore some lost movements. Motor-related cortical potentials (MRCPs) are small brain signals that appear when a person plans or initiates voluntary movements. They reflect the neural processes underlying decision-making and movement intention, even before the actual motion occurs. By decoding MRCPs, it is possible to infer intended actions, which is crucial for controlling BCIs effectively.

1.3 Problem Statement

People with cervical spinal cord injury lose most of their voluntary movement, particularly in their arms and hands. Brain-computer interfaces (BCIs), combined with motor neuroprostheses, have the potential to restore some degree of motor function. For natural and reliable control of such devices, it is essential to understand and decode how the brain generates movement signals, often by using electrical stimulation [12].

Previous studies, such as Ofner et al., have demonstrated promising approaches for decoding motor-related cortical potentials (MRCPs). However, several limitations remain:

- **Temporal resolution issues** - Many approaches rely on single time-point analyses. Neural activity is inherently dynamic, unfolding over time. Fixing the analysis to one specific time point introduces problems with Inter- and intra-subject variability and reduces generalizability [13, 14].
- **Restricted frequency sensitivity** - Current methods often work only with slow cortical waves. This assumption limits the adaptability of the pipeline to different frequency bands, making it less suitable for varied tasks or alternative analysis methods.
- **Integration challenges** - Existing pipelines are difficult to combine with advanced techniques, such as transfer learning [15] and other advanced methods [16, 17, 18], which are increasingly important for generalization across subjects and sessions.

To overcome these challenges, two methodological problems need to be addressed:

Problem 1: Classifier flexibility and robustness.
To support diverse use cases and implementations, a more stable and generalizable classifier is required. Riemannian geometry with second-order statistics provides a promising solution, as shown in previous study [18, 19] offering improved robustness and classification accuracy.

Problem 2: Inefficient covariance estimation in ERP-based BCIs.
In ERP paradigms, the sample covariance matrix is often estimated from ERPs–MRCPs in our scenario. However, ERP feature amplitudes are typically much smaller than the ongoing EEG background, which reduces encoding efficiency [19]. The xDAWN algorithm addresses this by enhancing evoked potentials, thereby improving both feature extraction and classification performance.

1.4 Objective of the Study

The main goal of this study is to train EEG-based BCIs to detect the planning and initiation of upper limb movements. With this aim and the research questions in mind, the following objectives are defined as the fundamental steps to achieve this goal.

The first objective is to distinguish each of the six imagined upper limb movements from the rest (no-action) condition individually. This classification is performed by applying the xDAWN filter as a novel method. Spatial features from the covariance matrix are analyzed in manifold and tangent spaces over different time windows using the minimum distance to mean (MDM) classifier and tangent space mapping (TSM + LDA). The second objective is to demonstrate that MRCP can be classified using xDAWN and to identify the presence of movement intention. The third objective is to evaluate how well the same pipeline performs on multiclass classification. The final objective is to apply the same pipeline in a cross-subject study without using transfer learning.

This thesis proposes a new method to extract features from MRCPs for early detection of movement intention. To our knowledge, it is the first work combining MRCP and xDAWN in Riemannian space for EEG-based movement preparation detection.

1.5 Research Questions

Based on the motivation and objectives described above, this study aims to answer the following four research questions:

1. How accurately can EEG-based BCIs detect the planning and initiation of upper limb movements using MRCP signals combined with the xDAWN filter?
2. Can the proposed method reliably distinguish between different imagined upper

limb movements and rest?

3. How well does the developed classification pipeline perform in multiclass scenarios involving multiple movement types?

4. Is the pipeline effective in cross-subject applications without applying transfer learning techniques?

1.6 Organization of the study

This thesis is organized as follows:

Chapter 1 Introduces the research problem, objectives, and research questions. Chapter 2 offers a literature review covering anatomy and physiology related to EEG-based BCIs, Riemannian geometry-based classifiers, and spatial filtering techniques such as xDAWN. Chapter 3 explains the methodology describing the dataset, preprocessing steps, feature extraction, and classification pipelines, experimental validation designs, including binary and multiclass decoding. Chapter 4 presents the resulty on decoding performance of MDM and TSM + LDA classifiers across different window sizes, frequency bands, and experimental conditions. Includes summary statistics, boxplots, heatmaps, histograms, and confusion matrices for both within-subject and cross-subject analyses. Chapter 5 discusses the effects of time window length, frequency band selection, and cross-subject generalization. Highlights the strengths and limitations of the proposed pipeline and implications for future BCI applications. Chapter 6 concludes the study by summarizing the key findings, outlining the contributions, and providing recommendations for future research on EEG-based movement intention decoding.

2. BACKGROUND

2.1 Brain

The brain is the main organ of the nervous system and controls all the functions of the body. It processes information from the senses, controls movements, and supports thinking, memory, and emotions. To understand how brain-computer interfaces (BCIs) can interpret brain activity, it is important to study the parts of the brain that produce measurable electrical signals. The next section focuses on the areas of the brain most relevant for EEG-based BCI systems.

2.2 Relevant Neuroanatomy for EEG

Brain-computer interfaces (BCIs) use the electrical activity of the brain to understand a person's intentions and to control devices outside the body. It is important to know which parts of the brain produce these signals for making EEG-based decoding systems. Because EEG mostly measures activity from the cortex, this study focuses on the cerebral cortex and its main functional areas.

2.2.1 Cerebral Cortex

The cerebral cortex is the primary source of EEG signals. It is divided into lobes with specialized functions:

- **Frontal Lobe:** Motor planning, decision-making, and executive functions.
- **Parietal Lobe:** Somatosensory processing.
- **Temporal Lobe:** Auditory processing and memory.

- **Occipital Lobe:** Visual processing.

The left and right hemispheres communicate via the corpus callosum [20, 21].

In particular, the **primary motor cortex** (precentral gyrus) and the **primary somatosensory cortex** (postcentral gyrus) are of special interest in motor imagery BCI studies. These regions generate strong, detectable EEG signals during movement planning and execution.

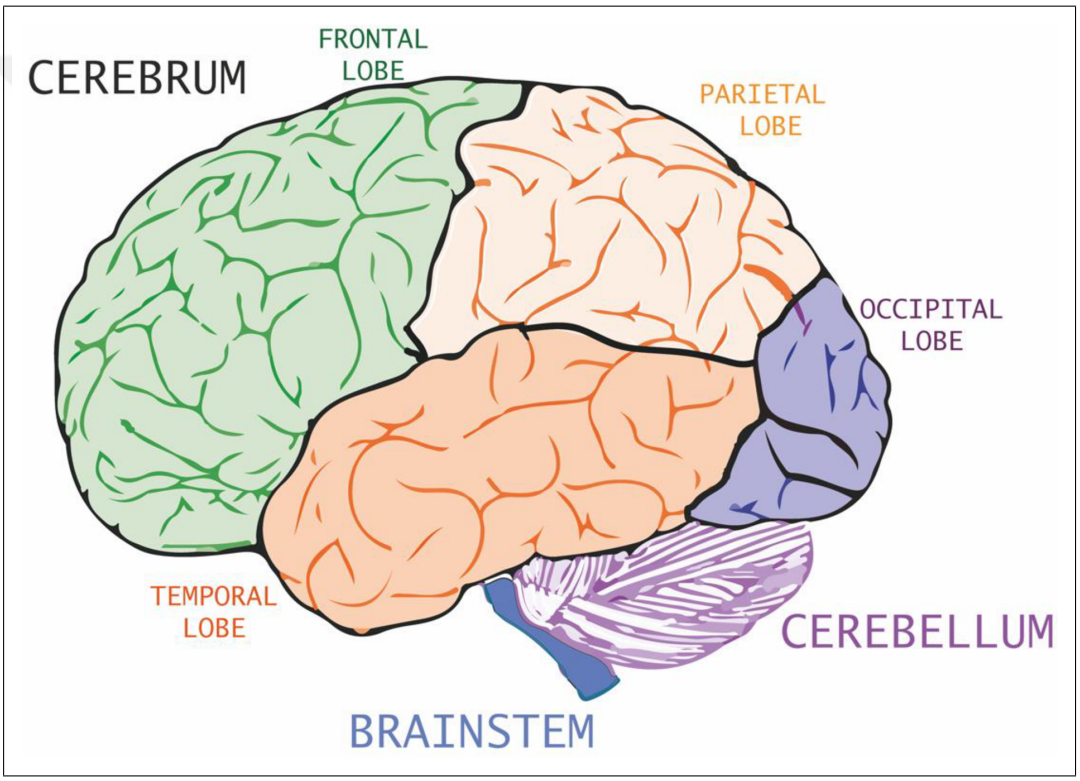

Figure 2.1 Brain structure illustration (adapted from [1]).

2.3 Relevant Neurophysiology for EEG

EEG signals primarily originate from the coordinated activity of cortical pyramidal neurons, particularly in layers of the cerebral cortex that are perpendicular to the scalp surface [8, 22]. These neurons generate postsynaptic potentials in response to synaptic input. When large groups fire synchronously, the resulting extracellular

currents summate to produce measurable oscillations at the scalp. Furthermore, the geometry and alignment of these neurons, as well as the thickness of cortical layers and the conductivity of intervening tissues (e.g., cerebrospinal fluid, skull, scalp), critically influence the amplitude and spatial distribution of EEG signals. The cortex is organized into functionally distinct regions, each contributing differently to EEG activity. The primary motor and somatosensory cortices generate prominent signals during movement and motor imagery. The occipital cortex, in contrast, produces strong alpha rhythms during relaxed wakefulness with eyes closed. The temporal and frontal cortices also contribute oscillatory activity associated with attention, memory, and higher cognitive functions [8, 22].

EEG primarily captures postsynaptic potentials rather than action potentials, reflecting summed dendritic currents from thousands to millions of neurons. The resulting signals exhibit rhythmic activity across various frequency bands, which correspond to different functional and cognitive states. In addition to normal physiological rhythms, EEG is sensitive to transient changes caused by sensory input, motor planning, and cognitive processes. Spatial resolution is limited by volume conduction, meaning that activity from deep or overlapping cortical sources can mix at the scalp electrodes. Temporal resolution, however, remains excellent, allowing precise tracking of fast neuronal dynamics.

Understanding these physiological mechanisms is essential for interpreting EEG recordings and designing brain-computer interface (BCI) systems, as it provides insight into how neural activity in specific cortical areas can be captured and decoded for external control.

2.4 EEG

2.4.1 Generation of EEG Signals

The human brain consists of billions of neurons that are continuously firing. Electroencephalography (EEG) records this electrical activity through sensors placed on the scalp [23, 7], as illustrated in Figure 2.2. This technique allows researchers to observe real-time brain responses. Because EEG signals fluctuate rapidly, they provide insight into brain processes occurring over very short durations. EEG has been widely used to study brain disorders, sleep patterns, and brain-computer interfaces [23, 7].

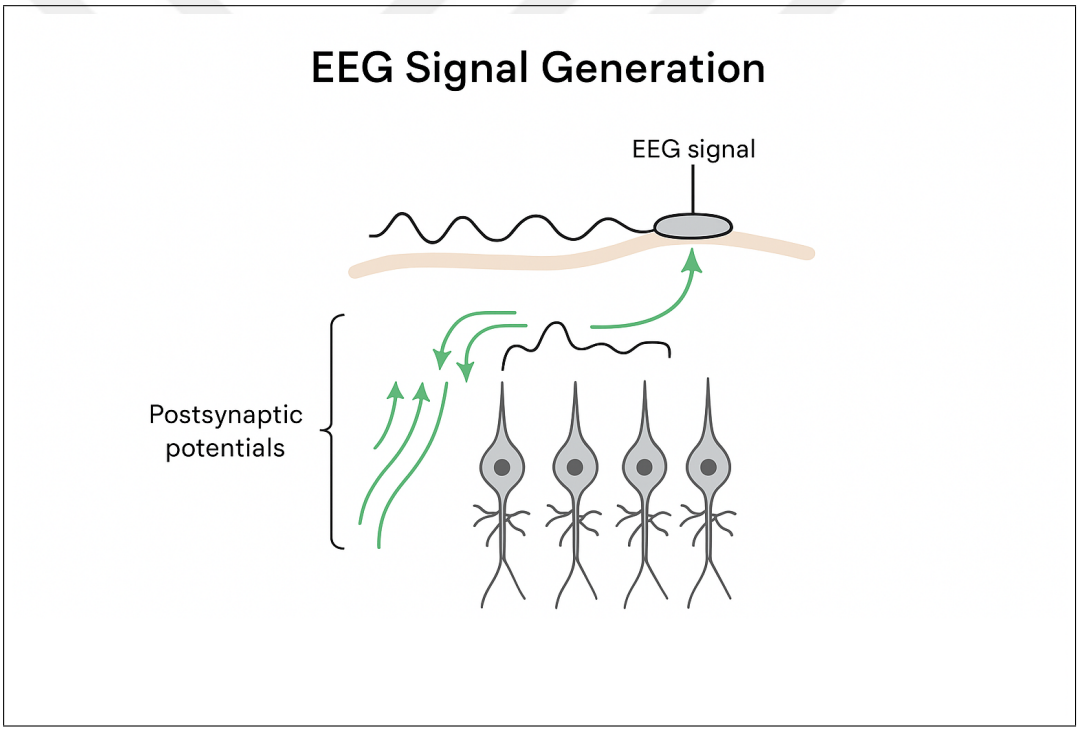

Figure 2.2 Schematic of cortical neurons generating postsynaptic potentials (PSPs) and the resulting signals recorded by a scalp EEG electrode.

In particular, EEG is a preferred method in Brain-Computer Interface (BCI) applications because it provides real-time signals that can be translated into commands for external devices [24]. This enables users, especially those with motor impairments, to control cursors, robotic arms, or communication tools, enhancing independence and interaction with their environment.

EEG signals are commonly divided into five frequency bands, ordered from low to high frequency: the δ (delta), θ (theta), α (alpha), β (beta), and γ (gamma) bands. Each band is associated with particular mental states and cognitive processes. Understanding these frequency bands is essential for interpreting EEG signals, as summarized in Table 2.1.

EEG offers several benefits and limitations that influence its application in research and clinical settings. Although it provides excellent temporal resolution and is relatively affordable and non-invasive, EEG is limited by low spatial resolution and its susceptibility to artifacts from muscles, eyes, and environmental noise. Table 2.2 summarizes the main advantages and disadvantages of EEG [9, 8, 23].

Table 2.1
Different EEG frequency bands, each defined by a particular frequency range and associated mental states (adapted from [7]).

Wave Band	**Frequency (Hz)**	**Mental State**
δ	0.5–4 Hz	deep sleep
θ	4–8 Hz	meditative, concentration, drowsiness, mental calculation
α	8–12 Hz	mental effort, inattention, daydreaming, closed eyes
β	12–30 Hz	attention, active thinking, motor activities
γ	>30 Hz	perception, higher-order motor functions

Table 2.2
Advantages and disadvantages of EEG as a neuroimaging technique (adapted from [8, 9, 10]).

Advantages	**Disadvantages**
High temporal resolution (milliseconds)	Low spatial resolution
Non-invasive and relatively inexpensive	Sensitive to noise and artifacts (e.g., eye blinks, muscle activity)
Portable and easy to set up	Requires conductive gel or specialized caps
Safe for repeated measurements	Limited ability to localize deep brain activity

2.5 EEG Montages

EEG montages define how electrodes are arranged on the scalp to record brain activity. The choice of montage affects how signals are interpreted and can emphasize either local or global brain activity patterns. Some commonly used montages include:

- **Bipolar Montage:** Each electrode is connected to its neighbor, recording the potential difference between them. This is effective for identifying localized brain activity.
- **Referential Montage:** Each electrode is compared to a single reference electrode, often placed centrally. This montage highlights general brain activity but may be less sensitive to small local changes.
- **Common Average Montage:** Each electrode is compared to the average of all electrodes, which can reduce the influence of widespread signals and enhance local activity detection.
- **Laplacian Montage:** Each electrode is referenced to the average of its neighboring electrodes, providing better spatial resolution of surface brain activity.

The American Clinical Neurophysiology Society (ACNS) recommends using more than one montage during EEG recordings to obtain complementary perspectives of brain activity [25].

In BCI experiments, selecting an appropriate montage is crucial because it directly impacts signal quality and the detection of task-relevant neural patterns [5]. Using the optimal montage can improve spatial resolution, reduce interference, and enhance the real-time decoding of EEG signals for accurate device control.

2.6 EEG Artifact

EEG recordings are often contaminated by various types of artifacts, which are unwanted signals that can obscure or distort brain activity. Artifacts can arise from physiological sources, such as eye movements, muscle activity, and cardiac signals, as well as from external factors like electrode movement or electrical interference. Proper identification and management of these artifacts are essential for reliable interpretation of EEG data. Table 2.3 summarizes the main types of EEG artifacts and their typical characteristics [26, 27].

2.6.1 Artifact Detection and Correction

There are different ways to detect and correct artifacts [27]:

- **Manual Inspection:** Look at EEG recordings and remove parts with artifacts. It works well but takes a lot of time.
- **Automated Algorithms:** Use computer methods like Independent Component Analysis (ICA), Principal Component Analysis (PCA), or machine learning to find and remove artifacts automatically.
- **Filtering Techniques:** Use filters to remove noise in certain frequencies, for example, line noise. Be careful not to remove brain signals by mistake.
- **Deep Learning Approaches:** Use neural networks to learn patterns of artifacts and clean the EEG data automatically.

Choosing the right method depends on the type of artifact, quality of EEG data, and the goal of the study. Often, using more than one method works better.

2.7 Motor Imagery

2.7.1 Motor Imagery Overview

Motor imagery (MI) involves the mental simulation of movement without actual execution. It engages neural circuits similar to those activated during real movement. EEG studies have identified specific patterns, such as event-related desynchronization (ERD) and synchronization (ERS) in the α and β frequency bands, which occur during MI tasks. These neural signatures are utilized in brain-computer interfaces (BCIs) to enable communication and control for individuals with motor impairments.

MI has also been explored as a rehabilitation technique. Evidence suggests that it can enhance motor recovery, mobility, balance, and psychological well-being, particularly when combined with physical practice. Overall, MI serves as a valuable tool in both clinical and research settings, providing insights into motor control and aiding the development of assistive technologies [28].

Table 2.3
Types of EEG artifacts.

Artifact Type	**Description**
Ocular Artifacts	From eye movements and blinks. Appear as large slow waves, mostly in frontal electrodes.
Muscle Artifacts	From facial or scalp muscle movements. Cause high-frequency noise in EEG.
Cardiac Artifacts	From the heart's electrical activity. Appear periodically near scalp electrodes.
Electrode/Environmental Artifacts	Caused by poor contact, electrode movement, or external electrical noise.
Sweat Artifacts	Caused by changes in skin impedance. Lead to baseline drifts or distortions.
Movement Artifacts	From head or body movements. Cause electrode shifts and short-term noise.
Line Noise	From power lines (50/60 Hz). Can contaminate EEG recordings.

2.7.2 EEG Signatures of Motor Imagery: ERD/ERS

Motor imagery (MI) engages neural circuits similar to those activated during actual movement, producing measurable changes in EEG oscillatory activity [29]. The most prominent rhythms related to MI are the μ rhythm (8–13 Hz) over the sensorimotor cortex and the β rhythm (14–30 Hz) over fronto-central regions. At rest, the μ rhythm is prominent and becomes desynchronized (event-related desynchronization, ERD) over the contralateral motor areas during imagined or executed movement. β activity shows task-related modulation: it typically decreases during movement preparation and execution and increases after movement offset (event-related synchronization, ERS).

During MI, ERD is generally observed contralateral to the imagined limb, while ipsilateral areas may show ERS. These patterns provide reliable electrophysiological markers of motor planning and intention, forming the basis for brain-computer interface (BCI) control. Power changes are typically referenced to a baseline resting period, which allows extraction of task-specific features. ERD/ERS dynamics have been extensively used to decode movement intentions in real time, supporting MI-based BCIs [29].

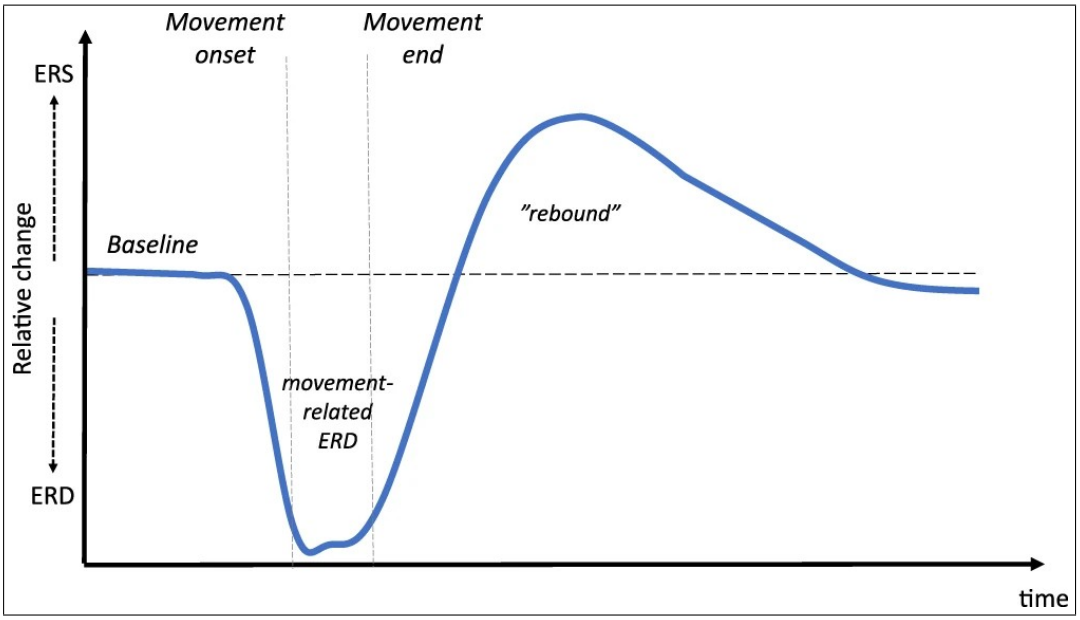

Figure 2.3 Typical ERD/ERS pattern observed in a healthy individual (adapted from [2]).

2.7.3 Movement-Related Cortical Potentials (MRCPs)

The BP (Bereitschaftspotential) represents the earliest stage of movement-related cortical potentials (MRCPs) and originates from the supplementary motor area, activating before the primary motor cortex [30]. It reflects preparatory activity for voluntary movements.

The BP consists of two components. The early component begins about 2 seconds before movement, is maximal at the central-parietal midline, and is symmetrically distributed across hemispheres. The late component appears roughly 400 ms before movement and is larger over the contralateral central region.

Characteristics of the BP, including timing and amplitude, are influenced by task preparation, movement type, learning, repetition, speed, difficulty, and neurological lesions [31, 30]. Detecting movement intention from MRCPs in real time is crucial for understanding decision-making, free will, and consciousness. It also supports the development of novel rehabilitation strategies and BCI applications.

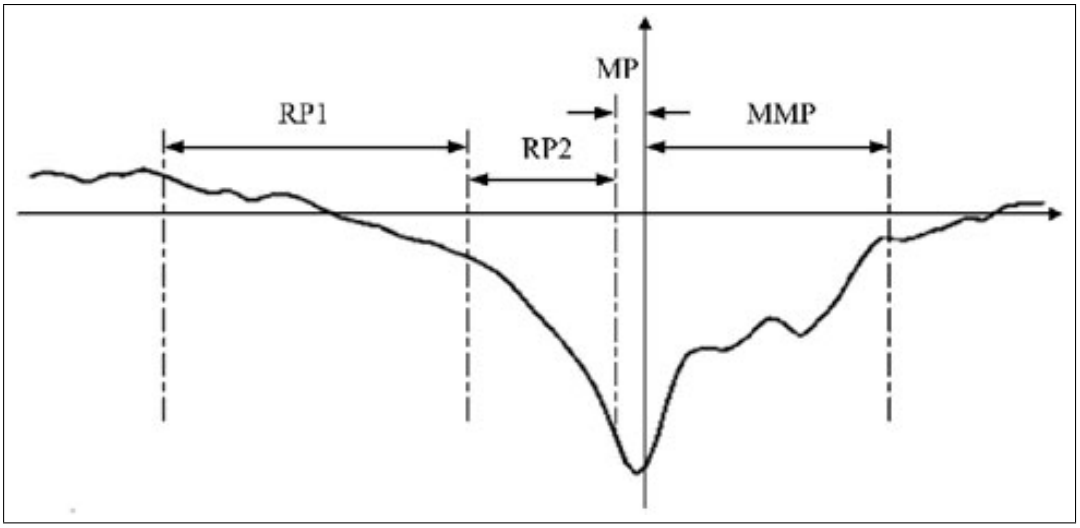

Figure 2.4 Movement-related cortical potential (MRCP) (adapted from [3]).

2.8 Brain-Computer Interfaces (BCI)

A Brain-Computer Interface (BCI) is a system that directly measures brain activity – such as EEG signals – associated with the user's intent. Through signal processing and pattern recognition techniques, it translates this activity into commands or messages to control a computer or interactive application [19].

For example, by imagining moving the left or right hand, a BCI can detect these intentions and move a ball either to the left or right on a computer screen (see Figure 2.5) [5]. Because of the features and advantages of EEG, as explained in Section 2.4, it is commonly used in BCI systems [6].

EEG-based BCIs take signals that carry encoded information from brain activity as input and translate them into computer commands. This translation usually relies on classification algorithms trained to recognize several different commands. During training, the user performs these commands to help the system learn. The goal is to create a system that can use this training to identify commands from the user in the future.

2.8.1 Classification in BCI Problems and Requirements for a Modern BCI

The success of machine learning in BCI largely depends on a well-planned experimental protocol [32]. This includes selecting appropriate paradigms and tasks that minimize interference or confusion. A strong and careful setup is essential for the entire ML pipeline, and common problems during model training must be addressed to ensure reliable performance.

Traditionally, EEG-based BCIs follow a sequence of steps: collecting signals, processing them, extracting important features, classifying the data, and providing feedback [32]. However, simply following these steps is not enough – many pitfalls can reduce system performance if they are not handled carefully.

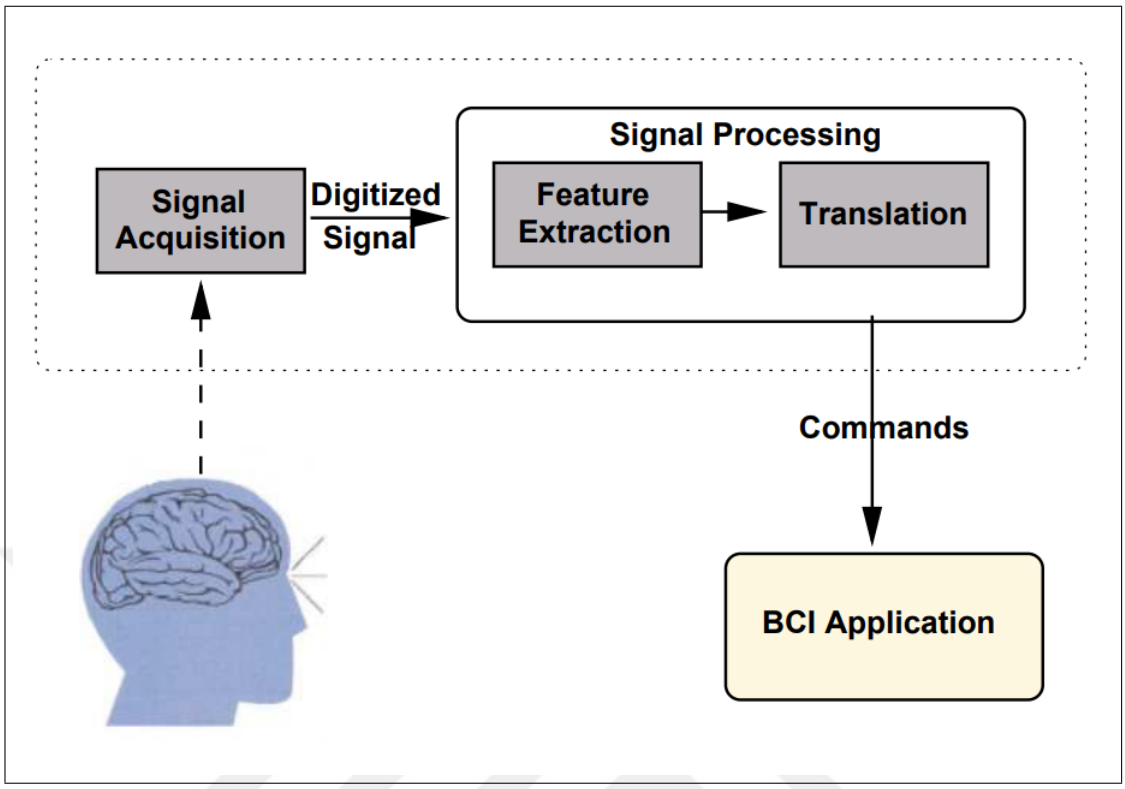

Figure 2.5 Basic design of a BCI system (adapted from [4]).

Designing a good BCI system is not simple, and many mistakes can occur [32]. First, the experiment needs careful planning to ensure both high-quality data and user comfort. It is also important to choose appropriate tasks and understand the brain signals involved.

Second, machine learning models must be trained and tested properly. This includes keeping training and testing data separate and independent for all steps, such as feature selection and parameter tuning.

Finally, when reporting results, using correct performance measures, sharing data and code whenever possible, avoiding confounding factors like muscle noise, and checking statistical validity are all critical. Ignoring these points can lead to incorrect or biased conclusions.

Besides design and evaluation challenges, BCI systems also face technical diffi-

culties related to brain signals [33]. For example, the recorded signals are often noisy and can include unusual or misleading data points.

In addition, because many sensors and features are used, the resulting data can be very large and complex, making analysis more difficult. Finally, brain signals change over time, so it is important to account for temporal information when processing the data.

These issues highlight the need for modern BCI systems that are not only accurate but also adaptable and practical in real-world conditions. Based on recent research, several key requirements have been suggested for a modern BCI decoder [19]:

- Accuracy — The system should perform well compared to established state-of-the-art approaches.
- Reliability — It should maintain stable function and accuracy in both routine and unexpected circumstances.
- Generalization — It should perform well with generic parameters for a new user and generalize across subjects and sessions.
- Adaptability — It should quickly learn the user's individual characteristics and adapt to changes in mental state and environment.
- Universality — It should be applicable to all BCI paradigms, including hybrid systems.
- Simplicity — The algorithms should be simple enough to allow robust, unsupervised online operation.
- Efficiency — It should be computationally efficient enough to run on small, portable electronic devices.
- Multi-User Capability — It should easily extend to scenarios involving multiple users, in line with the trend toward socially connected technologies.

2.8.2 BCI mode

Regarding the mode of operation, BCI systems are generally classified into two main types [34]:

Asynchronous Interfaces: This type of interface continuously monitors the user's voluntary activity and maintains a communication link with the system. The system regularly classifies the user's mental state and can measure changes over time in motor activity, such as changes in signal amplitude, which are then converted into commands. However, analyzing motor activity usually requires a long training period.

Spontaneous brain activity generates several types of signals that are used in BCI systems:

- Slow Cortical Potential Shifts (SCPS)
- Oscillatory activity (sensorimotor)
- Spontaneous EEG signals

Synchronous Interfaces: This type of interface analyzes EEG signals that are evoked by stimuli provided to the user, which can be visual, auditory, or tactile. In this case, the system performs most of the communication work, and the user only needs to respond (or not) to a series of stimuli. These interfaces do not use spontaneous brain activity but rather the brain's reactions to the stimuli, which are then converted into commands. This type of interface requires only a small amount of learning.

The main types of signals used in synchronous interfaces are:

- Steady-State Visual Evoked Responses (SSVERs)
- Event-Related Potentials (ERPs)

2.8.3 Paradigm

In BCI research, a **paradigm** defines the experimental protocol that elicits specific brain activity patterns for decoding. Paradigms are designed to maximize signal detectability while minimizing user fatigue and variability. The choice of paradigm directly affects BCI performance and usability [5].

2.8.3.1 Motor Imagery (MI) Paradigm. The MI paradigm requires users to mentally simulate specific movements, such as imagining moving the left or right hand, without producing actual movement. This elicits characteristic EEG patterns, such as $\mu$and βERD/ERS over sensorimotor areas, which can be detected and classified by the BCI system [30]. MI paradigms are widely used for both communication and neurorehabilitation applications.

2.8.3.2 P300 or Event-Related Potential (ERP) Paradigm. The P300 paradigm is a **synchronous** paradigm where users respond to rare or target stimuli within a sequence of frequent non-target stimuli. A positive EEG deflection approximately 300 ms after stimulus onset (P300) is used to detect the user's intention. This paradigm is popular in spellers and other communication BCIs [35].

2.8.3.3 Steady-State Visual Evoked Potential (SSVEP) Paradigm.
The SSVEP paradigm uses visual stimuli that flicker at specific frequencies. When a user focuses on a flickering stimulus, the EEG over occipital areas shows increased power at the stimulus frequency. SSVEP BCIs are highly accurate and require minimal training [36].

2.8.3.4 Other Paradigms. Other paradigms include slow cortical potential (SCP) BCIs, hybrid BCIs combining multiple paradigms, and auditory or tactile paradigms for users with sensory impairments [24].

Choosing the appropriate paradigm depends on the BCI application, user characteristics, required speed and accuracy, and the training effort available.

2.8.4 Riemannian Geometry in BCI Classification:

In traditional BCI pipelines, features are extracted from EEG signals and standard classifiers are applied. These methods usually treat the data as vectors in Euclidean space. However, EEG signals have a lot of structure, and their spatial covariance patterns contain important information that can be lost when the data is vectorized. Covariance matrices from multichannel EEG trials are symmetric positive definite (SPD), which means they lie in a curved, non-Euclidean space.

If we treat these matrices as points on a Riemannian manifold, their intrinsic geometry is preserved. This can make classification more robust and efficient. Figure 2.6 shows an analogy: the shortest path between two cities on a curved Earth is different from a straight line on a flat map. In the same way, distances between SPD matrices are better measured along the curved manifold, not with normal Euclidean distance. This is why Riemannian geometry is useful in modern BCI systems, as it allows us to use the structure of EEG covariance features [37, 19, 38].

In the next section, we explain the basics of Riemannian geometry and how it is used for EEG-based BCI classification.

2.9 Riemannian Geometry

Riemannian geometry is a branch of differential geometry that studies curved, non-Euclidean spaces [39]. Unlike Euclidean spaces, where the shortest path between two points is a straight line, Riemannian manifolds are smooth surfaces in which the shortest paths-called geodesics-are generally curves.

This mathematical framework was introduced to EEG-based BCIs over a decade ago and has grown in popularity for feature extraction and classification [19]. Covariance matrices derived from multichannel EEG signals are symmetric positive definite (SPD) and naturally reside on a Riemannian manifold. Treating these matrices as points on the manifold preserves their intrinsic geometry and allows for more robust and efficient classification.

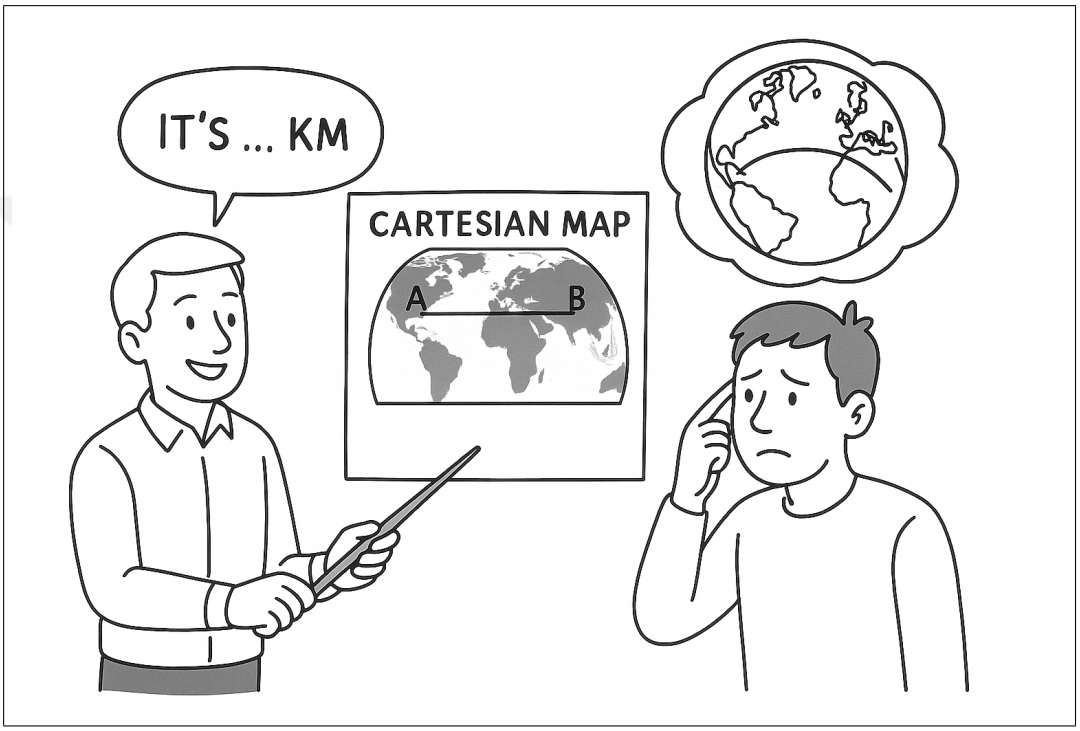

Figure 2.6 Illustration of the difference between Euclidean and Riemannian distances. On a flat map, distances between two points (cities) are measured as straight lines (Euclidean), whereas on the Earth's curved surface, the shortest path follows the geodesic along the sphere (Riemannian distance). This analogy helps understand distances between SPD matrices in EEG-based BCI classification.

2.9.1 From Motor Imagery EEG to Covariance Matrices

Motor imagery (MI) tasks require the user to mentally simulate movements, such as imagining moving the hand or foot, without physically moving. Different movements activate distinct regions of the motor cortex, producing characteristic EEG patterns [40].

Let a single EEG trial be represented as a matrix $X \in \mathbf{R}^{C \times N}$, where C is the

number of channels and N is the number of time samples. To capture the second-order statistics of the trial, the covariance matrix P is computed as:

$$P = \frac{1}{N-1} XX^\top \tag{2.1}$$

This matrix is square, symmetric, and empirically positive definite [6]. To improve robustness against noise, the Ledoit-Wolf shrinkage estimator can be applied, ensuring P remains SPD [41].

Each EEG trial is thus represented as a single point on the Riemannian manifold. The diagonal elements of P encode the variance of each channel, while the off-diagonal elements capture the synchronization between electrodes–features that are crucial for classification.

2.9.1.1 Symmetric Positive Definite (SPD) Matrices. A matrix $A \in \mathbf{R}^{n \times n}$ is said to be symmetric positive definite (SPD) if it satisfies two conditions:

- Symmetry: $A = A^\top$
- Positive definiteness: for any non-zero vector $\mathbf{v} \in \mathbf{R}^n$,

$$\mathbf{v}^\top A \mathbf{v} > 0 \tag{2.2}$$

Covariance matrices derived from EEG trials naturally satisfy these properties. Their SPD structure allows them to be treated as points on a Riemannian manifold, preserving the geometric relationships between trials. This manifold representation is essential for using Riemannian metrics and distances in classification tasks [6, 42].

2.9.2 Riemannian Metrics for SPD Matrices

Once EEG trials are represented as SPD covariance matrices on the manifold, we need a way to measure distances and averages between them. This is where **Riemannian metrics** become essential. The choice of metric affects both the accuracy and efficiency of classification on the manifold.

Two widely used metrics in BCI applications are the **Affine-Invariant Metric (AIM)** and the **Log-Euclidean Metric (LEM)** [6, 38].

2.9.2.1 Affine-Invariant Metric (AIM). For two SPD matrices $P_1, P_2 \in \mathbf{R}^{N\times N}$, the Riemannian (geodesic) distance is defined as:

$$\delta_R(P_1, P_2) = \left\| \log\left(P_1^{-1/2} P_2 P_1^{-1/2}\right) \right\|_F = \left(\sum_{i=1}^{N} \log^2 \lambda_i \right)^{1/2} \tag{2.3}$$

where λ_i are the eigenvalues of $P_1^{-1/2} P_2 P_1^{-1/2}$ and $\|\cdot\|_F$ denotes the Frobenius norm. This metric preserves the affine-invariant structure of the manifold, ensuring meaningful geometric comparisons between SPD matrices.

The Riemannian Fréchet mean of a set of SPD matrices $\{P_1, \ldots, P_I\}$ under AIM is defined as:

$$G = \arg\min_{P\in\mathcal{P}(N)} \sum_{i=1}^{I} \delta_R^2(P, P_i) \tag{2.4}$$

Since there is no closed-form solution, this mean is typically computed using iterative numerical optimization [6].

2.9.2.2 Log-Euclidean Metric (LEM). The Log-Euclidean metric provides a simpler alternative with similar performance in practice [43]. The distance between P_1

and P_2 is computed as:

$$\delta_L(P_1, P_2) = \|\log(P_1) - \log(P_2)\|_F \tag{2.5}$$

The Log-Euclidean mean of k matrices is given by a closed-form expression:

$$G = \exp\left(\frac{1}{k}\sum_{i=1}^{k}\log(P_i)\right) \tag{2.6}$$

This closed-form solution avoids iterative optimization and is computationally efficient, making it practical for real-time BCI applications.

2.9.3 xDAWN Spatial Filter

Before computing covariance matrices, EEG signals can be preprocessed with spatial filters to enhance task-relevant activity. The xDAWN algorithm is a supervised spatial filter that maximizes the signal-to-signal-plus-noise ratio (SSNR) of target evoked potentials [44, 45]. Originally developed for P300 ERPs, xDAWN can be applied to any ERP paradigm.

Let $X_i \in \mathbf{R}^{C\times N}$ denote the EEG data of trial i belonging to class k, where C is the number of channels and N is the number of time samples. The class-wise average of trials is:

$$P(k) = \frac{1}{|I(k)|}\sum_{i\in I(k)} X_i \tag{2.7}$$

where $I(k)$ is the set of indices of trials labeled as class k, and $|I(k)|$ is the number of trials in that class.

The spatial filter $\mathbf{w} \in \mathbf{R}^C$ is then obtained by solving:

$$\mathbf{w}^* = \arg\max_{\mathbf{w}} \frac{\mathbf{w}^\top P(k)P(k)^\top \mathbf{w}}{\mathbf{w}^\top XX^\top \mathbf{w}} \tag{2.8}$$

This can be solved using eigenvalue decomposition:

$$\left(P(k)P(k)^\top\right)\left(XX^\top\right)^{-1}\mathbf{w} = \lambda\mathbf{w} \tag{2.9}$$

Applying xDAWN improves the quality of covariance matrices, leading to better separability on the Riemannian manifold. Typically, xDAWN is applied to filtered EEG signals in the frequency range relevant to the evoked potentials study [46].

2.9.4 Riemannian Metrics and Classification

Representing EEG trials as SPD matrices on a Riemannian manifold allows classifiers to exploit the intrinsic geometry of the data. Distances computed with metrics such as AIM or LEM capture the true similarities between covariance matrices, unlike standard Euclidean distances which may distort these relationships. By using Riemannian metrics, classification algorithms-such as Minimum Distance to Mean (MDM) or tangent-space methods-can operate directly on the manifold. This leads to more robust and accurate decoding of mental states, especially in noisy or high-dimensional EEG data [40, 38]. In practice, these metrics enhance both the separation of classes and the stability of spatial features, making Riemannian geometry a powerful tool for modern BCI systems.

3. MATERIAL AND METHODS

This chapter describes the dataset used in the study and the EEG-based BCI decoding pipeline, which was developed using Riemannian geometry. The pipeline comprised three main stages: preprocessing, signal processing (including feature extraction), and classification [5, 19].

MATLAB and Python were used for both the preprocessing steps and the subsequent computational analyses. The following section describes the dataset to which the decoding pipeline was applied.

3.1 Dataset

This study utilized the motor imagery portion of the publicly available upper-limb movement BCI-based EEG dataset provided by Ofner *et al.* from the University of Graz [12]. The dataset contains electroencephalography (EEG) recordings from 15 healthy participants aged between 22 and 40 years (mean age: 27 ± 5 years), including nine females. All participants provided written informed consent and received monetary compensation for their participation.

During the motor imagery (MI) experiments, the participants imagined doing six movements: elbow flexion and extension, forearm supination and pronation, hand opening and closing, and also a rest condition. The experiment used a trial-based paradigm. Participants sat in front of a computer screen and cues for the tasks were shown on the screen. At the start of each trial, a beep sound played and a fixation cross appeared, which the participants had to look at. After 2 s, a cue for one of the movement tasks appeared. Each participant performed 10 runs, each run with 42 trials, giving a total of 420 trials per participant.

The EEG data were recorded using 61 active electrodes connected to four 16-channel amplifiers (g.tec Medical Engineering GmbH, Austria). The electrodes were placed over the frontal, parietal, and temporal regions of the brain. The reference and ground electrodes were placed on the right mastoid and AFz, respectively. For the MI experiments, EEG signals were recorded with an 8th-order Chebyshev band-pass filter (0.1–200 Hz) and a sampling rate of 512 Hz. A 50 Hz notch filter was also applied to reduce power line interference.

3.2 Preprocessing

This study was done offline, but the possibility of using it online was also considered. Therefore, the preprocessing steps were designed to require as few computational operations as possible. We partially followed the automated preprocessing methods described in *"EEG is Better Left Alone"* [47]. Two MATLAB toolboxes, *FieldTrip* [48] and *EEGLAB* [49], were used to preprocess the raw data and detect bad channels and trials.

For computational purposes, the data were downsampled to 256 Hz. We selected the main channels over the motor region: FC3, FC1, FCz, FC2, FC4, FCC3h, FCC1h, FCC2h, FCC4h, C3, C1, Cz, C2, C4, CCP3h, CCP1h, CCP2h, CCP4h, CP3, CP1, CPz, CP2, and CP4. A finite impulse response (FIR) band-pass filter with cutoff frequencies of 0.3–70 Hz was applied. This filter was designed using a Hamming window with a one-pass kernel and applied in a two-pass non-causal manner. The original reference to the right mastoid was kept. A 50 Hz notch filter was also applied, same as in the original recording settings. Bad channels were identified based on abnormal joint probabilities and abnormal kurtosis, using a threshold of five times the standard deviation (5σ). Before detecting bad trials, the data were reloaded as epochs of interest (trials), and the preprocessing steps above were applied to each trial individually. Previously marked bad channels were repaired. Then, bad trials were identified and removed based on the same criteria: abnormal joint probabilities and abnormal kurtosis (5σ threshold). Finally, scalp current densities were applied following [50] to reduce

the volume conduction effect. The preprocessed data were saved for further analysis.

The preprocessing steps described above were optimized to keep low-frequency components. To extract the movement-related cortical potentials (MRCPs), a low-pass filter with a 3 Hz cutoff was applied to the preprocessed data. The same steps were used to extract oscillatory event-related desynchronization (ERD) and event-related synchronization (ERS) by applying a band-pass filter with a 7–30 Hz cutoff.

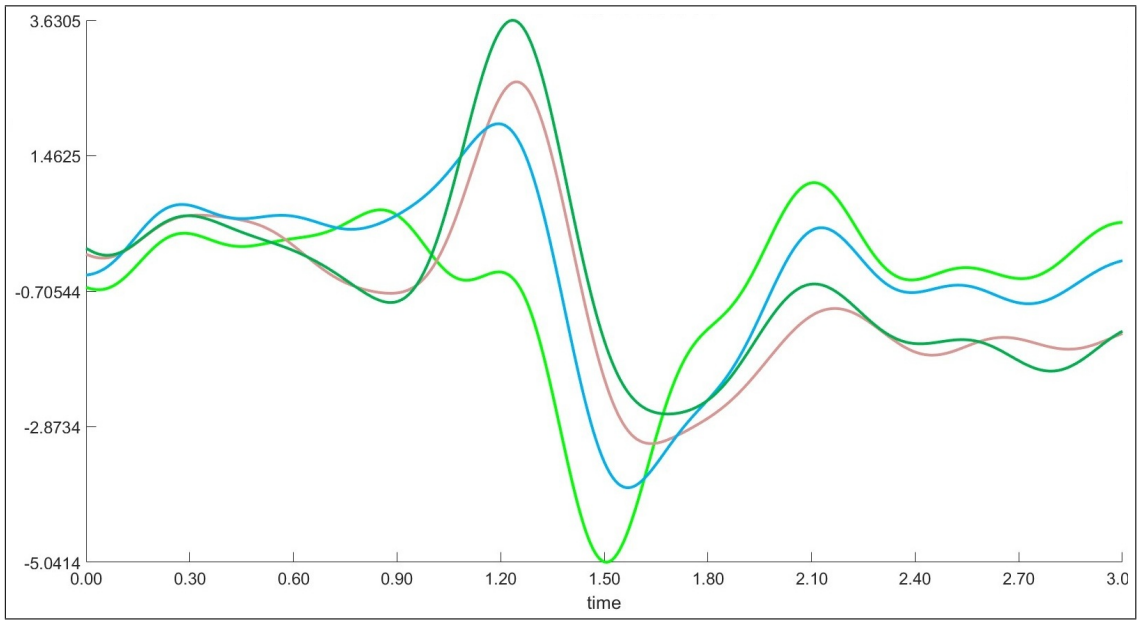

Figure 3.1 Grand average of movement-related cortical potentials (MRCPs) recorded from four motor-related EEG channels (FCz, Cz, C3, C4). The signals are stacked for better visualization of waveform patterns.

After extracting MRCPs or oscillatory features from the preprocessed EEG signals, these data were converted into covariance matrices using Eq. 2.1. Since covariance matrices do not lie in a standard Euclidean space but on a curved Riemannian manifold, conventional statistical methods may not describe them accurately. This motivates the use of Riemannian geometry, which provides the mathematical tools to handle and analyze the data in their natural space. The next step is to apply classification methods based on this framework.

Note: All covariance matrices were estimated using Ledoit and Wolf (LW) regu-

larization [41]. Mentioning the sample covariance matrix (SCM) still implies that LW regularization was applied during estimation.

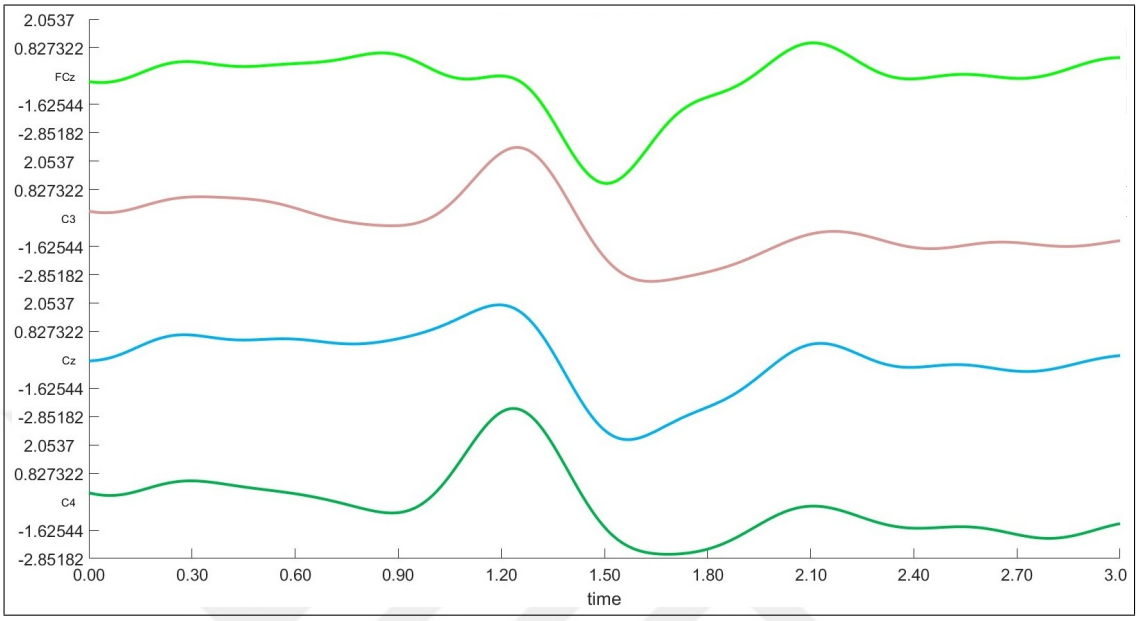

Figure 3.2 Grand average of MRCPs from the same four channels, displayed without stacking to illustrate their relative amplitudes.

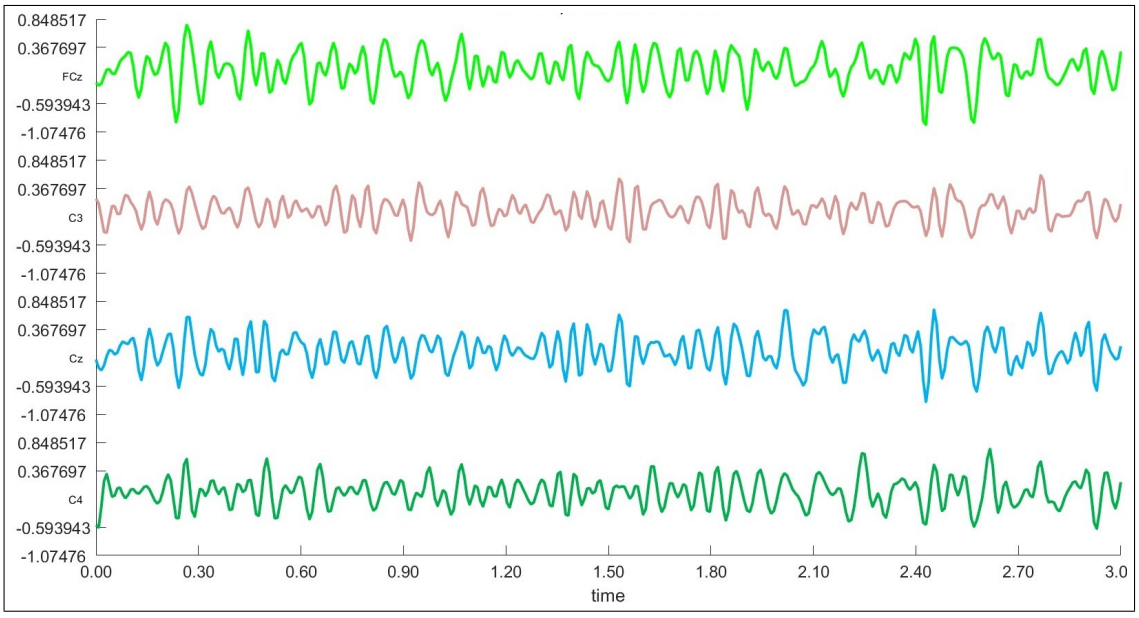

Figure 3.3 Grand average of ERD/ERS bands from the same four channels.

3.3 Classification

The main idea of the Riemannian Geometry Classifier (RGC) is to work directly with data on the Riemannian manifold. On this manifold, data can be manipulated naturally, including operations such as averaging, scaling, and classification. There are generally two main approaches in Riemannian geometry-based classification. The first is to classify the data directly on the manifold. The second is to map the data to the tangent space, where standard machine learning algorithms, from simple to advanced, can be applied. In this study, we implemented both approaches to compare their performance and evaluate their strengths and limitations.

3.3.1 Classification in Riemannian Space

The Riemannian Minimum Distance to Mean (RMDM), also referred to as the MDM classifier, is not the most accurate among Riemannian classifiers, but it is notable for its simplicity, computational efficiency, and applicability across different BCI paradigms [38].

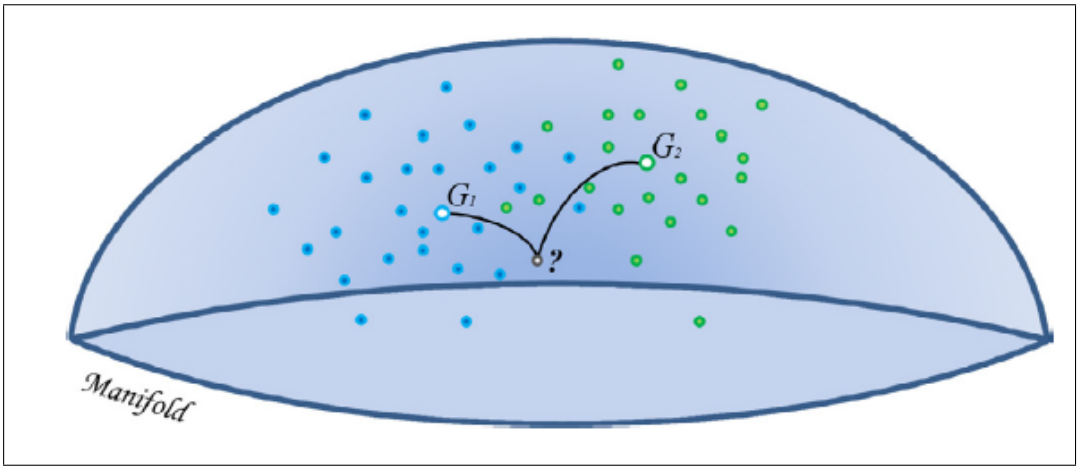

Figure 3.4 Illustration of the Riemannian Minimum Distance to Mean (RMDM) classifier applied to a two-class problem (adapted from [5]).

The RMDM is a fully deterministic, parameter-free classifier, so no parameters need to be tuned using cross-validation or other methods that might affect its general-

ization. Moreover, it can be easily extended to multi-user scenarios, has demonstrated reliable performance in individuals with clinical conditions, and is suitable for integrating transfer learning and adaptation strategies [38].

In this study, we adapted the non-adaptive (train–test) form of the RMDM classifier. During training, a set of SPD matrices representing BCI trials for the available classes is obtained. For each class, the center of mass of its trials is calculated using Eq. 2.4 [38]. In the testing phase, a BCI trial is encoded as an SPD matrix in the same way and assigned to the class whose center of mass is closest. The distance is measured using a suitable function defined on the manifold, as given in Eq. 2.3 [38] (see Figure 3.5 for an illustration of this process).

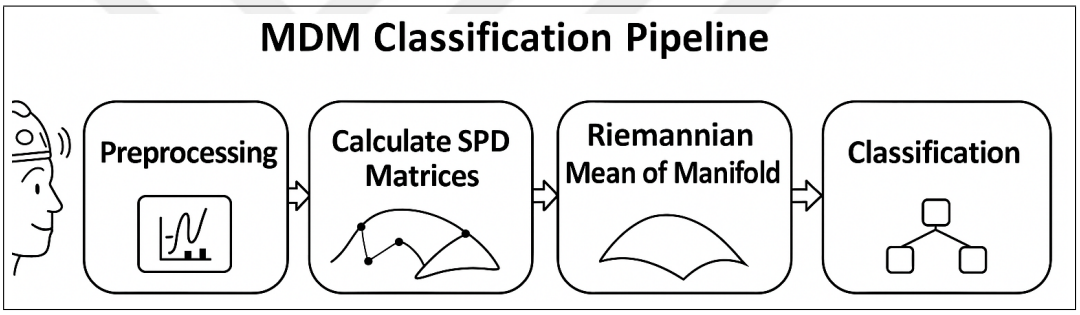

Figure 3.5 Main steps of MDM pipeline.

While the RMDM classifier operates directly on the manifold, an alternative approach projects the SPD matrices onto a tangent space, allowing standard machine learning algorithms to be applied.

3.3.2 Tangent Space

To understand the local structure of manifolds, the concept of tangent space is very helpful. It is particularly useful when working with data that lie on manifolds, such as the covariance matrices extracted from EEG signals. A tangent space allows operations like adding vectors or multiplying by scalars, which are not directly defined on the curved manifold. Because Riemannian manifolds are smoothly curved rather

than flat, classical Euclidean tools such as LDA, SVM, and others cannot be applied directly. Therefore, a set of SPD points $\{P_1, \ldots, P_L\}$ can be projected onto a flat space tangent to the manifold at a reference point G(see Figure 3.6). This tangent space projection, also called the logarithmic map, maps matrices onto the tangent space (TSM + LDA) and is defined in Eq. 3.1. Conversely, the tangent space retraction, or exponential map, defined in Eq. 3.2, maps the matrices back to the Riemannian space [6, 40, 43].

$$\log_G(P_i) = G^{-\frac{1}{2}} \log\left(G^{-\frac{1}{2}} P_i G^{-\frac{1}{2}}\right) G^{-\frac{1}{2}} \tag{3.1}$$

$$\exp_G(S_i) = G^{\frac{1}{2}} \exp\left(G^{-\frac{1}{2}} S_i G^{-\frac{1}{2}}\right) G^{\frac{1}{2}} \tag{3.2}$$

Once projected onto the tangent space, features are extracted by vectorizing the upper-triangular part of the symmetric matrix. This representation is referred to as the Log-Euclidean Tangent Vector:

$$s_i = \text{upper}\left(\log_G(P_i)\right) = \text{upper}\left(G^{-\frac{1}{2}} \log\left(G^{-\frac{1}{2}} P_i G^{-\frac{1}{2}}\right) G^{-\frac{1}{2}}\right) \tag{3.3}$$

Alternatively, the vector can be represented explicitly as:

$$s_i = \text{vec}(S_i) \in \mathbf{R}^{\frac{C(C+1)}{2} \times 1}, \quad \text{with } S_i = \log_G(P_i) \tag{3.4}$$

This vectorized representation serves as the feature input for standard machine learning algorithms. In this study, we adopted Linear Discriminant Analysis (LDA) as the classifier, using the tangent space feature vectors s_i for training and testing.

Linear Discriminant Analysis (LDA)

Linear Discriminant Analysis (LDA) is a supervised classification method that projects high-dimensional data onto a lower-dimensional subspace while maximizing class separability. It assumes normally distributed features with equal class covariance and seeks a linear decision boundary. Despite its simplicity, LDA remains highly competitive in many pattern recognition tasks [51, 52].

Given a dataset with two classes, the optimal projection vector $\mathbf{w}$ is obtained by maximizing the ratio of between-class variance to within-class variance:

$$\mathbf{w}^* = \arg\max_{\mathbf{w}} J(\mathbf{w}), \quad J(\mathbf{w}) = \frac{\mathbf{w}^\top S_B \mathbf{w}}{\mathbf{w}^\top S_W \mathbf{w}} \tag{3.5}$$

where

- $S_B = (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^\top$ is the **between-class scatter matrix**,
- $S_W = \sum_{i=1}^{2} \sum_{\mathbf{x} \in C_i} (\mathbf{x} - \boldsymbol{\mu}_i)(\mathbf{x} - \boldsymbol{\mu}_i)^\top$ is the **within-class scatter matrix**,
- $\boldsymbol{\mu}_i$ is the mean of class i.

The decision rule is then given by a linear discriminant function:

$$\delta_i(\mathbf{x}) = \mathbf{x}^\top \Sigma^{-1} \boldsymbol{\mu}_i - \tfrac{1}{2} \boldsymbol{\mu}_i^\top \Sigma^{-1} \boldsymbol{\mu}_i + \ln \pi_i \tag{3.6}$$

where Σ is the shared covariance matrix and π_i is the prior probability of class i. A sample $\mathbf{x}$ is assigned to the class with the largest discriminant value $\delta_i(\mathbf{x})$.

The main steps of the tangent space mapping and feature extraction are illustrated in Figures 3.7 and 3.8.

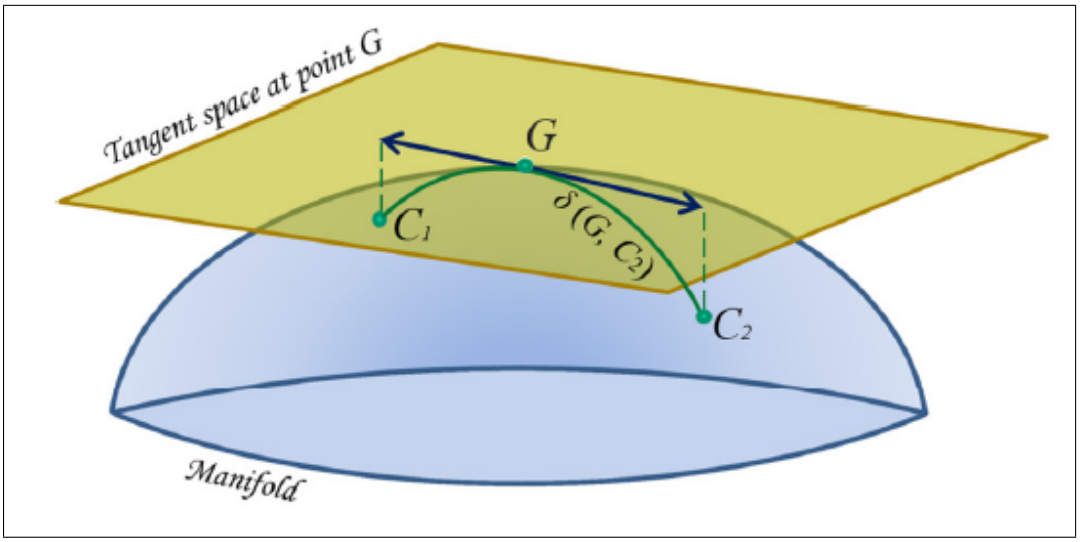

Figure 3.6 Points are mapped to the tangent space at the reference point G. (adapted from [5]).

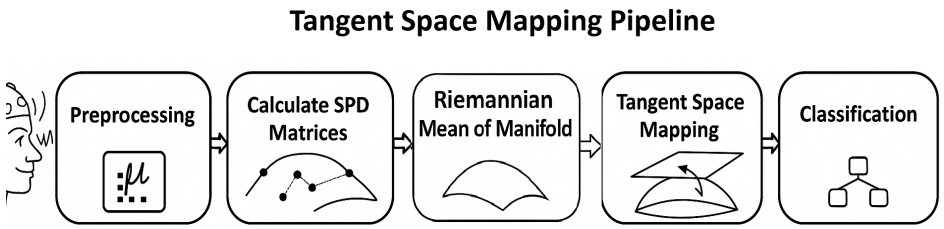

Figure 3.7 Main steps of TSM + LDA pipeline.

Building a classification system with reliable performance is an important step in machine learning. A common approach is to use a supervised pipeline, where the class labels are known. After training, the model can assign labels to new, unseen data [53].

In datasets like the one used in this study, which suffer from the curse of dimensionality and have a limited number of trials, a question arises: can we trust the classification accuracy obtained from such classifiers? Identifying predictive patterns that are statistically meaningful is important to ensure the model performs well not only on the training data but also on new data [53].

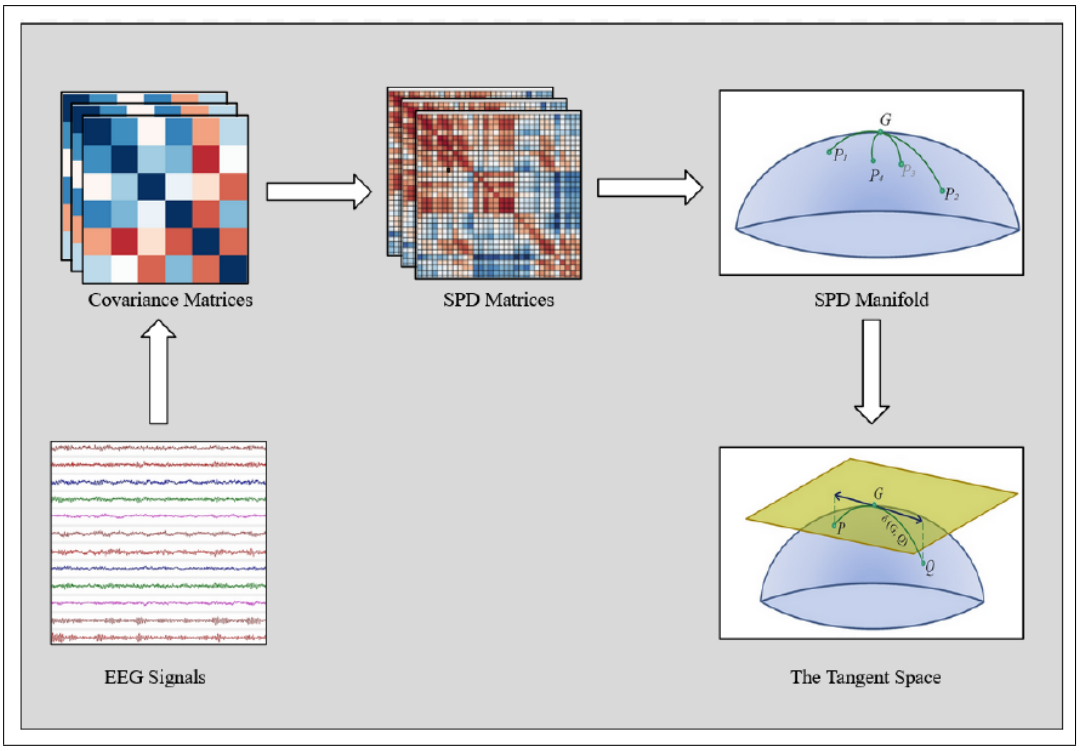

Figure 3.8 Pipeline on the SPD manifold (adapted from [6]).

3.3.3 Permutation

To examine whether the dataset has a real structure or if the observed accuracy could be due to chance, we applied the permutation testing method proposed by Markus Ojala [53]. The p-value is calculated as:

$$p = \frac{\left|\{D' \in \hat{D} \ : \ e(f, D') \leq e(f, D)\}\right| + 1}{k + 1} \tag{3.7}$$

where:

- D is the original dataset,
- $\hat{D} = \{D'_1, \ldots, D'_k\}$ is the set of k permuted datasets generated under the null hypothesis,
- f is the classifier being evaluated,
- $e(f, D)$ is the performance of classifier f on the original dataset,
- $e(f, D')$ is the performance of classifier f on a permuted dataset D'.

Under the null hypothesis that the labels are randomly assigned, Eq. 3.7 computes the empirical p-value, representing the probability of observing a classifier performance at least as high as that obtained from the original dataset. A small p-value indicates that the classifier performs significantly better than chance [53].

In practice, it is recommended to use at least 1,000 permutations for reliable statistical assessment [53]. However, due to computational constraints, we performed 250 permutations in this study.

3.3.4 False Discovery Rate (FDR)

When performing multiple statistical tests, such as applying the permutation test across many time windows or channels, the chance of obtaining significant results by random fluctuation increases. To address this, the False Discovery Rate (FDR) method can be used to control the expected proportion of false positives among all results declared significant.

Let $p_1, \ldots, p_m$ be the uncorrected p-values obtained from the permutation test (see Section 3.3.3), one for each of the m comparisons. First, the p-values are sorted in ascending order to give $p_{(1)}, p_{(2)}, \ldots, p_{(m)}$, where $p_{(1)}$ is the smallest. Following the Benjamini–Hochberg procedure [54], given an FDR threshold γ (e.g., $\gamma = 0.05$), the largest index k is found such that:

$$p_{(i)} \leq \gamma \frac{i}{m}, \quad \text{for } i \in \{1, 2, \ldots, m\}. \tag{3.8}$$

This ensures that the proportion of false discoveries among the reported significant results does not exceed γ, providing a more reliable interpretation of the permutation test outcomes.

3.3.5 Evaluation Metrics

In this study, we selected a simple and appropriate evaluation metric. Since all classes contained an equal number of trials, classification accuracy was used as the primary metric. In addition, for the multiclass classification scenario, we complemented accuracy with the confusion matrix, as it provides a more informative representation of class-wise performance.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{3.9}$$

$$\text{Confusion Matrix} = \begin{bmatrix} TP & FP \\ FN & TN \end{bmatrix} \tag{3.10}$$

$$\text{Confusion Matrix} = \begin{bmatrix} M_{11} & M_{12} & \cdots & M_{1C} \\ M_{21} & M_{22} & \cdots & M_{2C} \\ \vdots & \vdots & \ddots & \vdots \\ M_{C1} & M_{C2} & \cdots & M_{CC} \end{bmatrix} \tag{3.11}$$

Definition of labels in a confusion matrix:

- TP (True Positive): Number of samples correctly predicted as positive.
- TN (True Negative): Number of samples correctly predicted as negative.
- FP (False Positive): Number of negative samples incorrectly predicted as positive.
- FN (False Negative): Number of positive samples incorrectly predicted as negative.

For multi-class classification, let M_{ij} denote the number of samples with true class i and which are predicted as class j.

These metrics form the basis for evaluating the classification performance in the experiments described below.

3.4 Application

3.4.1 Binary Task Construction

The original dataset contains six movement classes and one rest class. To evaluate classifier performance in a binary motor imagery task, we pooled trials from all movement classes to form a single "movement" category. The number of trials from each movement class was randomly selected so that the total number of movement trials matched the number of rest trials. Specifically, we sampled an equal number of trials from each movement class to ensure class balance with the rest category. This procedure prevents class imbalance and allows a fair comparison between movement and rest conditions.

3.4.2 Movement vs. Movement Classification

For the multi-class scenario, each of the six movement classes was treated separately, and classifiers were trained to distinguish them directly. The number of trials was approximately equal across classes, so no additional balancing or pooling was required. This setup allows assessing class-wise accuracy and comparing confusion across movement types.

3.4.3 Evaluation Procedure

Before describing the results, we outline the evaluation strategy used for both binary and multi-class tasks.

Within-subject evaluation : For the binary classification of movement versus rest, we used 10-fold cross-validation within each subject. Trials from the movement classes were randomly pooled to match the number of rest trials, ensuring balanced classes. Each fold was used once as a test set while the remaining folds were used for

training, and performance metrics were averaged across folds. These same folds were also used in the permutation test to assess statistical significance.

Cross-subject evaluation : To assess generalization across subjects, we applied a leave-one-subject-out (LOSO) procedure. In each iteration, data from one subject were used as the test set, and the remaining subjects formed the training set. This allowed us to evaluate how well the classifiers generalize to unseen subjects.

Sliding window : We applied a sliding-window approach to extract features from EEG signals, following the time-resolved decoding framework [55]. The window center started at 1.0 s and moved until 2.00 s, with a step of 0.1 s. Each segment used a half-window of 0.05 s for non-overlapping windows, where each segment is independent (Figure 3.9). For overlapping windows, segments used half-windows of 0.2 s, 0.35 s, and 0.5 s, where consecutive windows share data (Figure 3.10). This setup allows evaluating the effect of window overlap and segment length on feature extraction and classification performance.

This section presents how the evaluation and classification procedures were applied to both binary and multi-class tasks, including feature extraction, preprocessing, and model training.

3.4.4 Feature Ablation

Feature ablation is a technique used to assess the importance of individual features in a predictive model. By systematically removing or altering each feature and observing the impact on model performance, we can identify which features contribute most significantly to the model's accuracy. This approach helps in understanding the relevance of each feature and aids in simplifying models by retaining only the most impactful variables [56].

To perturb the pipeline and evaluate feature importance, we applied two main

modifications:

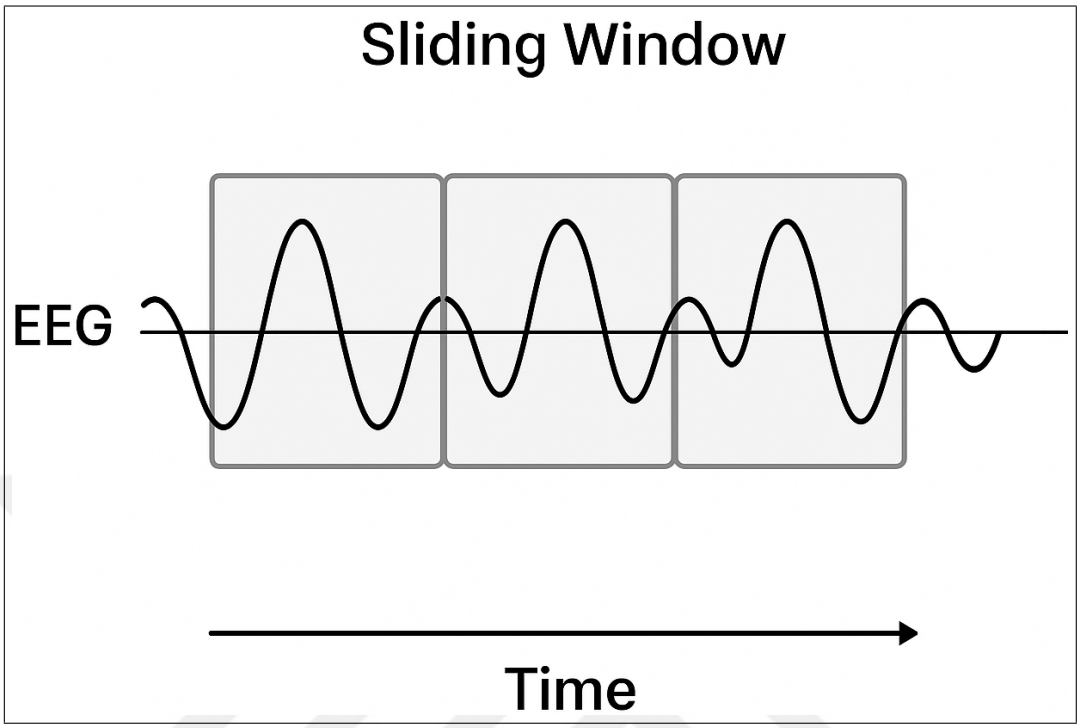

Figure 3.9 Illustration of the non-overlapping sliding window approach for EEG feature extraction. Each window segment is independent.

1. **MRCP-focused perturbation:** The xDAWN spatial filter was replaced by simple covariance matrices (SCM) computed from preprocessed EEG signals in the 0.3–3 Hz band.

2. **Oscillatory band perturbation:** The xDAWN filter was replaced by SCM and features were extracted from the 7–30 Hz frequency band.

These perturbations allow us to assess the contribution of spatial filtering and frequency band selection to the overall classification performance.

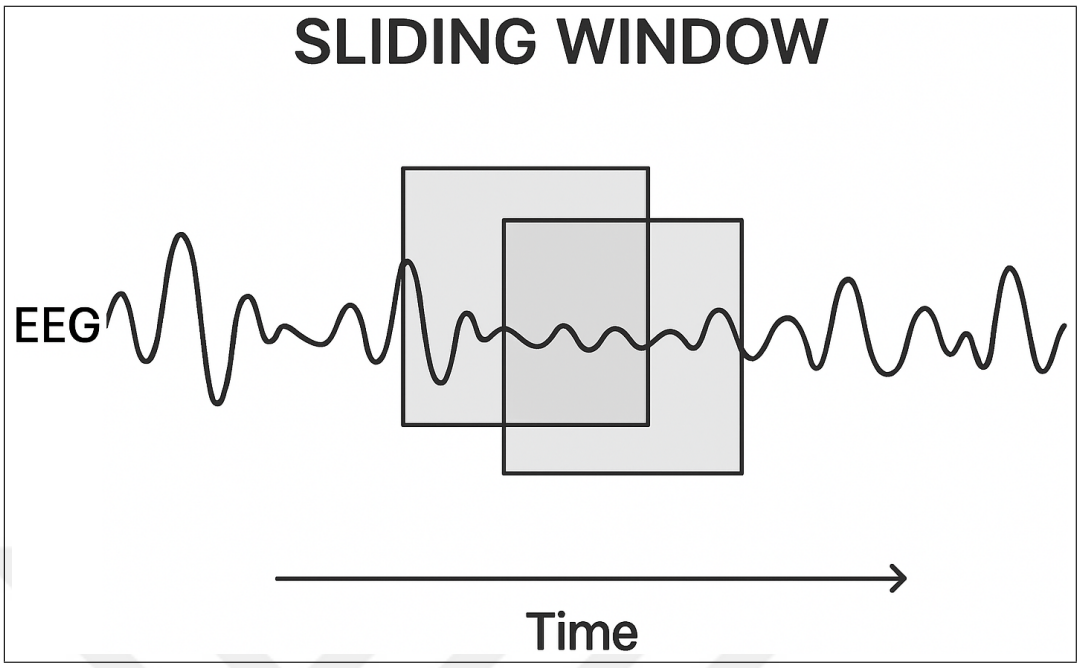

Figure 3.10 Illustration of the overlapping sliding window approach for EEG feature extraction. Consecutive windows share part of their data.

Feature ablation was applied only for the within-subject binary task. The procedures described above, including preprocessing, feature extraction, classification, and feature ablation, were applied to the dataset. The following chapter presents the outcomes of these analyses.

4. RESULTS

In this chapter, we present the results obtained from the proposed pipelines. We first report the performance of the models in terms of classification accuracy and confusion matrices. Then, we analyze the effect of different preprocessing strategies, including frequency band selection and covariance estimation methods. Lastly, we examined the outcomes across subjects to see the overall trend in performance.

4.1 Binary Classification

4.1.1 Within-Subject Analysis

The classification performance was first evaluated using the two proposed pipelines: the Minimum Distance to Mean (MDM) classifier and the Tangent Space Mapping with Linear Discriminant Analysis (TSM + LDA). Both methods were tested across multiple temporal windows of EEG data.

Table 4.1 summarizes classification accuracy within–subjects (mean, median, and standard deviation). Performance consistently increased with window length from 100 ms to 1 s for both classifiers, with TSM + LDA achieving slightly higher accuracy than MDM, particularly at longer windows. The latency section of Table 4.1 reports mean and standard deviation of classification latency. MDM showed slightly higher latency for smaller windows (up to 0.05s), Both classifer showed slightly higher latency with smaller windows.

Permutation tests with FDR correction confirmed that all reported differences were statistically significant ($p < 0.05$). Both MDM and TSM + LDA achieved robust decoding of motor imagery EEG demonstrating a modest advantage at longer windows.

Table 4.1
Summary statistics for each window using MDM and TSM + LDA classifiers.

Window	Mean		Median		STD	
	MDM	TSM + LDA	MDM	TSM + LDA	MDM	TSM + LDA
100 ms	0.6221	0.6148	0.6265	0.6136	0.0334	0.0301
400 ms	0.6830	0.6721	0.6841	0.6735	0.0562	0.0405
700 ms	0.6943	0.7018	0.6811	0.7114	0.0643	0.0637
1 s	**0.7060**	**0.7117**	**0.7000**	**0.7068**	0.0748	0.0709
Latency (s)	Mean		STD			
100 ms	1.4533	1.4000	0.2825	0.2928		
400 ms	1.3133	1.3400	0.2696	0.1993		
700 ms	1.3800	**1.2800**	0.2704	0.2077		
1 s	**1.2733**	1.3467	0.2685	0.2232		

Figures 4.1 illustrate subject-level variability in classifier performance across time centers and window lengths. Both MDM and TSM + LDA show higher median accuracies with slightly increased variability for longer windows. Accuracy tends to decrease at later time points, particularly after 1.5 s.

Pairwise Kruskal–Wallis post-hoc comparisons across window lengths indicated that classification accuracy differed between most window pairs, with longer windows generally achieving higher accuracy for both MDM and TSM + LDA. For MDM, the 100 ms and 400 ms windows had lower accuracy, 700 ms was slightly higher, and 1 s achieved the highest accuracy, though it did not differ significantly from 700 ms. A similar pattern was observed for TSM + LDA, with the longest windows showing higher accuracy. Overall, the 700 ms and 1 s windows did not differ significantly, indicating that accuracy plateaued at longer window lengths.

To investigate which features contributed most to this performance, we conducted a feature ablation analysis.

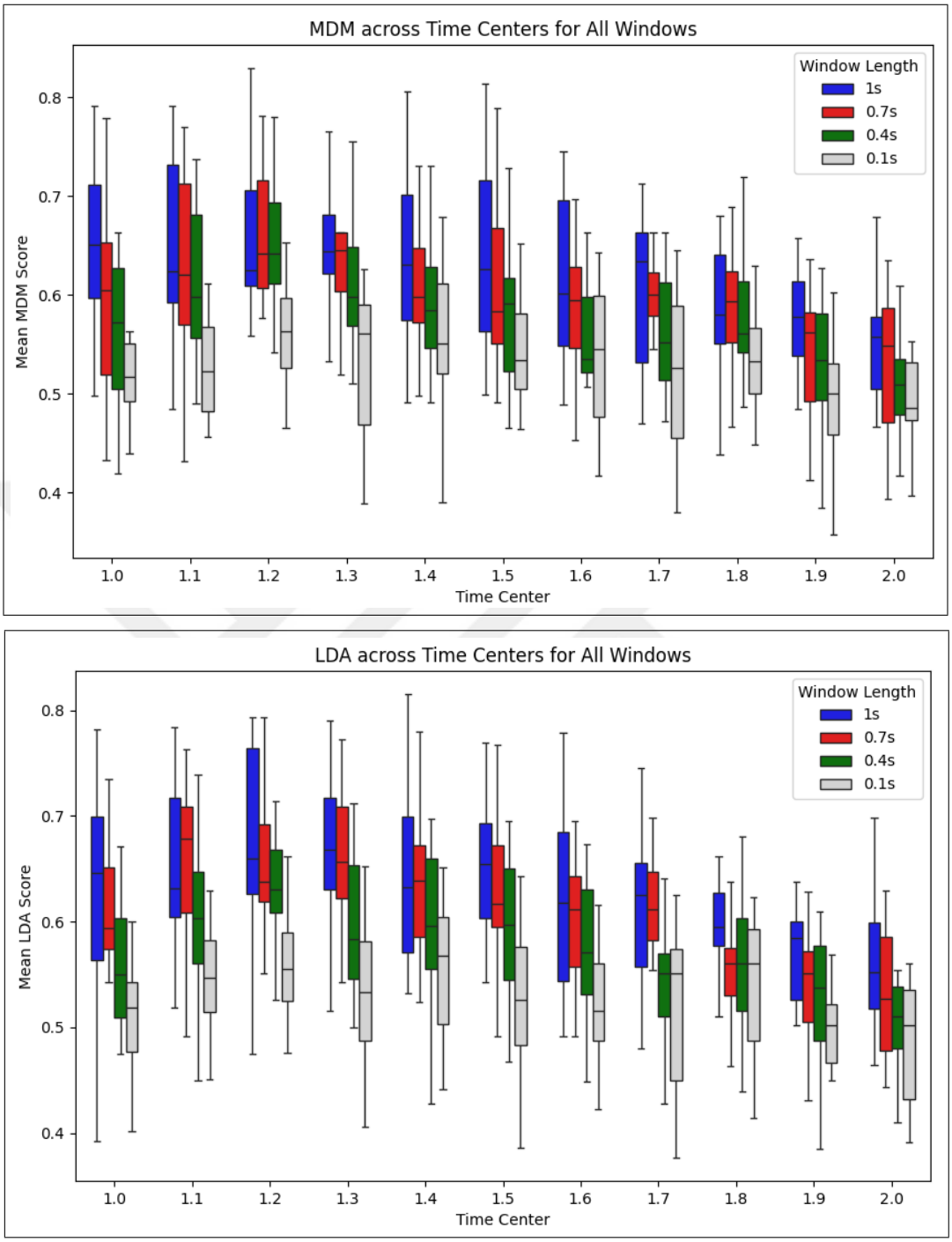

Figure 4.1 Boxplots showing classification performance across time centers for different window lengths, using (**Top**) MDM and (**Bottom**) TSM + LDA classifiers. Four window widths are compared: 100 ms, 400 ms, 700 ms, and 1 s. Each box represents the distribution of scores, highlighting both central tendency and variability as a function of window duration.

4.1.2 Feature Ablation

4.1.2.1 Effect of Removing xDAWN Spatial Filter. To evaluate the contribution of the xDAWN spatial filter, it was replaced with simple covariance matrices (SCM), while all other processing steps were unchanged. Table 4.2 reports the resulting classification accuracy and latency for MDM and TSM + LDA across different window lengths.

Table 4.2
Summary statistics for each window using MDM and TSM + LDA classifiers after replacing the xDAWN spatial filter with simple covariance matrices (SCM).

Window	Mean		Median		STD	
	MDM	TSM + LDA	MDM	TSM + LDA	MDM	TSM + LDA
100 ms	0.6207	0.6117	0.6091	0.6174	0.0413	0.0324
400 ms	0.6054	0.6054	0.5977	0.6152	0.0457	0.0478
700 ms	**0.6164**	**0.6259**	**0.5927**	**0.6523**	0.0616	0.0652
1 s	0.6137	0.6115	0.6098	0.6098	0.0511	0.0572
Latency (s)	Mean		STD			
100 ms	1.4600	1.4467	0.2230	0.2588		
400 ms	1.4400	1.3133	0.2324	0.2642		
700 ms	**1.3667**	**1.3000**	0.2410	0.2420		
1 s	1.4333	1.3667	0.3086	0.2554		

Removing xDAWN reduced mean and median accuracies for both classifiers (MDM: 0.7060 to 0.6137; TSM + LDA: 0.7117 to 0.6115), with slightly reduced variability (standard deviation). Latency remained mostly stable.

Figure 4.4 displays median FDR-corrected p-values across time centers for MDM and TSM + LDA, with and without xDAWN. Removing xDAWN reduced the number of significant time points, highlighting that spatial filtering enhances the detectability of discriminative EEG activity. The heatmaps in Figures 4.2 and 4.3 provide a complementary view of classifier performance across subjects, highlighting the temporal consistency of significant decoding rather than absolute accuracy. Removing

xDAWN reduced both the magnitude and stability of significant decoding, particularly for shorter windows. Overall, SCM preserves the general trend across windows, but xDAWN clearly improves accuracy and significance, emphasizing the importance of spatial filtering in extracting discriminative EEG features.

Next, we assessed the effect of using SCM in the 7–30 Hz frequency band. Table 4.3 summarizes classification accuracy and latency for MDM and TSM + LDA.

4.1.2.2 Feature Ablation in the 7–30 Hz Band. In this band, MDM reached 0.6727 (mean) at 1 s, while TSM + LDA consistently outperformed MDM, reaching 0.7311 (mean) and 0.7455 (median) at 1 s, with slightly higher variability across subjects. Latency increased, reaching approximately 1.7000 s for MDM and 1.8200 s for TSM + LDA. reflecting the characteristic ERD/ERS patterns, which occur around movement execution onset and shortly after.

Table 4.3
Summary statistics for each window using MDM and TSM + LDA classifiers after removing xDAWN in the 7–30 Hz frequency band.

Window	Mean		Median		STD	
	MDM	TSM + LDA	MDM	TSM + LDA	MDM	TSM + LDA
100 ms	0.6348	0.6503	0.6356	0.6348	0.0397	0.0556
400 ms	0.6639	0.6934	**0.6795**	0.6864	0.0487	0.0698
700 ms	0.6690	0.7152	0.6614	0.6970	0.0660	0.0646
1 s	**0.6727**	**0.7311**	0.6735	**0.7455**	0.0562	0.0748
Latency (s)	Mean		STD			
100 ms	1.5933	**1.6867**	0.2658	0.2200		
400 ms	**1.5800**	1.7400	0.2569	0.1682		
700 ms	1.7000	1.7600	0.1964	0.2063		
1 s	1.7000	1.8200	0.2726	0.1474		

Figure 4.5 illustrates the median FDR-corrected p-values across time centers for both classifiers and frequency bands, whereas heatmap in Figure 4.6 highlights the time latencies where the effects reach their maximum magnitude. Changing the frequency

band affects the number of significant time points, showing that feature selection and frequency content jointly influence discriminability.

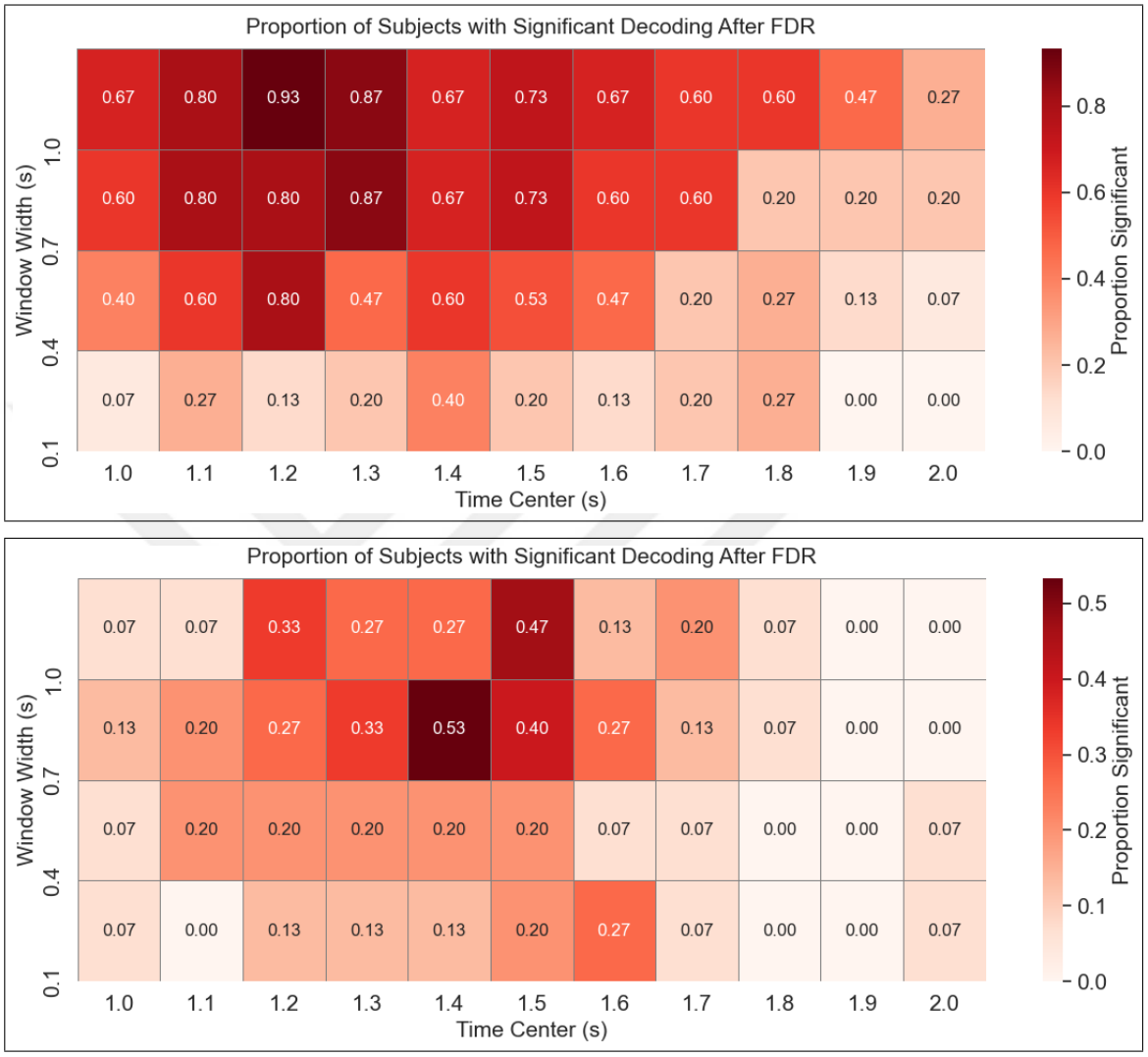

Figure 4.2 Proportion of subjects with significant decoding after FDR correction for the TSM+LDA pipeline. The x-axis indicates the center of the time window, and the y-axis shows the window width. Overall, the patterns illustrate changes in the extent of significant decoding. With xDAWN spatial filtering (**Top**), slow cortical potentials (0.3–3 Hz) yield stronger and more consistent decoding. Without xDAWN filtering (**Bottom**), decoding is weaker and less consistent, with smaller regions of significance.



Figure 4.3 Proportion of subjects with significant decoding after FDR correction for the MDM pipeline. The x-axis indicates the center of the time window, and the y-axis shows the window width. Overall, the patterns illustrate changes in the extent of significant decoding. With xDAWN spatial filtering (**Top**), slow cortical potentials (0.3–3 Hz) yield stronger and more consistent decoding. Without xDAWN filtering (**Bottom**), decoding is weaker and less consistent, with smaller regions of significance.

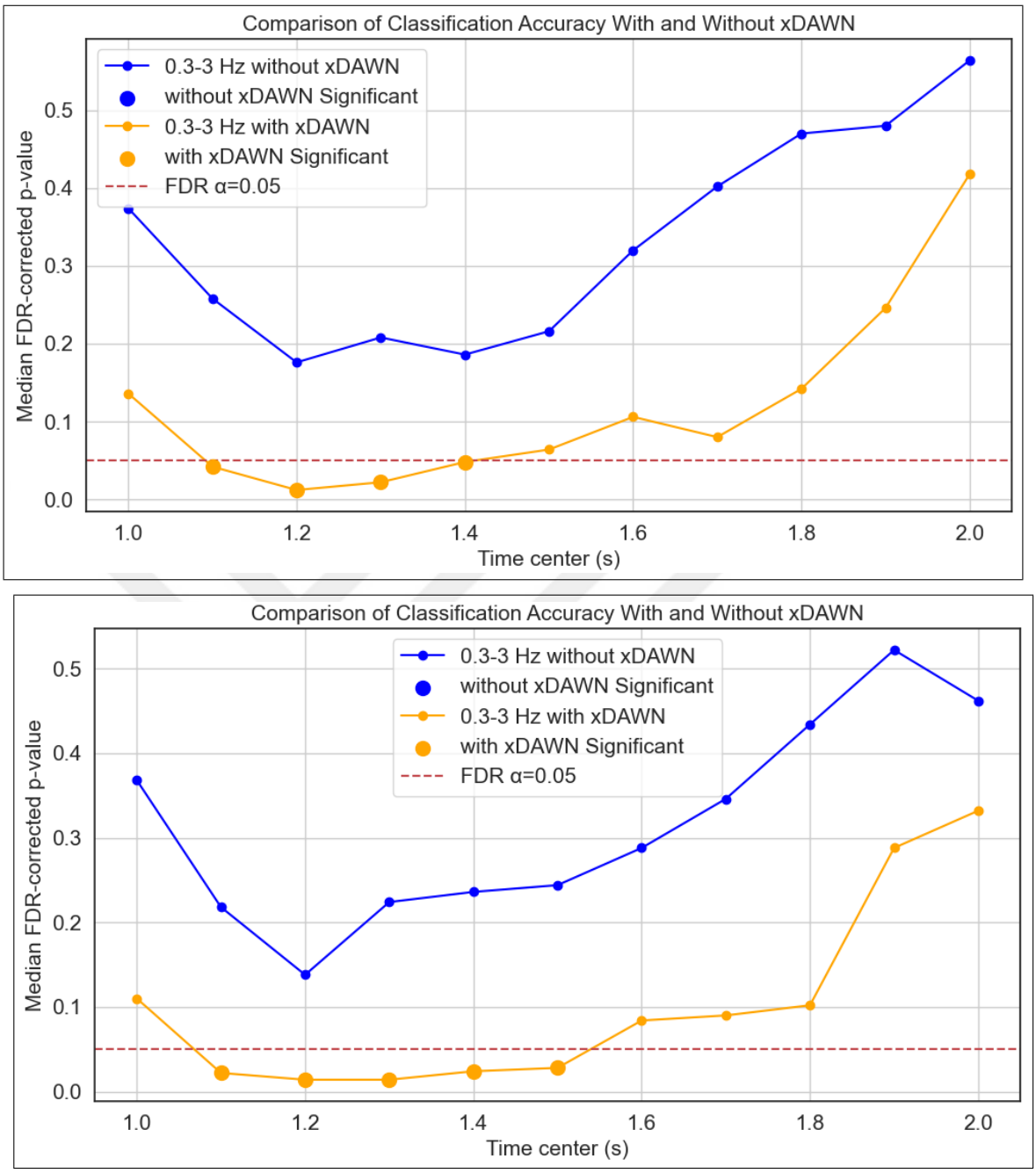

Figure 4.4 Median FDR-corrected p-values at each time center for MDM (**Top**) and TSM + LDA (**Bottom**) classification, pooling results from all four window lengths (100, 400, 700 ms, 1 s), with and without the xDAWN spatial filter (0.3–3 Hz). Markers indicate time points that reached statistical significance after FDR correction ($\alpha = 0.05$). The red dashed line indicates the FDR significance threshold.

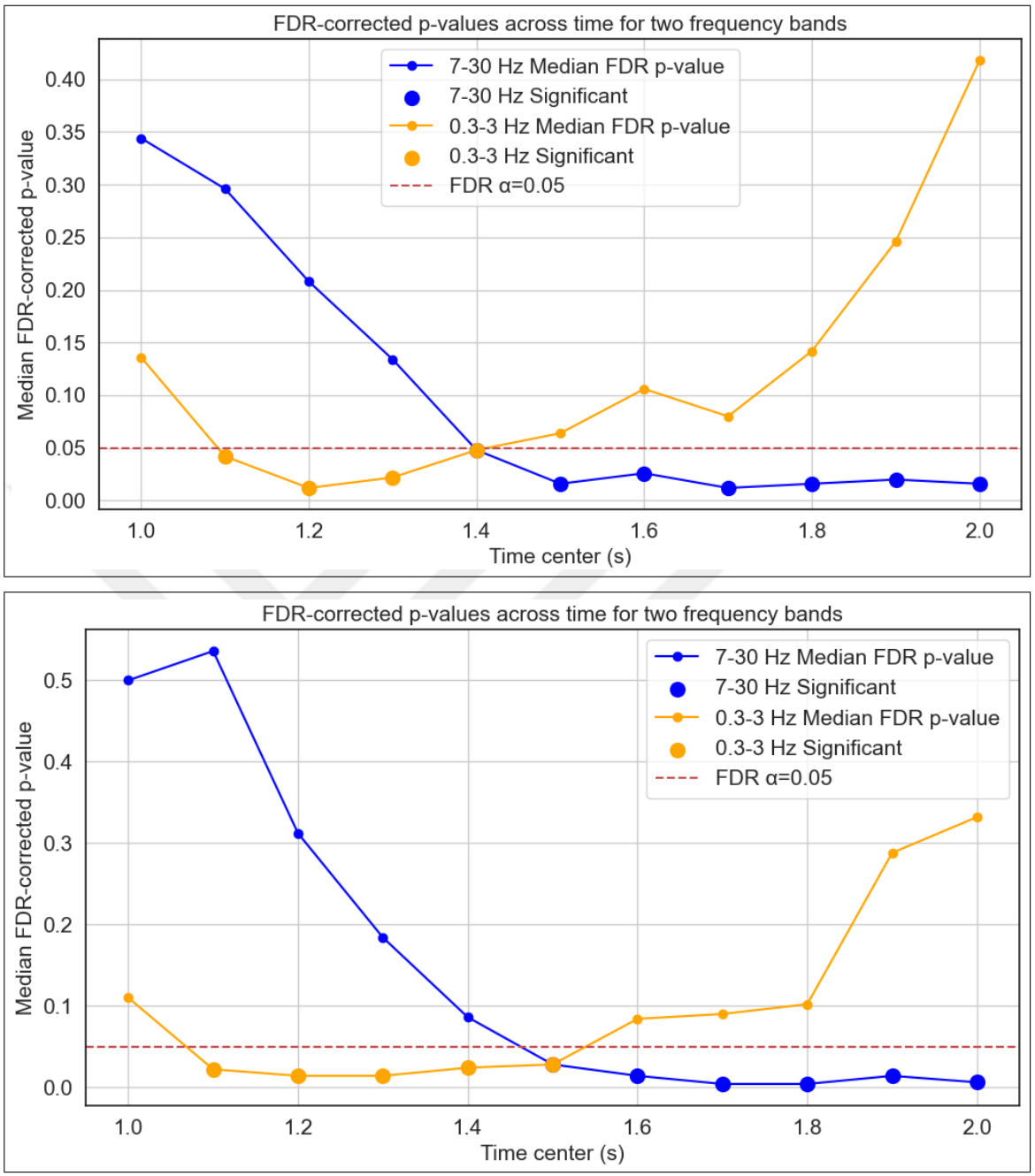

Figure 4.5 Median FDR-corrected p-values at each time center for MDM (**Top**) and TSM + LDA (**Bottom**) classification, pooled across all four window lengths (100, 400, 700 ms, 1 s) and two frequency bands (0.3–3 Hz and 7–30 Hz). Markers indicate time points reaching statistical significance after FDR correction ($\alpha = 0.05$). Red dashed line shows the FDR significance threshold.

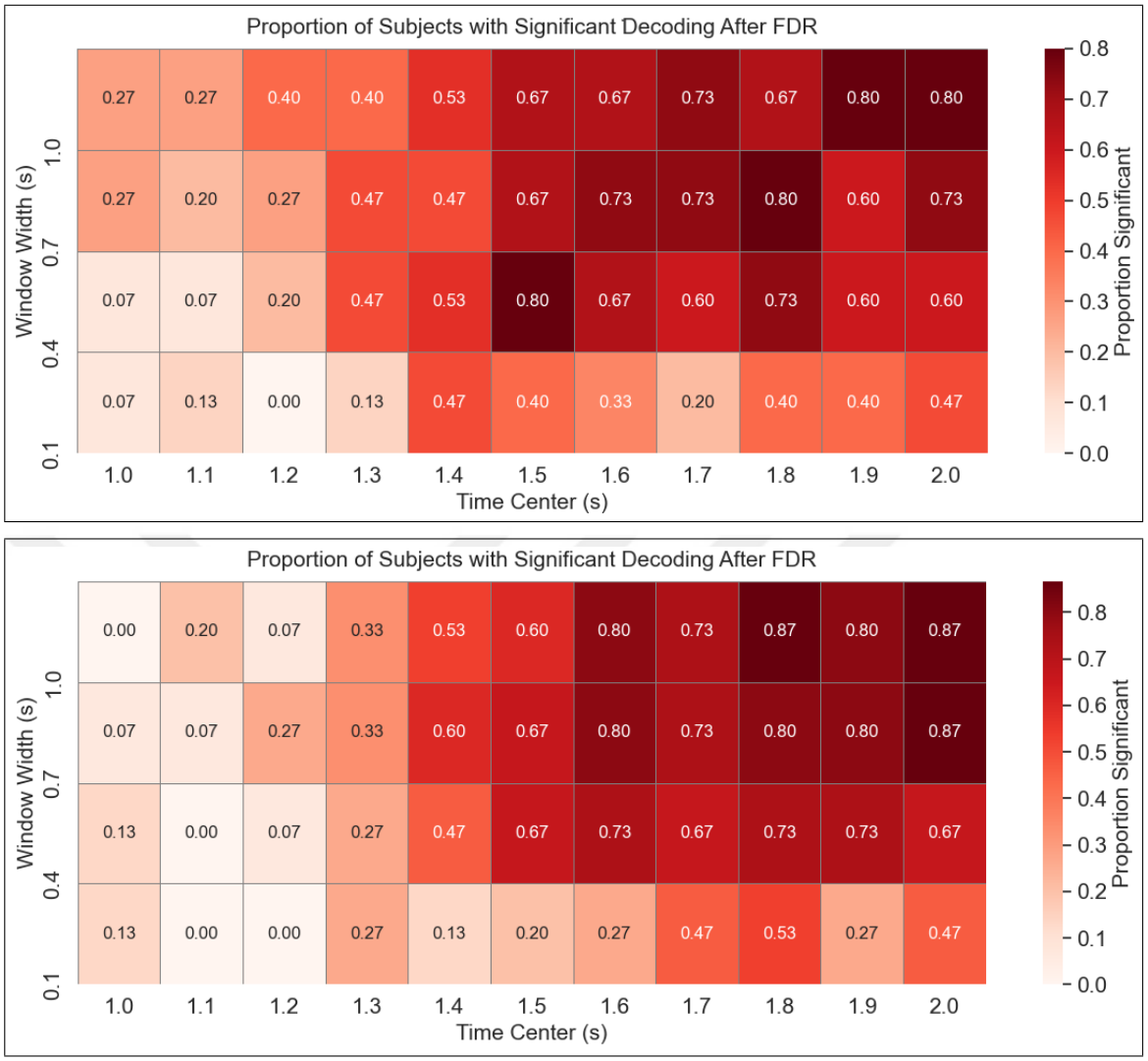

Figure 4.6 Proportion of subjects with significant decoding after FDR correction for the MDM (**Top**) and TSM + LDA (**Bottom**) pipelines in the oscillatory activity band (7–30 Hz). Compared to Figure 4.2 (**Top**) and Figure 4.3 (**Top**), this heatmap shows a temporal shift in the peak significance from around 1.2 s to approximately 1.9 s.

A histogram analysis of significant decoding centers further confirmed this temporal shift(Figure 4.7): xDAWN (0.3–3 Hz) produced earlier peaks around 1.2–1.4 s, whereas oscillatory features (7–30 Hz) peaked later, around 1.6–1.8 s, consistent with the latency differences reported in Table 4.3.

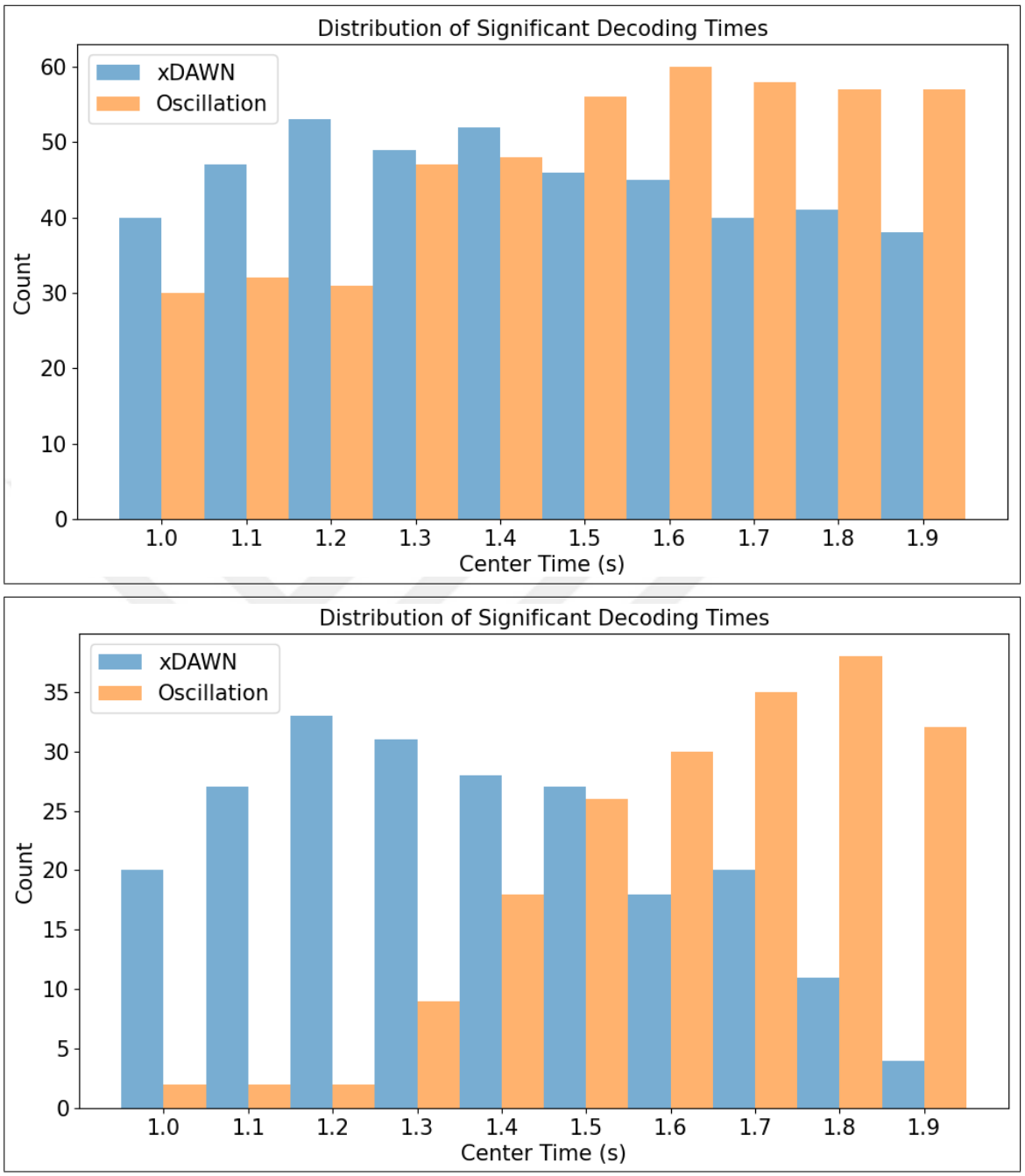

Figure 4.7 Histogram of significant decoding times across subjects. (**Top**): MDM; (**Bottom**): TSM + LDA. xDAWN peaks occur earlier (∼1.2–1.4 s) than oscillatory features (∼1.6–1.8 s).

For TSM + LDA, the results are consistent with the Wilcoxon signed-rank test ($p < 0.05$) and with the latency differences reported in Table 4.3. However, the Wilcoxon signed-rank test for MDM was not significant ($p > 0.05$), which may reflect its sensitivity to noise. Projecting the SPD matrices to the tangent space (TSM) can

potentially mitigate this sensitivity and improve classification stability.

Overall, these ablation studies highlight that the xDAWN spatial filter substantially improves classification accuracy and significance, especially for MDM. Using the 7–30 Hz frequency band further enhances performance for TSM + LDA, particularly at longer windows. However, this improvement comes at the cost of increased latency, approximately 1.75 s, compared to around 1.25 s when using the xDAWN spatial filter with the 0.3–3 Hz band.

This result is consistent with the known physiology: MRCPs primarily occur before movement onset, whereas activity in the 7–30 Hz oscillation band begins to reflect motor information slightly later, around the time of movement execution, which explains the increased latency observed in the ablation results.

We next compare all three pipelines to highlight these effects systematically.

Comparing Tables 4.1, 4.2, and 4.3, we can observe clear differences in classification performance and latency across three conditions: the main pipeline (xDAWN + 0.3–3 Hz), SCM with slow cortical potentials (SCM + 0.3–3 Hz), and SCM with the oscillatory band (SCM + 7–30 Hz).

Latency. In the main pipeline with xDAWN, both classifiers reached their minimum mean latencies earlier (1.2733–1.4533 s), indicating that discriminative activity was detected relatively quickly. Removing xDAWN while keeping the same slow cortical band (SCM + 0.3–3 Hz) slightly increased latencies (1.30–1.46 s), suggesting that spatial filtering accelerates early detection. Shifting to the oscillatory band with SCM further delayed discriminability, with latencies ranging from 1.58–1.82 s, indicating that oscillatory activity emerges later than slow cortical potentials.

Model performance. MDM benefited strongly from xDAWN in the slow cortical band, achieving its highest mean accuracy (0.7060) at the 1 s window. Removing xDAWN in the same band decreased performance moderately (MDM mean $\sim$0.6137),

reflecting the contribution of spatial filtering to robust low-frequency feature extraction. TSM + LDA showed a similar trend, with mean accuracy dropping from 0.7117 to 0.6115 when xDAWN was removed. Interestingly, switching to the oscillatory band with SCM restored or improved TSM + LDA performance, reaching a mean of 0.7311 and a median of 0.7455 at 1 s, while MDM performance remained slightly lower than the baseline (0.6727 mean).

Variability. Standard deviations were generally moderate across conditions. Removing xDAWN slightly increased inter-subject variability in the slow cortical band, particularly for TSM + LDA at shorter windows (100 ms, 0.0324 vs. 0.0301 with xDAWN). In the oscillatory band, variability increased further for TSM + LDA, while MDM remained relatively stable.

Overall. These findings suggest a complementary picture: xDAWN enhances early detection and boosts low-frequency feature performance, while SCM in the oscillatory band shifts decoding later in time but can improve accuracy for TSM + LDA. Thus, slow cortical activity with spatial filtering allows earlier but slightly less variable detection, whereas oscillatory dynamics provide strong discriminability at later times (see Figures 4.4 and 4.5).

Note on Feature Ablation for Other Cases. Feature ablation analyses were conducted only for the within-subject binary case in order to specifically evaluate the contributions of the xDAWN spatial filter and frequency band selection, where statistical power is highest. For cross-subject and multiclass analyses, feature ablation was not performed due to computational limitations and the exploratory nature of these experiments. Nevertheless, the patterns observed in the within-subject analysis are expected to qualitatively generalize to these other scenarios.

4.1.3 Cross-Subject Binary Case

In addition to the within-subject analysis, we extended this study to investigate the performance of the models across different subjects. Cross-subject analysis evaluates how well a model trained on data from some participants can generalize to other participants. In this section, we report the results of the cross-subject binary classification, focusing on the ability of the models to discriminate between movement and rest conditions across participants.

In the cross-subject binary analysis, it can be seen that both MDM and TSM + LDA classifiers increase decoding accuracy when the analysis window is longer, with the best performance around 700 ms to 1 s (Table 4.4). The peak latency stays almost the same for all windows (1.3–1.45 s), which shows that the timing of movement-related EEG features is mostly similar across participants. There is some variability between subjects, but overall the trends are consistent. Figure 4.8 shows these temporal changes, where short windows (100 ms) do not give enough information, while longer windows collect more discriminative signal and improve cross-subject generalization.

Table 4.4
Summary statistics for each window using MDM and TSM + LDA classifiers in a cross-subject study.

Window	Mean		Median		STD	
	MDM	TSM + LDA	MDM	TSM + LDA	MDM	TSM + LDA
100 ms	0.6177	0.6167	0.6140	0.6121	0.0319	0.0268
400 ms	0.6530	0.6472	0.6379	0.6525	0.0473	0.0444
700 ms	**0.6778**	0.6745	0.6607	0.6667	0.0513	0.0607
1 s	0.6732	**0.6763**	**0.6842**	**0.6780**	0.0621	0.0624
Latency (s)	Mean		STD			
100 ms	**1.3133**	1.3067	0.2356	0.2154		
400 ms	1.4533	1.4067	0.3563	0.3595		
700 ms	1.4200	**1.3600**	0.3212	0.3269		
1 s	1.4533	**1.3600**	0.3563	0.2694		

4.2 Multiclass Classification

In this section, we extend the analysis from binary motor imagery to multiclass decoding, where each movement class is distinguished from the others. Multiclass classification is more challenging, as the classifier must discriminate between multiple motor intentions rather than just movement versus rest. In the multiclass within–subject analysis, both MDM and TSM + LDA classifiers show low decoding accuracy even for longer analysis windows, including the 1 s window (Table 4.5). The accuracy remains around 22–24%, indicating that distinguishing multiple movement classes is considerably harder than binary classification. This is likely because all movements originate from the same limb, producing very similar EEG patterns that are difficult to separate spatially. In addition, the small number of trials per class (approximately 55) limits the classifiers' ability to reliably learn distinct patterns.

The peak latency also becomes slightly later with longer windows, suggesting that the most informative signals appear later due to inter-class variability (which is normalized when movements are pooled in the binary study). Confusion matrices (Figure 4.9) reveal that most errors occur between similar movement classes, confirming that EEG patterns for these movements overlap considerably. TSM + LDA shows slightly higher mean accuracy than MDM, but both classifiers follow similar trends, highlighting the inherent challenge of multiclass decoding even within subjects.

We also examined multiclass decoding in the cross–subject setting. Table 4.6 shows that both MDM and TSM + LDA achieve low accuracy (approximately 20%), with small standard deviations, indicating that performance is consistently limited across participants. Temporal misalignments and individual differences further reduce class separability. Confusion matrices (Figure 4.10) confirm frequent misclassifications between movement–related classes, emphasizing the difficulty of generalizing multiclass decoding to unseen subjects.

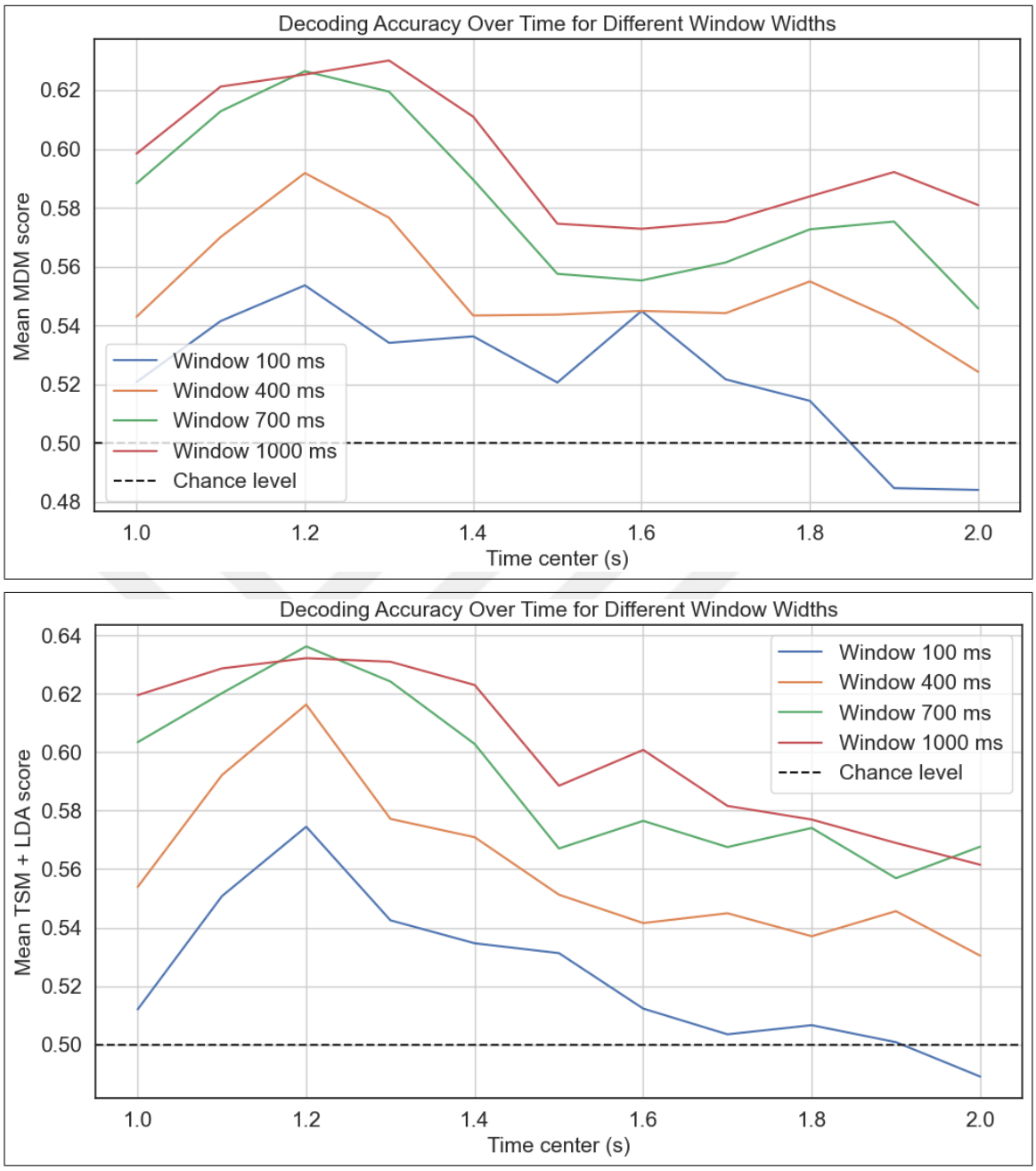

Figure 4.8 Decoding accuracy over time using MDM (**Top**) and TSM + LDA (**Bottom**) for four different window widths (100 ms, 400 ms, 700 ms, and 1 s). The x-axis represents the center of the analysis window in seconds, and the y-axis shows mean classification accuracy across subjects. The dashed line indicates chance level (50%). Both classifiers show that decoding accuracy increases to a maximum around 1.2 s for all window widths, then gradually decreases. Longer windows (700 ms and 1 s) achieve higher performance, whereas the 100 ms window shows the lowest accuracy.

We also examined multiclass decoding in the cross–subject setting. Table 4.6 shows that both MDM and TSM + LDA achieve low accuracy (approximately 20%), with small standard deviations, indicating that performance is consistently limited

across participants. Temporal misalignments and individual differences further reduce class separability. Confusion matrices (Figure 4.10) confirm frequent misclassifications between movement–related classes, emphasizing the difficulty of generalizing multiclass decoding to unseen subjects.

Table 4.5
Summary statistics for two windows using MDM and TSM + LDA classifiers in a within-subject multiclass study.

Window	Mean		Median		STD	
	MDM	TSM + LDA	MDM	TSM + LDA	MDM	TSM + LDA
100 ms	0.2221	0.2206	0.2224	0.2161	0.0172	0.0208
1 s	**0.2377**	**0.2446**	**0.2303**	**0.2399**	0.0373	0.0347
Latency (s)	Mean		STD			
100 ms	**1.3400**	**1.2933**	0.2947	0.2463		
1 s	1.4533	1.4933	0.2503	0.2939		

An interesting observation in the within-subject MDM confusion matrix (Figure 4.9) is the bias toward Pronation and Opening. Many trials from other true classes are predicted as these movements, possibly because these actions produce stronger or more consistent EEG signatures than others. This highlights that classifier tendencies can reveal underlying differences in feature strength across movement classes.

Overall, these results demonstrate that multiclass decoding is challenging when movements are similar, trial numbers are limited, and only one limb is recorded. Performance decreases further in cross-subject settings due to temporal and individual variability. Future studies could improve separability by increasing the number of trials, adding more EEG channels, or recording movements from both limbs to enhance spatial resolution.

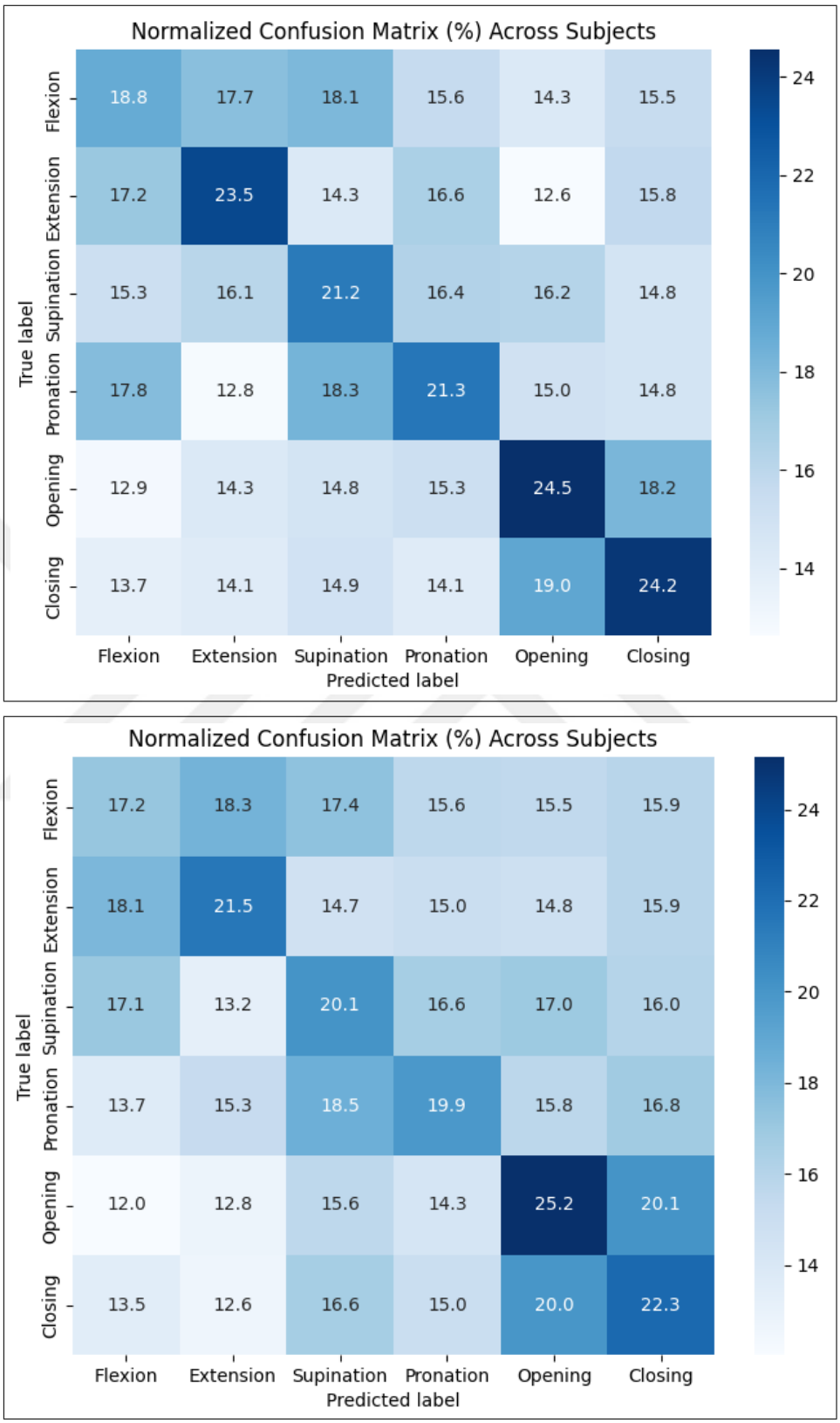

Figure 4.9 Normalized confusion matrices (in %) for the multiclass within-subject study using TSM + LDA (**Top**) and MDM (**Bottom**) classifiers with a 1 s analysis window. The matrices show the average classification performance across subjects for each movement class, where the diagonal represents correctly classified trials and off-diagonal entries indicate misclassifications.

Table 4.6
Summary statistics for each window using MDM and TSM + LDA classifiers in the cross-subject multiclass study.

Window	Mean		Median		STD	
	MDM	TSM + LDA	MDM	TSM + LDA	MDM	TSM + LDA
100 ms	0.1995	0.2067	0.2000	0.2018	0.0138	0.0209
1 s	**0.2133**	**0.2122**	**0.2152**	**0.2175**	0.0202	0.0194
Latency (s)	Mean		STD			
100 ms	**1.2133**	**1.2133**	0.2850	0.1356		
1 s	1.3933	1.3000	0.2963	0.2420		

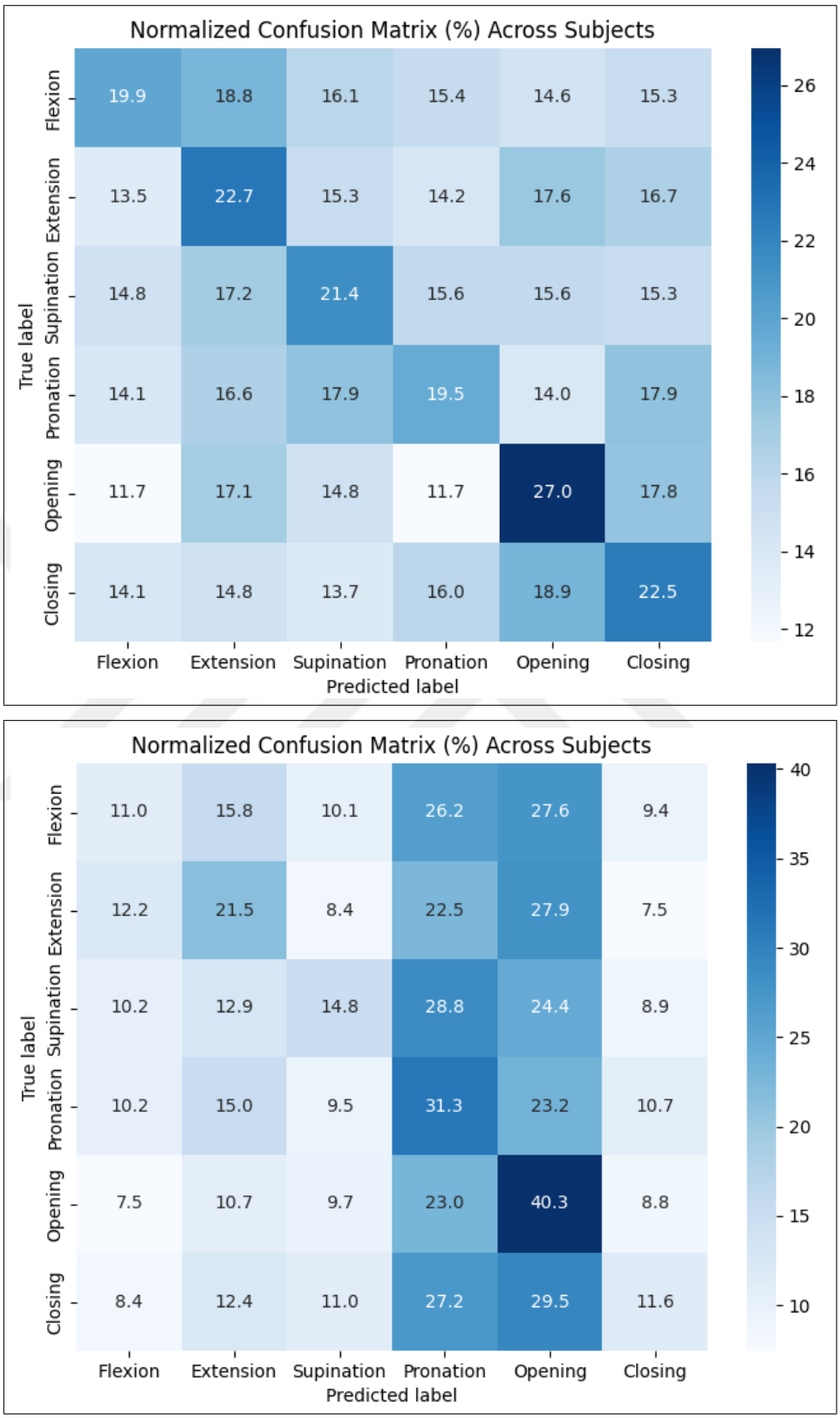

Figure 4.10 Normalized confusion matrices (in %) for the multiclass cross-subject study using TSM + LDA (**Top**) and MDM (**Bottom**) classifiers with a 1 s analysis window. The matrices show the average classification performance across subjects for each movement class, where the diagonal represents correctly classified trials and off-diagonal entries indicate misclassifications.

5. DISCUSSION

This study looked at the possibility of detecting movement intention from EEG signals using movement-related cortical potentials (MRCP) together with the xDAWN spatial filter and Riemannian geometry-based classifiers. The main aim was to see if these methods can reliably separate imagined upper limb movements from rest, and if they work across different movement classes and subjects. Overall, our results show that the proposed pipeline can detect movement intention above chance level. Performance varies across movement types and subjects, which shows both the potential and the current limitations of this approach. These findings contribute to understanding how spatial filtering and Riemannian classifiers can be combined for EEG-based brain-computer interfaces (BCIs).

Within-Subject Binary Classification

Our results show that both MDM and TSM + LDA classifiers were able to detect movement intention reliably in the within-subject binary task. Using xDAWN clearly improved performance compared to raw covariance matrices, which aligns with previous studies showing that xDAWN spatial filtering enhances evoked potentials and improves the signal-to-noise ratio [44]. By focusing on the most informative spatial components, xDAWN makes the EEG signal easier to decode. Looking at the time windows, longer segments (700 ms–1 s) gave better accuracy ($p < 0.05$), suggesting that more temporal information helps to capture the MRCP patterns. The latency analysis indicated that discriminative activity appeared around 1.2–1.3 s. Overall, these results show that combining xDAWN with Riemannian classifiers is an effective approach for detecting movement preparation, directly addressing the challenges described in Problems 1 and 2 in the introduction. This also suggests that for practical applications, selecting an optimal window size is important to balance early detection with classification performance.

Oscillatory vs. Slow Cortical Bands

When we changed the frequency band from slow cortical potentials (0.3–3 Hz) to oscillatory activity (7–30 Hz), the MDM classifier showed a slight decrease in performance, while the TSM + LDA classifier kept similar or slightly better accuracy. Interestingly, the peak decoding appeared later, around 1.7 s. This suggests that slow cortical potentials are better for detecting movement intention earlier espatially by using TSM + LDA ($p < 0.05$), whereas higher-frequency oscillations may provide additional information that emerges later in the motor preparation process [31]. These findings support the use of MRCPs for early intention detection, while also highlighting the complementary role of oscillatory activity. This comparison shows that including different frequency bands could help build more flexible BCIs that adapt to specific user needs.

Cross-Subject Generalization

The cross-subject analysis showed lower performance compared to the within-subject decoding, which agrees with previous studies pointing out the large differences between subjects in EEG signals [57]. The small standard deviations suggest that this decrease is systematic and not caused by a few outlier participants. These results indicate that, although the proposed pipeline can work as a proof-of-concept for cross-subject MRCP decoding, additional methods such as transfer learning or subject-specific calibration may be needed for practical use. In real applications, these strategies could help the system adapt to new users without retraining from scratch.

Multiclass Within-Subject and Cross-Subject Classification

Decoding multiple movement classes was more difficult than the binary task. Within subjects, mean accuracies were between 22% and 24% consistent with previous findings [58], while in the cross-subject evaluation they dropped to around 20%.

This shows the challenge of separating movements that come from nearby motor cortical areas. Even though the absolute accuracies were lower, longer time windows still tended to improve performance. Both MDM and TSM + LDA behaved in a similar way. The confusion matrices revealed that most mistakes happened between different movement types, rather than between movement and rest. This indicates that the pipeline can still detect movement intention reliably, even if it cannot perfectly classify each specific movement within or across subjects (above chance level approximately 16.7% for six classes). These observations suggest that while the system is effective for general movement detection, further refinements are needed for precise multiclass control, which could include more advanced feature extraction or hierarchical classification strategies.

Table 5.1
Mapping of Problems, Objectives, and Research Questions to Results and Discussion Sections.

Item	**Addressed in Section(s)**	**Key Findings / Notes**
Problem 1: Need for robust classifier → Riemannian geometry	Results: Binary & Multiclass; Discussion: Within-Subject & Cross-Subject	Both MDM and TSM + LDA reliably decode movement intention; Riemannian classifiers provide stability and good performance
Problem 2: MRCP amplitude low → xDAWN filter	Results: All classifications; Discussion: Within-Subject Binary	xDAWN improves SNR and enhances MRCP signals, boosting classification accuracy
Objective 1: Distinguish imagined movement vs rest (binary)	Results: Within-Subject Binary	Mean accuracy up to 71%; latency 1.2–1.3 s; longer windows improve performance
Objective 2: Proof-of-concept for MRCP classification	Results: All binary tasks	Movement intention reliably detected above chance using xDAWN + Riemannian classifiers
Objective 3: Evaluate multiclass classification	Results: Within- & Cross-Subject Multiclass	Within-subject: 22–24%; Cross-subject: 20%; confusion mainly between movement types
Objective 4: Cross-subject evaluation without transfer learning	Results: Cross-Subject Binary & Multiclass	Binary: 67%; Multiclass: 20%; moderate success; discussion notes potential for transfer learning to improve
RQ1: How accurately can EEG-based BCIs detect movement intention?	Results & Discussion: Binary tasks	Up to 71% within-subject; 67% cross-subject; above chance for multiclass
RQ2: Can movement vs rest be distinguished reliably?	Results: Binary; Discussion	Yes; reliable detection across time windows and subjects
RQ3: Multiclass decoding performance	Results & Discussion: Multiclass	Low but above chance; longer windows improve accuracy; confusion matrices analyzed
RQ4: Cross-subject generalization	Results & Discussion: Cross-Subject	Moderate performance; improvement possible with calibration or transfer learning

6. CONCLUSION

This thesis investigated motor imagery (MI) decoding using Riemannian geometry-based methods, focusing on the effects of time-window length, spatial filter selection, and within- versus cross-subject evaluation to characterize the early stages of movement preparation as reflected in MRCP signals. Several key findings emerged:

- **Effect of Time Window Length:** Longer analysis windows (up to 1 s) consistently improved classification performance. This trend was observed for maximum accuracy, suggesting that extended epochs allow more stable covariance estimation and benefit symmetric positive-definite (SPD) learning.
- **Binary Classification:** In within-subject binary tasks, mean accuracies reached up to 71% with the TSM + LDA pipeline, while MDM produced comparable results. Although differences were small, these findings confirm that time-resolved decoding can reveal informative windows of cortical activity. Cross-subject binary decoding also reached a competitive level (around 67%), indicating that shared neural patterns across individuals can be exploited despite inter-individual variability.
- **Multiclass Classification:** Decoding multiple movement classes proved more challenging, with accuracies of 22–24% within subjects and approximately 20% across subjects. Given six classes, the chance level was 16.7%. Confusion matrices, computed for multi-class movement decoding, revealed that most errors occurred between different movement classes, indicating that the pipeline still captured movement intention above chance.
- **Role of xDAWN Filtering:** Consistent with prior work on event-related potentials such as P300, applying xDAWN spatial filtering improved performance compared to raw covariance features, confirming its role in enhancing signal-to-noise ratio and facilitating decoding.
- **Cross-Subject Generalization:** Cross-subject decoding reached a competitive

level (around 67%), indicating that shared neural patterns exist across individuals. This suggests that advanced strategies such as transfer learning or domain adaptation could further improve generalization and enable more calibration-free BCIs.

6.1 Scope and Limitations

This study focused on classifying movement intention using multivariate methods in Riemannian space. The emphasis was on different imagined movements from only one upper limb (unilateral movements). Because these movements engage nearby brain areas, especially in the primary motor cortex, the EEG signals are spatially similar, making them difficult to separate.

There are some limitations in this study:

- EEG data contains noise, which can reduce accuracy.
- The dataset is relatively small, limiting multiclass classification.
- Detecting movement-related cortical potentials (MRCPs) is difficult due to small amplitudes and overlap with other brain activity.
- No similar external dataset was available to test the method beyond this work.
- Multiclass accuracies remain modest, reflecting the challenge of decoding spatially overlapping signals.
- Real-time implementation has not yet been tested.

Addressing these limitations could further improve the method and expand its applicability.

6.2 Significance of the Study

The significance of this study lies in reducing the time between when a person intends to move an upper limb and the machine response in BCI systems. Shortening this interval can improve system responsiveness and usability.

Such systems are particularly beneficial for individuals with cervical spinal cord injury (SCI) who have lost most motor control in their arms and hands. Using BCIs to restore some degree of movement can improve daily life quality.

This work also addresses a gap in the literature regarding the classification of movement-related cortical potentials (MRCP) using spatial features such as covariance. We demonstrate that the xDAWN method enhances MRCP signals, and, to our knowledge, this is the first study applying Riemannian geometry for MRCP classification.

Contribution and Implications

This study demonstrates, for the first time, that MRCP signals can be effectively enhanced using the xDAWN spatial filter and classified in Riemannian space. It provides insights into temporal dynamics, window-size effects, and frequency-specific decoding, forming a foundation for early detection of movement intention in EEG-based BCIs.

These findings are relevant for developing neuroprosthetic control systems for individuals with SCI, where earlier and robust detection can improve response time and usability. Understanding which features and frequency bands contribute most to decoding can guide the design of future adaptive BCI systems.

Future Directions

Future work could consider:

- Incorporating transfer learning or subject-specific calibration to improve cross-subject performance.
- Exploring additional feature extraction methods or combining MRCPs with oscillatory activity.
- Implementing real-time BCI systems to evaluate practical neuroprosthetic control.
- Expanding the dataset and including bilateral movements to enhance spatial separability.
- Optimizing epoch size and window overlap to enhance SPD learning and temporal decoding.
- Investigating alternative classifiers beyond MDM and TSM + LDA for improved accuracy.

Summary

In conclusion, xDAWN spatial filtering enhances MRCP signals, and Riemannian classifiers can reliably detect early movement intention in EEG. Longer time windows improve decoding performance, although multiclass and cross-subject decoding remain challenging. This study provides a strong foundation for flexible and robust BCIs and points to clear directions for future research aimed at practical neuroprosthetic applications.

REFERENCES

1. Servetto, G. P., C. M. Root, R. Giere, and R. Vigliaturo, "Mineralogy of the human brain: a review," *European Journal of Mineralogy*, Vol. 37, no. 2, pp. 279–304, 2025.

2. Vinding, M. C., P. Tsitsi, H. Piitulainen, J. Waldthaler, and et al., "Attenuated beta rebound to proprioceptive afferent feedback in parkinson's disease," *Scientific Reports*, Vol. 9, no. 1, p. 2604, 2019.

3. Dong, R., X. Zhang, H. Li, Z. Lu, and et al., "Cross-domain prediction approach of human lower limb voluntary movement intention for exoskeleton robot based on EEG signals," *Frontiers in Bioengineering and Biotechnology*, Vol. 12, 08 2024.

4. McCoy, D., and D. McCoy, "Pass-thoughts: authenticating with our minds," in *Proceedings of the 2005 workshop on New security paradigms*, pp. 45–56, ACM, 2005.

5. Lotte, F., L. Bougrain, A. Cichocki, M. Clerc, and et al., "A review of classification algorithms for EEG-based brain-computer interfaces: a 10 year update," *Journal of Neural Engineering*, Vol. 15, p. 031005, apr 2018.

6. Tibermacine, I. E., S. Russo, A. Tibermacine, A. Rabehi, and et al., "Riemannian geometry-based EEG approaches: A literature review," 2024.

7. Geirnaert, S., *Signal Processing Algorithms for EEG-based Auditory Attention Decoding*. Phd thesis, KU Leuven, Faculty of Engineering, 2022. Accessed: 2025-08-09.

8. Niedermeyer, E., and F. L. da Silva, *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*, Lippincott Williams & Wilkins, 5 ed., 2004.

9. Teplan, M., "Fundamentals of EEG measurement," *Measurement Science Review*, Vol. 2, no. 2, pp. 1–11, 2002.

10. Michel, C. M., and D. Brunet, "EEG source imaging: A practical review of the analysis steps," *Frontiers in Neurology*, Vol. 10, p. 325, 2019.

11. Wallisch, P., M. E. Lusignan, M. D. Benayoun, and T. I. et al., *MATLAB for Neuroscientists: An Introduction to Scientific Computing in MATLAB*, Academic Press, 2nd ed., 2014.

12. Ofner, P., A. Schwarz, J. Pereira, and G. R. Muller-Putz, "Upper limb movements can be decoded from the time-domain of low-frequency EEG," *PLOS ONE*, Vol. 12, pp. 1–24, 08 2017.

13. Wei, C.-S., C. J. Keller, J. Li, Y.-P. Lin, and et al., "Editorial: Inter- and intra-subject variability in brain imaging and decoding," *Frontiers in Computational Neuroscience*, Vol. 15, p. 791129, 2021.

14. Ouyang, G., A. Hildebrandt, W. Sommer, C. Zhou, and et al., "Exploiting the intra-subject latency variability from single-trial event-related potentials in the p3 time range: A review and comparative evaluation of methods," *Neuroscience and Biobehavioral Reviews*, Vol. 75, pp. 1–21, 2017.

15. Wu, D., X. Jiang, and R. Peng, "Transfer learning for motor imagery based brain-computer interfaces: A tutorial," *Neural Networks*, Vol. 153, pp. 235–253, Sept. 2022.

16. Yamamoto, M., K. Sadatnejad, T. Tanaka, M. Islam, and et al., "Modeling complex EEG data distribution on the riemannian manifold toward outlier detection and multimodal classification," *IEEE Transactions on Biomedical Engineering*, Vol. 71, no. 2, pp. 377–387, 2024.

17. Ivanov, N., and T. Chau, "Riemannian geometry-based metrics to measure and reinforce user performance changes during brain-computer interface user training," *Frontiers in Computational Neuroscience*, Vol. 17, 2023.

18. Andreev, A., and et al., "The riemannian means field classifier for EEG-based BCI data," *Sensors (Basel, Switzerland)*, Vol. 25, no. 7, p. 2305, 2025.

19. Congedo, M., A. Barachant, and R. Bhatia, "Riemannian geometry for EEG-based brain-computer interfaces; a primer and a review," *Brain-Computer Interfaces*, Vol. 4, no. 3, pp. 155–174, 2017.

20. Wineski, L. E., *Snell's Clinical Anatomy by Regions*, Wolters Kluwer, 10th ed., 2018.

21. Kandel, E. R., J. D. Koester, S. H. Mack, and et al., *Principles of Neural Science*, McGraw Hill, 6th ed., 2021.

22. Blinowska, K. J., and M. J. Kaminski, *Essentials of Clinical Neurophysiology*, Berlin, Germany: Springer, 1st ed., 2011.

23. Siuly, *Analysis and Classification of EEG Signals*. Doctoral dissertation, University of Southern Queensland, Australia, July 2012. Accessed: 2025-08-09.

24. Nicolas-Alonso, L. F., and J. Gomez-Gil, "Brain computer interfaces, a review," *Sensors*, Vol. 12, no. 2, pp. 1211–1279, 2012.

25. Acharya, J. N., and V. J. Acharya, "Overview of EEG montages and principles of localization," *Journal of Clinical Neurophysiology*, Vol. 36, pp. 325–329, Sept. 2019.

26. Saba-Sadiya, S., E. Chantland, T. Alhanai, T. Liu, and et al., "Unsupervised EEG artifact detection and correction," *Frontiers in Digital Health*, Vol. 2, 2021.

27. Sadiya, S., T. Alhanai, and M. Ghassemi, "Artifact detection and correction in EEG data: A review," in *2021 10th International IEEE/EMBS Conference on Neural Engineering (NER)*, pp. 495–498, May 2021.

28. Tedeschi, R., "Unlocking the power of motor imagery: a comprehensive review on its application in alleviating foot pain," *Acta neurologica Belgica*, Vol. 124, 03 2024.

29. Pfurtscheller, G., and C. Neuper, "Motor imagery and direct brain-computer communication," *Proceedings of the IEEE*, Vol. 89, no. 7, pp. 1123–1134, 2001.

30. Monge-Pereira, E., F. Molina-Rueda, F. Rivas-Montero, J. Ibanez, and et al., "Electroencephalography as a post-stroke assessment method: An updated review," *Neurologia*, Vol. 32, no. 1, pp. 40–49, 2017.

31. Savic, A. M., E. R. Lontis, N. Mrachacz-Kersting, and M. B. Popovic, "Dynamics of movement-related cortical potentials and sensorimotor oscillations during palmar grasp movements," *European Journal of Neuroscience*, Vol. 51, no. 4, pp. 1007–1019, 2020.

32. Riascos, J., M. Molinas, and F. Lotte, "Machine learning methods for BCI: challenges, pitfalls and promises," in *ESANN 2024 - European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*, (Bruges, Belgium), Oct. 2024.

33. Lotte, F., M. Congedo, A. Lecuyer, F. Lamarche, and B. Arnaldi, "A review of classification algorithms for EEG-based brain-computer interfaces," *Journal of Neural Engineering*, Vol. 4, pp. R1–R13, jun 2007.

34. Roman-Gonzalez, A., "EEG Signal Processing for BCI Applications," *Human - Computer Systems Interaction: Backgrounds and Applications 2, Advances in Intelligent and Soft Computing*, Vol. 98, pp. 51–591, Jan. 2012.

35. Farwell, L., and E. Donchin, "Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials," *Electroencephalography and Clinical Neurophysiology*, Vol. 70, no. 6, pp. 510–523, 1988.

36. Vialatte, F.-B., M. Maurice, J. Dauwels, and A. Cichocki, "Steady-state visually evoked potentials: Focus on essential paradigms and future perspectives," *Progress in Neurobiology*, Vol. 90, no. 4, pp. 418–438, 2010.

37. Congedo, M., A. Barachant, and A. Andreev, "A new generation of brain-computer interface based on riemannian geometry," 2013.

38. Congedo, M., P. L. Coelho Rodrigues, and C. Jutten, "The riemannian minimum distance to means field classifier," in *Proceedings of the 8th Graz Brain Computer Interface Conference 2019, Bridging Science and Application*, (Graz, Austria), Sep 2019.

39. Lee, J. M., *Riemannian Manifolds: An Introduction to Curvature*, Vol. 176 of *Graduate Texts in Mathematics*, New York: Springer, 1997.

40. Barachant, A., S. Bonnet, M. Congedo, C. Jutten, and et al., "Multiclass brain-computer interface classification by riemannian geometry," *IEEE Transactions on Biomedical Engineering*, Vol. 59, pp. 920–928, Mar 2012.

41. Ledoit, O., and M. Wolf, "A well-conditioned estimator for large-dimensional covariance matrices," *Journal of Multivariate Analysis*, Vol. 88, no. 2, pp. 365–411, 2004.

42. Bhatia, R., *Positive Definite Matrices*, Princeton, NJ: Princeton University Press, 2007.

43. Arsigny, V., P. Fillard, X. Pennec, and N. Ayache, "Log-euclidean metrics for fast and simple calculus on diffusion tensors," *Magnetic Resonance in Medicine*, Vol. 56, no. 2, pp. 411–421, 2006.

44. Rivet, B., A. Souloumiac, V. Attina, and G. Gibert, "xDAWN algorithm to enhance evoked potentials: application to brain-computer interface," *IEEE Transactions on Biomedical Engineering*, Vol. 56, pp. 2035–43, Aug 2009.

45. Rivet, B., H. Cecotti, A. Souloumiac, E. Maby, and et al., "Theoretical analysis of xDAWN algorithm: Application to an efficient sensor selection in a p300 BCI," in *2011 19th European Signal Processing Conference*, pp. 1382–1386, 2011.

46. Barachant, A., "MEG decoding using riemannian geometry and unsupervised classification: Technical report with the solution of the decmeg 2014 challenge," tech. rep., INRIA, 2014.

47. Delorme, A., "EEG is better left alone," *Scientific Reports*, Vol. 13, no. 1, p. 2372, 2023.

48. Oostenveld, R., P. Fries, E. Maris, and J.-M. Schoffelen, "Fieldtrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data," *Computational Intelligence and Neuroscience*, Vol. 2011, pp. 1–9, 2011.

49. Delorme, A., and S. Makeig, "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *Journal of Neuroscience Methods*, Vol. 134, no. 1, pp. 9–21, 2004.

50. Perrin, F., J. Pernier, O. Bertrand, and J. F. Echallier, "Spherical splines for scalp potential and current density mapping," *Electroencephalography and Clinical Neurophysiology*, Vol. 72, pp. 184–187, 1989.

51. Hastie, T., R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer, 2 ed., 2009.

52. Bishop, C. M., *Pattern Recognition and Machine Learning*, New York: Springer, 2006.

53. Ojala, M., and G. C. Garriga, "Permutation tests for studying classifier performance," in *2009 Ninth IEEE International Conference on Data Mining*, pp. 908–913, 2009.

54. Benjamini, Y., and Y. Hochberg, "Controlling the false discovery rate: A practical and powerful approach to multiple testing," *Journal of the Royal Statistical Society: Series B (Methodological)*, Vol. 57, pp. 289–300, 12 2018.

55. Carlson, T. A., T. Grootswagers, and A. K. Robinson, "An introduction to time-resolved decoding analysis for M/EEG," 2019.

56. Hameed, I., S. Sharpe, D. Barcklow, J. Au-Yeung, and et al., "BASED-XAI: Breaking ablation studies down for explainable artificial intelligence," 2022.

57. Apicella, A., P. Arpaia, G. D'Errico, D. Marocco, and et al., "Toward cross-subject and cross-session generalization in EEG-based emotion recognition: Systematic review, taxonomy, and methods," *Neurocomputing*, Vol. 604, p. 128354, Nov. 2024.

58. Zhang, M., J. Huang, and S. Ni, "Recognition of motor intentions from EEGs of the same upper limb by signal traceability and riemannian geometry features," *Frontiers in Neuroscience*, Vol. Volume 17 - 2023, 2023.