Development of Network Features for Brain-Computer Interfaces

Juliana Gonzalez-Astudillo

Sorbonne Université

Doctoral Thesis

Development of Network Features for Brain-Computer Interfaces

Author:
Juliana Gonzalez-Astudillo

Supervisors:
Fabrizio De Vico Fallani

Referees:
Sylvain Chevallier
Sophie Achard
Fabien Lotte

ARAMIS Lab
Sorbonne Université, Inria Paris Center, Institut du Cerveau - Paris Brain Institute (ICM), Inserm U 1127, CNRS UMR 7225, AP-HP Hôpital de la Pitié Salpêtrière

To be defended on October 6th, 2022

Abstract

A Brain-Computer Interface (BCI) is a system that can translate brain activity patterns into messages or commands for an interactive application. It enables a subject to send commands to a device only by means of brain activity, without requiring any peripherical muscular activity. These systems are increasingly explored for control and communication, as well as for treatment of neurological disorders, especially via the ability of subjects to voluntarily modulate their brain activity through mental imagery (MI).

To control a BCI, the user must produce different brain signal patterns that the system will identify and translate into commands. Even though this technique has been widely used, subjects performance, measured as the correct classification of the user's intent, still shows low scores. Much of the efforts to solve this problem have focused on the BCI classification block. While, the research of alternative features has been poorly explored. In most implemented systems, pattern recognition relies on power spectrum density (PSD) of a reduced number of sources, focusing on features that characterize a single brain region.

However, the brain is not a collection of isolated pieces working independently. It rather consists of a distributed complex network that integrates information across differently specialized regions. It turns out that examining signals from one specific region, while neglecting its interactions with others, oversimplifies the phenomenon. It would be preferable to have an understanding of the system's collective behavior to fully capture the brain functioning. Thus, we hypothesize that functional connectivity (FC) features could be more representative of the complexity of neurophysiological processes, since they measure interactions between different brain areas, reflecting the information exchange that is essential to decode brain organization. Then, these interactions can be quantified using network theoretic approaches, extracting few summary properties of the entire complex brain network. Thus, network analysis may also be more efficient by reducing the problem dimension and optimizing the computational cost.

Nevertheless, extracting topological properties of the network, while disregarding the intrinsic spatial nature of the brain, could overlook crucial information for understanding brain functioning. Recent neuroimaging studies demonstrated that brain connectivity reveals hemisphere lateralization during motor MI-related tasks. Covering these two concepts, we explored the dual contribution of brain network topology and space in modelling motor-related mental states through the concept of functional lateralization. Specifically, we introduced new metrics to quantify segregation and integration within and between the hemispheres, and we showed that they are highly relevant features for decoding a motor-imagery mental task. These network properties not only give competitive classification accuracy but also have the advantage of being neurophysiologically interpretable, compared to state-of-the-art approaches that are instead blind to the underlying mechanism.

Résumé

Une interface cerveau-machine (ICM) est un système capable de traduire les modèles d'activité cérébrale en messages pour une application. Il permet à un sujet d'envoyer des commandes à un appareil a travers l'activité cérébrale, sans nécessiter d'activité musculaire périphérique. Ces systèmes sont de plus en plus explorés pour le contrôle et la communication, ainsi que pour le traitement des troubles neurologiques, notamment via la capacité des sujets à moduler volontairement leur activité cérébrale grâce à l'imagerie mentale (IM).

Pour contrôler une ICM, l'utilisateur doit produire différents types de signaux cérébraux que le système identifiera et traduira en commandes. Même si cette technique a été largement utilisée, la performance des sujets, mesurée comme la correcte classification de l'intention de l'utilisateur, affiche toujours de faibles scores. Une grande partie des efforts pour résoudre ce problème s'est concentrée sur la classification. Alors que la recherche de features alternatives a été peu explorée. Dans la plupart des systèmes mis en œuvre, la reconnaissance des états mentaux repose sur la puissance spectrale d'un nombre réduit de sources, en se concentrant sur les caractérisation d'une seule région du cerveau.

Cependant, le cerveau n'est pas un ensemble de pièces isolées travaillant de manière indépendante. Il s'agit plutôt d'un réseau complexe qui intègre des informations dans des régions différemment spécialisées. Il s'avère que l'examen des signaux d'une région spécifique, tout en négligeant ses interactions avec les autres, simplifie à l'extrême le phénomène. Il serait préférable de comprendre le comportement collectif du système pour bien saisir le fonctionnement cérébral. Ainsi, nous pensons que l'étude a travers la connectivité fonctionnelle pourraient être plus représentatives de la complexité des processus neurophysiologiques, puisqu'elles mesurent les interactions entre différentes aires cérébrales, reflétant l'échange d'informations qui est essentiel pour décoder l'organisation cérébrale. Ensuite, ces interactions peuvent être synthétisées à l'aide d'estimateurs des réseaux complexes, modélisant le cerveau humain comme un réseau. Certes, l'analyse de réseau peut présenter une performance plus précise car elle optimise le coût de calcul et la dimensionnalité.

Néanmoins, la simple extraction des propriétés topologiques du réseau, sans tenir compte de la nature spatiale intrinsèque du cerveau, pourrait manquer des informations cruciales pour comprendre le fonctionnement du cerveau. Des études récentes ont démontré que la connectivité cérébrale révèle la latéralisation des hémisphères lors de tâches liées à l'IM moteur. Couvrant ces deux concepts, nous avons exploré la double contribution de la topologie et de l'espace dans la modélisation des états mentaux moteurs par la latéralisation fonctionnelle. Plus précisément, nous avons introduit de nouvelles métriques pour quantifier la ségrégation et l'intégration au sein et entre les hémisphères, et nous avons montré qu'il s'agit de caractéristiques très pertinentes pour décoder une tâche mentale d'imagerie motrice. Ces propriétés de réseau donnent non seulement des précisions de classification compétitives, mais ont également l'avantage d'être interprétables sur le plan neurophysiologique, par rapport aux approches de pointe qui sont plutôt aveugles au mécanisme sous-jacent.

Scientific production

First author journal papers

1. **Gonzalez-Astudillo, J.**, Cattai, T., Bassignana, G., Corsi, M.C., and De Vico Fallani, F., 2021. "Network-based brain–computer interfaces: principles and applications". *Journal of Neural Engineering*, 18, p.011001.

2. **Gonzalez-Astudillo, J.**, and De Vico Fallani, F.. "Network lateralization features for motor imagery-based brain-computer interfaces". *In Prep.*.

Book chapters

1. Couvy-Duchesne, B., Bottani, S., Camenen, E., Fang, F., Fikere, M., **Gonzalez-Astudillo, J.**, Harvey, J., Hassanaly, R., Kassam, I., Lind, P. and Liu, Q., Lu, Y., Nabais, M., Rolland, T., Sidorenko, J., Strike, L., Wright, M., 2022. "Main existing datasets for open data research on humans."

Conference abstracts

1. **Gonzalez-Astudillo, J.**, Cattai, T., Corsi, M.C. and De Vico Fallani, F., 2020. "On the classification of mental states by means of network-based features." In *NetSci 2020 - Network Science Society Conference 2020.*

Talks

1. **Gonzalez-Astudillo, J.** and De Vico Fallani, F., 2022. "Spatial lateralization in motor brain networks." In *MSCx 2022: Mediterranean School of Complex Networks.*

2. **Gonzalez-Astudillo, J.** and De Vico Fallani, F., 2022. "Spatial Networks features for Brain Computer Interfaces." In *FrCCS 2022: French Regional Conference on Complex Systems.*

3. **Gonzalez-Astudillo, J.** and De Vico Fallani, F., 2022. "A spatial network alternative for BCI inefficiency." In *Journées CORTICO 2022: COllectif pour la Recherche Transdisciplinaire sur les Interfaces Cerveau-Ordinateur.*

4. **Gonzalez-Astudillo, J.**, Ceballos-Dominguez, E.G., Cattai, T., Corsi, M.C. and De Vico Fallani, F., 2021. "Spatial network metrics for characterizing brain-computer interface mental states." In *Networks 2021: A Joint Sunbelt and NetSci Conference.*

5. **Gonzalez-Astudillo, J.**, Cattai, T., Corsi, M.C. and De Vico Fallani, F., 2020. "Towards the use of spatial networks for characterizing brain mental states." In *CCS: Conference on Complex Systems - Complex-Space 2020: Analysis and Modelling of Spatial Complex Systems (Satellite).*

Contents

List of Figures

List of Tables

List of Abbreviations

AUC	Area Under the Curve
BA	Barabási-Albert
BCI	Brain Computer Interface
CNN	Convolutional Neural Network
CSP	Common Spatial Pattern
CV	Cross-validation
dFC	dynamic functional connectivity
DMN	Default Mode Network
DLFC	Dorso-lateral prefrontal cortex
EEG	Electroencephalography
ER	Erdös-Rényi
ERD	Event-Related Desynchronization
ERS	Event-Related Synchronizationcan
EML	Extreme Learning Machine
FC	Functional Connectivity
GEVD	Generalized Eigen Value Decomposition
fMRI	functional Magnetic Resonance Imaging
HMMs	Hidden Markov Models
LDA	Linear Discriminant Analysis
LMI	Left hand Motor Imagery
M1	Primary Motor Area
MI	Motor Imagery
ME	Motor Execution
MEG	Magnetoencephalography
ML	Machine Learning
MRI	Magnetic Resonance Imaging
NFB	Neurofeedback
PMA	Premotor Area
PLV	Phase-Locking Value
PMA	Premotor Area
PSD	Power Spectrum Density
RG	Riemannian Geometry
RMI	Rigth hand Motor Imagery
ROI	Region of interest
ROC-AUC	Receiver Operating Characteristic - Area Under the Curve
S1	Primary Somatosensory Area
S2	Somatosensory Association Area
SBM	Stochastic Block Model
SLR	Sparse Logistic Regression
SMA	Supplementary Motor Area
SPD	symmetric positive definite
SVM	Support Vector Machines
WS	Watts-Strogatz

Introduction

Over the past decades, the way scientists have looked at the human brain has witnessed a paradigm shift. The view that cognition and behavior result from localized neuronal ensembles has progressively left room for the realization that their interaction is what really matters. Today, we know that the brain is not just a collection of isolated units working independently, but it rather consists of a complex network that integrates information across differently specialized regions via anatomical as well as functional connections (Bullmore and Sporns, 2009).

Such transition from a reductionist to a holistic perspective has been accompanied by the dawning of network science, i.e. a modern field drawing on graph theory that summarizes and quantifies organizational properties of complex interconnected systems. In human neuroscience, brain regions are treated as network nodes, and the connections between the nodes —inferred from structural or functional neuroimaging data— are represented as network edges (or links) (Rubinov and Sporns, 2010; De Vico Fallani et al., 2014; Stam and Reijneveld, 2007). Network properties including efficiency (Latora and Marchiori, 2001), modularity (Newman, 2006), node centrality (Borgatti, 2005), and laterality (Liu et al., 2009) have been demonstrated to support basic cognitive functions such as language and memory (Bullmore and Sporns, 2009). Critically, these network indexes are also sensitive to physiological and pathological alterations of the mental state and can capture brain organizational mechanisms across different spatiotemporal scales (Stam, 2014; Ganguly and Poo, 2013).

Such a fundamental relationship between network topology and brain function is a key element of modern neuroscience and offers a grounded tool for analyzing brain networks using few topological descriptors rather than high-dimensional connectivity matrices (Boccaletti et al., 2006). Network neuroscience has allowed answers to fundamental questions spanning consciousness, plasticity, and learning, but it can also play a role in engineering applications aiming to characterize different brain states and recognize mental intentions from functional neuroimaging recordings. This is the case of brain-computer interfaces (BCIs) which implement ideal communication pathways bypassing the traditional effector of the musculoskeletal system and directly interacting with external devices (Vidal, 1973; Bozinovski, Sestakov, and Bozinovska, 1988; Wolpaw et al., 2002). Based on the classification of mental states from brain activity, BCIs are increasingly explored for control and communication (Wolpaw et al., 2002; Carmena et al., 2003; Carlson and Millan, 2013), and for treatment of neurological disorders (Daly and Wolpaw, 2008; Vansteensel et al., 2016).

In this context, the first findings have shown that the modulation of brain activity elicited by motor imagery (MI) (Pfurtscheller and Da Silva, 1999) as well as by decision-making tasks (Donchin, Spencer, and Wijesinghe, 2000), generates detectable signal changes such

as Event-Related Desynchronization or Synchronizationcan (ERD/ERS) that corresponds to specific amplitude variations in the power of the signals for a particular frequency band. To minimize intra-class variance, power-based features are usually combined with filtering techniques, such as common spatial patterns (CSP), a particular type of supervised spatial filters (Blankertz et al., 2007). Other outstanding methods based their study on Riemannian geometry (RG). They enable the direct manipulation of the signal covariance matrices by using the topology of the manifold (Yger, Berar, and Lotte, 2016). Although these approaches exhibit high accuracy, there is still a non-negligible portion of subjects (∼30%) that show inefficient performance (Thompson, 2019). Besides, these methods lack of direct interpretability, which leaves open the possibility that artifacts lead classification results.

Notably, most of these current methods consider each sensor (ie, a brain region) as an isolated element, thus neglecting possibly existing interactions across them. This can be in part explained by the fact that the BCI community has mainly focused on improving the signal processing and classification block of the BCI pipeline, while neglecting the feature extraction part (Lotte et al., 2018). To capture the brain's collective behaviour, we hypothesize that functional connectivity (FC) could fill in this gap, by taking into account information exchange between different brain areas. More relevant to this PhD thesis, we investigate how network theory can be adapted to identify new spatial connection mechanisms subserving BCI tasks (De Vico Fallani and Bassett, 2019) (**Fig 1**).

Outline of the manuscript

The manuscript has the following structure:

- In **Chapter 1** we introduce the optimal ways to infer networks from brain activity. We present most relevant network properties to model brain interactions at different topological scales.
- In **Chapter 2** we describe the translation process from BCI-related data to brain network. We report the studies that have been done in the field, and the feasibility network features in a BCI pipeline.
- In **Chapter 3** we substantiate and develop lateralization network properties for identifying MI states (*laterality*, *segregation* and *integration*).
- In **Chapter 4** we select the most popular methods for feature extraction in BCI: CSP and RG. We explain how we have adapted the Riemannian approach to be comparable we our proposed method. We outline the supplementary procedures for the implementation and validation of our methods, including statistical analysis, feature selection and classification algorithms.
- In **Chapter 5** we demonstrate the discriminant power of network lateralization metrics in left and right hand-MI, for multiple BCI datasets. We validate them in a classification pipeline comparing its performance with state-of-the-art methods.
- Finally, in the Discussion, Conclusion and Perspectives chapters, we discuss our results and provide potential future research directions.

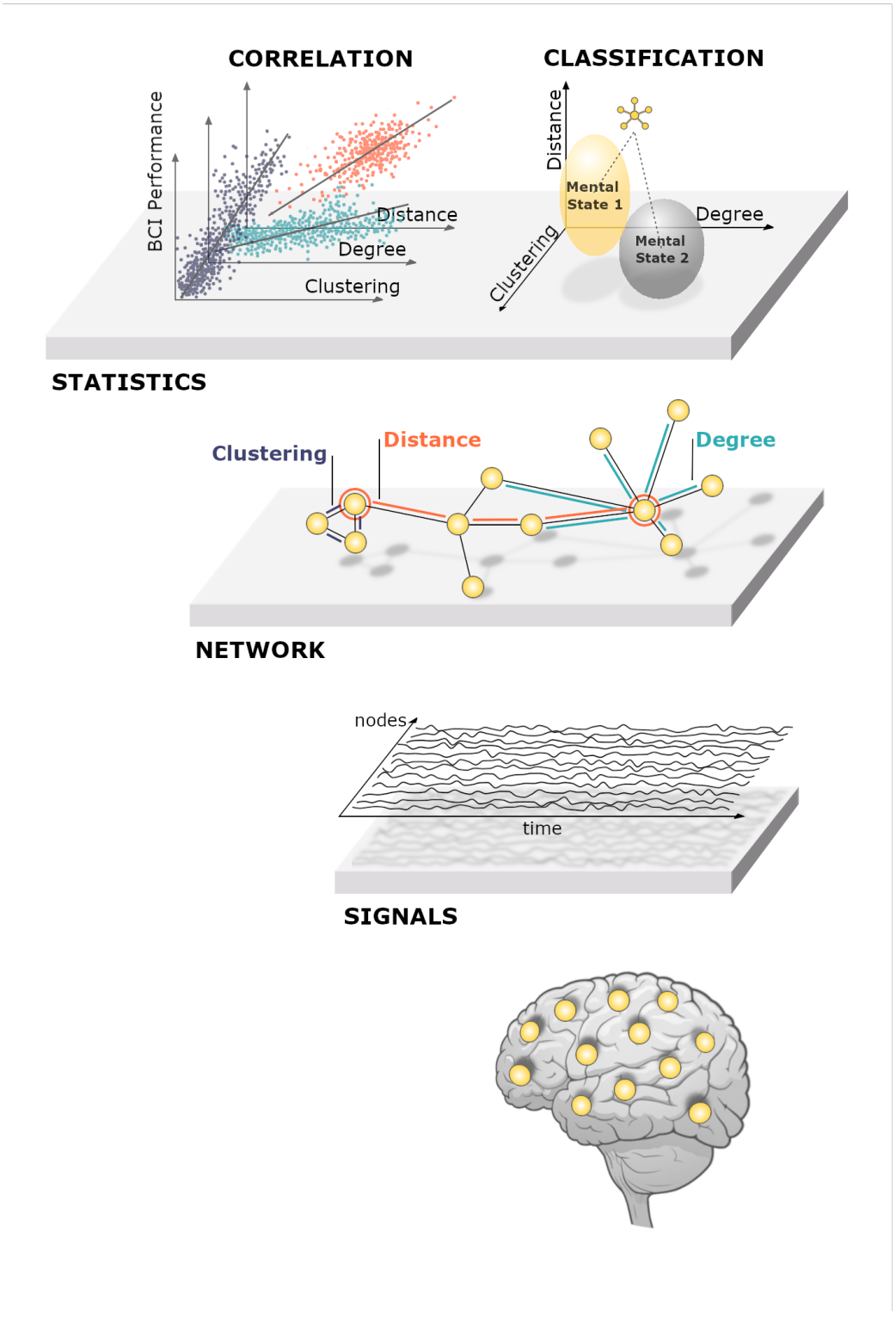


FIGURE 1: **From brain activity to BCI classification.** Principle scheme of a network-based brain–computer interface. From bottom to top: brain networks are reconstructed by computing functional connectivity between remote brain signals. The resulting connectivity networks are characterized by means of graph theoretic metrics, which extract summary indices quantifying different topological properties. These values correspond to specific network properties that can be used to identify predictors of BCI performance as well classify different BCI mental states.

Chapter 1

From brain activity to networks

Parts of this chapter has been published in *Journal of Neural Engineering*:

- **Title:** Network-based brain computer interfaces: principles and applications
- **Authors:** Juliana Gonzalez-Astudillo, Tiziana Cattai, Giulia Bassignana, Marie-Constance Corsi and Fabrizio De Vico Fallani
- **DOI:** doi:10.1016/j.media.2021.102219

1.1 Introduction

In this chapter, we first provide some general information about the construction of functional brain networks, we survey network theoretic measures, and illustrates their application to cognitive and motor BCI-related neuroimaging data. This constitutes a methodological reference and does not aim to provide new neurophysiological insights. Throughout the sections, we comment on the methodological limitations and the best practices for their application. This section goal is to provide an accessible introduction to the field.

1.2 From functional neuroimaging data to brain networks

The first step when studying brain networks is to decide which are the nodes and the edges. Typically, the definition of the nodes depends on the specific neuroimaging modality. For functional Magnetic Resonance Imaging (fMRI) and other voxel-based techniques, the most common approach consists in using anatomical atlas and each region of interest (ROI) corresponds to a node (Cohen et al., 2008; Salvador et al., 2005). For sensor-based modalities, such as EEG and MEG, each sensor typically corresponds to a node (Stam and Reijneveld, 2007; Bassett and Bullmore, 2009), although source-reconstruction techniques can be used to define nodes at the cortical level (Baillet, Mosher, and Leahy, 2001; Michel et al., 2004; Edelman, Baxter, and He, 2015). Because neuroimaging techniques only give access to regional activities, recorded as signals, the network edges must be inferred using statistical approaches. This is typically done by means of functional connectivity (FC)

TABLE 1.1: Selection of the most commonly used FC estimators. The different methods are organized according to their ability to capture directed or undirected interactions. Specific properties associated with some of the critical issues discussed in the section are reported on the right part of the table.

	Functional connectivity estimators	Properties		
		Non-linearity	Time-varying	Multivariate
Undirected	Spectral coherence (Carter, 1987)	-	-	-
	Imaginary coherence (Nolte et al., 2004)	-	-	-
	Phase-Locking Value (Aydore, Pantazis, and Leahy, 2013)	✓	-	-
	Weighted phase lag index (Vinck et al., 2011)	✓	-	-
	Partial coherence (Rosenberg et al., 1989)	-	-	✓
	Synchronization likelihood (Stam and Van Dijk, 2002)	✓	-	-
	Mutual information (Kaiser and Schreiber, 2002)	✓	-	-
	Wavelet coherence (Chavez and Cazelles, 2019)	✓	✓	-
Directed	Granger causality (Blinowska, Kuś, and Kamiński, 2004)	-	-	-
	Kernel Granger causality (Marinazzo et al., 2011)	✓	-	-
	Partial Granger causality (Barrett, Barnett, and Seth, 2010)	-	-	✓
	Partial directed coherence (Baccalá and Sameshima, 2001)	-	-	✓
	Transfer entropy (Kaiser and Schreiber, 2002)	✓	-	✓
	Directed Transfer Function (Kaminski and Blinowska, 1991)	✓	-	✓
	Adaptive partial directed coherence (Leistritz et al., 2013)	-	✓	✓

estimators which measure the temporal dependency between different brain signals. As a result, network edges correspond to FC estimates.

In the last decades, many methods have been developed to quantify functional interactions in the brain, relying on tools from signal processing and information theory. Even if each method is characterized by its own specific operations, the general procedure remains the same. Given a set of time series corresponding to the activity of different brain sites, the goal is to quantify the interaction between every signal pair. The literature is consistent in recognizing that the first distinction between FC estimators is between undirected and directed methods (De Vico Fallani et al., 2014; Bastos and Schoffelen, 2016). The former measures symmetric interactions, without considering the directionality of the information flow. The latter characterizes causal effects during activity propagation. Inside these categories, further distinctions can be done, according to their ability to describe linear or nonlinear interactions, bivariate or multivariate effect, time or frequency domain properties. **Table 1.1** shows a non-exhaustive list of the most used FC estimators in neuroscience, with their associated properties. In the following, we present some of the most challenging issues that significantly influence connectivity estimation.

1.2.1 Critical aspects

1.2.1.1 Spurious connectivity

The ultimate goal of FC methods is to quantify true signal interactions between different brain areas. However, several conditions can affect the correct estimation and introduce spurious contributions, thus giving a potentially distorted measure of the real interactions. This is in part due to the fact that most of the experiemental techniques for recording

noninvasve human brain signals, such as EEG, MEG or fMRI (Hwang et al., 2013; Harrison and Connolly, 2013; Fouad et al., 2015) can only indirectly capture the real neuronal source activity. For example, EEG and MEG measure respectively the electrical activity and magnetic flux produced by neurons within the brain. The electromagnetic signals propagate through the head tissues from the cortex - i.e., the source space - to the scalp - i.e., the sensor space. During this propagation, the different electrical conductibility of the tissues generates a spatial smearing of the signals on the scalp (Broek et al., 1998; Nunez et al., 1997). As a consequence, the signal measured in one electrode does not reflect the activity of one single source and this phenomemon, also known as volume conduction effect, can lead to spurious instantaneous interactions (Nolte et al., 2004). One possible solution consists of computing FC in the source domain, after having reconstructed the signals of the cortical space by means of inverse procedures. While source reconstruction techniques do alleviate the volume conduction effect, they do not entirely solve the problem and results can strongly depend on the implemented algorithm (Mahjoory et al., 2017). Furthermore, individual head models obtained from structural MRI are often necessary to have best high-quality results (Michel et al., 2004; Edelman, Baxter, and He, 2015).

Because volume conduction effects exclusively affect instantanous interactions, an alternative solution is the use of FC estimators that purposely remove lag-zero contributions from the estimates, such as *imaginary coherence* (Nolte et al., 2004), or *weighted phase lag index*. While these approaches significantly limit the bias introduced by the volume conduction smearing, they might remove possibly existing instantaneous neurophysiological signal interactions (Vicente et al., 2008). Spurious FC can also be introduced by third-party influences when multiple signals are available. When estimating the interaction between two signals, a portion of the interaction might be merely given by the presence of a third signal interacting with them. In some cases, it is therefore crucial to isolate this contribution and eventually remove it from the estimate (Kus, Kaminski, and Blinowska, 2004; Jalili and Knyazeva, 2011). While the large part of the FC methods have tended to neglect third-party influences, there are now several methods in literature, based on *partial coherence* (Rosenberg et al., 1989; Makhtar et al., 2020) or *partial directed coherence* (Baccalá and Sameshima, 2001), which have been designed to circumvent and alleviate those spurious effects.

1.2.1.2 Non-linear interactions

The neural system at a microscopic scale is characterized by nonlinear dynamics such as those of neuronal responses to stimuli or synaptic transmission (Haken, 2000). A crucial question is whether the brain activity at a macroscopic scale can be instead approximated by linear dynamics and take advantage of the efficacy of linear methods (Gourévitch, Bouquin-Jeannès, and Faucon, 2006). The findings related to this subject are controversial (Winterhalder et al., 2004). Several studies have investigated nonlinearities in brain signals using the largest Lyapunov exponent, the correlation integral or the method of data surrogate. The obtained results show that in healthy subjects there is a weak signal nonlinearity (Theiler and Rapp, 1996; Paluš, 1996). Other works have reported nonlinear behavior in epileptic

patients explained by the transitions between ordered and disordered states and the low-dimensional chaos (Babloyantz and Destexhe, 1986; Iasemidis and Sackellares, 1996). The latter evidence was nevertheless contradicted by more recent endeavors showing that even in diseased subjects, nonlinear methods perform as well as linear ones (McSharry, Smith, and Tarassenko, 2003; De Clercq et al., 2003).

More in general, nonlinearity also concerns the statistical interdependence between different brain signals. This typically means that FC is not proportional to either magnitude or phase of the signal frequency contents. In the early 1980s, the concept of synchronization was already extended and explained as a result of the adjustment of the oscillators caused by the presence of weak interactions (Pikovsky, Rosenblum, Kurths, et al., 2001). In these situations, the use of linear FC can fail to provide a complete description of the temporal properties of the signal interactions. Despite such limitation, the majority of FC studies still rely on linear-based interaction methods because of their simplicity and intuitive interpretation. In the case of *spectral coherence* and related estimators (*partial coherence*, *imaginary coherence*, etc..) it has been shown that they are relatively robust to nonlinear fluctuations in the signal amplitudes but not in phases (Sakkalis and Zervakis, 2009). However, if there is a precise for non-linearity, several estimators can be used to capture non-linear FC taking into undirected (*mutual information* (Kraskov, Stögbauer, and Grassberger, 2004), *phase locking value*, *synchronization likelihood*) or directed relationships (e.g., *transfer entropy* (Schreiber, 2000), *kernel Granger-causality*) **Table 1.1**.

1.2.1.3 Time-varying dynamic connectivity

FC estimators have been typically applied to extract connectivity patterns characterizing relatively long time periods (from dozens of seconds to minutes). In the last decade, the focus has shifted to shorter time scales that can be studied with dynamic functional connectivity (dFC) (Hutchison et al., 2013). Indeed, the possibility to determine how FC fluctuates during specific tasks is particularly appealing for BCI applications, where the mental state of the subjects rapidly varies to control the effector or accommodate the feedback. To this end, the simplest approach consists of reducing the length of the time window, letting it slide along the entire period of interest, with or without overlapping. On the one hand, reducing the size of the time window has also the effect of ensuring the signal (quasi)stationarity hypothesis required by many FC estimators (Cestari and Rosa, 2017; Kwiatkowski et al., 1992; Horwitz, 2003). On the other hand, the statistical reliability of the estimates strongly depends on the available temporal data points. That is, the larger is the number of available data points, the better is the ability of the FC estimator to capture the underlying connection mechanism. This situation is further exacerbated in the case of multivariate and non-linear estimators, which typically require more data points to give reliable estimates (Netoff et al., 2006; Pereda, Quiroga, and Bhattacharya, 2005). Standard solutions consist in concatenating the temporal windows associated with multiple repetitions of the same experimental task or averaging the FC estimates obtained in each repetition (Wibral et al., 2013). Methods based on multi-window spectrum estimation can be also used (Thomson, 1982; Baraniuk and Bayram, 2000). They allow the analysis of short data segments by using smoothing over orthogonal windows and they can be defined in the Fourier (Babadi

and Brown, 2014) and Wavelet domain (Brittain et al., 2007). Another elegant approach to estimate time-varying FC would consist in the use of methods formally designed to deal with non-stationary signals, such as detrended fluctuation analysis (Márton et al., 2014) or wavelet decomposition (Santoso et al., 1997). Among others, time-frequency methods such as *wavelet coherence* (Babloyantz and Destexhe, 1986; Lachaux et al., 2002) and *adaptive partial directed coherence* (Leistritz et al., 2013; Sanei and Chambers, 2013) represent particularly appealing solutions

More in general, the development of FC methods able to capture time-varying interactions is a fertile research field. For instance, tracking algorithms of brain correlation dynamics have been recently exploited (Monti et al., 2017), also considering low-rank subspaces (Ozdemir, Bernat, and Aviyente, 2017). Other approaches are based on model assumptions on the nature of signals (Romero, Ioannidis, and Giannakis, 2017), time-varying autoregressive models and variation of standard connectivity estimators (Chavez and Cazelles, 2019; Kraut, Scharf, and Butler, 2005).

1.2.2 Choosing the best FC estimator

We reported some of the most common FC estimators and their associated ability to solve one or more criticalities. It is important to notice that in general none of them is able to simultaneously solve all the raised issues. While it may be expected that applying all the possible methods would lead to consistent results, this approach lacks a precise rationale because different estimators intrinsically capture different signal properties and address different methodological questions (eg, causality versus synchronization). Instead, the choice of the "best" estimator mainly depends on the specific scientific question (Bastos and Schoffelen, 2016). If the scientific hypothesis that guides the analysis is clear, the choice of the estimator should be a natural consequence. For example, if the goal of the study is to determine information flows between two brain areas, a directed estimator should be used in a bivariate framework. In this scenario, under the assumption of linear dynamics, linear methods such as *Granger causality* should be used, otherwise non-linear estimators such as *transfer entropy* should be preferred.

Particular attention should be paid when studying brain signals with rich frequency dynamics. The use of estimators defined in the frequency domain is well-suited if the goal is to determine FC at specific frequency bands. The frequency transformation implemented by these estimators is typically obtained either via parametric techniques, such as autoregressive models, or non-parametric techniques such as Fourier or Hilbert transformations. In the case of temporal-domain FC estimators, it is still possible to derive estimates in the frequency domain by pre-bandpassing the signals, e.g. *phase-locking value*.

Another element involved in the choice of the estimator is the temporal resolution of the neuroimaging technique. In fact, EEG and MEG signals are characterized by high temporal dynamics, in the order of milliseconds, while fMRI data exhibit low temporal dynamics, in the order of seconds. Thus, EEG and MEG signals can exhibit signal changes in a very broad frequency range, from infra-slow ($<$1 Hz) to ultra-fast ($>$100 Hz) dynamics depending on the task and on the presence of pathological conditions (Palva and Palva, 2007; McFarland et al., 2000; Jacobs et al., 2009). For this reason, frequency-domain methods are more

appropriate with EEG/MEG signals as they allow to isolate FC in specific frequency bands of interest. On the contrary, time-domain methods, such as Pearson correlation and partial correlation (Marrelec et al., 2006), can be more appropriate with fMRI data, where the available frequency range is rather limited (i.e. < 1 Hz) (Fox and Raichle, 2007).

1.2.3 From brain connectivity to networks

After computing FC for each pair of signals, the corresponding values can be collected in the so-called connectivity matrix A, i.e. a $N \times N$ matrix, where N is the number of nodes (sensors, ROIs, ...) and the entry a_{ij} contains the FC value for the connection, or edge, between the nodes i and j (**Fig 1.1**).

Diagonal elements a_{ii} correspond to FC of a node with itself. Because their interpretation is not trivial, the main diagonal of the connectivity matrix is typically set to null values. In addition, in presence of directed FC the direction of the connection must be specified to correctly read the connectivity matrix. In fact, while for undirected FC there is a symmetric relation between the elements of the connectivity matrix (i.e., $a_{ij} = a_{ji}$) for directed FC the relation becomes asymmetric (i.e., $a_{ij} \neq a_{ji}$).

The values contained in the connectivity matrix depend on the nature of the employed FC estimator. While the majority of the methods give normalized values within the [0,1] interval, there might be in general different ranges or scales. In these situations, it is often preferable to transform the data, taking into account the nature of the FC estimator, so to re-scale them within the normative interval. For example, Pearson correlation gives values that span the interval $-1 \leq a_{ij} \leq 1$, i.e. from perfect anticorrelation (anti-phase) to perfect correlation (in-phase). However, since it might be difficult to interpret the negative values from a neurophysiological perspective (i.e., true anti-phase behavior or simple delayed interaction), a common procedure is to consider the absolute values in the corresponding connectivity matrix and interpret their magnitude as general correlation.

Statistical approaches based on known properties of the estimators or on data surrogates can be eventually used to remove non-significant FC values (De Vico Fallani et al., 2014).

1.3 Network science to model functional connectivity

Together, nodes and edges form a new type of networked data that cannot be studied with standard tools, but needs appropriate techniques from *network science*, i.e. a modern field that draws on graph theory, statistical mechanics, data mining and inferential modeling (Albert and Barabási, 2002; Vespignani, 2018). Network science allows to analyze *complex systems* at different spatial scales – from molecular biology to social sciences – and to quantify organizational mechanisms by extracting indices that characterize specific topological properties (Boccaletti et al., 2006; Newman, 2012).

In this framework complex networks are modeled as graphs, i.e. mathematical objects defined by nodes and edges (Newman, 2012). After being constructed, the resulting brain network corresponds to a weighted graph whose edges code for the magnitude of the FC between different nodes. Common courses in brain network analysis typically use thresholding procedures to filter the raw networks by retaining, and eventually binarizing, a certain

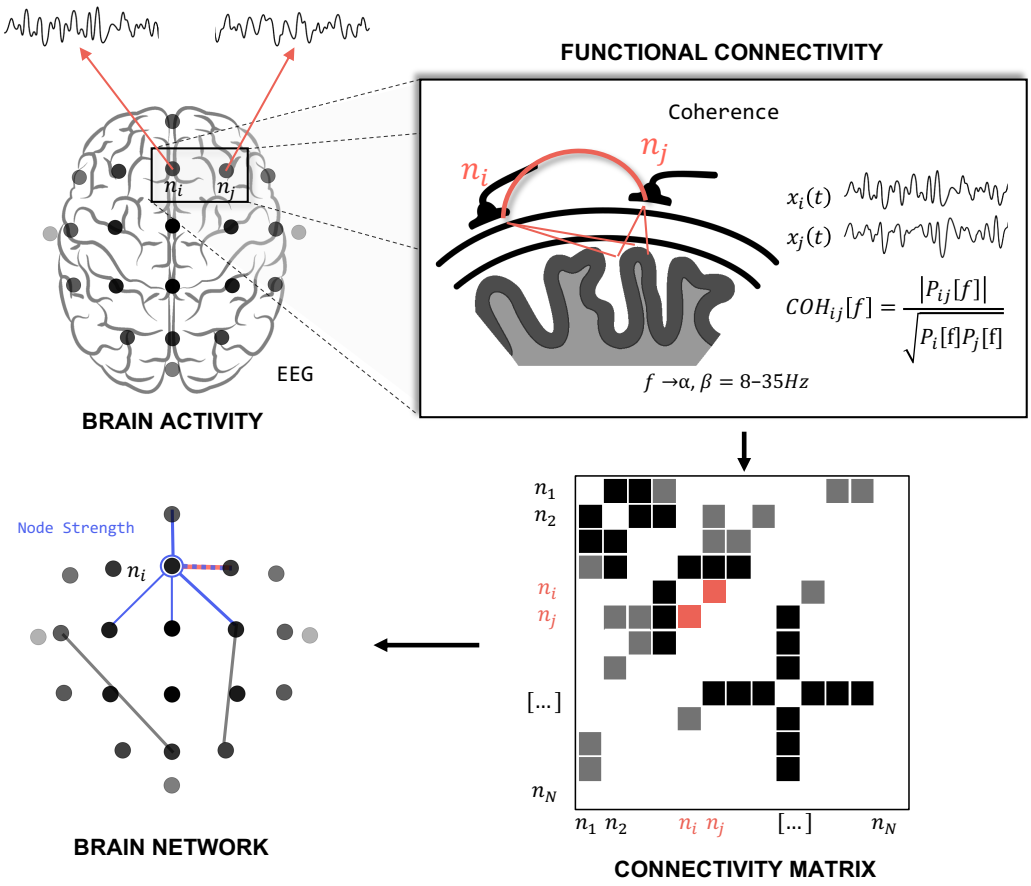


FIGURE 1.1: **Brain network construction.** Nodes correspond to specific brain sites according to the used neuroimaging technique. In an EEG system, each electrode is translated into a node. Links are estimated by measuring the FC between pairs of nodes; this information is summarized in a connectivity matrix. By means of filtering procedures, based on thresholds, only the most important links constitute the brain graph. The topology of the resulting brain network can be quantified by different graph metrics (or indices) (e.g. node strength).

percentage of the available links. These procedures typically result in sparse networks with a relatively low connection density (Box 1.3). Despite the consequent information loss, thresholding is often adopted to mitigate the uncertainty of the estimated weakest edges, reduce the false positives, and facilitate the interpretation of the inferred network topology (De Vico Fallani, Latora, and Chavez, 2017; Smith et al., 2013; Sherbondy, Rowe, and Alexander, 2010).

The simplest way to proceed is to fix a threshold on the number of strongest links to retain or on the FC value. However, these approaches are parametric and researchers are often required to repeat the analysis for a broad range of different thresholds and eventually select the one belonging to an interval for which results remain relatively stable.

Since these approaches might be considered suboptimal, they can be alternatively replaced by theoretically-grounded non-parametric methods based on different criteria including statistical contrasts with data surrogates (Winterhalder, Schelter, and Timmer, 2006; Valencia et al., 2009), topological optimization (Tumminello et al., 2005; Serrano, Boguná, and Vespignani, 2009; Tewarie et al., 2015; De Vico Fallani, Latora, and Chavez, 2017) and

population-based consensus (Roberts et al., 2017; Betzel et al., 2019).

After thresholding, network properties can be extracted from the resulting sparse networks, which can be weighted or unweighted depending on whether the remaining weights are maintained or binarized. For the sake of simplicity, we will describe in the following graph theoretic metrics in the case of undirected and unweighted networks, mentioning how they can be extended in the general cases.

Box 1.2 - Basic characteristics of graphs or networks
Density: ratio of actual number of edges and the number of total possible edges in the network. Brain networks tend to be relatively sparse (i.e. density $< 50\%$) (Ringo, 1991), although there is a high variability due to thresholding procedures.
Walk, cycle and path: a walk is a sequence of successive edges which joins a sequence of nodes. A cycle is a closed walk where the first and last nodes coincide. A path is a walk in which all edges and nodes are distinct. A graph is said to be connected if there exists a path between any possible node pair.
Distance: length of the shortest path between two nodes. In weighted graphs the shortest path is the one that minimizes the sum of the edge weights along the path (**Fig 1.2**). In brain networks, weights should be inverted when computing distances as the highest weights correspond to the strongest, most reliable, links (De Vico Fallani et al., 2014; Goñi et al., 2014).

1.3.1 Network metrics

In this section, starting from general notions, we present the main network metrics to quantify local-, meso-, and global- scale topological properties of brain networks or graphs. Local scale properties are at the level of a single node, and quantify its importance in the network according to different criteria.

Meso-scale properties refer to grouping of nodes based on distinctive interaction patterns. Global-scale properties characterize the network as a whole and represent a summary index.

As a reminder, we refer to A as to the connectivity or adjacency matrix of the filtered brain network containing N nodes and L links, or edges.

1.3.1.1 Local-scale properties

Degree: The most intuitive metric for a node is the so-called node degree which counts the number of connections with the rest of the network. For binary, undirected networks the degree of node i can be computed as

$$k(i) = \sum_{j=1, j\neq i}^{N} A_{ij} \tag{1.1}$$

The analog of node degree in weighted networks is known as node strength, which simply sums the weights of the connections of node i to the rest of the network. In the case of directed graphs, it's possible to both count the number of incoming edges of node i, and the number of outgoing edges considering the sum of the rows or columns of A (**Fig 1.2**).

The node degrees are generally used to identify the most connected nodes in the graph that hold a large part of the overall system's connectivity and therefore represent candidate hubs of the brain network (**Fig 1.2**).

Betweenness: Apart from the node degree, there are in general several ways in which a node can be considered central or important in a network. Betweenness centrality measures the extent to which a node lies "between" other pairs of nodes by considering the proportion of shortest paths (Box 1.3) in the network passing through it (Freeman, 1977; Anthonisse, 1971). In practice, the betweenness centrality of a node $i$reads as

$$C_B(i) = \frac{1}{(N-1)(N-2)} \sum_{h=1,h\neq j}^{N} \sum_{j=1,j\neq i}^{N} \frac{\sigma_{hj}(i)}{\sigma_{hj}}, \tag{1.2}$$

where $\sigma_{hj}(i)$ the number of shortest paths between nodes h and j that pass through i, σ_{hj} is the number of shortest paths between nodes h and j. Betweenness centrality can be computed in the same way for weighted and directed networks, i.e. calculating the shortest paths following the direction of the edges.

Assuming that information flow along shortest paths, the betweenness centrality can be used to identify those nodes which are crucial for the information transfer between topologically distant brain regions.

Communicability: Differently from betweenness centrality, communicability takes into account the contribution of all possible walks between node pairs (Estrada and Hatano, 2008). By doing so, communicability reflects a network's capacity for parallel information transfer.

Formally, the communicability of a node i is given by

$$C_C(i) = \sum_{j=1}^{N} [e^A]_{ij}, \tag{1.3}$$

where e^A denotes the matrix exponential of the matrix A that takes into account for each pair of nodes the total number of walks between them (Benzi and Klymko, 2013). Communicability in weighted networks can be computed by normalizing the connectivity matrix with appropriate transformations (Crofts and Higham, 2009). In the case of directed networks, heuristic approaches can be used to identify all the possible paths of a specified maximum length (Vico Fallani et al., 2012).

Communicability can be particularly suitable for identifying brain areas that are central for the diffusion of information across the network (Crofts and Higham, 2009; Chavez et al., 2013).

Eigenvector: The eigenvector centrality of a node is a metric which considers the importance of its neighbors, i.e. the nodes directly connected, or adjacent, to it. Hence, it can be thought as being equivalent to the summed centrality of its neighbors (Bonacich, 1972). The eigenvector centrality of a node i is obtained by computing graph spectrum and reads as

$$C_E(i) = \frac{1}{\lambda} \sum_{j=1}^{N} A_{ij} v_j, \tag{1.4}$$

where λ is the largest eigenvalue of A and v is the associated leading eigenvector. Eigenvector centrality can be extended to weighted networks, subject to certain conditions (Newman, 2012; Newman, 2004). In this case, A must be positive definite and this condition might not be satisfied for correlation-based networks which also contain negative entries. One solution is to remap edge weights to a positive range, by taking for instance the absolute value of the correlation coefficients. In directed networks, the adjacency matrix A is asymmetric and there are two leading eigenvectors, which can be therefore used to isolate the contribution of either incoming or outgoing edges.

Eigenvector centrality can be used to identify brain areas which do not necessarily have a high number of links, but that are connected to other central regions (Lohmann et al., 2010).

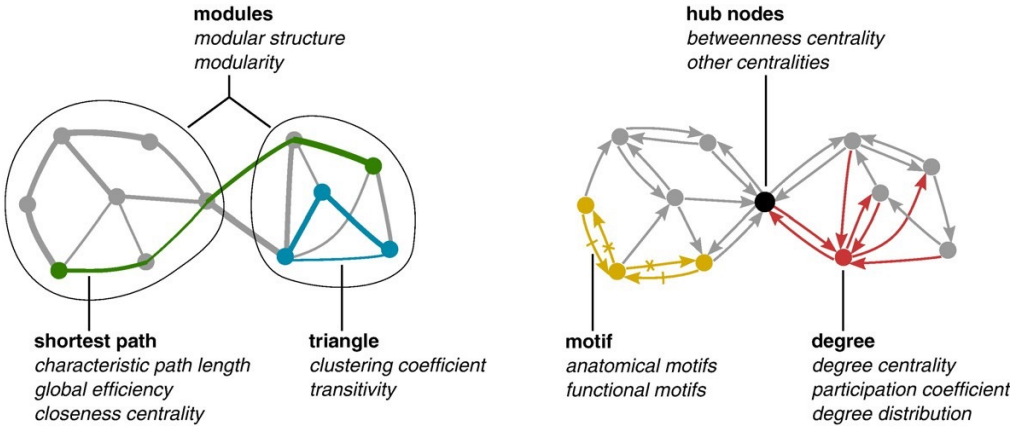


FIGURE 1.2: **Measures of network topology.** An illustration of key complex network measures (in italics). These measures are typically based on basic properties of network connectivity (in bold type). Thus, measures of integration are based on shortest path lengths (green), while measures of segregation are often based on triangle counts (blue) but also include more sophisticated decomposition into modules (ovals). Measures of centrality may be based on node degree (red) or on the length and number of shortest paths between nodes. Hub nodes (black) often lie on a high number of shortest paths and consequently often have high betweenness centrality. Patterns of local connectivity are quantified by network motifs (yellow). An example three-node and four-link anatomical motif contains six possible functional motifs, of which two are shown—one motif containing dashed links, and one motif containing crossed links. Extracted from Rubinov and Sporns, 2010.

1.3.1.2 Meso-scale properties

Motifs: Network motifs are subgraphs that repeat themselves in a network. Each of these subgraphs, defined by a particular pattern of interactions between nodes, often reflects a mode in which particular functions are realized by the network.

The motif detection can be done under various paradigms including exact counting, sampling, and pattern growth methods (Masoudi-Nejad, Schreiber, and Kashani, 2012).

After calculating the frequency F -as the number of occurrences- of a subgraph G the assessment of its significance is given by

$$Z(G) = \frac{F(G) - \mu(G)}{\sigma(G)} \tag{1.5}$$

Where μ and σ indicate respectively the mean and standard deviation of the frequency of the subgraphs in an ensemble of random networks corresponding to a null-model associated to the empirical network (see next subsection). The resulting Z-score indicates if the motif G is occurring either more or less than expected by chance. While motif detection naturally applies to binary networks, the extension to weighted ones can be achieved by replacing the motif occurrence with its intensity (Onnela et al., 2005).

Motifs represent the basic building blocks of a network and may provide a deep insight into the brain network's functional abilities (Sporns, Kötter, and Friston, 2004; De Vico Fallani et al., 2008b), albeit their detection is computationally challenging as the number of nodes becomes higher than six (Milo et al., 2002) (**Fig 1.2**).

Communities and modularity: Communities, or modules, are often defined in terms of network partitions where each node is assigned to one and only one module (**Fig 1.2**). Community detection structure is not trivial and many algorithms to identify community structures are available. For instance, they may be based on hierarchical clustering, spectral embedding, statistical inference, and more recently machine learning approaches (Fortunato and Hric, 2016; Liu et al., 2020).

The quality of the identified partition can be measured by the so-called modularity index

$$Q = \frac{1}{2L} \sum_{i=1}^{N} \sum_{j=1}^{N} (A_{ij} - R_{ij}) \delta(m_i, m_j), \tag{1.6}$$

where R_{ij} is the probability to observe an edge as expected by chance and the Kronecker delta $\delta(m_i, m_j)$ equals one if nodes i and j belong to the same module (i.e., $m_i = m_j$) and zero otherwise. When Q is positive, the network tends to have high intra-module connectivity and low inter-module connectivity; when Q is less than or equal to zero then the network lacks a modular structure. The above equation can be extended to the analysis of weighted (Newman, 2004), and directed networks (Leicht and Newman, 2008).

In brain networks, topological modules tend to be spatially localized, and they typically include cortical areas that are known to be specialized for visual, auditory, and motor functions (Sporns and Betzel, 2016).

Core-periphery structure: Core-periphery is a peculiar partition of the network consisting of a group of tightly connected nodes (i.e. the core), and a group made by the remaining weakly connected nodes (i.e. the periphery) (Borgatti and Everett, 2000). Identifying the core of a network can be achieved through methods optimizing a fitness function or via statistical null models (Csermely et al., 2013). These methods rely on subjective fine-tuning of one or more free parameters and tend to be relatively complex with consequent scalability issues.

Here we report an alternative method that only requires the degree sequence and no prior knowledge on the network (Ma and Mondragón, 2015). The basic idea is to separate

the nodes in two groups based on their rank, as determined by their node centrality (e.g. the degree). The optimal separating rank position is then given by

$$r^* = \underset{r}{\text{argmax}}(k_r^+), \tag{1.7}$$

where k_r^+ is the number of links that a node of rank r shares with nodes of higher rank. This method has the advantage of being fast, highly scalable and it can be readily applied to weighted and directed networks.

In brain networks, core-periphery organization is thought to emerge as a cost-effective solution for the integration of distributed regions in the periphery (Battiston et al., 2018). A related concept is that of rich-club behavior, where the brain network hubs tend to be mutually interconnected (Van Den Heuvel and Sporns, 2011).

1.3.1.3 Global-scale properties

Characteristic path length and global-efficiency: The characteristic path length is a scalar that measures the global tendency of the nodes in the network to integrate and exchange information. Assuming that the information flows through the shortest paths, the characteristic path length is given by (Watts and Strogatz, 1998)

$$P = \frac{1}{N(N-1)} \sum_{i=1, i\neq j}^{N} d_{ij}, \tag{1.8}$$

where d_{ij} is the distance between nodes i and j. Because the distance between two nodes that are not connected through any path is equal to infinity, P is ill-defined for disconnected networks.

To overcome this issue, the $efficiency$ between two nodes as the reciprocal of their distance was introduced. With this measure the contribution of two disconnected nodes becomes zero. Hence, the global-efficiency of a network is a normalized scalar given by [5]

$$E_{glob} = \frac{1}{N(N-1)} \sum_{i=1, i\neq j}^{N} \frac{1}{d_{ij}}, \tag{1.9}$$

Both P and E_{glob} can be easily applied to directed and weighted networks taking into account the appropriate way to compute the distance (Box 1.3). Characteristic path length and global-efficiency represent two of the most widely used measures of integration in brain networks because of the simplicity of their interpretation (Rubinov and Sporns, 2010).

An average short distance between the nodes may constitute a biological mechanism to minimize the energetic cost associated with long-range connectivity, and could provide more efficient and less noisy information transfer (Bullmore and Sporns, 2009; Achard and Bullmore, 2007).

Clustering coefficient and local-efficiency: Clustering is an important feature in complex networks that measures the extent to which nodes' neighbors are mutually interconnected. Strongly related to the presence of triangles in the network (i.e. triads of nodes fully connected), the clustering coefficient is a normalized scalar given by

$$C = \frac{1}{N}\sum_{i=1}^{N}\frac{2l_i}{k_i(k_i-1)}, \tag{1.10}$$

where l_i is the number of links between the neighbors of node i and k_i its node degree. The extension to weighted and directed networks was proposed in (Onnela et al., 2005; Fagiolo, 2007).

Alternatively, the overall tendency of a network to form a clustered group of nodes can be obtained in terms of network global-efficiency. The so-called local-efficiency is given by averaging the global-efficiencies of the network's subgraphs

$$E_{loc} = \frac{1}{N}\sum_{i=1}^{N}E_{glob}(G_i), \tag{1.11}$$

where G_i denotes the subgraph comprising all nodes that are immediate neighbors of the i^{th} node. In brain networks, the clustering coefficient and local-efficiency are often interpreted as a measure of functional segregation or specialization (Wig, 2017).

Together with distance-based metrics (P and E_{glob}), clustering metrics are used to quantify the small-world properties of a network, i.e. the tendency to optimize simultaneously integration and segregation of information (Watts and Strogatz, 1998). Because the strong parallel with a plausible model of neural functioning these metrics are the most widely used in the field of network neuroscience (Latora, Nicosia, and Russo, 2017).

In practice, the small-world propensity can be computed by normalizing the values of the empirical network with those obtained from network surrogates, such as random graphs (Humphries, Gurney, and Prescott, 2006). Then, a small-world index can be obtained, for example, by combining the normalized P and C values

$$w = \frac{C}{\mu(C_{rand})}\frac{\mu(P_{rand})}{P}, \tag{1.12}$$

where P_{rand} and C_{rand} are vectors containing the values obtained for the network surrogates. Notably, other types of small-world indexes can be obtained by opportunely substituting P and C, with E_{glob} and E_{loc} (De Vico Fallani et al., 2013), or by adopting normalization with other types of network surrogates (Telesford et al., 2011).

We report in **Table 1.2** the time complexity of the above metrics for unweighted and weighted networks (Latora, Nicosia, and Russo, 2017).

1.3.2 Normalizing network metrics

Most measures of network organization scale with the number of nodes and edges in a graph. Thus, to compare the values of the metrics extracted from different size and connection densities, it is often necessary to account for basic properties of the underlying network. As mentioned before, normalization with respect to null, or reference, models provides a practical benchmark to determine the extent to which a network property deviates from what would be expected by chance and to compare network properties across different conditions (Rubinov and Sporns, 2010; Van Wijk, Stam, and Daffertshofer, 2010; Newman, 2012).

TABLE 1.2: Computational complexity of network metrics. N = number of nodes; L = number of links; g = size of the motif; $\bar{k}$ = average node degree; k^2 = node degree variance.

		Computational complexity	
	Network metric	Unweighted	Weighted
Local-scale	Degree	$O(L)$	$O(L)$
	Betweenness	$O(N(N+L))$	$O(N(L+NlogN))$
	Communicability	$O(N^3)$	$O(N^3)$
	Eigenvector	$O(N^2)$	$O(N^2)$
Meso-scale	Motifs	$O(Ng)$	-
	Communities	$O(NlogN)$	$O(NlogN)$
	Modularity	$O(L)$	$O(L)$
	Core-periphery	$O(L+NlogN)$	-
Global-scale	Characteristic path length	$O(N(N+L))$	$O(N(N+L))$
	Global-efficiency	$O(N(N+L))$	$O(N(N+L))$
	Clustering coefficient	$O(Lk^2/\bar{k})$	$O(Lk^2/\bar{k})$
	Local-efficiency	$O(N(k^2-\bar{k}))$	$O(N(k^2-\bar{k}))$

Generating reference networks that match all properties of an actual network except for the one that has to be normalized is difficult in practice, since most properties are interrelated. It is therefore usual to match only basic properties, such as network size, connection density, and degree distribution. This kind of null network is typically obtained using randomization strategies, where the actual network is randomly rewired according to a set of rules. In particular, the rewiring may be performed either preserving the degree distribution or not, the former being a more conservative choice (Maslov and Sneppen, 2002).

Because the rewiring process is stochastic, a certain number of network samples - typically higher than 100 - should be generated in order to constitute an ensemble of reference networks with similar characteristics.

The normalized value of a metric can then be computed as the ratio of the value measured on the observed network (M_{obs}) and the mean obtained from the randomized network ensemble

$$M' = \frac{M_{obs}}{\mu(M_{rand})}, \tag{1.13}$$

While the ratio is the preferred way to normalize network metrics, Z-scores procedures can be used as well (Eq. 1.5). Notably, rewiring procedures that preserve the degree distribution have been extended to weighted and signed networks (Rubinov and Sporns, 2011). While generating purely random network ensembles is the most intuitive way of normalizing, alternative strategies that generate more complex null models might be adopted too (Box 2.2).

1.3.3 Advanced network approaches

The previous paragraphs introduced some of the well-established graph metrics used in network neuroscience that might be particularly relevant to BCI applications. Nonetheless,

the field of network science is quickly advancing and new research directions are currently in development to address the open challenges.

First, the above mentioned graph metrics have been mainly conceived as topological descriptors of static networks, whose links do not change in time. This is an oversimplification of the real phenomena as brain networks are intrinsically dynamic and functional connectivity can change across multiple time scales (e.g., within and between BCI sessions). Hence, time must be formally considered as a part of the network problem and not merely as a repeated measure (Holme and Saramäki, 2012). In neuroscience, many network metrics have been rethought temporally by considering the nature of time-respecting paths (Tang et al., 2010) and the persistence of specific motifs (De Vico Fallani et al., 2008b) and modules (Bassett et al., 2011). The theoretical development of temporal networks appears therefore particularly relevant for future BCI-related studies.

Second, the characteristics of the brain network strongly depend on the neuroimaging technique (i.e., the nodes) and on the type of functional connectivity estimator used (i.e., the edges). That means that multiple brain networks simultaneously characterize the same subject. Multilayer networks have been recently introduced to provide theoretically grounded metrics integrating the available information from multiple sources (De Domenico et al., 2013; Boccaletti et al., 2014). In multilayer brain networks, different types of connectivity are represented on different layers (e.g., neuroimaging modality (Battiston et al., 2018) and frequency bands (De Domenico, Sasai, and Arenas, 2016; Guillon et al., 2017)) and connectivity can span both within and between layers (e.g., cross-frequency coupling (Jirsa and Müller, 2013)). Notably, multilayer network metrics are able to extract higher-order information that cannot be obtained by simply aggregating connectivity across layers. Therefore, this innovative framework for integrating different connectivity levels might be particularly useful for the development of multimodal BCI systems (Corsi et al., 2019).

Together with the descriptive nature of the network metrics (which are intrinsically data-driven) the development of network models could greatly advance the study of brain networks in BCI by providing complementary statistical information. Since brain networks, as in other real networks, are typically inferred from experimental data their edges are subject to statistical uncertainty. Stochastic network models based on spatial, topological or Bayesian rules have been recently introduced to tackle those aspects and obtain a more robust understanding of the organizational properties of complex brain networks (Betzel et al., 2016; Obando and De Vico Fallani, 2017; Faskowitz et al., 2018). Finally, approaches based on network controllability (Liu, Slotine, and Barabási, 2011) could be used in brain networks to identify the driver nodes that could be experimentally targeted by BCI feedback to elicit specific mental states or behaviors (Stiso et al., 2020).

Chapter 2

Network-based BCI

Parts of this chapter has been published in *Journal of Neural Engineering*:

- **Title:** Network-based brain computer interfaces: principles and applications
- **Authors:** Juliana Gonzalez-Astudillo, Tiziana Cattai, Giulia Bassignana, Marie-Constance Corsi and Fabrizio De Vico Fallani
- **DOI:** doi:10.1016/j.media.2021.102219

2.1 Network properties underlying BCI motor tasks

BCIs involve a complex mixture of cognitive processes not necessarily directly linked with the targeted task (Golub et al., 2016; Farahani, Karwowski, and Lighthall, 2019). Among them are attention and task engagement (Walz et al., 2013), working memory and decision-making (Hampson et al., 2006; Stanley et al., 2015; Gong et al., 2016; Markett et al., 2018), but also error-potential have been shown to occur during BCI tasks (Buttfield, Ferrez, and Millan, 2006; Ferrez and Millán, 2008; Chavarriaga and Millán, 2010). These higher-order cognitive processes result from interactions between different areas that engender brain network reorganization. Here, we will specifically focus on the network changes underlying motor (executed and imagined) performance, which is largely studied in the literature and directly associated with one of the most used BCI paradigms.

2.1.1 Short-term dynamic network changes during motor tasks

Performing motor imagery-based BCI experiments consists of the voluntary modulation of α/β activity to control an object (Guillot and Collet, 2010). The analysis of event-related desynchronization and event-related synchronization enables the detection of mental states (Pfurtscheller and Da Silva, 1999; Pfurtscheller and Aranibar, 1977; Pfurtscheller et al., 1997; Pfurtscheller et al., 2006; Neuper and Pfurtscheller, 2001). Notably, motor imagery (MI) and execution (ME) tasks have been shown to share similar characteristics such as the spatial and frequency localization of the evoked brain activity (Wilson et al., 2009; Munzert, Lorey, and Zentgraf, 2009; Lotze and Halsband, 2006).

Meta-analyses, mainly based on fMRI and PET studies, recently revealed a group of regions involved during ME (McDougle, Ivry, and Taylor, 2016) and MI (Hétu et al., 2013), including premotor area (PMA), primary sensorimotor area (S1), supplementary motor area (SMA), posterior parietal lobe. Notably, Hardwick et al., 2018, made a comparison between imagery, observation and execution. They identified two main clusters involved in both MI and ME: bilateral cortical sensorimotor and premotor clusters. They also performed contrast analyses to elicit regions more consistently involved in MI than in ME. It appeared that MI tends to recruit more often premotor regions and left inferior and superior parietal cortex. These results seem to be corroborated by studies performed from a network perspective. By using betweenness centrality, Xu et al., 2014 showed that in ME the most important region lies in the SMA cortex whereas during MI the most central area was located in the right PMA. In the case of ME, it would suggest that SMA could enable an efficient communication between brain areas, especially motor ones (Luppino and Rizzolatti, 2000; Cauda et al., 2011) during sequential execution. In the case of MI, PMA could integrate both sensorimotor information from motor areas (e.g. SMA) and spatial information of movements from regions such as posterior parietal lobe to enable motor planning (Luppino and Rizzolatti, 2000; Hoshi and Tanji, 2007; Kantak et al., 2012).

Complementary to the previous studies, another approach consists of studying time-varying network properties while performing tasks (De Vico Fallani et al., 2008b; Valencia et al., 2008; De Vico Fallani et al., 2008a). In the specific case of motor tasks, a work based on the use of time-varying partial direct coherence (PDC) revealed that the cingulate motor areas could be seen as a hub of outgoing flows during dorsal flexions of the right foot (De Vico Fallani et al., 2008a). Based on experiments performed with five subjects via a 64-EEG channel system, the authors observed changes of network patterns at different stages of the task. The preparation of the movement presented a high level of efficiency, associated with an increase of clustering coefficient and a reduction of the characteristic path length. During the movement, strong functional links between the cingulate motor and the supplementary motor areas were obtained but also a lower network efficiency at the global level. These results illustrate the existence of a dynamic network reorganization process during the preparation and execution of a simple motor task.

2.1.2 Long-term longitudinal network changes during BCI learning

Understanding how we learn to use a BCI is crucial to adapt to individual variability and improve performance. Learning is a complex phenomenon that can be categorized in different types such as instructed (supervised (Knudsen, 1994) or reinforced (Dayan and Niv, 2008)) or unsupervised (Barlow, 1989), explicit or implicit (Seger, 1994).

Regardless of the type of learning, it is characterized by changes in brain associations from microscale, with the synapse strengthening for example, to macroscale levels, including changes of functional brain connectivity. In this section, we present some of the recent studies using network science approaches to characterize large-scale neural processes of human learning at the macroscale (Seger and Miller, 2010; Bassett and Mattar, 2017).

2.1.2.1 Motor learning

In the past years, studies focusing on functional connectivity demonstrated changes induced by motor skill learning. Comparisons made before and after a locomotor attention training revealed an alteration of the connectivity in the sensorimotor areas potentially modulated by focusing attention on the movements involved in ambulation (Katiuscia et al., 2009). Sensorimotor adaptation tasks involve notably prefrontal cortex, premotor and primary motor and parietal cortices (McDougle, Ivry, and Taylor, 2016) and once acquired, motor skills are encoded in fronto-parietal networks (Rizzolatti and Luppino, 2001). However, little is known about its evolution through training.

In Taubert et al., 2011, fourteen healthy subjects performed a dynamic balance task once a week during six consecutive weeks. They underwent four fMRI scans: before the first, the third, the fifth sessions and one week after the training program. The authors observed an increased fronto-parietal network connectivity in one week. Training sessions progressively modulated these modifications. Changes induced by motor imagery learning have been observed, notably in resting-state functional connectivity of the default mode network (DMN) (Ge et al., 2014). These results prove that motor learning relies on areas beyond those directly involved during the task performance and illustrate the need to study how communication between brain regions evolves during the training.

From a network perspective, a large number of metrics characterizing the topological properties have been considered to capture the motor acquisition process. Heitger et al., 2012 showed that the motor performance improvement of a complex bimanual pattern was associated with an increase of clustering coefficient and a shorter communication distance. However, it should be considered that the latter one was possibly influenced by the reported higher connection density and strength.

Network modularity has been used as a marker in the case of age-related changes (Meunier et al., 2009) but also in the case of induced brain plasticity (Gallen and D'Esposito, 2019). Therefore, it seems particularly of interest in the study of learning process as it captures changes in the modular organization of the brain (Bassett et al., 2011). In the specific case of motor skill acquisition based on the practice of finger-movement sequences over six weeks, the use of modularity revealed that learning induced an autonomy of sensorimotor and visual systems and individual differences in amount of learning could be predicted by the release of cognitive control hubs in frontal and cingulate cortices (Bassett et al., 2015).

Based on the temporal extension of network modularity, Bassett et al., 2011 defined the "flexibility" as the number of times a node changes its module allegiances between two consecutive time steps. This measure was used to study the evolution of brain network properties during a motor learning task. Twenty-five healthy subjects were instructed to generate responses to a visually cued sequence by using the four fingers of their non-dominant hand. They participated in three training sessions in a five-day period, performed inside the fMRI. The flexibility predicted the relative learning rate, particularly in frontal, pre-supplementary motor, posterior parietal and occipital cortices.

2.1.2.2 Neurofeedback and human learning

To master closed-loop systems such as neurofeedback (NFB) or BCIs, several training sessions are typically needed. Recent studies suggest that the involved learning process is analogous to cognitive or motor skill acquisition (Hiremath et al., 2015). NFB could induce behavioral modifications and neural changes within trained brain circuits that last months after training (Sitaram et al., 2017). At microscale, changes at the neuronal level have been observed and simulated during BCI learning (Ito et al., 2020). At larger spatial scales, the recruitment of areas beyond those targeted by BCI has been observed during the skill acquisition (Wander et al., 2013; Orsborn and Pesaran, 2017). For example, the decrease of the global-efficiency in the higher-beta band indicated the involvement of a distributed network of brain areas during MI- based BCI training (Pichiorri et al., 2011). These findings motivated a deeper understanding of the brain network reorganization, at the macroscale, underlying the BCI/NFB learning process.

In a recent study, Corsi et al., 2020 studied how the brain network reorganizes during a MI-based BCI training. Twenty healthy, and BCI-naïve, subjects followed a four-session training over two weeks. The BCI task consisted of a standard 1D two-target task (Wolpaw and McFarland, 2004). To hit the up-target, the subjects had to perform a sustained MI of right-hand grasping and to hit the down-target they remained at rest. MEG and EEG signals were simultaneously recorded during the sessions.

Results obtained from the relative node strength showed a progressive reduction of integration among, primary visual areas and associative regions, within the α and β frequency ranges. This metric could also predict the learning rate more specifically in the anterior part of the cingulate gyrus and the orbital part of the inferior frontal gyrus, both known to be involved in human learning (Euston, Gruber, and McNaughton, 2012), and the frontomarginal gyrus and the superior parietal lobule, which is associated with learning and motor imagery tasks (Stephan et al., 1995; Solodkin et al., 2004). To fully take advantage of the behavioral and MEG information to predict learning, a multimodal network approach has been adopted by Stiso et al., 2020. The authors used a non-negative matrix factorization to identify regularized, covarying subgraphs of functional connectivity to estimate their similarity to BCI performance and detect the associated time-varying expression. From their observations, they deduced a model tested via the network control theory in which specific subgraphs support learning via a modulation of brain activity in areas associated with sustained attention.

Despite the promising evidence, brain network reorganization needs to be further investigated to better understand learning mechanisms underlying the use of BCI devices and enhance the usability in clinical applications (De Vico Fallani and Bassett, 2019; Orsborn and Pesaran, 2017).

2.1.3 Clinical applications: the case of stroke

It is well known that neurological or psychiatric disorders lead to changes in terms of communication between brain regions (Stam, 2014). For example, connectivity between high-degree hub nodes has been observed in schizophrenia (Van Den Heuvel et al., 2013) and comatose

patients (Achard et al., 2012). Decreased global- and local-efficiencies has been reported in Parkinson disease (Skidmore et al., 2011), while modifications of the core-periphery structure (Guillon et al., 2019) and a loss of inter-frequency hubs has been found in Alzheimer disease (Guillon et al., 2017). In the case of attention-deficit/hyperactivity disorder in children the increase of local-efficiency and lower global efficiency suggested a disorder-related tendency toward regular organization (Wang et al., 2009). In addition, modifications in nodal properties have been observed in both children and adults in the attention, sensorimotor and DMN (Raichle et al., 2001) and striatum (Wang et al., 2009; Di Martino et al., 2013; Hart et al., 2013).

Brain network changes in stroke patients are particularly relevant for BCI clinical applications and neurofeedback rehabilitation strategies. Recent studies showed that stroke recovery is accompanied by an increased smallworldness, which supports increased efficiency in information processing (Grefkes and Fink, 2011; Westlake and Nagarajan, 2011). Laney et al., 2015 performed a study with ten stroke patients that participated in six-weeks training sessions dedicated to improve voluntary motor control. fMRI data were collected, before and after training, while patients performed an auditory-cued grasp and release task of the affected hand. Finger extensions were assisted by an MRI compatible exoskeleton. Two opposite effects were observed: an increased node closeness-centrality (Boccaletti et al., 2006) with sensorimotor and cerebellum networks and a decreased closeness-centrality in the DMN and right frontal-parietal components. The authors associated the former to an improved within-network communication and the latter to a reduced dependence on cognition as motor skill enhanced (Laney et al., 2015). In another study (Termenon et al., 2016), authors aimed to characterize the brain network reorganization after stroke in the chronic stage in a group of twenty patients. Brain networks were constructed by estimating wavelet correlation from fMRI signals. They showed an overall reduction of connectivity in the hubs of the contralesional hemisphere as compared to healthy controls. Most of these studies are based on a static representation of the brain plasticity and partially inform on the individual ability of stroke patients to recover motor or cognitive functions. Recently, an approach based on temporal network models that aimed at tackling these issues indicated that both the formation of clustering connections within the affected hemisphere and interhemispheric links enabled to characterize the longitudinal network reorganization from the subacute to the chronic stage (Obando et al., 2019). These mechanisms could predict the chronic language and visual outcome respectively in patients with subcortical and cortical lesions.

MI has been proved to be a valuable tool in the study of upper-limb recovery after stroke (Sharma, Pomeroy, and Baron, 2006). It enabled observations of changes in ipsilesional intrahemispheric connectivity (Pichiorri et al., 2015) but also modifications in connectivity in prefrontal areas, and correlations between node strengths and motor outcome (Sharma, Baron, and Rowe, 2009). Within the β frequency band, performing a MI task of the affected hand induced lower small-worldness and local-efficiency compared to the MI of the unaffected hand (De Vico Fallani et al., 2013). Based on previous observations in resting-state (Dubovik et al., 2012), a recent double-blind study revealed that node strength, computed from the ipsilesional primary motor cortex in the α band, could be a target for a MI-based NFB and

lead to significant improvement on motor performance (Mottaz et al., 2018).

2.2 Network features for improving BCI performance

The use of network approaches in BCI is a relatively young and unexplored area, yet, the existing publications show encouraging results. In this section, we first provide a proof-of-concept on simulated data to illustrate the theoretical benefit of using network metrics from a classification perspective. Then, we present some of the recent classification results obtained with neuroimaging data during real BCI experiments.

Box 2.2 - Network generative models

Random networks are generated with the Erdös-Rényi (ER) model. They are constructed by fixing a parameter p which fixes the probability to have a link between two randomly selected nodes in the graph. By construction, p coincides with the connection density of the resulting networks. In general, ER networks do not exhibit any particular structure but typically low characteristic path lengths (Erdős, Rényi, et al., 1960).

Small-world networks are generated with the Watts-Strogatz (WS) model. Starting from a ring lattice graph, where each node is connected to its first k neighbors, WS networks are generated by rewiring the links with a probability p_{WS} i.e. the model parameter. With relatively low values of p_{WS}, the resulting networks exhibit both high clustering coefficient and low characteristic path length. This is a feature observed in many real-world interconnected systems and it optimizes both segregation and integration of information (Watts and Strogatz, 1998).

Scale-free networks are generated with the Barabási-Albert (BA) model. Its construction starts with m_0 nodes. Then, new nodes are iteratively added with m links ($m \leq m_0$) that connect them to existing nodes with a probability p_{BA} proportional to their node degree. As a result of such preferential attachment rule, BA networks show highly heterogeneous node degrees, few strongly connected hubs as well as low characteristic path length and null clustering coefficient. These features have been found in many real networks as a sign of resilience (Barabási and Albert, 1999).

Modular networks are generated with the stochastic block model (SBM). This model partitions the nodes in M groups of arbitrary size. Then it assigns edges between nodes with a probability that fixes the expected connection density within- (p_{intra} and between-groups (p_{inter}). By construction, SBM networks have high modularity values as well as typical small-world properties (Holland, Laskey, and Leinhardt, 1983).

2.2.1 Simulating brain network changes

In current settings, different mental strategies are used to control the MI-based BCI. The resulting brain states are translated into features that need to be properly recognized by

the classifier. To reproduce this scenario, we associated different brain states with ideal networks having distinct topological properties. Specifically, we generated synthetic networks exhibiting four different topologies, or classes, which have been extensively reported in neuroscience, i.e. small-world, modular, scale-free and random networks (Stam, 2014; Bassett and Bullmore, 2006). These networks were generated with the models described in **Box 2.2**. We fixed the same number of nodes (N=100) and links (L=600) for all of them. The specific model parameters values were: p=0.12 for random networks; p_{ws}=0.1 for small-world; m_0=m=6 for scale-free; and finally M=4 of equal size for modular ones, with p_{intra}0.46 and p_{intra}=0.01. These networks qualitatively exhibit disparate properties in terms of integration, segregation and heterogeneity of information (**Fig 2.1**A). To quantify these differences, we computed four relevant network metrics, i.e. global- efficiency, local-efficiency, modularity and degree-variance. In order to sample the distribution of these properties across models, we generated a large ensemble of 1000 networks per class (**Fig 2.1**B).

We then evaluated the performance of network metrics in discriminating the four classes as compared to the use of the entire connectivity matrix. We specifically tested 2-classes and 4-classes scenarios according to the typical number of mental states used in BCIs. To reproduce the fact that nodes might not correspond exactly to the same brain areas across different subjects -because of a natural individual spatial and functional variability (Sheng et al., 2020)- we further performed a random permutation of the node labels. Notably, this procedure did not alter the intrinsic topology of the generated networks.

Classification accuracies were finally obtained from a repeated random sub-sampling validation with 100 random balanced-splits. Specifically, the training set consisted of 80% of all the networks, while 20% of the networks were used as testing set. Results showed that when we applied node permutation classification accuracy of connectivity matrices progressively decreased down to chance levels, while network metrics always exhibited a perfect classification. More precisely, from 50% of node permutation, the accuracy obtained by using connectivity matrices was significantly lower than network metrics (Wilcoxon test, $p < 1.6 \times 10^{-10}$, Bonferroni corrected for multiple comparisons) (**Fig 2.1**C).

All network analysis and classification have been performed with the freely available networkX[1] and scikit-learn[2] packages in Phyton.

Taken together, these results indicated the theoretical benefit of including network metrics into the classification of BCI-related mental states. The development of sophisticated machine learning techniques, which operate on the entire connectivity matrices (Wu et al., 2020), could lead to similar performance in the next future, too. Finally, it is important to mention that the advantage of network metrics also lies in their relatively low computational cost and dimensionality, as well as in an easier direct interpretation.

2.2.2 State-of-the-art of network-based BCI

The first study using network metrics in MI-based BCI classification was Daly, Nasuto, and Warwick, 2012. Authors assessed the discrimination ability of mean clustering coefficient to differentiate between tap and no-tap, in real and imagined finger tapping task. They

[1] https://networkx.org/
[2] https://scikit-learn.org/

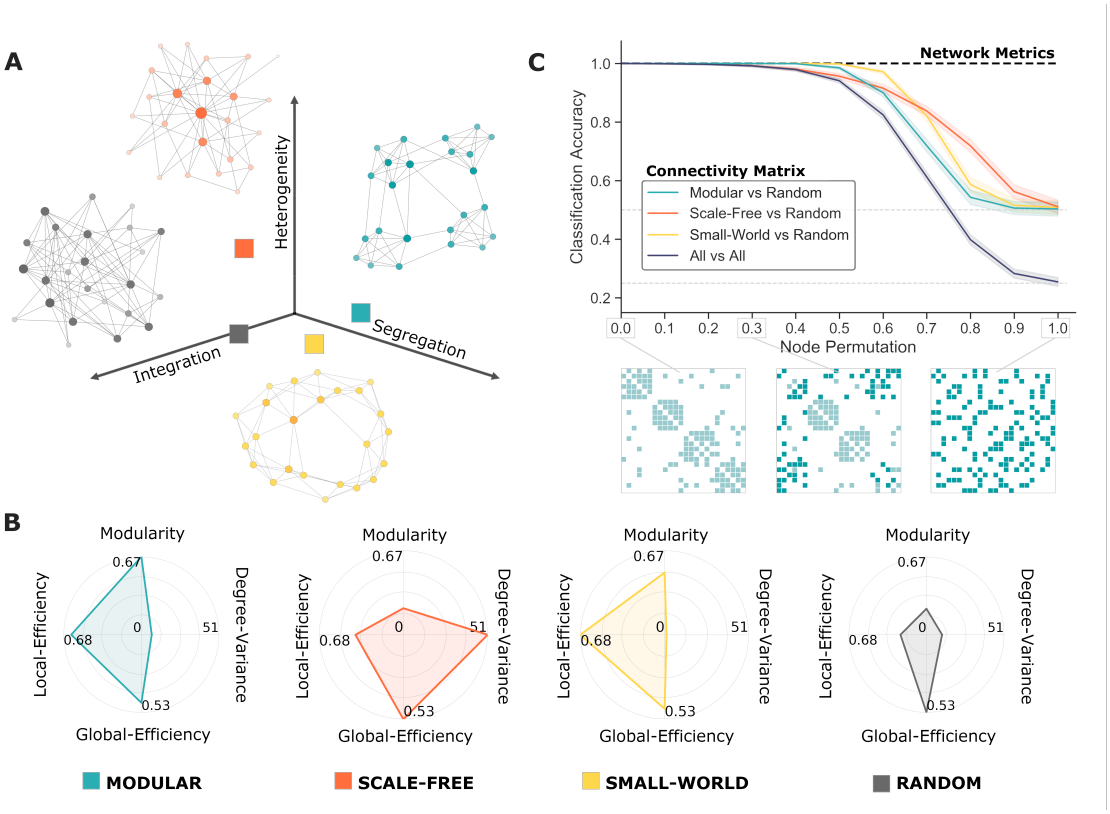


FIGURE 2.1: **Classification of networks via graph metrics**. Panel A) illustrates the graphs associated with each network class. For illustrative purposes graphs contain here N=24 nodes. Their position in the three-dimensional space qualitatively emphasizes their intrinsic properties in terms of segregation, integration and heterogeneity of information. Panel B) shows the radar plots for the mean values of the network metrics (section 1.3.1) obtained from 1000 synthetic networks generated with different network models (**Box 2.2**). Each model corresponds to a different "class" of networks. Panel C) shows the accuracy results for the classification of the synthetic networks. Notably, 2-class and 4-class scenarios were performed, which is in line with the typical number of mental states used in BCI applications. Both connectivity matrices and network metrics were fed separately as input features into the classifier. Specifically, connectivity matrices were vectorized taking into account only the upper triangular matrix. Thus, the size of the feature vectors was 4950 for connectivity matrices and 4 for network metrics, respectively. To deal with the resulting complexity, we used singular value decomposition-based linear discriminant analysis (LDA) classifiers, which implement appropriate dimensionality reductions. To challenge the classifier, we increasingly permuted in a random fashion the nodes in the connectivity matrices. This corresponded to an increasing ratio of random relabeling of the nodes in the networks (x-axis). The line plots show the average value of the classification accuracy, while standard deviation is represented as shading patches around the average (obtained from a repeated random sub-sampling validation). For illustrative purposes, we also show an example for a modular network, where the darker colors correspond to the links of the nodes which have been permuted.

recorded EEG data from twenty-two subjects performing the different task modalities. Then, to model the dynamics of inter-regional communication within the brain, they built FC networks by setting up phase synchronization links between each pair of electrodes. This resulted in a set of variable networks across time and frequency, potentially analyzable via graph theoretic tools. In order to characterize this synchronization dynamics, they computed

mean clustering coefficients over the whole collection of networks. The result was a time-frequency map of mean clustering coefficients for each trial. The statistically significant differences between conditions, tap versus no-tap, suggested the potential of using time series of clustering coefficient as classification features. Thus, satisfying the fact that these features are not temporally independent, they used Hidden Markov Models (HMMs) to model and classify the temporal dynamics of these patterns. The discriminatory capability was superior when compared to traditional band power-based features, achieving accuracies above 70% for all subjects, which was not reproduced by band power approach.

Based on the same variations in phase synchronization during MI, Stefano Filho, Attux, and Castellano, 2018 also tested the potential of graph metrics to characterize these changes. In an offline study, EEG signals were recorded from eight participants during imagination of right and left hands movements using 64 electrodes. In the same direction as (Daly, Nasuto, and Warwick, 2012), networks were built for every 1 second window of left and right MI, but in this case they filtered the time series in two frequency bands of interest, α- and β-bands. Then they computed five different graph theory metrics and used them as inputs for a least-squares based linear discriminant analysis classifier (LSLDA). At the same time, they extracted power spectrum density (PSD) features to perform a fair comparison. Using a leave-one-out cross-validation, the accuracies for single network metric classification were substantial, being around 80%, but when compared with PSD estimates its results were superior, being closer to perfect rate (99%). Nonetheless, the authors proposed a pair-wise combination of metrics which was enough to reproduce similar rates reached by PSD. Notably, the performance achieved by combined metrics involved a significantly smaller number of features, due to a selection of electrodes according to its individual classification rates. It is important to highlight that during the classification process this would be translated into less computational cost, which is encouraging when considering the implementation of network features in real time applications.

With a similar dataset, Uribe et al., 2019 investigated the potential of centrality measures to discriminate between left and right hand MI. They considered the difference between each pair of symmetric electrodes across hemispheres for every graph metric. They used degree, betweenness and eigenvector centrality to provide information regarding node's importance within the network. Two different classification methods were implemented, LDA and EML (Extreme Learning Machine), and feature selection was likewise based on classification rate improvement. Their results, expressed in terms of average classification error, showed better performance in α- band when using degree centrality and EML. In a more ambitious attempt, the authors tested their approach on the BCI Competition IV 2a database (Tangermann et al., 2012). Using a wrapper feature selection their results were ranked in the third place, while the best known performance was obtained with PSD and CSP (Common Spatial Patterns) feature selection (Ang et al., 2008).

The introduction of network-based BCI should not necessarily imply the exclusion of traditional features. Instead, it should be seen as a complementary approach to improve performance by integrating multiple neuronal mechanisms. In Cattai et al., 2021b they proposed different types of features combination. After revealing brain signal amplitude/phase synchronization mechanisms during EEG-based MI vs rest tasks, authors detected specific

brain network changes associated with MI. Based on these findings, they computed spectral-coherence and imaginary-coherence connectivity matrices. The computation was performed for frequency bins in the 4 to 40Hz band with 1 Hz resolution, considering 9 electrodes in the sensorimotor area contralateral to the imagined movement. For every MI and rest trial, they extracted three types of features, coherence-based node strength, imaginary coherence-based node strength and power spectrum density. Then they tested all possible combinations with a cross-validated LDA classifier. While single node strength discrimination gave poorer results than power spectrum, their combination led to classification improvements in most of the subjects.

Zhang et al., 2019 also demonstrated the success of multimodal features fusion. Their cross-validated classification showed that the combination of node strength, or clustering coefficient, with CSP power selection, achieved higher accuracy than single feature. Getting accuracies over 70% for certain subjects. Noteworthy, they chose the participants relying on their PSD-based MI-BCI inefficiency, i.e. its accuracy was under 70% (Kübler et al., 2001) when using power spectrum. Similar to the previous study, it is interesting to point out that they also used spectral-coherence as a connectivity estimator. Their frequency selection was reduced to α band and, in order to avoid volume conduction effects, they selected 20 spatially separated electrodes. This is a potential explanation of the fact that they even got better accuracies than CSP when using single network metrics.

In a recent study, Gu et al., 2020 explored lower limbs MI. They did a detailed analysis of synchronization patterns in α and β rhythms, to distinguish between left and right foot MI. Their study revealed a subset of sensorimotor networks exhibiting a cortical lateralization in the β band with the respect to the imagined movement. Then the assessment with multiple network metrics showed a dynamic behavior between integration and segregation across each task repetition. Exploiting these results, they used and compared three variations of sparse logistic regression (SLR) to perform feature selection combined with support vector machine (SVM) classifier. The best accuracy was up to 75%, with all participants scores above the chance level, which is notable for foot MI discrimination. Furthermore, they contrasted the classification accuracy with features extracted with CSP method, but results were not able to outperform those obtained with network metrics.

As seen in section 2.1, network analysis can also be implemented in the study of other mental processes commonly evoked in usual BCI tasks, as for example cognition. In a preliminary study conducted by Buch et al., 2018, a single subject with 122 intracranial EEG electrodes performed a test where reaction time was studied as an index of cognitive assessment. The experimental procedure consisted in a randomly chosen waiting period followed by a go signal after which the subject had to indicate its perception with a keypress; defining the reaction time as the delay between these two. Their premise was that dynamic changes in functional brain networks before and after the cue, could reflect temporal expectancy. Thus, they measured phase-locking value from sliding 500ms windows for the high γ activity (70-100 Hz) of all pairs of electrodes, i.e. nodes, constituting the weighted links between them. They found that for fast reaction time trials, the immediate pre-cue period network (500ms before the cue) was characterized by a high node strength value compared to slower reaction times. When contrasting with traditional spectral based features, they did not find

any pattern associated with reaction time variations. Going deeper in the network analysis, they computed communicability and showed a potential prediction ability based on the significant correlation between fast reaction time and high communicability in the left anterior cingulate. Motivated by these results, a SVM classifier was trained to discriminate between fast and slow trials, and then evaluated with a 10-fold cross-validation and permutation t-test. More precisely, they arbitrarily generated 2-classes labels and then randomized them 1000 times to create a null distribution of area under the curve AUCs. Results exhibited a reliable performance of the classifier (AUC = 0.72, p = 0.03). These results demonstrate the potential of network features as control signals for alternative cognitive-based BCIs.

Taken together, these results indicated the potential of network metrics as complementary features in BCIs. Future works should assess the robustness of these new features during online and real-time classification scenarios, where the reliability of the estimated brain networks becomes more challenging.

Chapter 3

Spatial brain network lateralization

3.1 Introduction

Motor imagery has been defined as the conscious mental simulation of actions involving our brain's motor representations in a way that is similar to when we actually perform movements (Jeannerod, 1995; Lotze and Halsband, 2006). MI-based Brain-Computer Interface (MI-BCI) especially relies on the imagination of kinesthetic movements of large body parts such as hands, feet, and tongue, which result in modulations of brain activity (Guillot et al., 2009). This paradigm has a wide range of applications, such as controlling devices, virtual reality, or even neurorehabilitation (Lotte, Bougrain, and Clerc, 2015).

MI is associated with event-related desynchronization (ERD) in α (8-13 Hz) and β rhythms (13-30 Hz). By contrast, resting state results in event-related synchronization (ERS). The ERD and ERS modulations are most prominent in EEG signals acquired from electrodes located above the sensorimotor cortex (Pfurtscheller and Aranibar, 1977; Beisteiner et al., 1995; Pfurtscheller and Neuper, 1997; Neuper, 1999). Previous studies further confirmed that MI activates brain regions that are linked to actual movement generation (Porro et al., 1996), regions that intervene in planning and preparation of such movements (Jeannerod, 1995; Pfurtscheller and Neuper, 1997). More especially the posterior supplementary motor area (SMA) and the premotor cortex, which seem to be predominant areas (Hétu et al., 2013). Since neurons in the SMA are involved in the preparation of movements, then it is reasonable that preparatory aspects of movement may be closely related to MI (Stephan et al., 1995). In this spatial layout of MI, another prominent characteristic is lateralization. It is well known that the motor cortex is principally involved in controlling the contralateral side of the body (Beisteiner et al., 1995). Most motor-based BCI paradigms rely on this spatial lateralization to decode MI from brain signals (Pfurtscheller and Da Silva, 1999).

These dynamics of brain oscillations in the motor cortex, associated with sensory, cognitive and motor processing, form complex spatial patterns. This is reflected in changes in the functional connectivity (FC) within the implied areas. Recent neuroimaging studies demonstrated that FC is able to capture the contralateral asymmetry of brain activity during motor-related tasks (Cattai et al., 2021b). It has also been shown that functional brain lateralization exhibits greater preference for within-hemisphere interactions in the left hemisphere, particularly for cortical regions involved in fine motor coordination. While the right hemisphere behaves in a more integrative fashion with stronger bilateral interactions,

focused in regions related to visuospatial and attentional processing (Gotts et al., 2013). All these asymmetries can be quantified by means of network science estimators. Gotts et al., 2013 and Liu et al., 2009 quantified functional lateralization in the brain by studying interactions across homotopic regions between hemispheres.

Here, we explore the dual contribution of brain network topology and space in modeling MI-related mental states through functional lateralization. We introduce network estimators to examine the degree to which brain networks preferentially interact with ipsilateral or contralateral nodes. Specifically, we formulate the notions of *laterality*, *segregation*, and *integration* to quantify patterns of biased interactions in brain asymmetry.

3.2 Brain lateralization

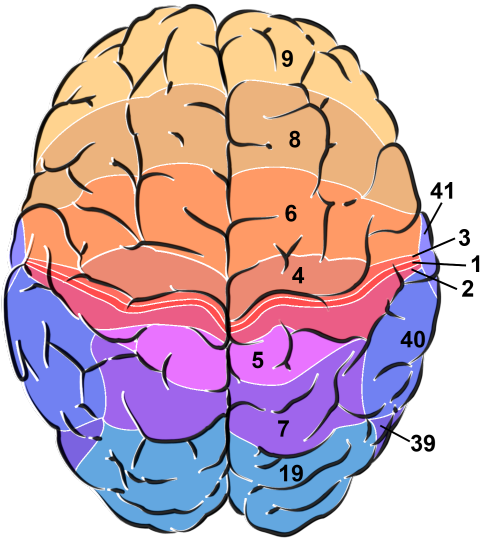
Functional lateralization of the brain refers to the tendency for some neural functions or cognitive processes to be specialized in one hemisphere. Although the macro-structure of the two hemispheres seems almost identical, intrinsic brain organization relies on specialized functions that may differ across hemispheres. One documented example is motor function. Each brain hemisphere controls the contralateral side of the body, and the natural asymmetry of the brain is the cause of left/right side dominance, i.e., better performance or preference for using one hand (Amunts et al., 1996).

3.2.1 Lateralization in motor imagery

MI activates many of the same cortical areas as those involved in the planning and execution of motor movements (supplemental motor area, premotor cortex, dorsolateral prefrontal cortex, and posterior parietal cortex) (**Fig 3.1**) (Miller et al., 2010; Porro et al., 1996). The premotor cortex (PMA) and supplementary motor areas (SMA) appear to be higher-level areas that encode complex patterns of motor output and that select appropriate motor plans to achieve desired results. The relevance of primary motor cortex (M1) may be considered as an unresolved issue, since it exists evidence both supporting and against a role for M1 during MI (Beisteiner et al., 1995; Porro et al., 1996; Roth et al., 1996; Roth et al., 1996; Pfurtscheller and Neuper, 1997; Lotze and Halsband, 2006).

Authors have found that increases in motor cortex activity primarily occur contralateral to the imagined hand movement (Beisteiner et al., 1995; Roth et al., 1996). Nonetheless, the programming of motor behavior of each body side cannot be completely segregated. It has also been shown that MI primarily engages the left (dominant) motor cortex for both right and left MI tasks (Beisteiner et al., 1995; Stinear, Fleming, and Byblow, 2006). This suggests that MI involves the dominant cortex in higher-order function, possibly related to movement planning and/or generation of the expected sensory components of task performance (Naito et al., 2002).

The asymmetry of MI may have important implications for the usefulness of rehabilitation. In the case of stroke, it may depend on the lateralization of the lesion. Sabaté, González, and Rodríguez, 2004 found that MI of the right hand following right (non-dominant) hemisphere stroke was largely unaffected. Conversely, if the dominant hemisphere is affected, MI may be of little benefit.



1-3 - Primary Somatosensory Cortex (S1)
4 - Primary Motor Cortex (M1)
5 - preparietal Somatosensory Association Cortex (S2)
6 - Premotor cortex (PMA)
and Supplementary Motor Cortex (SMA)
7 - sup. parietal Somatosensory Association Cortex (S2)
8 - Frontal eye fields
9 - Dorsolateral prefrontal cortex (DLFC)
19 - Tertiary or Associative visual cortex
39 - angular
40 - supramarginal
41 - ant. transverse temporal

FIGURE 3.1: **Brodmann areas top view.** The functional units of the cerebral hemispheres have been separated into what are called Brodmann areas, and include areas 1 through 47. This numbering is still used as a shorthand for describing the functional regions of the cortex, particularly those related to sensory functions. Some overlap exists among functional areas. For example, the motor cortex is area 4; the primary sensory cortex includes areas 3, 1, and 2.

3.3 Lateralization properties

Brain lateralization can be modeled by means of graph theory. As seen in **section 1.2**, interactions across different brain regions can be quantified with multiple network properties.

In weighted networks, edges can assume a range of different values, capturing variations in the strength of connectivity between pairs of nodes (Boccaletti et al., 2006). This spectrum can vary according to the connectivity estimator (**section 1.2**). If we look at a single node, its connectivity sum will also be influenced by the size of the graph. For example, in the case of a spectral coherence-based undirected EEG network, one link will vary between 0 and 1, but the sum of edges of a node will scale proportionally with the number of sensors.

These variations are perfectly capture by node *strength* property introduce in **section 1.3.1**. If we consider W as the weighted connectivity matrix of the non-filtered N nodes brain network, then we can compute the strength of node i as

$$s(i) = \sum_{j=1, j\neq i}^{N} W_{ij}, \tag{3.1}$$

where W_{ij} is the strength or weight of the edge linking nodes i and j (**Fig 3.2**).

3.3.1 Laterality index

It is possible to implement strength to study connection patterns related to brain lateralization. For this, we consider pairs of homotopic nodes, which means mirror channels across the hemispheres. E.g., if we consider 10-20 international system EEG configuration, then nodes C3 and C4 are defined as homotopic. Then we can estimate the laterality index (λij)

Box 3.3 - Within and inter-hemisphere connections

Lateralization properties are based on functional connectivity within or across hemispheres. For a node i in the left hemisphere, strength within-hemisphere (LL_i) is measured by summing the connectivity values between it and the rest of nodes located in the left hemisphere. For the particular case of EEG-based networks, the connections that node i established with the central line electrodes (LC_i) are also included in the within-hemisphere connections. On the other side, strength across-hemisphere (LR_i) is estimated by summing the connectivity between it and all the nodes located in the right hemisphere.

$$LL_i = \sum_{l \neq i}^{L} W_{il(LL)}, \qquad LC_i = \sum_{c}^{C} W_{ic(LC)}, \qquad LR_i = \sum_{r}^{R} W_{ir(LR)} \tag{3.2}$$

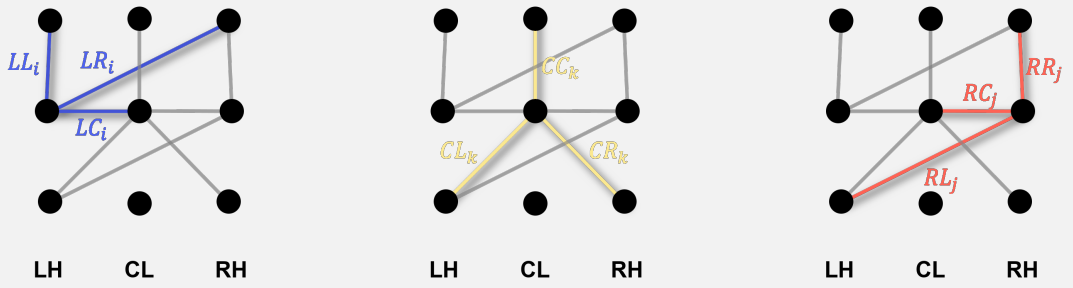
Subsequently, for a node j in the right hemisphere, RR_j, RC_j and RL_j are obtained using the same reasoning.

$$RR_j = \sum_{r \neq j}^{R} W_{jr(RR)}, \qquad RC_j = \sum_{c}^{C} W_{jc(RC)}, \qquad RL_j = \sum_{l}^{L} W_{jl(RL)} \tag{3.3}$$

The same is repeated for node k located in the EEG central line obtaining CC_k, CR_k and CL_k.

$$CC_k = \sum_{c \neq k}^{C} W_{kc(CC)}, \qquad CR_k = \sum_{r}^{R} W_{kr(CR)}, \qquad CL_k = \sum_{l}^{L} W_{kl(CL)} \tag{3.4}$$

For obvious reasons, the concepts of segregation and integration do not apply to nodes located in the central line.



LH: Left Hemisphere, **CL:** Center Line, **RH:** Right Hemisphere

To clarify the notation, notice that each capital letters term respectively denotes the locations of node i and the nodes it establishes connections with (e.g. LR_i means that node i belongs to the left hemisphere and we consider the connections that link it to the right hemisphere nodes).

by measuring the strength difference between the homotopic pair i and j, normalized by the strength of the closest middle line node k.

$$\lambda_{ij} = \frac{LL_i - RR_j}{CC_k}, \tag{3.5}$$

For a clarified notation, check **Box 3.3**. In **Fig 3.2** we illustrate how these interactions are distributed for a toy example network.

3.3.2 Segregation

The concept of functional lateralization can be further developed by analyzing the influence of interactions within and across hemispheres. Adapting the metrics proposed by Gotts et al., 2013 and Liu et al., 2009, we can define segregation as the tendency for greater within-hemisphere interactions compared to between-hemisphere interactions, calculated as the difference of intra- and inter-hemispheric strength (i.e., $LL_i+LC_i-LR_i$ or $RR_j+RC_j-RL_j$) (**Fig 3.2**). Particularly, the lateralization of segregation for the pair of homotopic nodes i and j is calculated as their segregation difference:

$$\sigma_{ij} = \frac{(LL_i + LC_i - LR_i) - (RR_j + RC_j - RL_j)}{(CL_k + CR_k + CC_k)} \tag{3.6}$$

The analysis of this metric sign could lead to some misinterpretation. Then, it is necessary to underline two main aspects. First, the strength values involved in the equation are strictly positive since we are working with undirected networks. Secondly, to guarantee a true sided σ_{ij} we empirically proved that $LL_i + LC_i > LR_i$ and $RR_j + RC_j > RL_j$ for every node (see **section 3.2.1**, **Fig A.5**). It means that a negative σ_{ij} value reflects higher lateralization of segregation in the right homotopic node of the pair. In other words, the within-hemisphere interactions are stronger in the right hemisphere. The opposite situation occurs for a positive value.

3.3.3 Integration

Integration seeks the contribution of contralateral connections, characterizing how the information flows across hemispheres. Then it is defined as the summed effect of intra- and inter-hemispheric interactions (e.g., $LL_i + LC_i + LR_i$ or $RR_j + RC_j + RL_j$) (**Fig 3.2**). Therefore, the lateralization of integration for a node i in the left hemisphere as compared to node j in the right is calculated as:

$$\omega_{ij} = \frac{(LL_i + LC_i + LR_i) - (RR_j + RC_j + RL_j)}{(CL_k + CR_k + CC_k)} \tag{3.7}$$

As a general remark, it is important to highlight that all these properties are local, ie they characterize each node. From a classification perspective, it means that the number of nodes is equivalent to the number of features. Yet, the lateralization metrics reduce the number of features to the half minus the central line, since each pair of homotopic nodes have the same feature value but with opposite signs (e.g. $\lambda ij = -\lambda ji$).

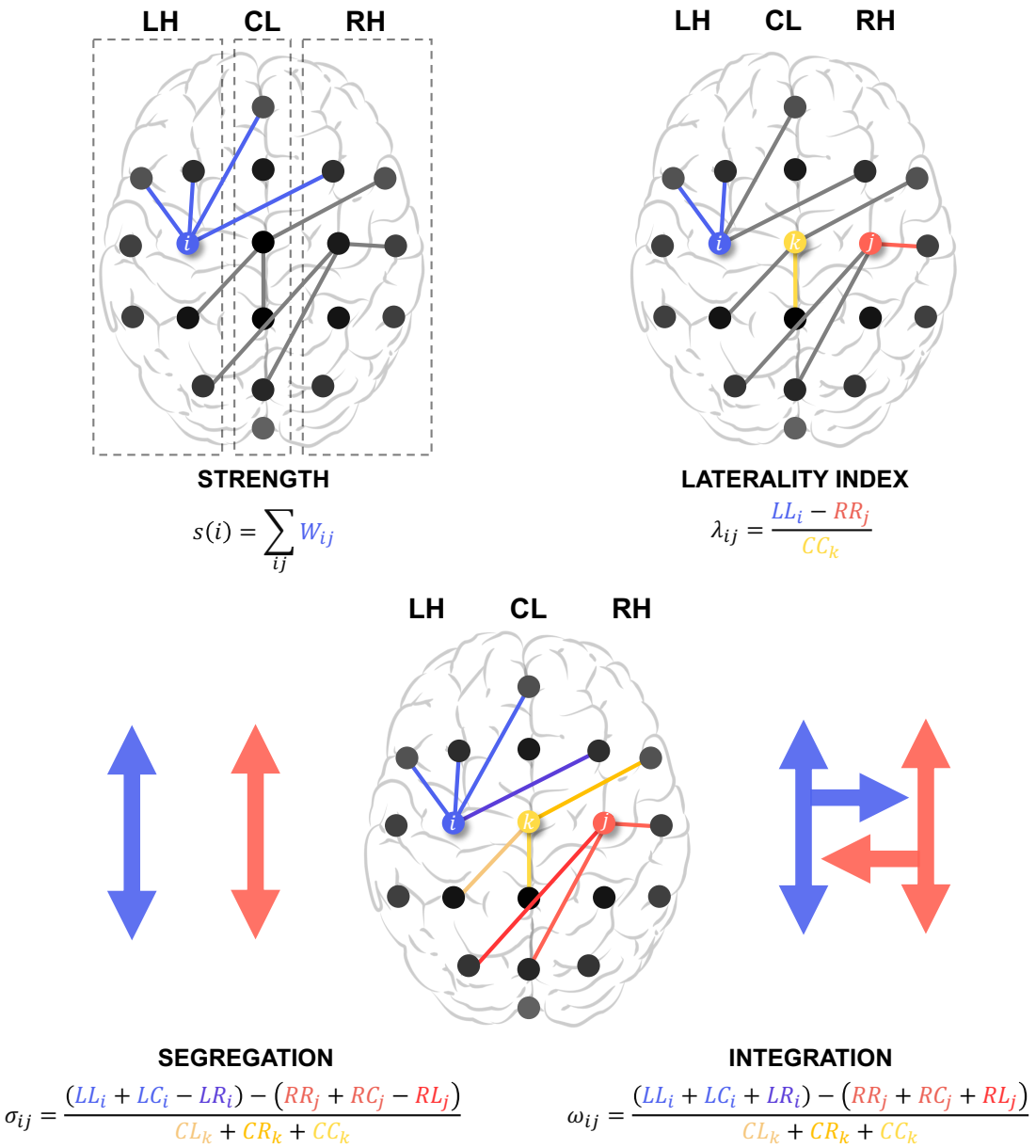


FIGURE 3.2: **Network properties**. It is possible to identify functional lateralized nodes by comparing the strength between homotopic pairs. Through these three pictures, we illustrate how links of the same network are considered for the computation of the studied properties. In the top left figure, we represent the computation of the *strength* of node i. On the top right, we introduce *lateralityindex* for the homotopic pair i-j (λij). The bottom figure represent the distinction between *segregation* (σ) and *integration* (ω) at the same pair ij. The key difference in the interpretation remains on the influence of inter-hemispheric edges (LR_i and RL_j). While ω adds the strength of bilateral interactions, σ challenges the strength of within interactions. That is to say, a large positive value for σ would suggest that the bias for stronger within-hemisphere is stronger for the left hemisphere. In contrast, a large negative value would indicate that the bias for within-hemisphere interactions is stronger for the right. Notations are the same as in **Box 3.3**

Chapter 4

Conventional methods and classification

4.1 Introduction

Identifying mental intentions from brain signals requires working in different domains, temporal, frequency, and space. Since EEG, as many other neuroimaging techniques, are often characterized by noisy measurements and low spatial resolution it is necessary to implement methods that enhance characteristics that define each mental task. During a MI task, brain activation is well localized in the sensorimotor cortex. Then one smart solution is to apply spatial filtering (Lotte, 2014). These methods try to collect this information while reducing the influence of the surrounding activity. In the last decades, the most reproduced in the BCIs fields is Common Spatial Patterns (CSP) (Ramoser, Muller-Gerking, and Pfurtscheller, 1999; Pfurtscheller and Neuper, 2001; Blankertz et al., 2005; Blankertz et al., 2007). This filter works as a data-driven dimension reduction method that aims to extract the signal sources by maximizing the variance ratio between two conditions. It is based on the simultaneous diagonalization of two covariance matrices of the band-pass filtered signal for the two classes. In this way, covariance matrices are just handled in the Euclidean space.

Another technique that has gained large space in the field is Riemannian geometry. Basically, it enables direct manipulation of the signal covariance matrices and subspaces (Yger, Berar, and Lotte, 2016; Congedo, Barachant, and Bhatia, 2017). The core idea behind these algorithms is to work with covariance matrices in the manifold of symmetric positive-definite (SPD) matrices and use them as features in a classifier that respects their intrinsic geometry. Matrices with such property form a manifold $\mathcal{M}$, in which the tangent space at each point is a finite-dimensional Euclidean space. In particular, the approaches that use tangent space projection have been shown to out-perform most other conventional methods (Barachant et al., 2013; Jayaram and Barachant, 2018). However, these methods undergo two major disadvantages, high computational complexity and lack of interpretation. Since they work in the space of sensor covariance matrices, their size scales quadratically with the number of sensors. Then when projecting to the tangent space, this easily becomes an overfitting problem when the vector dimension is higher than the available training trials (Rodrigues et al., 2017; Congedo et al., 2017). This translates into unfeasible application in high-density BCI systems. Further, the issue of interpretation is a significant commonly forgotten problem. Riemannian methods do not count with a direct way to determine what

parts of a signal are being used to build a tangent space classifier. Neither the classifiers in the manifold contemplate for this issue (Barachant et al., 2010b; Barachant et al., 2011).

To tackle both problems, we propose the feature selection in the manifold introduced by Barachant and Bonnet, 2011, where we keep the covariance components that best maximize the Riemannian distance between classes (**section 4.3**). In other words, the selected rows/columns of the covariance matrix are equivalent to the most discriminative sensors. Using this approach, we guarantee a proper dimensionality and features can be validated by visualizing the selected sources. Then the reduced SPD matrices can be projected onto the tangent space.

Finally, to homogeneously assess the classification performance of all the described methods in this thesis, we implement a typical classification algorithm used in BCI, a Support Vector Machines (SVM). More details are presented in **section 4.5**.

4.2 Common Spatial Pattern

CSP is a supervised spatial filter that allows addressing the two classes paradigm in multichannel data. This technique aims to maximize one condition's variance while minimizing the other (see **Fig 4.1** and **Fig 4.2**). It generates N spatial filters by generalized eigen value decomposition (GEVD) of the average covariance matrices, being N equivalent to the number of electrodes (Ramoser, Muller-Gerking, and Pfurtscheller, 1999; Blankertz et al., 2005; Blankertz et al., 2007). Mathematically, the matrix $W \in \mathbb{R}^{N\times N}$ containing the spatial filters, projects the EEG signal $x(t) \in \mathbb{R}^N$ from the original sensor space into the signal $z(t) \in \mathbb{R}^N$ which lives in the surrogate sensor space:

$$z(t) = [z_1(t)...z_N(t)]^{\mathrm{T}} = W^{\mathrm{T}}x(t) \tag{4.1}$$

Each column vector $w_j \in \mathbb{R}^N$ of W, constitutes a spatial filter and $z_j(t) = w_j^{\mathrm{T}}x(t)$ the corresponding spatial filtered signal at instant t ($j = 1, ..., N$). Yet, it is necessary to select an optimal subset of filters to best capture the difference between classes but simultaneously avoid overfitting. In this thesis, we work with eight components selected based on decreasing mutual information (Barachant et al., 2010a; Gramfort et al., 2013). Then we project the original signals by the selected filters and compute their logarithmic power. Finally, these eight-dimensional log-variances are linearly combined as features for a linear classifier (Lotte et al., 2007).

From the spatial filters decomposition, it is also possible to extract the corresponding patterns of brain activation by taking the inverse of the transposed full filters matrix W (Blankertz et al., 2005; Blankertz et al., 2007; Barachant et al., 2010a).

$$A = (W^{\mathrm{T}})^{-1} \tag{4.2}$$

Spatial patterns are the column vectors a_j of the matrix A. These patterns are used to represent the projected sources to the scalp, hence they can be used to validate neurophysiological plausibility.

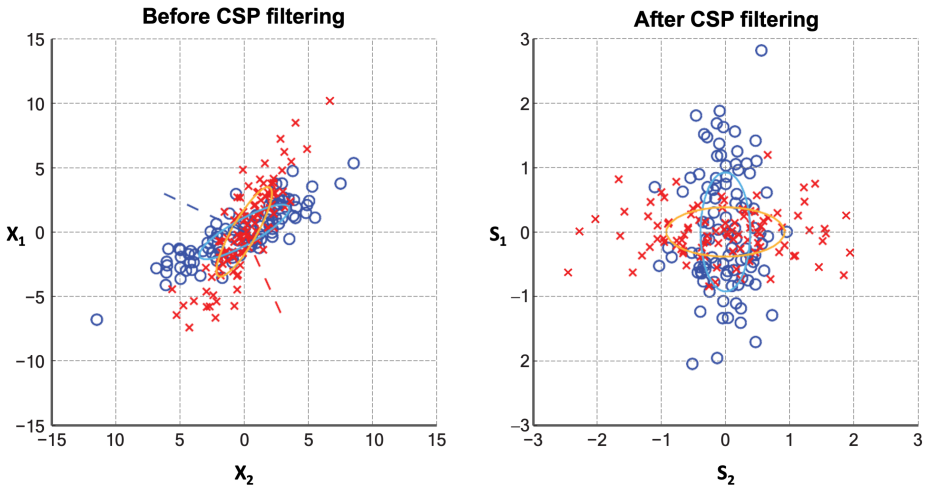


FIGURE 4.1: **CSP filtering**. On the left, the distribution of samples before filtering. Two ellipses show the estimated covariances and dashed lines show the direction of CSP projections. On the right, the distribution of samples after the filtering. The horizontal (vertical) axis gives the largest variance in the red (blue) class and the smallest in the blue (red) class, respectively. Extracted from Blankertz et al., 2007

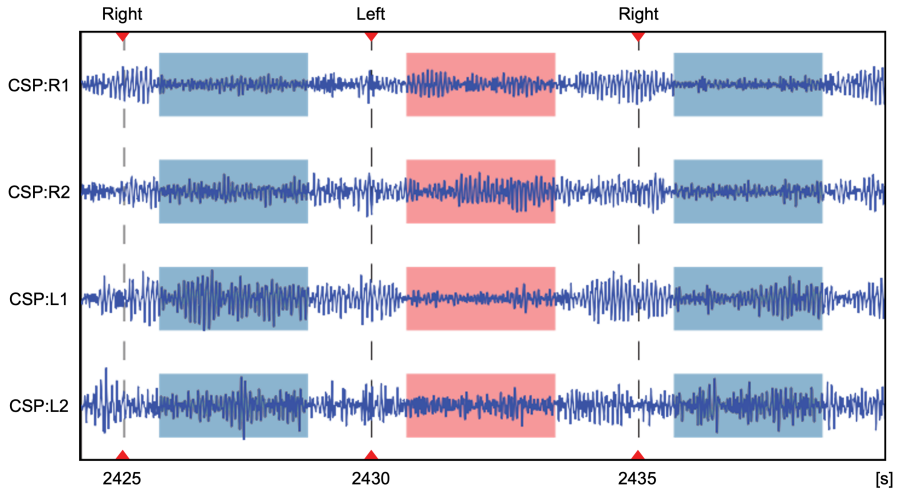


FIGURE 4.2: **Effect of spatial CSP filtering.** Example of applying four CSP filters to EEG MI data. Intervals of right hand MI are shaded green and show larger variance in the CSP:L1 and CSP:L2 filters, while during left MI (shaded red) variance is larger in the CSP:R1 and CSP:R2 filters. Extracted from Blankertz et al., 2007

Despite its popularity and efficiency, CSP is also known to be highly sensitive to noise (Blankertz et al., 2007; Grosse-Wentrup et al., 2009). Artifacts are more likely to be captured by a component with a high variance. Then if an artifact is more pronounced for one class, the class's variance is maximized by extracting the artifactual component that does not provide information on the subject's intention.

4.3 Riemannian method

Riemannian methods enable the direct manipulation of EEG signal covariance matrices. These matrices are SPD that live in a Riemannian manifold (Yger, Berar, and Lotte, 2016; Congedo, Barachant, and Bhatia, 2017). The classification accuracy of these approaches has shown to outperform other reported methods in BCI (Jayaram and Barachant, 2018).

However, their results lack neurophysiological interpretability. To address this issue and get comparable results with our proposed method (**section 3.3**), we decided to implement the Riemannian-based feature selection introduced by Barachant and Bonnet, 2011. This algorithm takes as a selection criterion the Riemannian distance δ_R between the class-conditional mean covariance matrices.

$$Crit = \delta_R(\bar{\mathrm{C}}^{(1)}, \bar{\mathrm{C}}^{(2)}), \tag{4.3}$$

where $\bar{\mathrm{C}}^{(1)}$ denotes the mean covariance matrix of class 1 and $\bar{\mathrm{C}}^{(2)}$ for class 2 (**Fig 4.3**). This criterion is implemented in a backward feature selection where, in an iterative loop, we keep the $N*$ electrodes that best maximize δ_R. Note that $N*$ is a predefined value lower than N (set to 10 for our analysis) and that a particular electrode i represents the i-th row and column in the covariance matrix. This enables to list the anatomical location of the retained channels.

With the objective to explore multi-type feature fusion, we transformed the reduced covariance matrices into suitable inputs to a vector-based classifier. We mapped the them onto the tangent space of the Riemannian manifold Barachant et al., 2011. There they can be vectorized and used as input to an SVM (**Fig 4.3**). This vector has dimension $N(N+1)/2$,

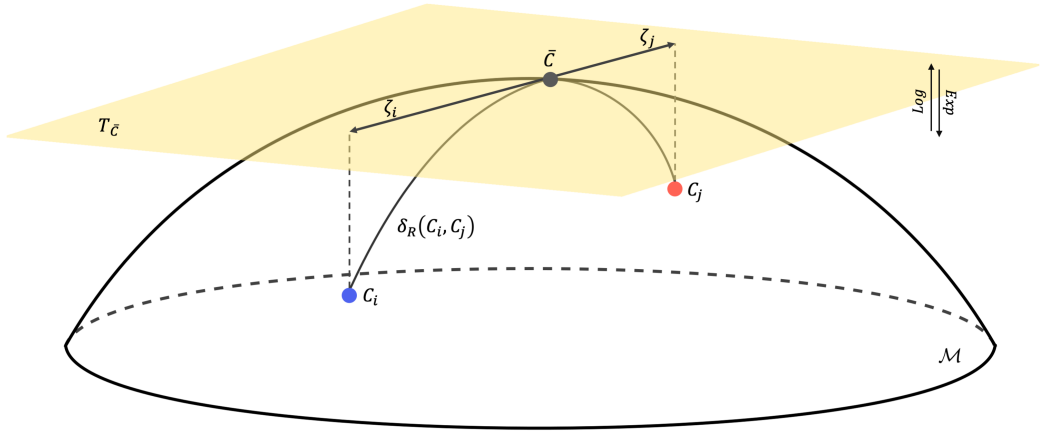


FIGURE 4.3: **Manifold $\mathcal{M}$.** Schematic representation of manifold $\mathcal{M}$, the geometric mean $\bar{C}$ of two points and the tangent space at C. Consider two points (e.g., two covariance matrices) C_i and C_j on $\mathcal{M}$. The geometric mean of these points is the midpoint on the geodesic connecting C_i and C_j that minimizes the sum of the squared distances ($\bar{C} = \mathrm{argmin}_C \sum_{i=1}^{N} \delta_R^2(C, C_i)$). If we construct the tangent space $T_{\bar{C}}$ at $\bar{C}$, there exists only one tangent vector ζ_i (respectively ζ_j) corresponding to the geodesic departing from $\bar{C}$ and arriving at C_i (respectively C_j) on the $\mathcal{M}$. The map from the $\mathcal{M}$ to the $T_{\bar{C}}$ is a logarithmic map. This type of mapping is used to vectorized the selected features for the SVM classifier.

which might risk overfitting by exceeding the number of training trials. Nonetheless, this issue is tackled by dimensionality reduction done with feature selection.

We evaluate these methods performance through a 5-fold cross-validation procedure. Since both methods use class label information, the calculation of the CSP filters and the Riemannian channels selection are performed within the cross-validation on samples of the respective training set. Then the spatial filters and the selected electrodes are applied to the samples of the test set (see **section 4.5**).

4.4 Statistical analysis

It has been explained how lateralization properties quantify asymmetries in the brain. Since we already know that hand-MI is reflected as a lateralized task in the motor cortex (Beisteiner et al., 1995; Pfurtscheller and Da Silva, 1999; Xu et al., 2014; Cattai et al., 2021a), we hypothesize that our proposed metrics have great potential in differentiating between left and right hand-MI.

To statistically evaluate this ability, we performed a 5000 permutation t-test for each of the metrics. Comparing at the subject level, we assumed a null hypothesis that the metric means for the two conditions were equal. Analysis was carried out by using SciPy[1] python package. A level of $P<0.05$ was considered critical for assigning statistical significance. Since we carried out this analysis for each node, it enabled us to detect the most discerning electrodes. We repeated this test for the entire population of subjects (**Table 5.1**).

4.5 Feature selection and classification

A common problem in BCI systems is the small sample size datasets (Lu et al., 2010; Lotte et al., 2018). EEG sensors need gel or saline liquid to improve contact and reduce impedance with the scalp. Consequently, their drying limits long time recording. In addition to this, not all available datasets count with a testing set. To circumvent this issue, we decide to assess the model with a k-fold cross-validation (CV) procedure (Buitinck et al., 2013). Data is divided into 5 stratified folds, and classification is performed with a support vector machine (SVM) classifier (the splits are kept identical for all pipelines in a given subject). Finally, classification performance is measured in terms of ROC-AUC (receiver operating characteristic - area under the curve). In order to return a single score per subject, the scores from each session are averaged when multiple sessions were present. All our analyses are inspired and based on MOABB- *WithinSessionEvaluation*[2] and scikit-learn[3] python packages.

4.5.1 Feature selection for network properties

The properties proposed in this thesis characterize the network at a local scale (see **section 1.3.1**). In the classification scenario, this is translated into a number of features equal

[1] https://scipy.org
[2] http://moabb.neurotechx.com/docs/generated/moabb.evaluations.WithinSessionEvaluation
[3] https://scikit-learn.org/

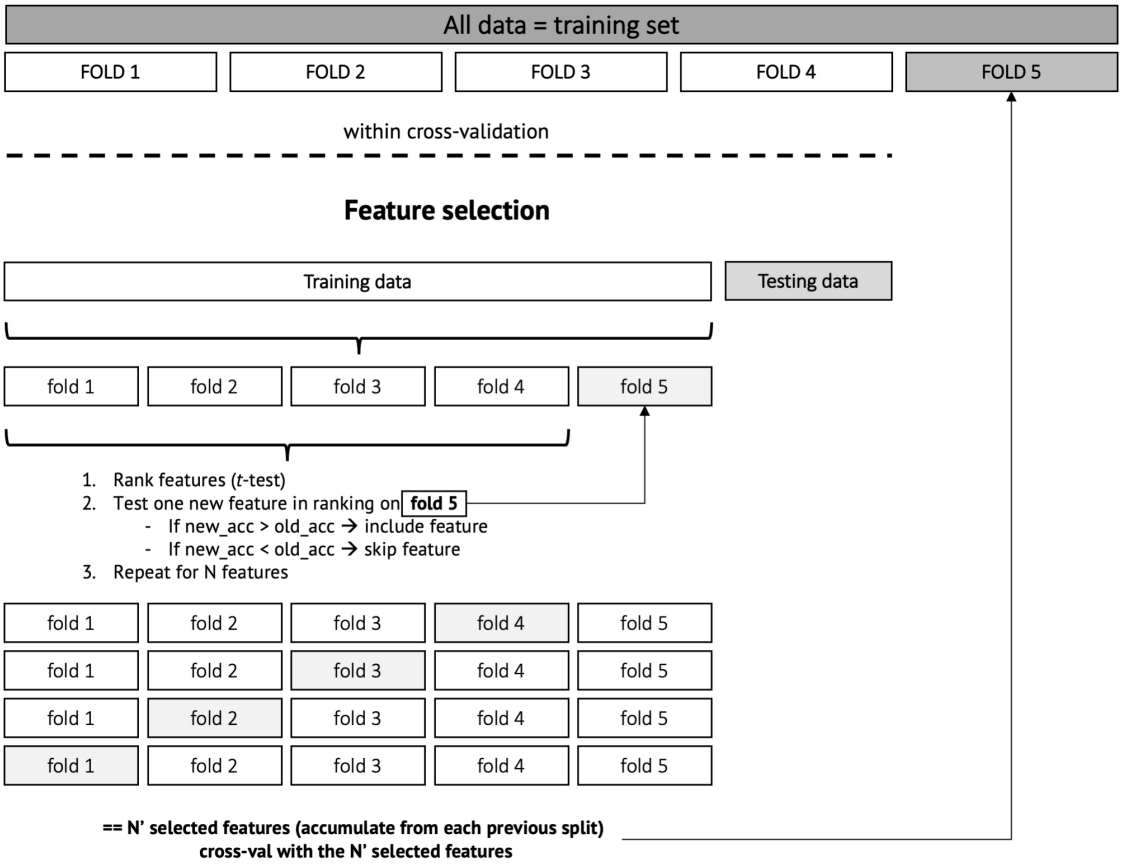


FIGURE 4.4: **Feature selection algorithm.** Within a cross-validation framework (CV_{orig}), this approach uses a nested 5-fold cross-validated SVM (CV_{nested}) on a sub-training set (80% of the CV_{orig} training) (Dominguez, 2009), to obtain a subset of selected features, $N' = 10$. For eachCV_{nested} iteration, features are ranked according to their discriminant power between classes (t-test). In a forward sequential order, a feature is going to be retained and accumulated in the selected set, if its accuracy is higher than the previous set. The output of this CV_{nested} is a group of 10 selected features on which the CV_{orig} validation set (FOLD 5) is going to be tested. This is repeated for each iteration in CV_{orig}.

to the number of nodes. Then, if the data set size is small and the number of channels is large, we risk an overfitting problem. To prevent this issue, we add a feature selection step, that limits the number of selected features to 10. This procedure benefited from fewer parameters to be optimized by the classifier and the possibility to neurophysiologically interpret the selected features.

It is important to clarify that feature selection does not mean node removal. In other words, the interactions of a non-selected node with the selected ones, still account for the latter.

We implemented an embedded approach to select the best discriminant features. We use a sequential forward feature selection. Within a nested cross-validation framework, this algorithm adds features to form a feature subset. To limit the research complexity, at each stage we rank our features from the training folds according to their discrimination power (t-test).Then we perform a bottom up search procedure which conditionally includes new features to the selected set based on the cross-validation score (**Fig 4.4**).

Chapter 5

Application to real data

Part of this chapter will be published in (in prep.):

- **Title:** Network lateralization features for motor imagery-based brain-computer interfaces
- **Authors:** Juliana Gonzalez-Astudillo and Fabrizio De Vico Fallani

---

5.1 Introduction

Most BCI systems rely on electroencephalography (EEG) as the acquisition system (Lotte, Bougrain, and Clerc, 2015). EEG measures small electrical currents that reflect brain activity. Like other recording techniques used in this field, such as MEG or fMRI, EEG has a high risk of being contaminated by undesirable non-neural sources (Jackson and Bolger, 2014; Niedermeyer and Silva, 2005; Nunez, Srinivasan, et al., 2006). Dealing with this problem, BCI must provide predictive classifiers, based purely on brain-derived features. These divergent requirements prompted the field to develop in two different directions: spatial filtering and Riemannian manifold techniques (Ramoser, Muller-Gerking, and Pfurtscheller, 1999; Yger, Berar, and Lotte, 2016). These methods have proven outstanding accuracies but rarely report the corresponding feature interpretation. Then their neurophysiological validation is still an open question.

On the contrary, the network-based method has as starting point the direct modeling of the complexity of neurophysiological processes (Bullmore and Sporns, 2009; De Vico Fallani et al., 2014). They synthesize interactions between different brain areas using graph properties. Then by simply looking at these properties we can identify the most prominent interactions and how each node or sensor contributes to the process. Still their validation as discriminant classification features in a BCI pipeline is an emerging field (Gonzalez-Astudillo et al., 2021).

BCI essentially requires real-time brain recording and classification. However, this is not the best scenario to first validate a new pipeline, since it adds variability that cannot be controlled by the system like the influence of the subject and the experimenter (Attina et al., 2008; Saha and Baumert, 2020). As a result, validating approaches has always been

a difficult task. Because of this, we decide to validate the methods proposed in this project in an offline scenario. This allows us to test in many publicly available EEG datasets, that have shared characteristics but defer in number of electrodes (Jayaram and Barachant, 2018). We present results in a classification scenario of two-class imagined movement, as that is the most widely used MI paradigm.

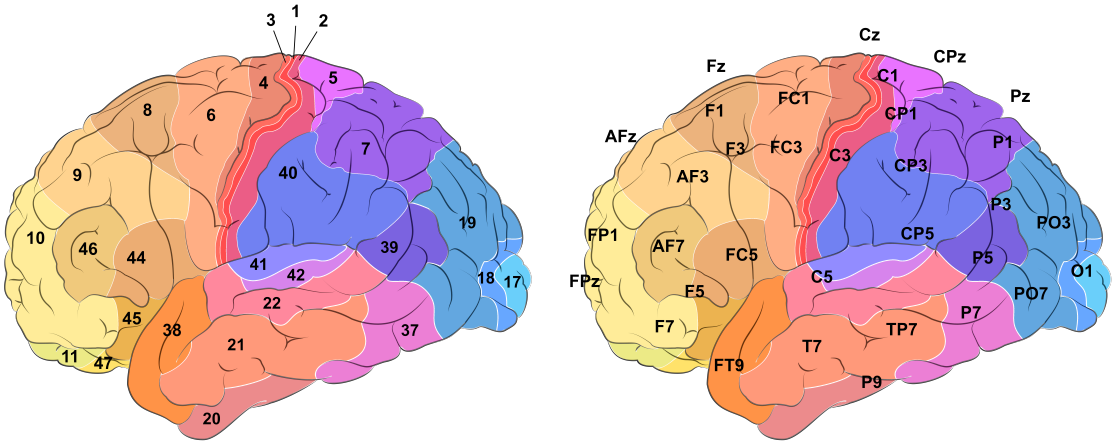
Through this section we explore feature interpretation as well as feature performance in terms of classification score. Our goal is not to achieve the best performance, but to validate the neurophysiological plausibility of lateralization properties in hand MI, comparing it with the most distinguished benchmark methods in the field.

5.2 EEG

Electroencephalography (EEG) measures the electrical activity of the brain (Jackson and Bolger, 2014; Niedermeyer and Silva, 2005; Nunez, Srinivasan, et al., 2006). Signals are captured through sensors distributed over the scalp (non-invasive) or by directly placing the electrodes on the brain surface (invasive), a procedure that requires a surgical intervention (Engel et al., 2005). This technique is characterised by its high temporal resolution, enabling the study of dynamic processes such as cognition and motor task. Yet, EEG signals are non-stationary and have a non-linear nature, which makes it difficult to get useful information directly in the time domain. Nonetheless, specific patterns can be extracted using advanced signal processing techniques.

During signal recording, undesirable potential coming from sources other than the brain may alter the quality of the signals. These artifacts should be detected and removed to improve pattern recognition. Multiple methods could be applied depending on the artifact to be eliminated (Bashashati et al., 2007; Lotte, 2014). Here we use pass-band filtering to keep the physiological rhythms of interest (Cheveigné and Nelken, 2019; Michel and Brunet, 2019). MI task triggers different frequency oscillations as a consequence of changes in the signal's amplitude, known as event-related desynchronization (ERD) (Pfurtscheller and Aranibar, 1977; Neuper, Wörtz, and Pfurtscheller, 2006). Hence, this phenomenon represents frequency specific changes in ongoing EEG (Pfurtscheller and Da Silva, 1999). Amplitude increase and decrease in the temporal domain translates into a power increase and decrease of a particular frequency band in the spectral domain. The most important bands in the context of MI-based BCI are α (8–13 Hz) and β (13–30 Hz) (Pfurtscheller and Neuper, 2001).

This frequency responses are not necessarily located at the same electrodes. EEG oscillations can also be used to identify the brain areas activated during a task, which could offer greater insight into cortical dynamics (Başar and Düzgün, 2016). Electrodes can spatially approximate the activity localization in terms of Brodmann areas, enabling the functional mapping of the brain (**Fig 5.1**) (Brodman, 1909). Many Brodmann areas are closely associated with diverse cortical functions. For example, Brodmann areas control several sensory and motor functions (Başar, 2012; Johnson et al., 2002; Faymonville et al., 2000). In this thesis, it should be contemplated that we presume a good manipulation and placement of the electrodes since we do not apply source reconstruction techniques. We assume that



1-3 - Primary Somatosensory Cortex (S1)	20 - inferior temporal
4 - Primary Motor Cortex (M1)	21 - middle temporal
5 - preparietal Somatosensory Association Cortex (S2)	22 - superior temporal
6 - Premotor cortex (PMA) and Supplementary Motor Cortex (SMA)	37 - occipitotemporal
7 - sup. parietal Somatosensory Association Cortex (S2)	38 - temporopolar
8 - Frontal eye fields	39 – angular
9 - Dorsolateral prefrontal cortex (DLFC)	40 - supramarginal
10 - frontopolar (DLFC)	41-42 – ant.&post. transverse temporal
11 - prefrontal	44 - opercular
17 - Primary visual cortex	45 - triangular
18 - Secondary visual cortex	46 - middle frontal
19 - Tertiary or Associative visual cortex	47 - orbital

FIGURE 5.1: **EEG electrodes and its nearest Brodmann areas.** On the left side, we represent the brain's left hemisphere profile of Brodmann area division. This segmentation is based on their neuronal organization, correlated to diverse cortical functions. On the right side, we represent the electrodes distribution of a 10-20 EEG system in these areas. These locations interpolates to the right hemisphere for the homotopic electrodes. This figure seeks for illustrative purpose. For a more exhaustive association between channels and Brodmann areas, refer to **Table 5.2**

the EEG cap is always correctly placed according to the anatomical landmarks Nasion and Inion.

5.2.1 Dataset cohorts

EEG is the principal acquisition tool in BCI due to its time resolution, cost and portability. This type of data can be found on open-access warehouses, as well as via collaborative projects such as the BNCI Horizon 2020[1]. These repositories are valuable since they contribute to establishing harmonisation procedures in processing and creating benchmarks. Typically, all datasets guarantee informed consent and anonymization to protect the participants privacy. Data come in different formats according to the acquisition system or the pre-processing software. The most common formats for EEG are .edf, .gdf, .eeg, .csv or .mat files. These different formats can create challenges when working with multiple datasets.

Luckily, some tools have been developed to handle this problem. In this thesis we rely all our analysis on data available from the python package MOABB[2] (*Mother of all BCI*

[1] http://bnci-horizon-2020.eu
[2] http://moabb.neurotechx.com/docs/

Benchmark) (Jayaram and Barachant, 2018). This open source project helps to overcome the problem of limited number of subjects, converting all the data to an MNE-Python[3] exploitable format. MOABB also counts with a set of algorithms and utility functions for analysis and visualisation.

We have selected six open-access datasets of healthy participants. This data contains non-invasive EEG signals measured during MI experiments focusing on left and right hand grasping motions. **Table 5.1** provides a description of the selected datasets and **Fig 5.2** shows the spatial layout of EEG montages. Each trial is band-passed filtered in a broad α-β band (8-35Hz), where we typically observed characteristic signal changes while subjects are performing MI (Pfurtscheller and Aranibar, 1977; Pfurtscheller and Da Silva, 1999; Neuper and Pfurtscheller, 2001).

TABLE 5.1: **Dataset attributes.** Overview of all included datasets with EEG recordings in a left versus right hand MI paradigm. #: number, sub: subjects, ch: channels.

Dataset	#sub	#ch	#trial/class	epoch[s]	#sessions	ref.
001-2014	9	22	144	4	2	Tangermann et al., 2012
Cho2017	49	64	100	3	1	Cho et al., 2017
Lee2019MI	54	62	100	4	2	Lee et al., 2019
Schirrmeister2017	14	128	120	4	1	Schirrmeister et al., 2017
Weibo2014	10	60	80	4	1	Yi et al., 2014
Zhou2016	4	14	160	5	3	Zhou et al., 2016

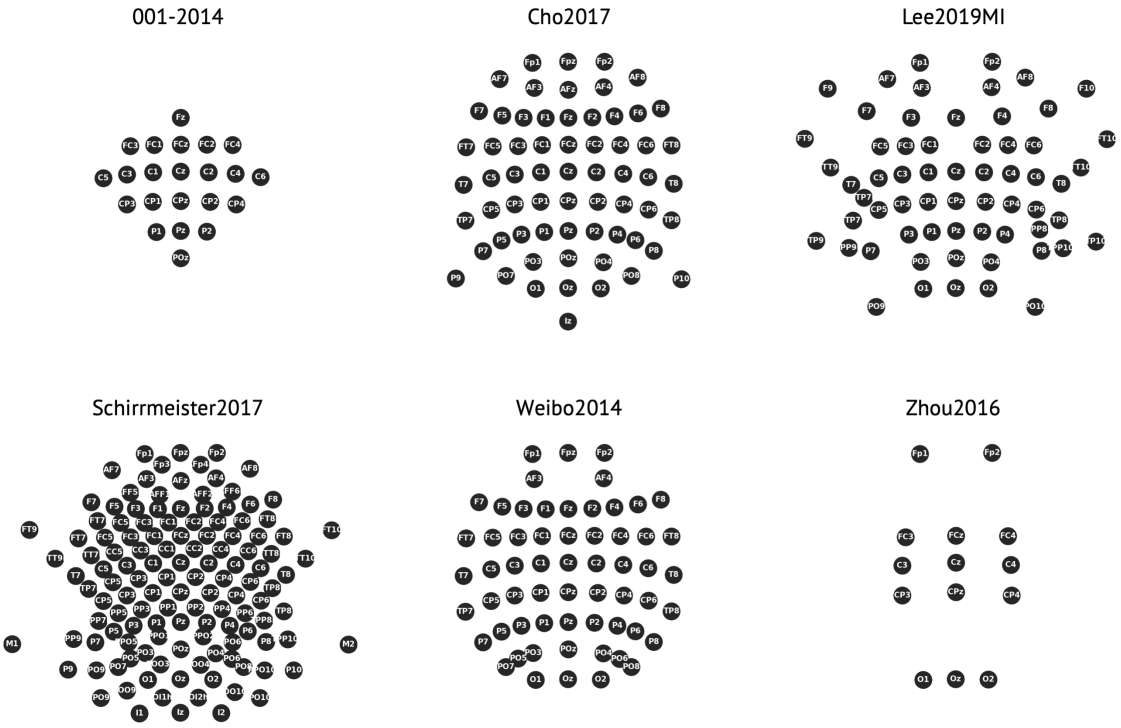


FIGURE 5.2: **Datasets EEG montages.**

[3] https://mne.tools/

5.3 Building functional brain networks

Functional connectivity (FC) derives the existence of interaction between distant brain areas, if there is a statistical synchronization between them (Bastos and Schoffelen, 2016; De Vico Fallani et al., 2014). There are several FC estimators (see **section 1.2**), here we use undirected spectral coherence (W) (Carter, 1987) that has already been well documented in the MI-BCI domain (Hamedi, Salleh, and Noor, 2016; Cattai et al., 2021b; Corsi et al., 2022). Given two signals i and j from two electrodes, this estimator computes their normalized cross-spectral density for a particular frequency f:

$$W_{ij}[f] = \frac{|P_{ij}[f]|}{\sqrt{P_i[f]P_j[f]}}, \tag{5.1}$$

where $P_{ij}[f]$ express the cross-spectrum and $P_i[f]$ the auto-spectrum at frequency f.

We estimate the cross-spectral density of each pair of EEG signals at the trial level, using multitapers (Slepian, 1978) with time windows of 1 second length with an overlap of 0.5 seconds and 1Hz frequency resolution. We average the resulting FC matrices over α-β bands (8-35Hz). Summing up, for each trial, we obtain a W symmetric adjacency matrices of shape $N \times N$, where N is equivalent to the number of EEG channels. These matrices correspond to fully connected and weighted networks. Now the complexity of their interactions can be quantitatively analyzed using network science tools(De Vico Fallani et al., 2014; Gonzalez-Astudillo et al., 2021).

5.4 Network lateralization patterns during motor imagery

In this section we investigate the lateralization properties introduced in **section 3.3**, with the objective of identifying their discrimination power between the two mental tasks under study, left and right hand-MI. We apply the statistical analysis described in **section 4.4** to each of the metrics. For the sake of simplicity and to avoid any mixture of MI classes with hemisphere sides, we refer to left MI as LMI and right MI as RMI.

In **Fig 5.3** we show the node *strength* t-values obtained across trials and averaged across subjects. These results reveal interesting patterns for a subset of nodes. Notably, the largest changes tend to concentrate on motor-related areas. But what is more remarkable is the predominance of positive t-values in the right hemisphere. This trend confirms that LMI evokes higher *strength* in the contralateral motor cortex. The inverse situation occurs for RMI but with lower t-values, suggesting that this task requires fewer connectivity resources. This might be explained by the predominance of right-handed subjects.

If we look at each dataset separately (**Fig A.1**, **Table A.1**), we observe that those that show significant t-values (*001-2014*, *Cho2017*, *Schirrmeister2017*, *Zhou2016*) concentrate the most discriminant *strength* at nodes related to the motor cortex, particularly on the right hemisphere.

This evidence of sided-contrast connections across tasks encourages us to perform a lateralization analysis. To do so, we consider the spatial locations of the electrodes, differentiating between intra- and inter-hemispheric interactions. We implement laterality (λ) to quantify the strength difference across homotopic pairs of nodes (see **Eq. 3.5**). Then, we

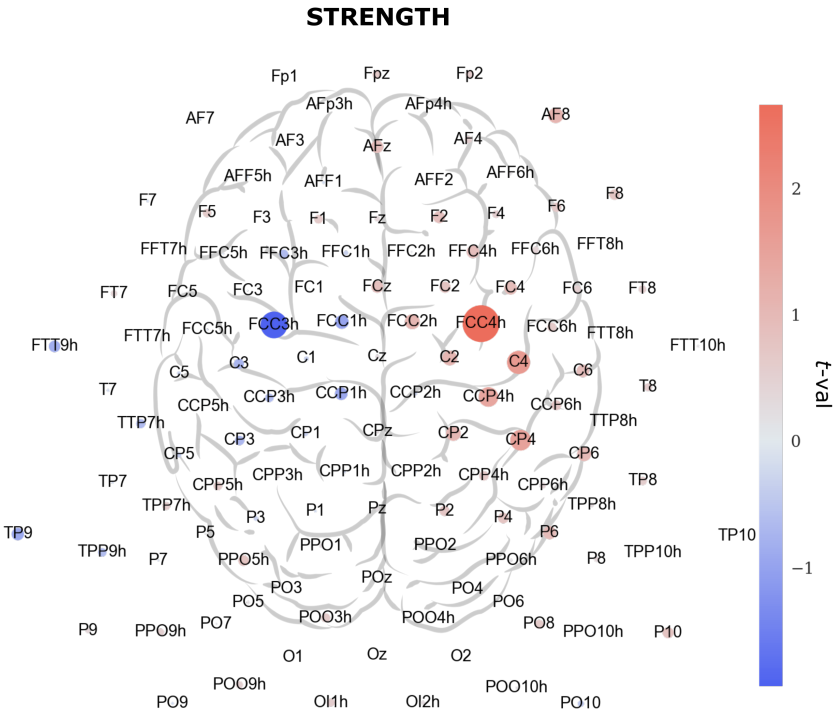


FIGURE 5.3: **Strength lateralization in MI tasks.** Group-averaged t-values, contrasting s_{LMI} versus s_{RMI} in the α-β band. Evidence of lateralization is observed in channels located in motor-related areas.

also study integration (ω) and segregation (σ), that respectively account for the contribution of across- and within-hemispheric connections (see **Eq. 3.7** and **3.6**).

Lateralization metrics give symmetric inverse values for each pair of homotopic nodes. Then in the following results, we represent just one hemisphere. Also notice that the t-values sign is strictly related to the task, i.e. a positive value means stronger lateralization for LMI and a negative for RMI.

Laterality index. When repeating the same statistical analysis done for *strength* on lateralization metrics, we obtain a comparable behaviour between *strength* and λ (**Fig 5.4**). Most of the highest t-values are located in MI related areas (M1, PMA, SMA and S1) (Jeannerod and Frak, 1999; Grezes and Decety, 2001; Guillot et al., 2009; Hétu et al., 2013), but with accentuated t-values particularly in M1 and S1 (see **Fig 5.1** and **Table 5.2**). This suggests that just combining homotopic information helps to increase the difference among MI tasks.

Integration. In the same line, by considering the contribution of inter-hemispheric interactions, ω increases this difference over nodes related to the parietal S2 cortex and supramarginal area (see **Fig 5.5**). Distinctively, this metric maintains significant t-values for channels in the M1 and S1 areas, while reducing the rest (see **Fig 5.1** and **Table 5.2**). The same behavior is observed for each individual dataset (except *Schirrmeister2017*) (**Fig A.4**, **Table A.1**).

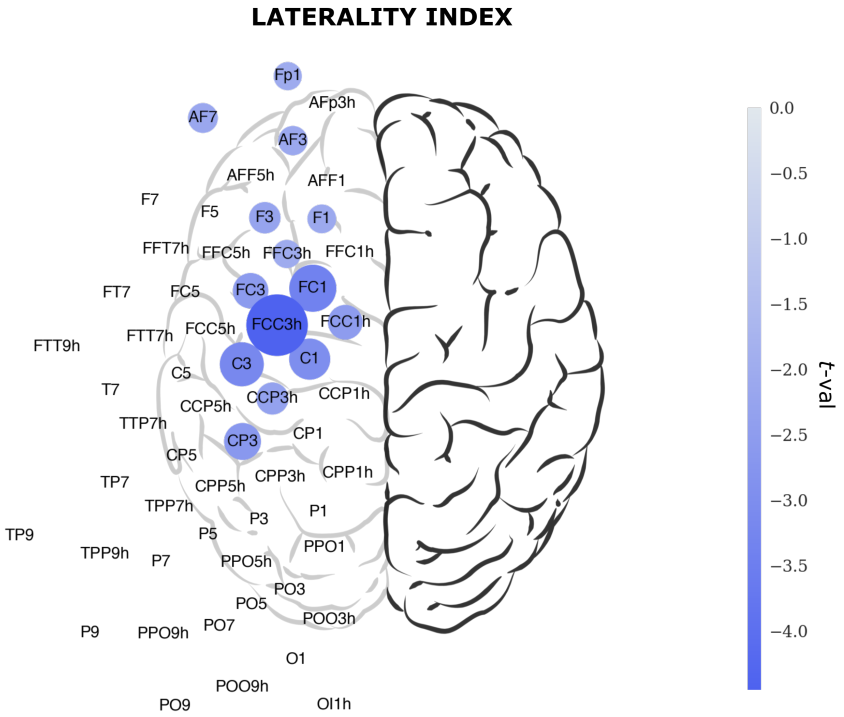


FIGURE 5.4: **Laterality index in MI.** Group-averaged t-values in differentiating between LMI and RMI states in the α-β band. For illustrative purposes we combine results across datasets for each channel. Results at the dataset-level are presented in Appendix A. For a simpler visualization, just significant t-values are shown ($p < 0.05$).

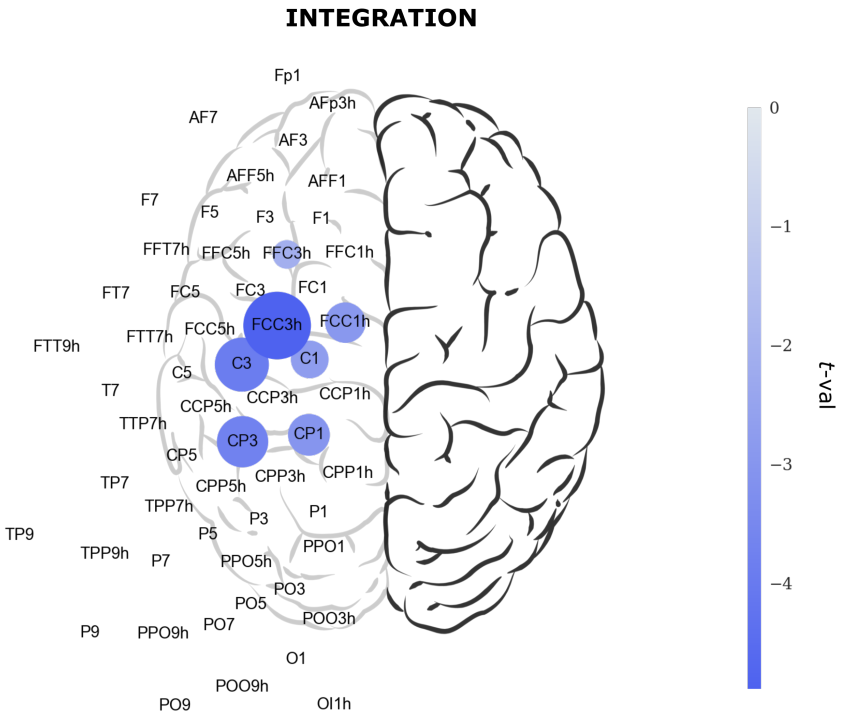


FIGURE 5.5: **Integration in MI.** Group-averaged significant t-values in differentiating between LMI and RMI states ($p < 0.05$).

Segregation. For the particular case of σ we first query if the connections with the middle line nodes (LC_i and RC_j) were deemed as within-hemisphere or not. Thus, we analyze their role in the two possible scenarios by statistically comparing the differences. By looking at the behaviors shown in **Fig A.5**, we conclude that middle line connections play a more reasonable role in within-hemisphere. We base this judgment on the fact that the nodes with predominant connections of this type are the ones closer to the middle line, and some of them are strategic for the MI tasks under study. Then if we reduce their influence by subtracting LC_i and RC_j links, this may alter the neurophysiological nature of the results. Besides, we also avoid any misinterpretation of the sign of σ.

When analyzing the impact of subtracting the now well-defined inter-hemispheric connections (LR_i, RL_j), σ shows the highest impact in the frontal-central electrodes. These nodes are primarily linked with SMA and PMA cortex, along with the dorsolateral prefrontal cortex (DLFC) associated with action planning (Jeannerod and Frak, 1999; Curtis and D'Esposito, 2003; Gao, Duan, and Chen, 2011; Mokienko et al., 2013; Hétu et al., 2013).

Most of the significant nodes give negative t-values. This is due to higher σ values for RMI. Since these nodes exhibit negative values for this property for both classes (see **Fig A.6**), it indicates that the bias for within-hemisphere connections is stronger in the right hemisphere when performing LMI.

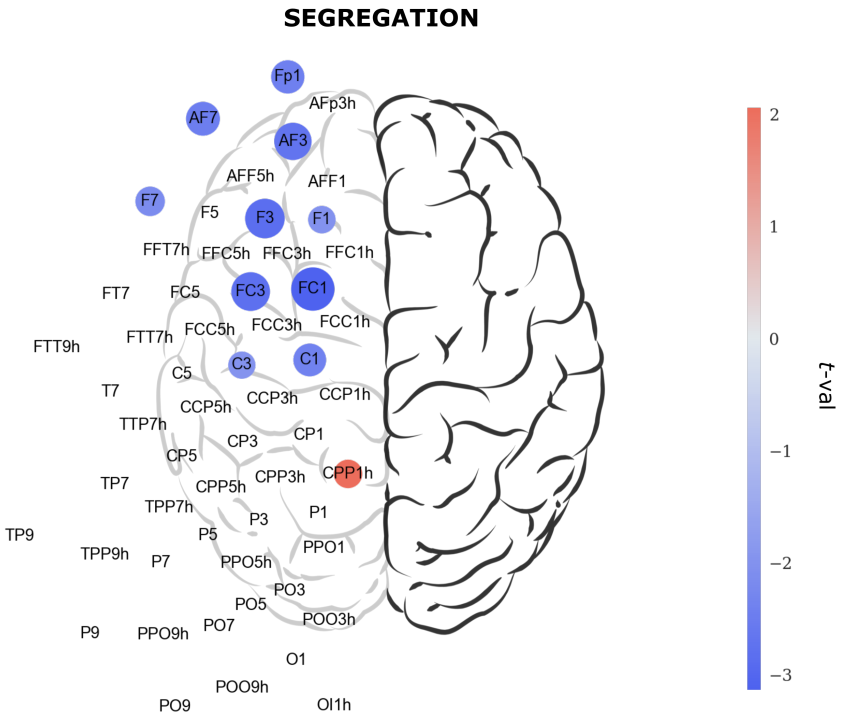


FIGURE 5.6: **Segregation in MI.** Group-averaged significant t-values in differentiating between LMI and RMI states ($p < 0.05$).

These findings indicate that MI of the hand grasping elicits detectable brain network changes that might be useful to characterize and discriminate MI-based BCI tasks. These changes revealed the existence of two parallel lateralization behaviors (i.e. fronto-central σ,

while central-parietal ω) that primarily involved sensorimotor areas. Moreover, the location of the selected channels within the classification procedure (see **section 4.5.1**) follow the same distribution as the previous t-values (see **Fig A.7**). This reinforces the idea that our approach enables to capture relevant processes in the MI task.

TABLE 5.2: **EEG channels and corresponding Brodmann areas**, for a 10-10 montage. L: left hemisphere, R: right hemisphere (Başar, 2012; Johnson et al., 2002; Faymonville et al., 2000).

EEG channel	Brodmann areas	EEG channel	Brodmann areas
FP1	10L,9L,46L,11L,32L	C2	5R,1R,2R,4R,6R
FPz	10L,10R,9L,9R,11R	C4	4R,2R,1R,3R,40R
FP2	10R,9R,46R,11R,45R	C6	42R,41R,22R,3R,40R
AF7	46L,10L,45L,9L,11L	T8	21R,22R,41R,42R,20R
AF3	9L,46L,8L,10L,45L	TP7	21L,42L,22L,37L,41L
AFz	9L,9R,32R,32L,8L	CP5	40L,39L,41L,42L,22L
AF4	9R,46R,8R,10R,45R	CP3	40L,2L,1L,7L,3L
AF8	46R,10R,45R,9R,47R	CP1	5L,7L,2L,1L,31L
F7	45L,47L,46L,44L,38L	CPz	5L,5R,7R,7L,31L
F5	45L,46L,44L,47L,9L	CP2	5R,7R,1R,2R,40R
F3	8L,6L,44L,45L,46L	CP4	40R,1R,2R,5R,3R
F1	8L,6L,9L,32L,24L	CP6	40R,42R,39R,22R,41R
Fz	8L,8R,32L,24L,6R	TP8	37R,21R,22R,42R,20R
F2	8R,6R,9R,32R,24R	P9	20L,37L,21L,22L,42L
F4	8R,9R,6R,44R,46R	P7	37L,39L,19L,21L,41L
F6	46R,44R,45R,9R,8R	P5	39L,19L,37L,40L,41L
F8	45R,47R,46R,44R,38R	P3	39L,7L,19L,40L,2L
FT9	38L,20L,21L,47L,22L	P1	7L,31L,5L,7R,39L
FT7	44L,47L,22L,38L,21L	Pz	7R,7L,31R,31L,5R
FC5	44L,45L,22L,42L,41L	P2	7R,31R,5R,19R,39R
FC3	6L,4L,3L,44L,2L	P4	39R,40R,7R,19R,1R
FC1	6L,4L,5L,1L,2L	P6	39R,19R,37R,40R,42R
FCz	6R,5L,6L,4L,5R	P8	37R,39R,19R,42R,22R
FC2	6R,4R,8R,2R,1R	P10	20R,37R,21R,22R,41R
FC4	6R,4R,44R,3R,2R	PO7	19L,18L,37L,39L,17L
FC6	44R,45R,41R,22R,42R	PO3	19L,39L,18L,17L,7L
FT8	47R,38R,44R,21R,22R	POz	17R,17L,18R,18L,7R
FT10	38R,20R,21R,47R,22R	PO4	19R,39R,18R,17R,7R
T7	21L,42L,22L,41L,20L	PO8	19R,18R,39R,37R,17R
C5	42L,41L,22L,40L,3L	O1	18L,17L,19L,17R,39L
C3	2L,4L,1L,3L,40L	Oz	17R,17L,18L,18R,19R
C1	5L,2L,4L,1L,6L	O2	18R,17R,19R,17L,39R
Cz	5L,5R,1R,1L,4L		

5.5 Common spatial pattern

We compare the previous results with one of the state-of-art-methods based on spatial filtering, common spatial pattern (CSP). As explained in **section 4.2**, this approach allows feature interpretation by looking at the resulting filters (w_j) and patterns (a_j). Spatial patterns represent a correlation map between the original electrode signal and the spatially filtered signal.

To compare this transformation with our network approach, we analyze the pair of vectors (w_j, a_j) that correspond to the largest and the smallest eigenvalues for each subject. Even though we work with eight filters, for the neurophysiological interpretation we just show these pairs because they correspond to the most discriminant components for each MI condition. For illustrative purpose, **Fig 5.7** and **Fig 5.8** display the interpolation to sensor space of the group-averaged filter and pattern that best minimize each class's variance. For each subject, we include the absolute normalized topographic maps. These values are normalized to compensate for the difference between datasets. Signs are not considered, since they are irrelevant in our analysis.

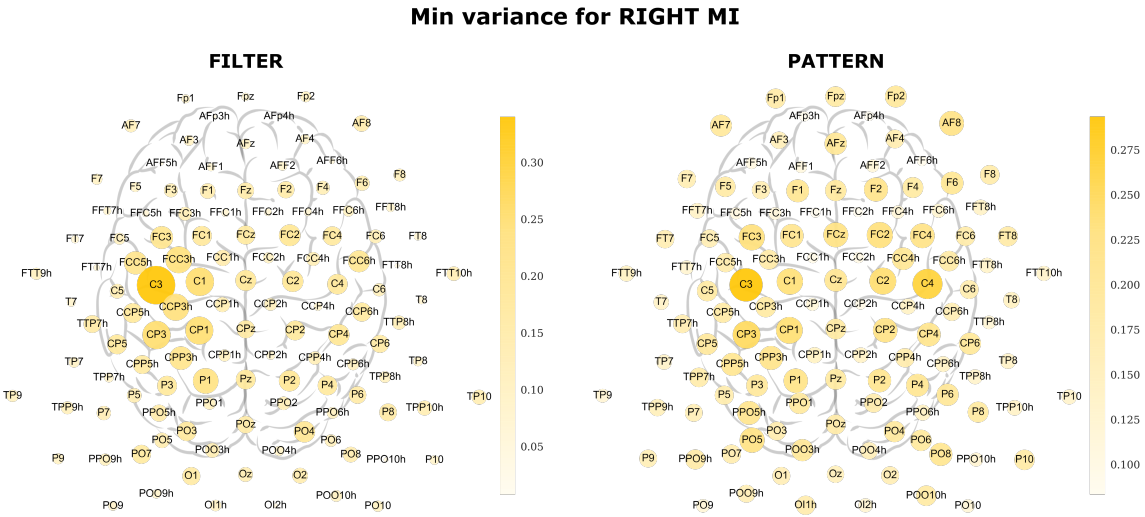


FIGURE 5.7: **CSP filter and pattern in *RMI*.** Filters apply the highest weights to electrodes related to motor task on the corresponding contralateral side (C3, C1, CP3, CP1, CCP3h), while the resulting patterns show a bilateral behaviour (C3, C4, CP3, C2, CP1).

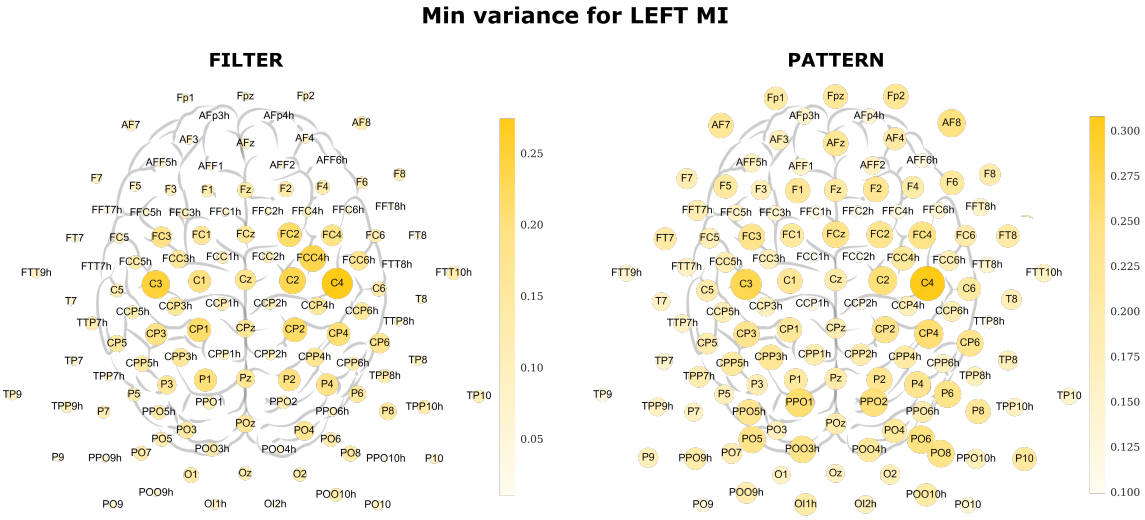


FIGURE 5.8: **CSP filter and pattern in *LMI*.** Here the bilateral involvement is present in both, filetrs (C4, C3, C2, FCC4h, CP2) and patterns (C4, C3, PPO1, CP4, C2).

In RMI (**Fig 5.7**), the maximum filter weights apply to electrodes that belong to areas related to hand movement. More precisely, the contralateral M1, S1, and S2 cortex (see Fig 5.1 and **Table 5.2**). However, the patterns exhibit a bilateral activation. On the contrary, filters in LMI (**Fig 5.8**) reveal bilateral recruitment of motor-related areas, with predominance on the contralateral side. This is also consistent with the patterns. This dual activation of contralateral and ipsilateral areas is usually observed for non-dominant hand task (Kim et al., 1993). Nonetheless, the patterns also show involvement of parieto-occipital areas (channels in PO line), that are principally related to the associative visual cortex.

It is important to point out that CSP is not a source separation or localization method (Blankertz et al., 2007). Contrarily, each filter is optimized for the maximization of one class's variance while minimizing the other's. In the case of RMI versus LMI paradigm, if we consider a filter that maximizes variance for class LMI and minimizes it for RMI, then an expected high weight on the left hemispherical motor area can have two plausible causes. It can either originate from an ERD during RMI, or from an ERS during LMI (RMI areas are more relaxed if concentration focuses on LMI, therefore the idle rhythm may increase). Or it can be a mixture of both effects, since LMI elicits both hemispheres. Regardless this possible misinterpretation, for the discrimination task this mixing effect is irrelevant. This becomes a significant limitation for neurophysiological interpretation.

5.6 Riemannian geometry

We performed a channel selection in the manifold to corroborate a valid interpretation of Riemannian features. In the 5-fold CV framework, the backward selection procedure selected 10 sensors that best maximized the Riemannian distance between classes for each subject (see **section 4.3**). **Fig 5.9** resumes the group-cumulative occurrences in a sensor plot. For each electrode, the number of selection times is normalized by the maximum possible occurrences. For example, C4 is the most selected electrode, with an occurrence of 58% over the total times it could have been selected.

In general, we observed a concentration of features in the M1, S1, and S2 on the right hemisphere and punctually on the left M1. These channels prove that Riemannian features are directly associated with the sensorimotor cortex. Nonetheless, we observe a subset of channels located in the parieto-occipital area that are not strictly related to MI, being P08 the 5^{th} most selected channel. If we look more in detail, we notice that datasets *Schirrmeister2017* and *Weibo2014* have a relevant influence of this area (see **Fig A.12**).

Even though this manipulation approaches the Riemannian method to feature interpretation, it still lacks a clear understanding of which features are associated with each class. For example, it is not possible to figure out if occurrences located in the motor cortex are related to the contralateral hand MI or the consequence of bilateral recruitment of these areas. We also observe that there is a bias for selections on the right hemisphere. We might think that it is due to higher resource consuming on the non-dominant hemisphere by establishing a correlation with results obtained with previous methods. But this is still a speculation that could not be assessed by only examining Riemannian selection.

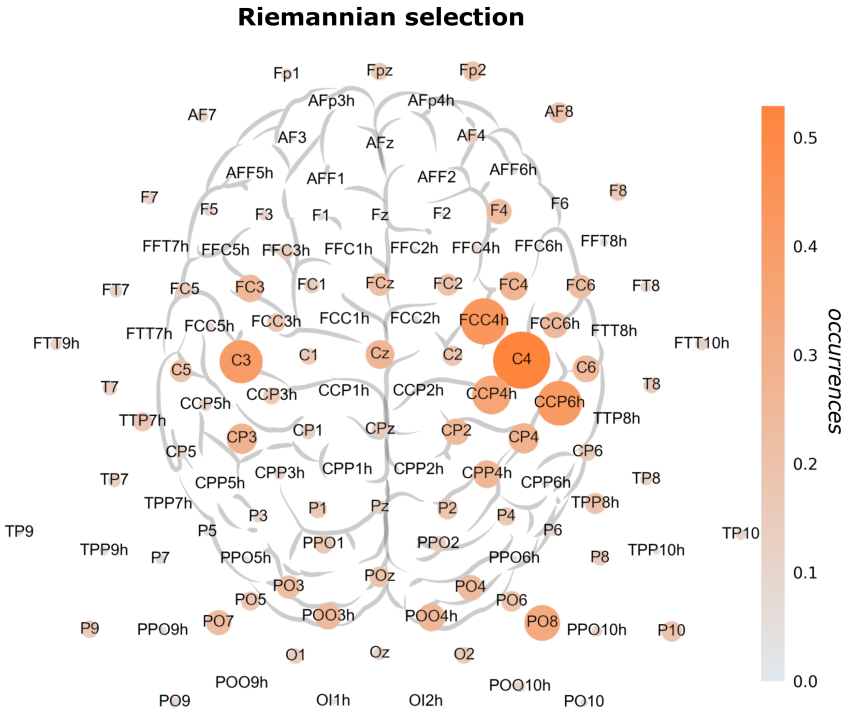


FIGURE 5.9: **Riemannian-based feature selection.** Group-averaged normalized occurrences. It illustrates the number of times that a specific feature in the manifold has been chosen. Most selected electrodes are C4, FCC4h, CCP6h, C3, CCP4h, PO8 (*occurrence* $> 30\%$).

5.7 Classification performance

We finally evaluate the three methods' performance through a 5-fold cross-validation (CV) procedure, using an SVM classifier (see **section 4.5**). Each feature extraction method follows a particular arrangement to end up in a proper input for the classifier. Network features follow a within-CV selection to reduce their dimensionality and guarantee the most discriminant nodes. CSP method first projects the signal by the selected spatial filters, and then it takes the logarithm of the power of the projected signal. Lastly, reduced-Riemannian SPD matrices are projected onto the tangent space (TS) of the manifold, where they can be vectorized. All types of features converge in separate SVM classifiers.

Table 5.3 and **Fig 5.10** show all the results generated by this entire processing chain. Decoding accuracies for multiple sessions datasets are calculated independently, and averaged for this representation. Besides the gaps in neurophysiological interpretation, we observe that for almost all the datasets, state-of-the-art methods (CSP and Riemannian) still outperform network-based features.

Average scores across datasets are quite heterogeneous independent of pipeline. The differences in hardware, strategy paradigm, and subject can generate large variation in the outcome of a BCI task, making it very difficult to generalize findings. When we look at the case of *Zhou2016*, we observe outstanding accuracies for all the methods. Nonetheless, this dataset has pre-trained subjects (Zhou et al., 2016), which might be the reason for such remarkable difference.

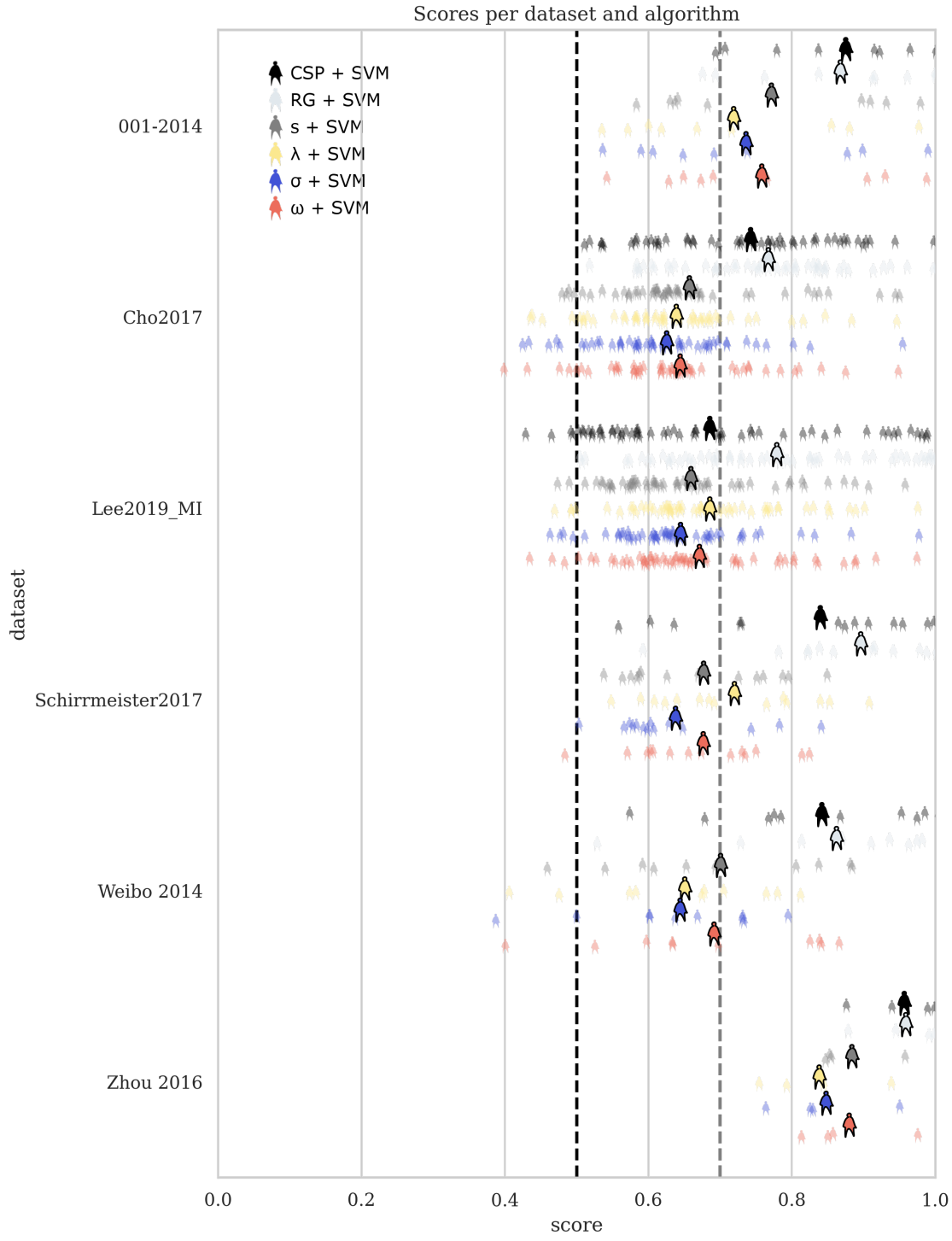


FIGURE 5.10: **Classification performances.** Visualization of all classification scores for each method across datasets applying a 5-fold cross-validated SVM (**section 4.5**). Each transparent silhouette represents a single subject and the bigger contoured silhouette presents the mean across subjects. Note that there is only one score per subject, representing the mean between sessions (if it applies, see **Table 5.1**). The black dotted line corresponds to a chance level performance (0.5), and the grey one to the begging of efficient performances (0.7) (Thompson, 2019). Network-based features applied the feature selection algorithm described in **section 4.5.1**. RG: Riemannian geometry method, s: *strength*, λ: *laterality index*, σ: *segregation*, ω: *integration*.

TABLE 5.3: **Classification performances:** Average accuracies across methods for each dataset. RG: Riemannian geometry method, s: *strength*, λ: *laterality index*, σ: *segregation*, ω: *integration*.

Dataset	s+SVM	λ+SVM	σ+SVM	ω+SVM	CSP+SVM	RG+SVM
001-2014	75.68 ± 15.8	70.45 ± 15.56	72.16 ± 15.75	74.33 ± 15.46	86.04 ± 12.05	85.31 ± 12.62
Cho2017	64.24 ± 11.54	62.45 ± 10.78	61.09 ± 10.24	62.95 ± 11.03	72.80 ± 13.21	75.24 ± 11.93
Lee2019MI	64.46 ± 11.27	67.11 ± 11.11	63.05 ± 9.63	65.63 ± 11.96	67.10 ± 16.60	76.46 ± 15.08
Schirrmeister2017	66.26 ± 10.37	70.54 ± 11.03	62.31 ± 9.31	66.17 ± 9.65	82.53 ± 15.43	88.16 ± 11.58
Weibo2014	68.63 ± 15.02	63.62 ± 13.37	62.96 ± 12.32	67.69 ± 15.64	82.75 ± 14.49	84.78 ± 13.99
Zhou2016	86.91 ± 5.33	82.36 ± 7.98	83.34 ± 7.81	86.52 ± 7.04	94.20 ± 5.63	94.44 ± 5.51

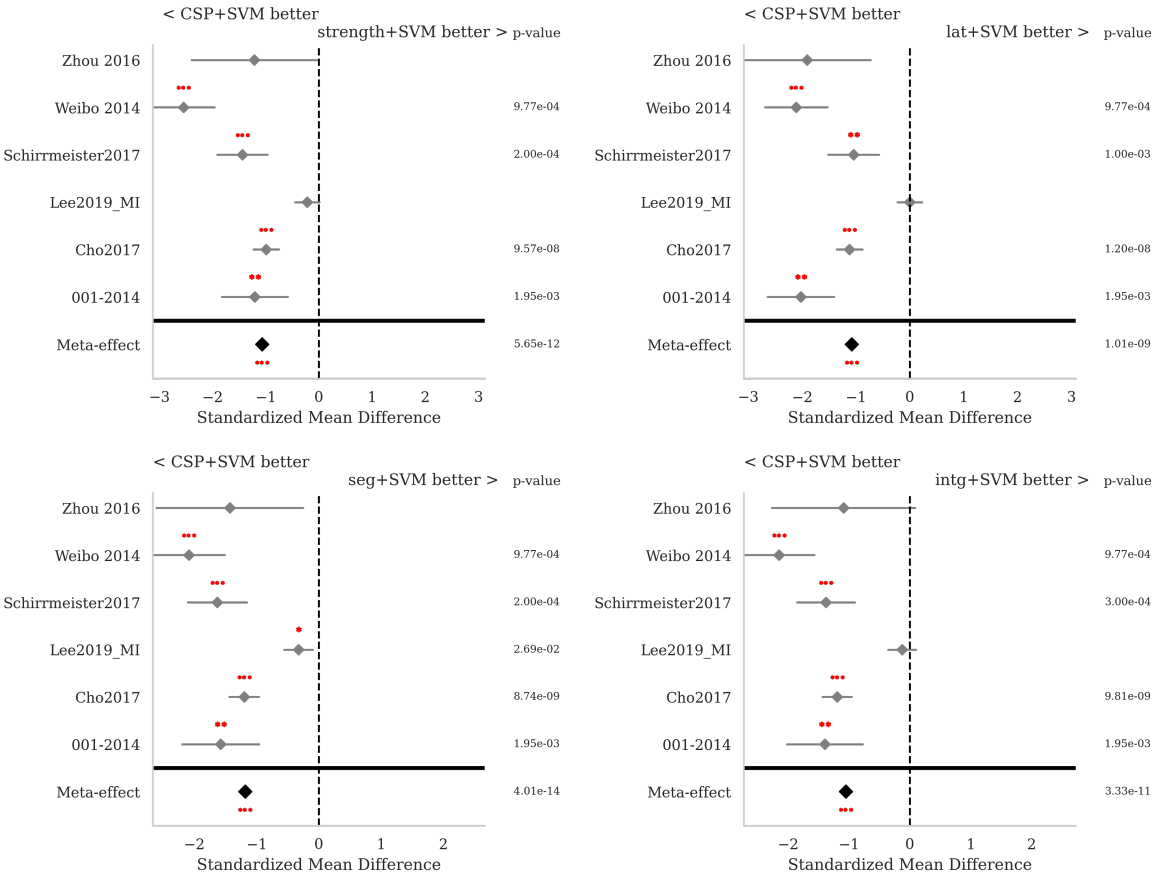


FIGURE 5.11: **Classification statistical analysis: lateralization versus CSP.** Meta-analysis style plots showing the performance of lateralization network metrics versus CSP. The effect sizes shown are standardized mean differences, with p-values corresponding to the one-tailed Wilcoxon signed-rank test for the hypothesis given at the top of the plot and 95% interval denoted by the grey bar. Stars correspond to *** = p $<$ 0.001, ** = p $<$ 0.01, * = p $<$ 0.05. The meta-effect is shown at the bottom of the plot. even though there is a significant amount of variance between datasets, the overall trend shows that CSP outperforms the other algorithms in this setting.

On the other side, CSP and TS methods are all well-known approaches and have been compared against each other often in the past, then a comparison between them would be of great interest. Repeating the same statistical analysis, we find that TS outperforms CSP across datasets. However, the score difference is not reliably significant for two of the datsets, *Zhou2016* and *Weibo2014*, and there is also one dataset for which the opposite trend is shown, *001-2014* (see **Fig 5.12**). The confidence intervals also show that this is

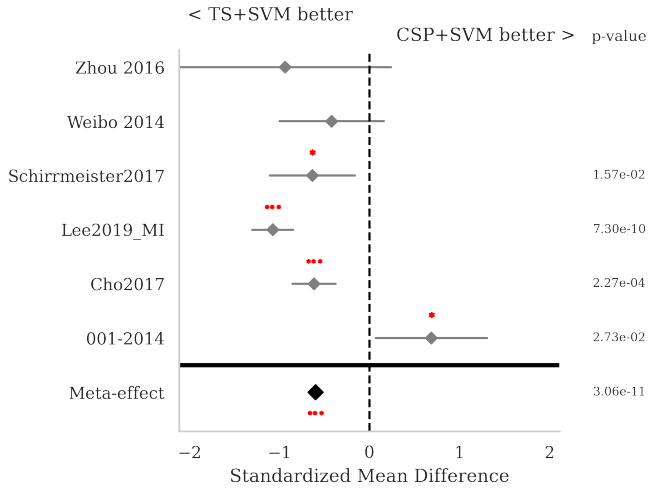


FIGURE 5.12: **Classification statistical analysis: Riemannian method versus CSP.** The same interpretation as in **Fig 5.11** is applied. The overall performance shows that TS works better than CSP.

likely the case for datasets with few subjects (see **Table 5.1**).

In order to improve classification scores for network metrics, we have tried other techniques like feature fusion and ensemble classifiers. For the first, we hypnotized that the complementarity of lateralization properties could also be reflected in an accuracy improvement. Unfortunately, this is not the case, showing scores closed to the already obtained. Another strategy that we have tested is ensemble classifiers. We have worked with soft voting, a technique that combines multiple classifiers and uses the average predicted probabilities to predict the class labels. Such a classifier is worthwhile for a set of equally well-performing models to balance out their individual weaknesses. Since CSP and Riemannian methods outperform network metrics by a significant difference, the resulting scores are ruled by the state-of-the-art methods.

5.7.1 Classification in *Lee2019MI* dataset: a promising case

It is interesting to note that in the case of *Lee2019MI* dataset, λ shows almost equal mean score than CSP (**Fig 5.10**). Even though this difference is not significant (see **Fig 5.11**), for the interest of this project, it incites a deeper analysis. Across single subjects scores, we observe that for CSP there is a subset that accumulates between 0.5 and 0.6 scores, and another smaller subset between 0.9 and 1. On the other side, network properties accuracies are more homogeneous with a normal-like distribution around the means on each metric (between 0.6 and 0.7). Thus, it means that there must be a group of subjects with low CSP scores, that get better performance with network features.

To confirm this tendency, we implement a paired plot that compares scores across methods for each individual (**Fig 5.13**). Note that there is only one score per subject, regardless of the number of sessions. This plot confirms the trend for better results with network features for the subset of subjects that get low scores with CSP. More precisely, if we consider the group with accuracy below 0.7 (33 subjects), that is to say, CSP inefficient subjects

(Thompson, 2019), we notice that their global accuracy with network features is significantly higher (**Fig 5.14**).

Looking at the feature interpretation associated with this dataset (see **section A.1** and see **section A.4**), we observe that there are some discrepancies across types of features. Even though *RMI* and *LMI* CSP patterns exhibit emphasized contralateral motor areas, patterns for *RMI* seems to be governed by possible artifacts in the frontal area (**Fig A.9** and **Fig A.11**). A high influence of frontal nodes is also observed for the case of *strength* (**Fig A.1**), but it is important to point out that they are located on the contralateral side and not on the ipsilateral as it is for *RMI* CSP. On the contrary, λ reduces this influence on frontal sensors and presents a higher distinction between classes in MI-associated areas (PMA, SMA, and S1) (**Fig A.2**). Yet, the Riemannian method shows cleaner features, with almost zero frontal selections (**Fig A.12**), giving also the best scores.

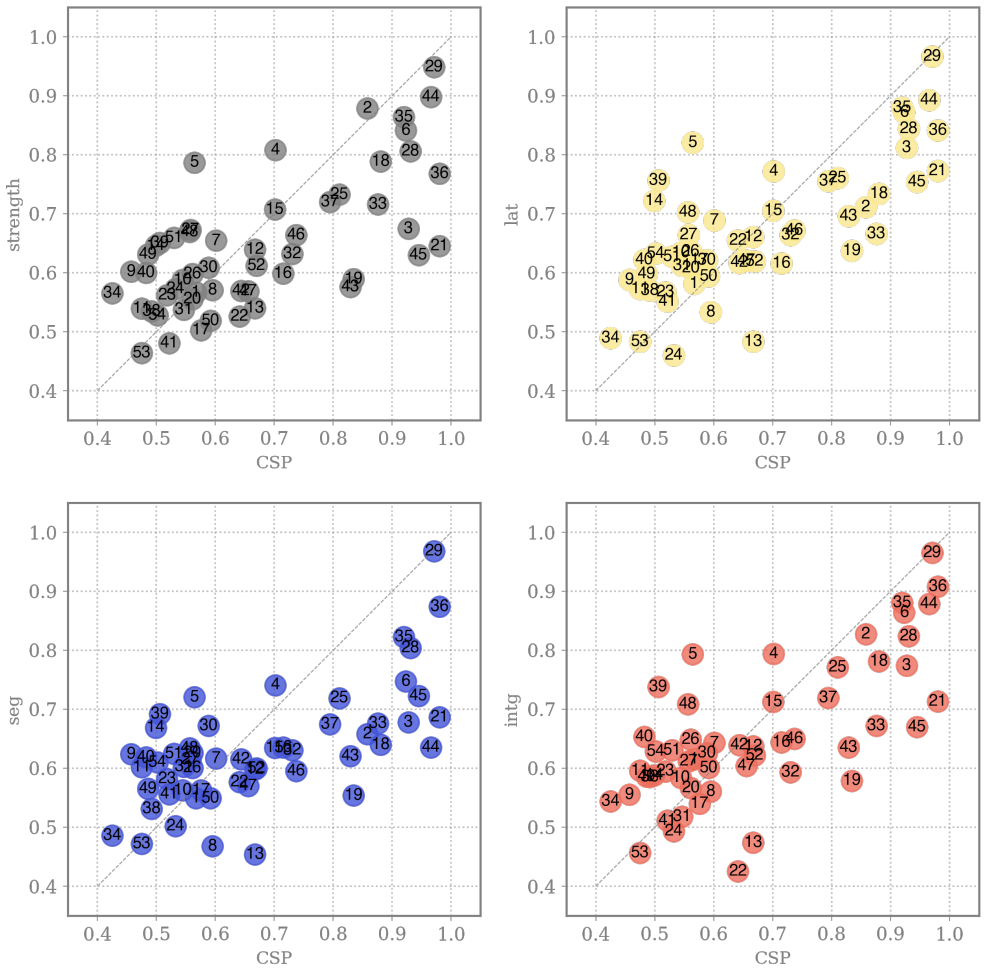


FIGURE 5.13: **Score paired-plot for *Lee2019MI* dataset.** Comparison between CSP and network properties scores. Each numbered point represents the mean across sessions for a single subject. Subjects that get a low score with CSP, tend to get better results with network metrics.

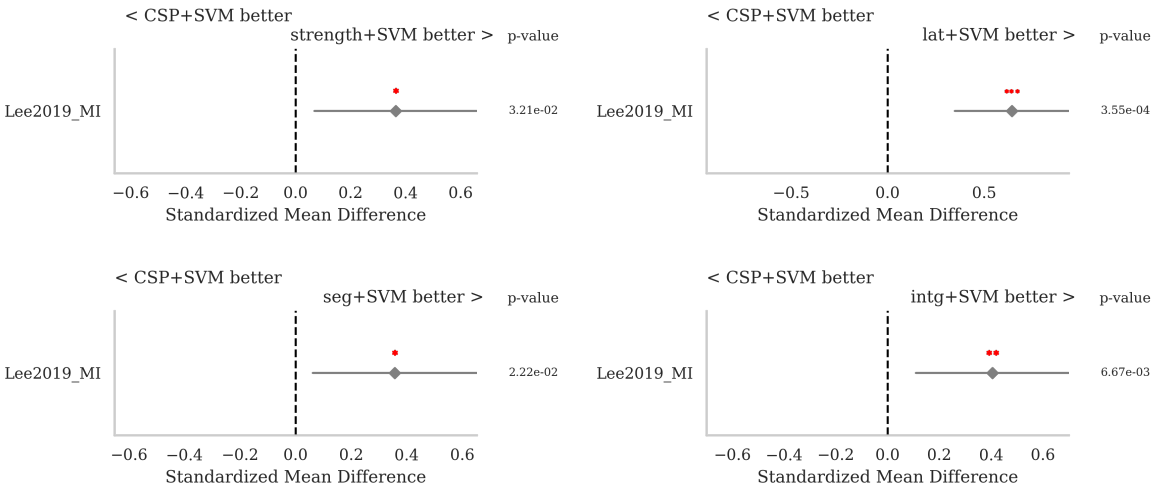


FIGURE 5.14: **Classification statistical analysis for *Lee2019MI* dataset:** lateralization versus CSP, for the subset of subjects (33) that get low CSP scores (<0.7). The same interpretation as in **Fig 5.11** is applied. We observe a significant bias for better results with network features, particularly with *strength* and λ.

Discussion

This PhD project was motivated by the hypothesis that brain network properties might have a beneficial role in the discrimination of different mental states associated with BCIs. In particular, we hypothesized that including the spacial component in the mathematical formulation of the network metrics, might give more interpretable, and possibly accurate, results as compared to standard network topological indices.

The obtained results showed brain network lateralization is a distinct attribute in hand MI, making it particularly appropriate for a classification scenario. Moreover, we evaluated the replicability of our approach over 140 subjects via an open-access toolkit (Jayaram and Barachant, 2018) and demonstrated its reliability in identifying the underlying brain connectivity mechanisms of MI.

Features interpretation

Our main contribution is the proof that brain network lateralization properties can reveal the intrinsic mechanisms that underlay hand MI, converting them into promising features for a MI-BCI system. The relevance of these features can be easily assessed by the spatial position of the most discriminant nodes, mainly covering sensorimotor-related areas, ensuring that artifactual sources are not leading the differentiation.

The t-values topographical maps reveal that the electrodes with the highest discriminant *strength* were located in the contralateral hemisphere of the imagined movement, showing that lateralization is reflected in the connectivity of MI organization. Considering the anatomical symmetry of sensorimotor areas, this boosted the development of properties that compares functional lateralization on homotopic brain regions.

Interestingly, each introduced network metric emphasized different groups of nodes that are related to the motor task at different stages (e.g. planning and coordination). *Laterality index* principally highlighted differences at nodes associated with M1 and S1 cortex, essential areas in motor execution. Notably, this metric considers both hemispheres as completely isolated modules, i.e. it does not include inter-hemispheric links. Hence, one possible conclusion is that areas related to pure motor tasks mostly rely on within-hemisphere connections. On the contrary, inter-hemispheric interactions, are involved in the significance and the complexity of information exchange in high-order functions (Liu, Spagna, and Bartolomeo, 2021). In fact, *integration* increased differentiation on the parietal S2 and supramarginal cortex, more related to spatial orientation (Vanderah and Gould, 2020), visuospatial awareness (Mesulam, 2000) respectively. Then *segregation* emphasized PMA, SMA, and DLFC areas, typically involved in motor planning (Gao, Duan, and Chen, 2011; Hétu et al., 2013).

These results should be also discussed with respect to standard reference methods such as CSP and Riemannian. For the case of CSP, the obtained spatial patterns were also consistent with the MI task by giving the highest weight to electrodes over the motor cortex. Nevertheless, the expected contralateral engagement could be observed since they show a bilateral activation of the motor cortex and influence of other channels, notably in parieto-occipital regions. Apart from this global conclusion, no specific areas could be referred since the patterns showed the involvement of all the sensors.

The Riemannian geometry-based method showed a relevant occurrences concentration over the sensorimotor areas, mainly on right M1 and S1. However, an unexpected subset of selected nodes was located over the parieto-occipital area (electrodes in the PO line) that could be associated with artifacts since these channels are located in the associative visual cortex which has not been proven to be involved in the motor task. Another inconvenience was the impossibility to associate patterns to each mental task since the focus is not put on characterizing each mental state, but on maximizing the distance between SPD matrices.

Classification contrast

Despite the promising discriminant network lateralization patterns, the classification scores of network metrics were still under the reference established by state-of-the-art methods. While we are still exploring the reasons for this evidence, there is a number of considerations that is important to make.

First, we need to recognize that our method does not add any manipulation strictly committed to improving classification as CSP and Riemannian methods do. For instance, CSP aims to directly maximize the variance ratio between two conditions and not to purely recognize the neural sources that generate that variance. In addition, CSP is not robust to outliers, in fact, one single trial with high variance can have a strong impact on the resulting filters (Blankertz et al., 2007).

Second, Riemannian methods are the best in terms of accuracy, but their principal limitation is their lack of interpretability. Most implemented techniques based their results on Minimum Distance Mean (MDM) (Barachant et al., 2011). They based their results on the distance between class mean SPD matrices, without any intermediate interpretation, being blind to the true feature that generates this distance. Another popular technique consists in projecting to the tangent space. There we have to pay special attention to the dimensionality of SPD matrices. Features derived from high-dimensional covariance matrices are prone to overfitting because of the limited number of trials usually available in BCI datasets (Lotte et al., 2007; Rodrigues et al., 2017). The Riemannian solution used in this project (Barachant and Bonnet, 2011), tackles both issues by selecting a limited number of sensors in the manifold based on their class discriminability, enabling their posterior interpretation.

Recent publications have demonstrated their interest in validating Riemannian-based accuracies with a suitable neurophysiological interpretation. Larzabal et al., 2021 took the same Riemannian selection approach used in our project (Barachant and Bonnet, 2011). But instead of working with accumulated occurrences, they attributed Riemannian distances to

electrodes. Within each backward iteration (for more details see **section 4.3**) they assigned the Riemannian distance between classes to the removed electrode. Then there is an inverse relationship between distance and the contribution of the electrode in separating the classes. In congruence with our work, they reported better results in terms of interpretability and classification performance, than CSP. Other authors have tried the combination of both techniques, CSP and Riemannian, looking for better interpretation. Xu, Grosse-Wentrup, and Jayaram, 2020 studied spatial filters in the tangent space that enables CSP-like pattern analysis while improving accuracies. Compared to the CSP method the resulting patterns were less prone to artifacts and could extract additional neurophysiological activity.

Methodological considerations

Even though we showed the reliability of our approach in finding consistent neurophysiological sources over a considerable number of datasets, this study presents clear caveats that need to be acknowledged and addressed in the future. A first limitation is related to the signal preprocessing steps included in our pipeline. Indeed, only pass-band filtering is included. Other filtering techniques like Common Average Reference (CAR) for re-referencing or Independent Component Analysis (ICA) for artifacts suppression would have been beneficial for this study and might have helped in improving accuracy (Bashashati et al., 2007). Nonetheless, looking at the topographical t-test scalp maps on the network side, artifacts do not seem to show a leading role in the results.

With respect to pass-band filtering, we have worked with the assumption that MI generates distinguishing ERD/ERS in the α and β frequency bands. But a more thorough study within sharper bands may be worthwhile since α and β components differ with temporal behavior. Pfurtscheller and Neuper, 2001 have demonstrated the existence of at least three different types of oscillations at the same electrode location over the sensorimotor cortex in voluntary hand movement. Then working at different band levels may generate different and potentially more precise results. One possibility is to test the characteristics of each frequency band before immersing into feature extraction (Corsi et al., 2022) or work it out at the feature selection level by looking at precise single frequency bins (Cattai et al., 2021b).

In our MI lateralization study, we have worked at the network level. Other studies have directly explored the possibility to use FC as feature for MI classification (Brunner et al., 2006; Hamner et al., 2011; Zhang et al., 2012; Li et al., 2016; Feng et al., 2020). Here we have decided not to test FC because it implies concatenating the connectivity matrix which can drive to an overfitting problem. Two possible solutions to this problem would be channel selection and dimensionality reduction techniques. Indeed, Li et al., 2016 used principal component analysis (PCA) on concatenated FC matrices to reduce feature dimension after feature extraction. Their results on the BCI competition II dataset (*001-2014*) of SVM classification were consistently high (82%, 2% under the winning accuracy reported by the competition (Blankertz, Curio, and Müller, 2001)). This suggests that a direct classification at the FC level can give competitive results.

Lastly, all these results assume an approximated correspondence between the EEG channel locations and the brain areas beneath. Further analysis in the source space, could be

of interest to provide a more accurate description of the neural mechanisms detected with our method (Jatoi et al., 2014; Barzegaran and Knyazeva, 2017). To address this solution two main limitations must be considered. First, individual magnetic resonance images (MRIs) are needed to have a realistic model of the brain but are not available for the studied datasets. Second, FC estimators can be sensitive to signal transformations and results can strongly depend on the selected reconstruction algorithm. Future research is necessary to investigate the stability of our results at the source space level.

Conclusion and Perspectives

In this PhD project, we aimed to introduce a novel approach to improve BCI performance. By acting on the feature extraction block of a typical BCI pipeline, we based our original contribution on the development of network-based metrics extracted from functional brain connectivity. More specifically, we hypothesized that the spacial brain organization (i.e. the fact that brain nodes are spatially embedded) might have a role in the discrimination of different BCI-related mental states. The obtained results showed that brain network lateralization is a distinct attribute in hand MI, making it particularly appropriate for a classification scenario. The ensemble of introduced lateralization indexes proved its efficiency in identifying the key components that intervene in MI.

Several BCI studies tend to overlook the feature interpretation, focusing only on classification scores to validate their designs. In this project, we compared our approach to two typically used methods, CSP and Riemannian geometry, and by looking at the most discriminant features, we found out that they were not all strictly related to the MI task. This questions the interpretability of the classification performance and the extend to which their scores can be actually associated to the neural processes of the MI task.

In the BCI community, we have high expectations for the development of tools that will help in decoding mental states. Two major conditions have to be simultaneously reached, high accuracy and neurological plausibility. The latter has been validated by our method, but research is still needed to improve its scores.

Many ideas emerge from this project in several research directions, intended to develop new techniques or possible applications, that could be pursued in the future.

The multiple brain areas involved in MI elicits a complex ordered dynamic of activation. Pfurtscheller and Neuper, 2001 claimed that movement preparation and execution go through multiple stages in its whole process, generating different patterns of oscillation over time. These fluctuations in the mental state are related to dynamic variations in the functional network (Zalesky et al., 2014; Shine et al., 2015). Therefore, analyses of time-varying network reconfiguration during MI may provide a new tool to capture the dynamics of the task and corroborate the involvement of different sensorimotor areas at different stages (De Vico Fallani et al., 2008a).

Building on the advances in information geometry for BCI, Corsi et al., 2022 proposes a novel framework that combines FC estimators and covariance-based pipelines. By ensuring that connectivity matrices are SPD, they worked with a two-step classification in which FC matrices are handled in the Riemannian manifold, obtaining outstanding and neurophysiologically validated results. Encouraged by this work, we presume that network-based classification in the Riemannian manifold could be a promising approach. Given a vector

of local properties characterizing an EEG-based network (e.g. strength), it is possible to work with the covariance between each pair of electrode properties in the vector. Then if the resulting matrix is SPD, it can be projected onto the Riemannian manifold and follow a similar path as the Riemannian method presented in this thesis ((see **section 4.3**)). The appropriate mythological feasibility of this hypothesis must be demonstrated.

Another interesting avenue is clinical application. BCI is a recognized technique in the field of stroke rehabilitation due to the power of MI to enhance motor recovery (Ang and Guan, 2013). Brain lesions typically affect only one brain hemisphere, then the brain distribution of MI task may have practical consequences. In an attempt to maximize the residual cortical output, two patterns are observed during recovery, bilateral recruitment or lateralization toward the perilesional tissue and remaining motor regions of the lesioned hemisphere (Westlake and Nagarajan, 2011; Grefkes and Fink, 2011; Sabaté, González, and Rodríguez, 2004). These plasticity and compensation mechanisms can be capture by network lateralization properties, and exploited as features for MI-BCI rehabilitation.

Appendix A

Features analysis

A.1 Network properties per dataset

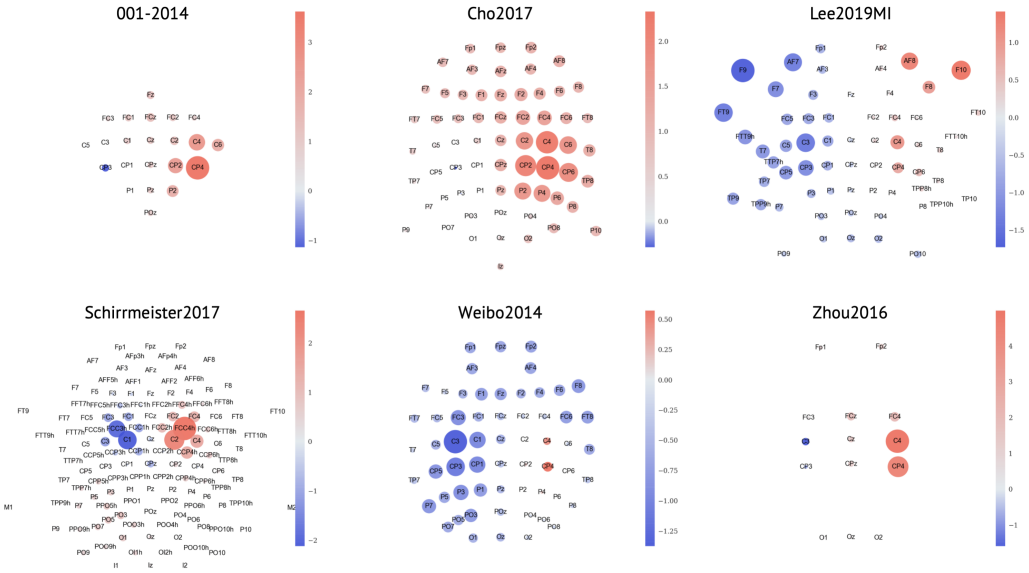


FIGURE A.1: **Strength lateralization in MI tasks per dataset.**

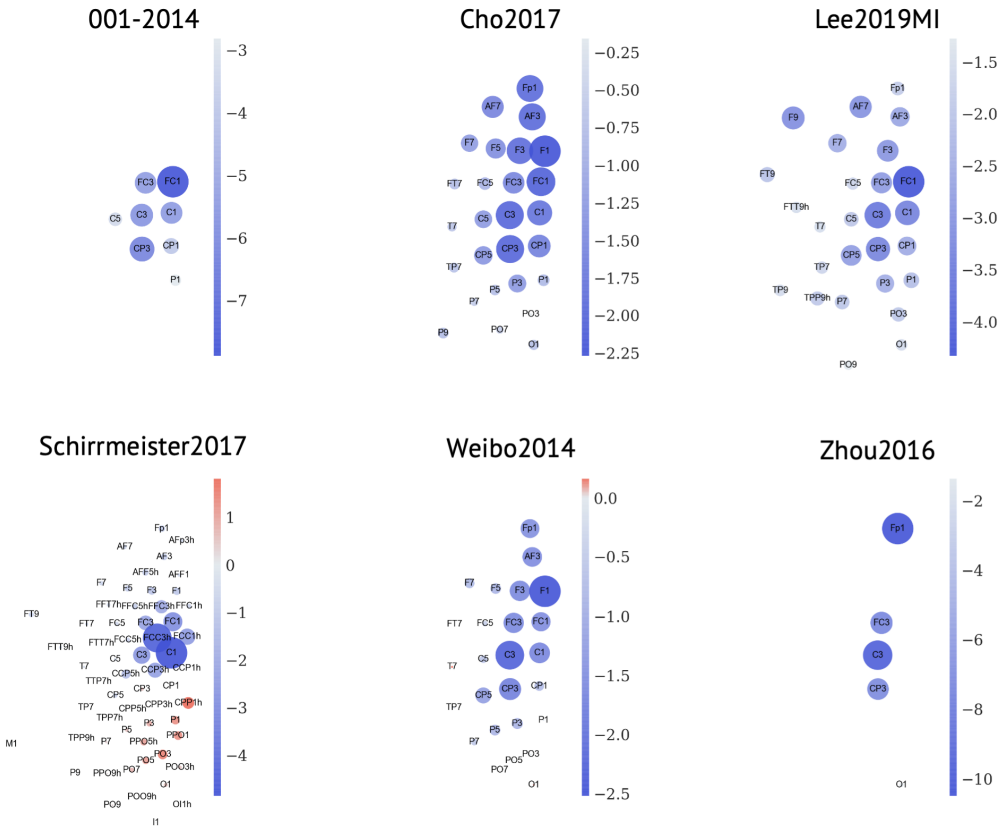


FIGURE A.2: **Laterality index in MI tasks per dataset.**

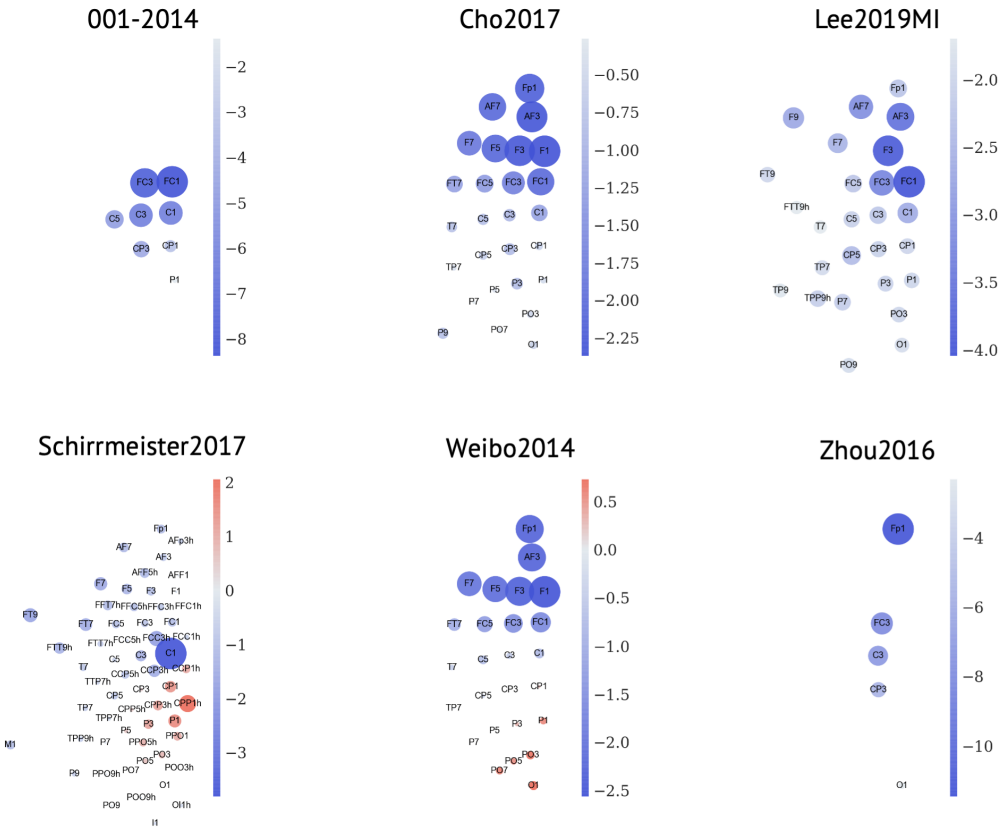


FIGURE A.3: **Segregation in MI tasks per dataset.**

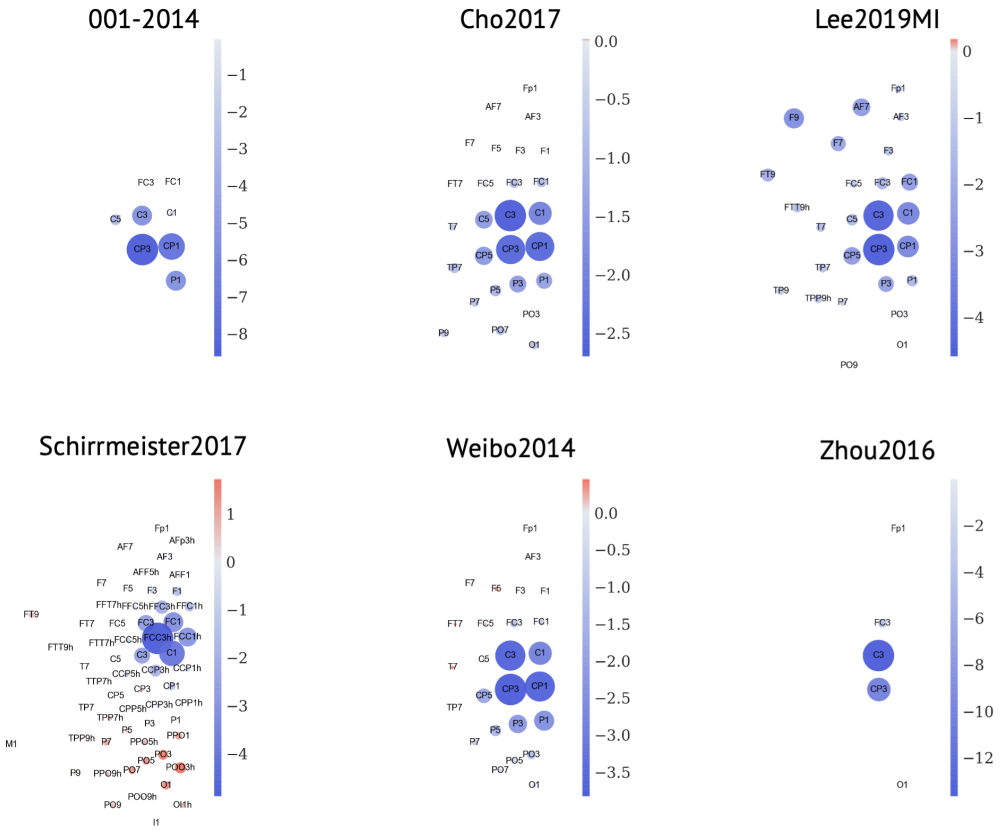


FIGURE A.4: **Integration in MI tasks per dataset.**

TABLE A.1: **Minimum and maximum t-values obtained across datasets.** t-val: t-value, ch: channel.

Dataset	*strength*				*laterality index*				*segregation*				*integration*			
	max		min		max		min		max		min		max		min	
	t-val	ch	t-val	ch	t-val	ch	t-val	ch	t-val	ch	t-val	ch	t-val	ch	t-val	ch
001-2014	3.62	CP4	-0.03	C3	-7.88	FC1	-2.81	P1	-8.38	FC1	-1.38	P1	-8.61	CP3	-0.04	FC1
Cho2017	2.33	CP4	0.01	O1	-2.27	F1	-0.16	PO3	-2.37	F1	-0.26	P7	-2.70	C3	0.01	F5
Lee2019MI	-1.73	F9	0.03	TP10	-4.33	FC1	1.27	PO10	-4.04	FC1	1.69	FTT10h	-4.59	CP3	-0.05	PO10
Schirrmeister2017	2.66	FCC4h	0.01	TTP8h	-4.85	C1	-0.01	TPP7h	-3.81	C1	-0.03	POO3h	-4.89	FCC3h	0.00	AF7
Weibo2014	-1.37	C3	0.01	P2	-2.51	F1	-0.03	PO7	-2.56	F1	-0.00	P5	-3.82	CP3	-0.03	TP7
Zhou2016	4.99	C4	0.40	O1	-10.48	Fp1	-1.37	O1	-11.46	Fp1	-2.31	O1	-13.66	C3	-0.01	O1

A.2 Segregation

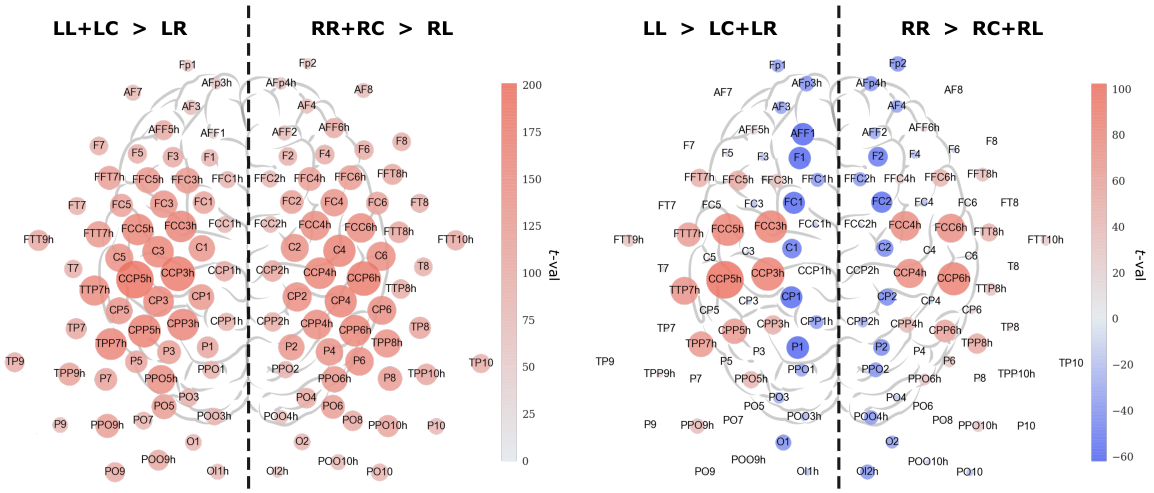


FIGURE A.5: **Influence of middle line links.** Group average t-values that experimentally demonstrate the influence of middle line edges on each hemisphere. At each node, we statistically compare the within connection versus the inter-hemispheric in two possible scenarios. On the left, we show the results when we consider the influence of including middle line links (LC_i, RC_j) as within-hemisphere. On the right, the results of considering them as inter-hemispheric. Excluding LC_i and RC_j from the within-connections has a localized negative impact on nodes closer to the central line. On the other hand, $LL_i + LC_i > LR_i$ and $RR_j + RC_j > RL_j$ guarantees positives values for each hemisphere segregation. Then when analyzing the lateralization of σ, a negative value implies stronger segregation on the right hemisphere.

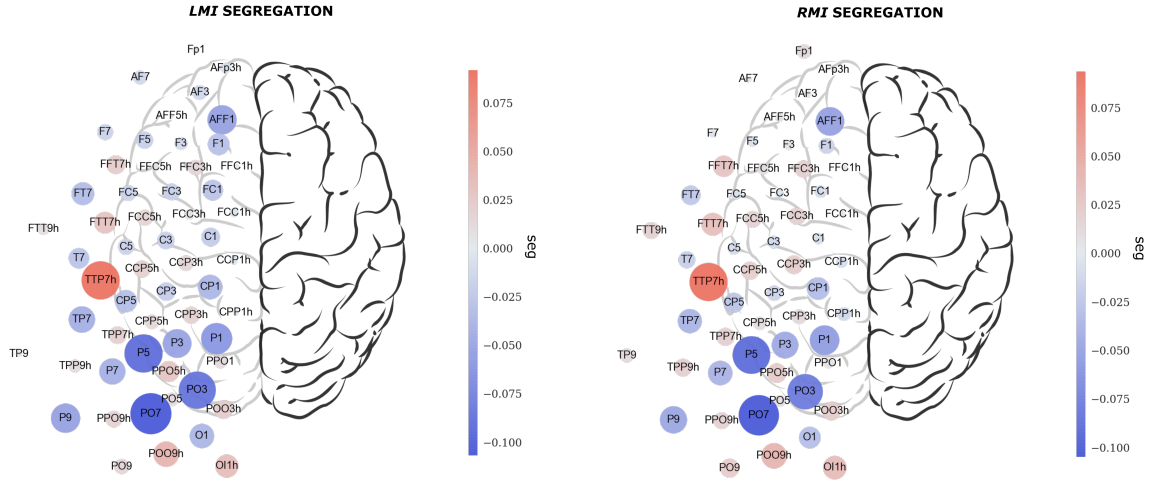


FIGURE A.6: **Segregation in LMI and RMI.**

A.3 Network feature selection

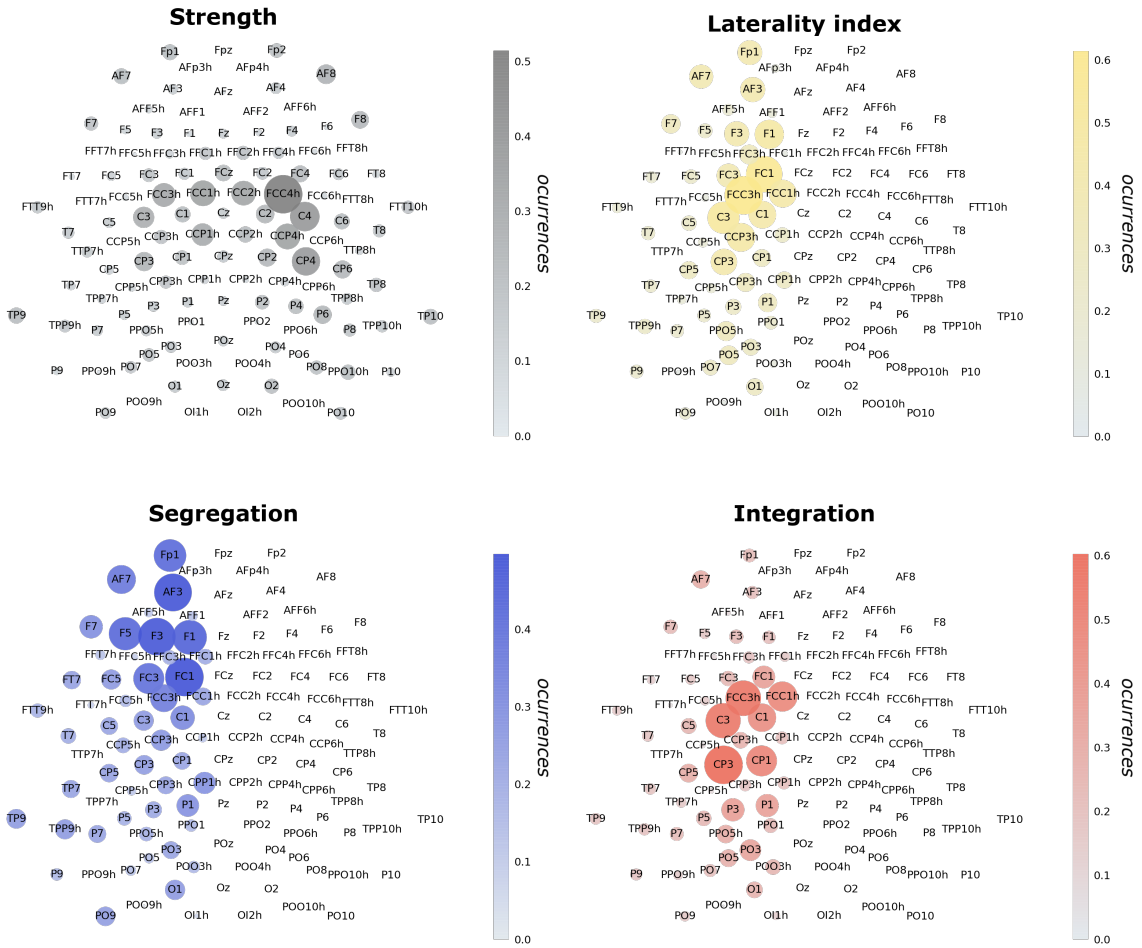


FIGURE A.7: **Network feature selection.**

A.4 Common spatial pattern per dataset

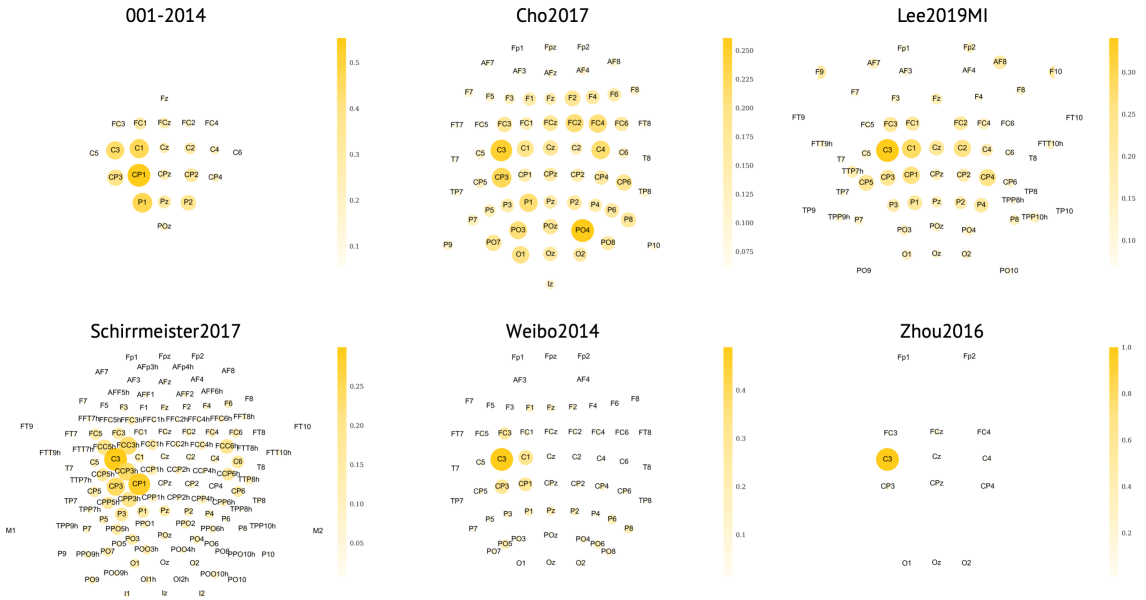


FIGURE A.8: **CSP filters in *RMI* per dataset.**

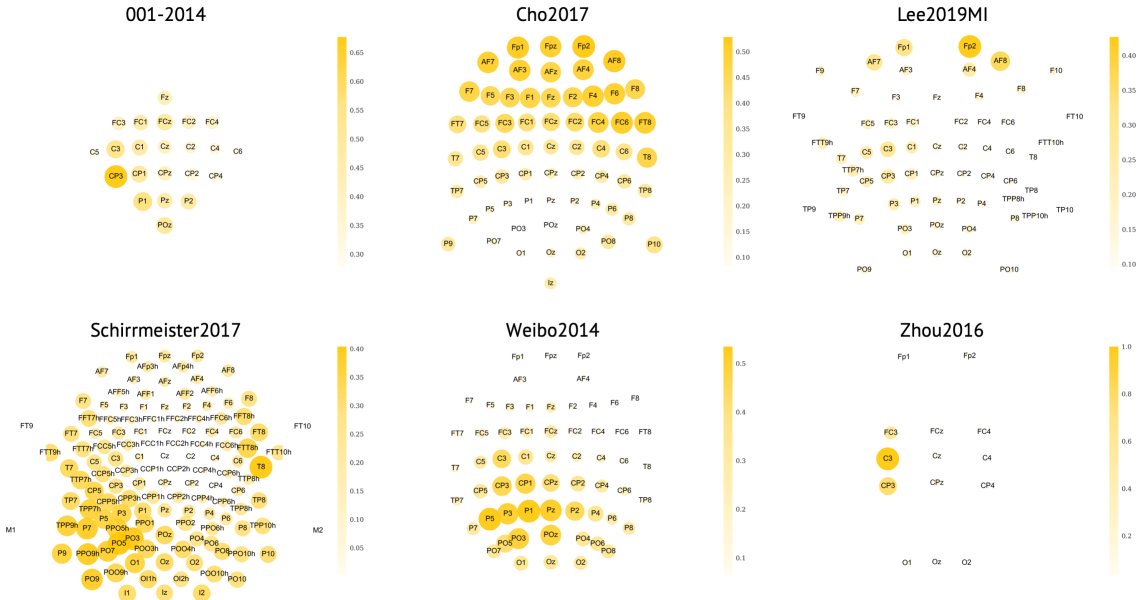


FIGURE A.9: **CSP patterns in *RMI* per dataset.**

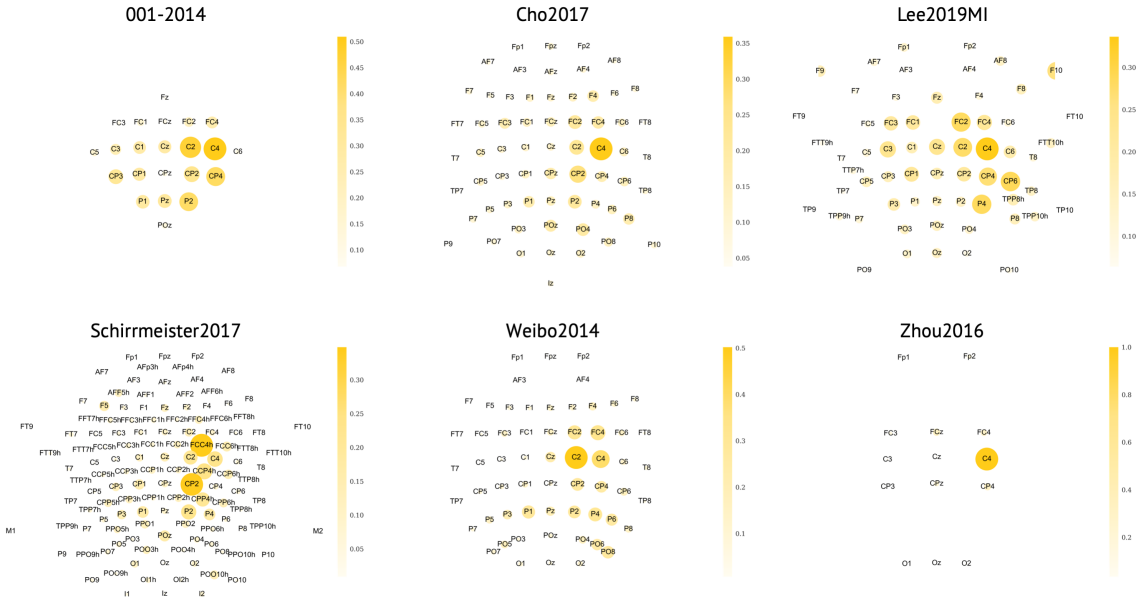


FIGURE A.10: **CSP filters in *LMI* per dataset.**

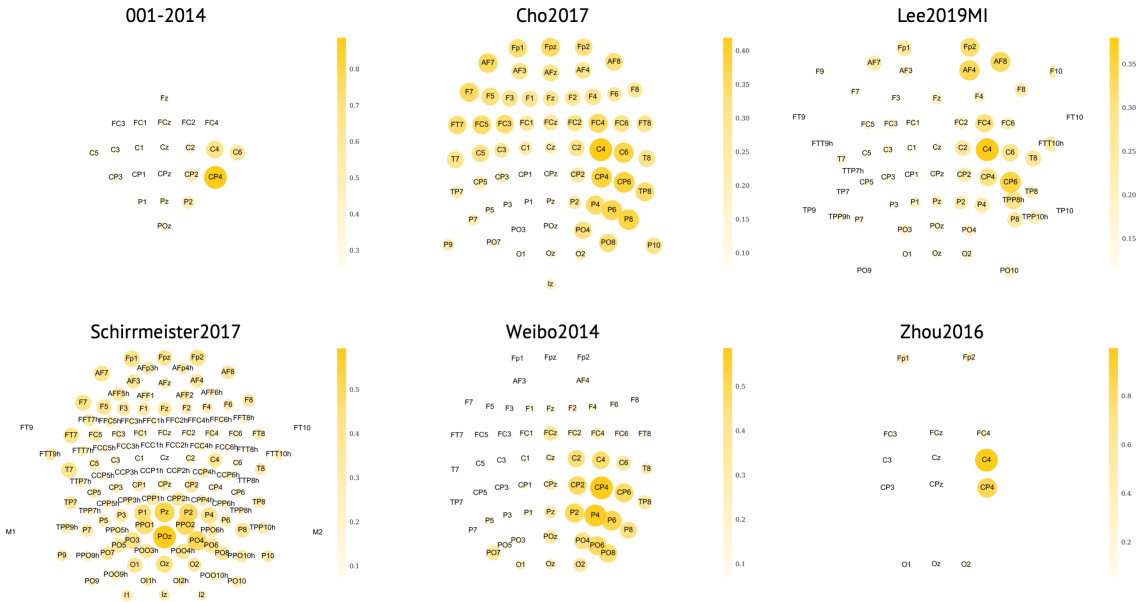


FIGURE A.11: **CSP patterns in *LMI* per dataset.**

A.5 Riemannian feature selection per dataset

FIGURE A.12: **Riemannian-based feature selection per dataset.**

Bibliography

Achard, S. and E. Bullmore (2007). "Efficiency and cost of economical brain functional networks". In: *PLoS computational biology* 3.2, e17.

Achard, S. et al. (2012). "Hubs of brain functional networks are radically reorganized in comatose patients". In: *Proceedings of the National Academy of Sciences* 109.50, pp. 20608–20613.

Albert, R. and A.-L. Barabási (2002). "Statistical mechanics of complex networks". In: *Reviews of modern physics* 74.1, p. 47.

Amunts, K. et al. (1996). "Asymmetry in the human motor cortex and handedness". In: *Neuroimage* 4.3, pp. 216–222.

Ang, K. K. and C. Guan (2013). "Brain-computer interface in stroke rehabilitation". In: *Journal of Computing Science and Engineering* 7.2, pp. 139–146.

Ang, K. K. et al. (2008). "Filter bank common spatial pattern (FBCSP) in brain-computer interface". In: *2008 IEEE international joint conference on neural networks (IEEE world congress on computational intelligence)*. IEEE, pp. 2390–2397.

Anthonisse, J. M. (1971). "The rush in a directed graph". In: *Stichting Mathematisch Centrum. Mathematische Besliskunde* BN 9/71.

Attina, V et al. (2008). *The importance of individual features for motor-imagery based bci*. 4th International Brain-Computer Interface Workshop and Training Course.

Aydore, S., D. Pantazis, and R. M. Leahy (2013). "A note on the phase locking value and its properties". In: *Neuroimage* 74, pp. 231–244.

Babadi, B. and E. N. Brown (2014). "A review of multitaper spectral analysis". In: *IEEE Transactions on Biomedical Engineering* 61.5, pp. 1555–1564.

Babloyantz, A. and A. Destexhe (1986). "Low-dimensional chaos in an instance of epilepsy." In: *Proceedings of the National Academy of Sciences* 83.10, pp. 3513–3517.

Baccalá, L. A. and K. Sameshima (2001). "Partial directed coherence: a new concept in neural structure determination". In: *Biological cybernetics* 84.6, pp. 463–474.

Baillet, S., J. C. Mosher, and R. M. Leahy (2001). "Electromagnetic brain mapping". In: *IEEE Signal processing magazine* 18.6, pp. 14–30.

Barabási, A.-L. and R. Albert (1999). "Emergence of scaling in random networks". In: *science* 286.5439, pp. 509–512.

Barachant, A. and S. Bonnet (2011). "Channel selection procedure using Riemannian distance for BCI applications". In: *2011 5th International IEEE/EMBS Conference on Neural Engineering*. IEEE, pp. 348–351.

Barachant, A. et al. (2010a). "Common spatial pattern revisited by Riemannian geometry". In: *2010 IEEE International Workshop on Multimedia Signal Processing*. IEEE, pp. 472–476.

Barachant, A. et al. (2010b). "Riemannian geometry applied to BCI classification". In: *International conference on latent variable analysis and signal separation*. Springer, pp. 629–636.

Barachant, A. et al. (2011). "Multiclass brain–computer interface classification by Riemannian geometry". In: *IEEE Transactions on Biomedical Engineering* 59.4, pp. 920–928.

Barachant, A. et al. (2013). "Classification of covariance matrices using a Riemannian-based kernel for BCI applications". In: *Neurocomputing* 112, pp. 172–178.

Baraniuk, R. G. and M. Bayram (2000). "Multiple window time varying spectrum estimation". In: *Cambridge University Press*.

Barlow, H. B. (1989). "Unsupervised learning". In: *Neural computation* 1.3, pp. 295–311.

Barrett, A. B., L. Barnett, and A. K. Seth (2010). "Multivariate Granger causality and generalized variance". In: *Physical Review E* 81.4, p. 041907.

Barzegaran, E. and M. G. Knyazeva (2017). "Functional connectivity analysis in EEG source space: the choice of method". In: *PLoS One* 12.7, e0181105.

Başar, E. (2012). *Brain function and oscillations: volume I: brain oscillations. Principles and approaches*. Springer Science & Business Media.

Başar, E. and A. Düzgün (2016). "The CLAIR model: Extension of Brodmann areas based on brain oscillations and connectivity". In: *International Journal of Psychophysiology* 103, pp. 185–198.

Bashashati, A. et al. (2007). "A survey of signal processing algorithms in brain–computer interfaces based on electrical brain signals". In: *Journal of Neural engineering* 4.2, R32.

Bassett, D. S. and E. T. Bullmore (2009). "Human brain networks in health and disease". In: *Current opinion in neurology* 22.4, p. 340.

Bassett, D. S. and M. G. Mattar (2017). "A network neuroscience of human learning: potential to inform quantitative theories of brain and behavior". In: *Trends in cognitive sciences* 21.4, pp. 250–264.

Bassett, D. S. et al. (2011). "Dynamic reconfiguration of human brain networks during learning". In: *Proceedings of the National Academy of Sciences* 108.18, pp. 7641–7646.

Bassett, D. S. et al. (2015). "Learning-induced autonomy of sensorimotor systems". In: *Nature neuroscience* 18.5, pp. 744–751.

Bassett, D. S. and E. Bullmore (2006). "Small-world brain networks". In: *The neuroscientist* 12.6, pp. 512–523.

Bastos, A. M. and J.-M. Schoffelen (2016). "A tutorial review of functional connectivity analysis methods and their interpretational pitfalls". In: *Frontiers in systems neuroscience* 9, p. 175.

Battiston, F. et al. (2018). "Multiplex core–periphery organization of the human connectome". In: *Journal of the Royal Society Interface* 15.146, p. 20180514.

Beisteiner, R et al. (1995). "Mental representations of movements. Brain potentials associated with imagination of hand movements". In: *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section* 96.2, pp. 183–193.

Benzi, M. and C. Klymko (2013). "Total communicability as a centrality measure". In: *Journal of Complex Networks* 1.2, pp. 124–149.

Betzel, R. F. et al. (2016). "Generative models of the human connectome". In: *Neuroimage* 124, pp. 1054–1064.

Betzel, R. F. et al. (2019). "Distance-dependent consensus thresholds for generating group-representative structural brain networks". In: *Network neuroscience* 3.2, pp. 475–496.

Blankertz, B., G. Curio, and K.-R. Müller (2001). "Classifying single trial EEG: Towards brain computer interfacing". In: *Advances in neural information processing systems* 14.

Blankertz, B. et al. (2005). "The Berlin brain-computer interface: Report from the feedback sessions". In: *FIRST Reports*.

Blankertz, B. et al. (2007). "Optimizing spatial filters for robust EEG single-trial analysis". In: *IEEE Signal processing magazine* 25.1, pp. 41–56.

Blinowska, K. J., R. Kuś, and M. Kamiński (2004). "Granger causality and information flow in multivariate processes". In: *Physical Review E* 70.5, p. 050902.

Boccaletti, S. et al. (2006). "Complex networks: Structure and dynamics". In: *Physics reports* 424.4-5, pp. 175–308.

Boccaletti, S. et al. (2014). "The structure and dynamics of multilayer networks". In: *Physics reports* 544.1, pp. 1–122.

Bonacich, P. (1972). "Factoring and weighting approaches to status scores and clique identification". In: *Journal of mathematical sociology* 2.1, pp. 113–120.

Borgatti, S. P. (2005). "Centrality and network flow". In: *Social networks* 27.1, pp. 55–71.

Borgatti, S. P. and M. G. Everett (2000). "Models of core/periphery structures". In: *Social networks* 21.4, pp. 375–395.

Bozinovski, S., M. Sestakov, and L. Bozinovska (1988). "Using EEG alpha rhythm to control a mobile robot". In: *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, pp. 1515–1516.

Brittain, J.-S. et al. (2007). "Single-trial multiwavelet coherence in application to neurophysiological time series". In: *IEEE Transactions on Biomedical Engineering* 54.5, pp. 854–862.

Brodman, K (1909). *Vergleichende Lokalisationslehre der Grosshirnrinde, in ihren Prinzipien dargestellt auf Grund des Zellenbaues*. Barth Leipzig.

Broek, S. P. van den et al. (1998). "Volume conduction effects in EEG and MEG". In: *Electroencephalography and clinical neurophysiology* 106.6, pp. 522–534.

Brunner, C. et al. (2006). "Online control of a brain-computer interface using phase synchronization". In: *IEEE Transactions on Biomedical Engineering* 53.12, pp. 2501–2506.

Buch, V. P. et al. (2018). "Network brain-computer interface (nBCI): An alternative approach for cognitive prosthetics". In: *Frontiers in neuroscience*, p. 790.

Buitinck, L. et al. (2013). "API design for machine learning software: experiences from the scikit-learn project". In: *arXiv preprint arXiv:1309.0238*.

Bullmore, E. and O. Sporns (2009). "Complex brain networks: graph theoretical analysis of structural and functional systems". In: *Nature reviews neuroscience* 10.3, pp. 186–198.

Buttfield, A., P. W. Ferrez, and J. R. Millan (2006). "Towards a robust BCI: error potentials and online learning". In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 14.2, pp. 164–168.

Carlson, T. and J. d. R. Millan (2013). "Brain-controlled wheelchairs: a robotic architecture". In: *IEEE Robotics & Automation Magazine* 20.1, pp. 65–73.

Carmena, J. M. et al. (2003). "Learning to control a brain–machine interface for reaching and grasping by primates". In: *PLoS biology* 1.2, e42.

Carter, G. C. (1987). "Coherence and time delay estimation". In: *Proceedings of the IEEE* 75.2, pp. 236–255.

Cattai, T. et al. (2021a). "Improving J-divergence of brain connectivity states by graph Laplacian denoising". In: *IEEE transactions on Signal and Information Processing over Networks* 7, pp. 493–508.

Cattai, T. et al. (2021b). "Phase/amplitude synchronization of brain signals during motor imagery BCI tasks". In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 29, pp. 1168–1177.

Cauda, F. et al. (2011). "Discovering the somatotopic organization of the motor areas of the medial wall using low-frequency bold fluctuations". In: *Human brain mapping* 32.10, pp. 1566–1579.

Cestari, D. M. and J. L. G. Rosa (2017). "Stochastic and deterministic stationarity analysis of EEG data". In: *2017 International Joint Conference on Neural Networks (IJCNN)*. IEEE, pp. 63–70.

Chavarriaga, R. and J. d. R. Millán (2010). "Learning from EEG error-related potentials in noninvasive brain-computer interfaces". In: *IEEE transactions on neural systems and rehabilitation engineering* 18.4, pp. 381–388.

Chavez, M. and B. Cazelles (2019). "Detecting dynamic spatial correlation patterns with generalized wavelet coherence and non-stationary surrogate data". In: *Scientific reports* 9.1, pp. 1–9.

Chavez, M. et al. (2013). "Node accessibility in cortical networks during motor tasks". In: *Neuroinformatics* 11.3, pp. 355–366.

Cheveigné, A. de and I. Nelken (2019). "Filters: when, why, and how (not) to use them". In: *Neuron* 102.2, pp. 280–293.

Cho, H. et al. (2017). "EEG datasets for motor imagery brain–computer interface". In: *GigaScience* 6.7, gix034.

Cohen, A. L. et al. (2008). "Defining functional areas in individual human brains using resting functional connectivity MRI". In: *Neuroimage* 41.1, pp. 45–57.

Congedo, M., A. Barachant, and R. Bhatia (2017). "Riemannian geometry for EEG-based brain-computer interfaces; a primer and a review". In: *Brain-Computer Interfaces* 4.3, pp. 155–174.

Congedo, M. et al. (2017). "A closed-form unsupervised geometry-aware dimensionality reduction method in the Riemannian Manifold of SPD matrices". In: *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, pp. 3198–3201.

Corsi, M.-C. et al. (2019). "Integrating EEG and MEG signals to improve motor imagery classification in brain–computer interface". In: *International journal of neural systems* 29.01, p. 1850014.

Corsi, M.-C. et al. (2020). "Functional disconnection of associative cortical areas predicts performance during BCI training". In: *NeuroImage* 209, p. 116500.

Corsi, M.-C. et al. (2022). "Functional connectivity ensemble method to enhance BCI performance (FUCONE)". In: *IEEE Transactions on Biomedical Engineering*.

Crofts, J. J. and D. J. Higham (2009). "A weighted communicability measure applied to complex brain networks". In: *Journal of the Royal Society Interface* 6.33, pp. 411–414.

Csermely, P. et al. (2013). "Structure and dynamics of core/periphery networks". In: *Journal of Complex Networks* 1.2, pp. 93–123.

Curtis, C. E. and M. D'Esposito (2003). "Persistent activity in the prefrontal cortex during working memory". In: *Trends in cognitive sciences* 7.9, pp. 415–423.

Daly, I., S. J. Nasuto, and K. Warwick (2012). "Brain computer interface control via functional connectivity dynamics". In: *Pattern recognition* 45.6, pp. 2123–2136.

Daly, J. J. and J. R. Wolpaw (2008). "Brain–computer interfaces in neurological rehabilitation". In: *The Lancet Neurology* 7.11, pp. 1032–1043.

Dayan, P. and Y. Niv (2008). "Reinforcement learning: the good, the bad and the ugly". In: *Current opinion in neurobiology* 18.2, pp. 185–196.

De Clercq, W. et al. (2003). "Anticipation of epileptic seizures from standard EEG recordings". In: *The Lancet* 361.9361, p. 971.

De Domenico, M., S. Sasai, and A. Arenas (2016). "Mapping multiplex hubs in human functional brain networks". In: *Frontiers in neuroscience* 10, p. 326.

De Domenico, M. et al. (2013). "Mathematical formulation of multilayer networks". In: *Physical Review X* 3.4, p. 041022.

De Vico Fallani, F. and D. S. Bassett (2019). "Network neuroscience for optimizing brain–computer interfaces". In: *Physics of life reviews* 31, pp. 304–309.

De Vico Fallani, F., V. Latora, and M. Chavez (2017). "A topological criterion for filtering information in complex brain networks". In: *PLoS computational biology* 13.1, e1005305.

De Vico Fallani, F. et al. (2008a). "Cortical network dynamics during foot movements". In: *Neuroinformatics* 6.1, pp. 23–34.

De Vico Fallani, F. et al. (2008b). "Persistent patterns of interconnection in time-varying cortical networks estimated from high-resolution EEG recordings in humans during a simple motor act". In: *Journal of Physics A: Mathematical and Theoretical* 41.22, p. 224014.

De Vico Fallani, F. et al. (2013). "Multiscale topological properties of functional brain networks during motor imagery after stroke". In: *Neuroimage* 83, pp. 438–449.

De Vico Fallani, F. et al. (2014). "Graph analysis of functional brain networks: practical issues in translational neuroscience". In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 369.1653, p. 20130521.

Di Martino, A. et al. (2013). "Shared and distinct intrinsic functional network centrality in autism and attention-deficit/hyperactivity disorder". In: *Biological psychiatry* 74.8, pp. 623–632.

Dominguez, L. G. (2009). "On the risk of extracting relevant information from random data". In: *Journal of neural engineering* 6.5, p. 058001.

Donchin, E., K. M. Spencer, and R. Wijesinghe (2000). "The mental prosthesis: assessing the speed of a P300-based brain-computer interface". In: *IEEE transactions on rehabilitation engineering* 8.2, pp. 174–179.

Dubovik, S. et al. (2012). "The behavioral significance of coherent resting-state oscillations after stroke". In: *Neuroimage* 61.1, pp. 249–257.

Edelman, B. J., B. Baxter, and B. He (2015). "EEG source imaging enhances the decoding of complex right-hand motor imagery tasks". In: *IEEE Transactions on Biomedical Engineering* 63.1, pp. 4–14.

Engel, A. K. et al. (2005). "Invasive recordings from the human brain: clinical insights and beyond". In: *Nature Reviews Neuroscience* 6.1, pp. 35–47.

Erdős, P., A. Rényi, et al. (1960). "On the evolution of random graphs". In: *Publ. Math. Inst. Hung. Acad. Sci* 5.1, pp. 17–60.

Estrada, E. and N. Hatano (2008). "Communicability in complex networks". In: *Physical Review E* 77.3, p. 036111.

Euston, D. R., A. J. Gruber, and B. L. McNaughton (2012). "The role of medial prefrontal cortex in memory and decision making". In: *Neuron* 76.6, pp. 1057–1070.

Fagiolo, G. (2007). "Clustering in complex directed networks". In: *Physical Review E* 76.2, p. 026107.

Farahani, F. V., W. Karwowski, and N. R. Lighthall (2019). "Application of graph theory for identifying connectivity patterns in human brain networks: a systematic review". In: *frontiers in Neuroscience* 13, p. 585.

Faskowitz, J. et al. (2018). "Weighted stochastic block models of the human connectome across the life span". In: *Scientific reports* 8.1, pp. 1–16.

Faymonville, M. E. et al. (2000). "Neural mechanisms of antinociceptive effects of hypnosis". In: *The Journal of the American Society of Anesthesiologists* 92.5, pp. 1257–1267.

Feng, Z. et al. (2020). "Functional connectivity for motor imaginary recognition in brain-computer interface". In: *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, pp. 3678–3682.

Ferrez, P. W. and J. d. R. Millán (2008). "Error-related EEG potentials generated during simulated brain–computer interaction". In: *IEEE transactions on biomedical engineering* 55.3, pp. 923–929.

Fortunato, S. and D. Hric (2016). "Community detection in networks: A user guide". In: *Physics reports* 659, pp. 1–44.

Fouad, M. M. et al. (2015). "Brain computer interface: a review". In: *Brain-computer interfaces*, pp. 3–30.

Fox, M. D. and M. E. Raichle (2007). "Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging". In: *Nature reviews neuroscience* 8.9, pp. 700–711.

Freeman, L. C. (1977). "A set of measures of centrality based on betweenness". In: *Sociometry*, pp. 35–41.

Gallen, C. L. and M. D'Esposito (2019). "Brain modularity: a biomarker of intervention-related plasticity". In: *Trends in cognitive sciences* 23.4, pp. 293–304.

Ganguly, K. and M.-m. Poo (2013). "Activity-dependent neural plasticity from bench to bedside". In: *Neuron* 80.3, pp. 729–741.

Gao, Q., X. Duan, and H. Chen (2011). "Evaluation of effective connectivity of motor areas during motor imagery and execution using conditional Granger causality". In: *Neuroimage* 54.2, pp. 1280–1288.

Ge, R. et al. (2014). "Motor imagery learning induced changes in functional connectivity of the default mode network". In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 23.1, pp. 138–148.

Golub, M. D. et al. (2016). "Brain–computer interfaces for dissecting cognitive processes underlying sensorimotor control". In: *Current opinion in neurobiology* 37, pp. 53–58.

Gong, D. et al. (2016). "Functional integration between salience and central executive networks: a role for action video game experience". In: *Neural plasticity* 2016.

Goñi, J. et al. (2014). "Resting-brain functional connectivity predicted by analytic measures of network communication". In: *Proceedings of the National Academy of Sciences* 111.2, pp. 833–838.

Gonzalez-Astudillo, J. et al. (2021). "Network-based brain–computer interfaces: principles and applications". In: *Journal of Neural Engineering* 18.1, p. 011001.

Gotts, S. J. et al. (2013). "Two distinct forms of functional lateralization in the human brain". In: *Proceedings of the National Academy of Sciences* 110.36, E3435–E3444.

Gourévitch, B., R. L. Bouquin-Jeannès, and G. Faucon (2006). "Linear and nonlinear causality between signals: methods, examples and neurophysiological applications". In: *Biological cybernetics* 95.4, pp. 349–369.

Gramfort, A. et al. (2013). "MEG and EEG Data Analysis with MNE-Python". In: *Frontiers in Neuroscience* 7.267, pp. 1–13. DOI: 10.3389/fnins.2013.00267.

Grefkes, C. and G. R. Fink (2011). "Reorganization of cerebral networks after stroke: new insights from neuroimaging with connectivity approaches". In: *Brain* 134.5, pp. 1264–1276.

Grezes, J. and J. Decety (2001). "Functional anatomy of execution, mental simulation, observation, and verb generation of actions: A meta-analysis". In: *Human brain mapping* 12.1, pp. 1–19.

Grosse-Wentrup, M. et al. (2009). "Beamforming in noninvasive brain–computer interfaces". In: *IEEE Transactions on Biomedical Engineering* 56.4, pp. 1209–1219.

Gu, L. et al. (2020). "EEG-based classification of lower limb motor imagery with brain network analysis". In: *Neuroscience* 436, pp. 93–109.

Guillon, J. et al. (2017). "Loss of brain inter-frequency hubs in Alzheimer's disease". In: *Scientific reports* 7.1, pp. 1–13.

Guillon, J. et al. (2019). "Disrupted core-periphery structure of multimodal brain networks in Alzheimer's disease". In: *Network Neuroscience* 3.2, pp. 635–652.

Guillot, A. and C. Collet (2010). *The neurophysiological foundations of mental and motor imagery*. Oxford University Press.

Guillot, A. et al. (2009). "Brain activity during visual versus kinesthetic imagery: an fMRI study". In: *Human brain mapping* 30.7, pp. 2157–2172.

Haken, H. (2000). "Nonlinearities in biology: the brain as an example". In: *Nonlinear Science at the Dawn of the 21st Century*. Springer, pp. 427–445.

Hamedi, M., S.-H. Salleh, and A. M. Noor (2016). "Electroencephalographic motor imagery brain connectivity analysis for BCI: a review". In: *Neural computation* 28.6, pp. 999–1041.

Hamner, B. et al. (2011). "Phase-based features for motor imagery brain-computer interfaces". In: *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, pp. 2578–2581.

Hampson, M. et al. (2006). "Brain connectivity related to working memory performance". In: *Journal of Neuroscience* 26.51, pp. 13338–13343.

Hardwick, R. M. et al. (2018). "Neural correlates of action: Comparing meta-analyses of imagery, observation, and execution". In: *Neuroscience & Biobehavioral Reviews* 94, pp. 31–44.

Harrison, A. H. and J. F. Connolly (2013). "Finding a way in: a review and practical evaluation of fMRI and EEG for detection and assessment in disorders of consciousness". In: *Neuroscience & Biobehavioral Reviews* 37.8, pp. 1403–1419.

Hart, H. et al. (2013). "Meta-analysis of functional magnetic resonance imaging studies of inhibition and attention in attention-deficit/hyperactivity disorder: exploring task-specific, stimulant medication, and age effects". In: *JAMA psychiatry* 70.2, pp. 185–198.

Heitger, M. H. et al. (2012). "Motor learning-induced changes in functional brain connectivity as revealed by means of graph-theoretical network analysis". In: *Neuroimage* 61.3, pp. 633–650.

Hétu, S. et al. (2013). "The neural network of motor imagery: an ALE meta-analysis". In: *Neuroscience & Biobehavioral Reviews* 37.5, pp. 930–949.

Hiremath, S. V. et al. (2015). "Brain computer interface learning for systems based on electrocorticography and intracortical microelectrode arrays". In: *Frontiers in integrative neuroscience* 9, p. 40.

Holland, P. W., K. B. Laskey, and S. Leinhardt (1983). "Stochastic blockmodels: First steps". In: *Social networks* 5.2, pp. 109–137.

Holme, P. and J. Saramäki (2012). "Temporal networks". In: *Physics reports* 519.3, pp. 97–125.

Horwitz, B. (2003). "The elusive concept of brain connectivity". In: *Neuroimage* 19.2, pp. 466–470.

Hoshi, E. and J. Tanji (2007). "Distinctions between dorsal and ventral premotor areas: anatomical connectivity and functional properties". In: *Current opinion in neurobiology* 17.2, pp. 234–242.

Humphries, M. D., K. Gurney, and T. J. Prescott (2006). "The brainstem reticular formation is a small-world, not scale-free, network". In: *Proceedings of the Royal Society B: Biological Sciences* 273.1585, pp. 503–511.

Hutchison, R. M. et al. (2013). "Dynamic functional connectivity: promise, issues, and interpretations". In: *Neuroimage* 80, pp. 360–378.

Hwang, H.-J. et al. (2013). "EEG-based brain-computer interfaces: a thorough literature survey". In: *International Journal of Human-Computer Interaction* 29.12, pp. 814–826.

Iasemidis, L. D. and J. C. Sackellares (1996). " REVIEW: Chaos Theory and Epilepsy". In: *The Neuroscientist* 2.2, pp. 118–126.

Ito, H. et al. (2020). "Self-reorganization of neuronal activation patterns in the cortex under brain-machine interface and neural operant conditioning". In: *Neuroscience Research* 156, pp. 279–292.

Jackson, A. F. and D. J. Bolger (2014). "The neurophysiological bases of EEG and EEG measurement: A review for the rest of us". In: *Psychophysiology* 51.11, pp. 1061–1071.

Jacobs, M. P. et al. (2009). "Curing epilepsy: progress and future directions". In: *Epilepsy & Behavior* 14.3, pp. 438–445.

Jalili, M. and M. G. Knyazeva (2011). "Constructing brain functional networks from EEG: partial and unpartial correlations". In: *Journal of integrative neuroscience* 10.02, pp. 213–232.

Jatoi, M. A. et al. (2014). "A survey of methods used for source localization using EEG signals". In: *Biomedical Signal Processing and Control* 11, pp. 42–52.

Jayaram, V. and A. Barachant (2018). "MOABB: trustworthy algorithm benchmarking for BCIs". In: *Journal of neural engineering* 15.6, p. 066011.

Jeannerod, M. (1995). "Mental imagery in the motor context". In: *Neuropsychologia* 33.11, pp. 1419–1432.

Jeannerod, M. and V. Frak (1999). "Mental imaging of motor activity in humans". In: *Current opinion in neurobiology* 9.6, pp. 735–739.

Jirsa, V. and V. Müller (2013). "Cross-frequency coupling in real and virtual brain networks". In: *Frontiers in computational neuroscience* 7, p. 78.

Johnson, S. C. et al. (2002). "Neural correlates of self-reflection". In: *Brain* 125.8, pp. 1808–1814.

Kaiser, A. and T. Schreiber (2002). "Information transfer in continuous processes". In: *Physica D: Nonlinear Phenomena* 166.1-2, pp. 43–62.

Kaminski, M. J. and K. J. Blinowska (1991). "A new method of the description of the information flow in the brain structures". In: *Biological cybernetics* 65.3, pp. 203–210.

Kantak, S. S. et al. (2012). "Rewiring the brain: potential role of the premotor cortex in motor control, learning, and recovery of function following brain injury". In: *Neurorehabilitation and neural repair* 26.3, pp. 282–292.

Katiuscia, S. et al. (2009). "Reorganization and enhanced functional connectivity of motor areas in repetitive ankle movements after training in locomotor attention". In: *Brain research* 1297, pp. 124–134.

Kim, S.-G. et al. (1993). "Functional magnetic resonance imaging of motor cortex: hemispheric asymmetry and handedness". In: *Science* 261.5121, pp. 615–617.

Knudsen, E. I. (1994). "Supervised learning in the brain". In: *Journal of Neuroscience* 14.7, pp. 3985–3997.

Kraskov, A., H. Stögbauer, and P. Grassberger (2004). "Estimating mutual information". In: *Physical review E* 69.6, p. 066138.

Kraut, S., L. L. Scharf, and R. W. Butler (2005). "The adaptive coherence estimator: A uniformly most-powerful-invariant adaptive detection statistic". In: *IEEE Transactions on Signal Processing* 53.2, pp. 427–438.

Kübler, A. et al. (2001). "Brain-computer communication: self-regulation of slow cortical potentials for verbal communication". In: *Archives of physical medicine and rehabilitation* 82.11, pp. 1533–1539.

Kus, R., M. Kaminski, and K. J. Blinowska (2004). "Determination of EEG activity propagation: pair-wise versus multichannel estimate". In: *IEEE transactions on Biomedical Engineering* 51.9, pp. 1501–1510.

Kwiatkowski, D. et al. (1992). "Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?" In: *Journal of econometrics* 54.1-3, pp. 159–178.

Lachaux, J.-P. et al. (2002). "Estimating the time-course of coherence between single-trial brain signals: an introduction to wavelet coherence". In: *Neurophysiologie Clinique/Clinical Neurophysiology* 32.3, pp. 157–174.

Laney, J. et al. (2015). "Quantifying motor recovery after stroke using independent vector analysis and graph-theoretical analysis". In: *NeuroImage: Clinical* 8, pp. 298–304.

Larzabal, C. et al. (2021). "The Riemannian spatial pattern method: mapping and clustering movement imagery using Riemannian geometry". In: *Journal of Neural Engineering* 18.5, p. 056014.

Latora, V. and M. Marchiori (2001). "Efficient behavior of small-world networks". In: *Physical review letters* 87.19, p. 198701.

Latora, V., V. Nicosia, and G. Russo (2017). *Complex networks: principles, methods and applications*. Cambridge University Press.

Lee, M.-H. et al. (2019). "EEG dataset and OpenBMI toolbox for three BCI paradigms: an investigation into BCI illiteracy". In: *GigaScience* 8.5, giz002.

Leicht, E. A. and M. E. Newman (2008). "Community structure in directed networks". In: *Physical review letters* 100.11, p. 118703.

Leistritz, L et al. (2013). "Time-variant partial directed coherence for analysing connectivity: a methodological study". In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 371.1997, p. 20110616.

Li, J. et al. (2016). "Decoding EEG in cognitive tasks with time-frequency and connectivity masks". In: *IEEE Transactions on Cognitive and Developmental Systems* 8.4, pp. 298–308.

Liu, F. et al. (2020). "Deep learning for community detection: progress, challenges and opportunities". In: *arXiv preprint arXiv:2005.08225*.

Liu, H. et al. (2009). "Evidence from intrinsic activity that asymmetry of the human brain is controlled by multiple factors". In: *Proceedings of the national academy of sciences* 106.48, pp. 20499–20503.

Liu, J., A. Spagna, and P. Bartolomeo (2021). "Hemispheric asymmetries in visual mental imagery". In: *Brain Structure and Function*, pp. 1–12.

Liu, Y.-Y., J.-J. Slotine, and A.-L. Barabási (2011). "Controllability of complex networks". In: *nature* 473.7346, pp. 167–173.

Lohmann, G. et al. (2010). "Eigenvector centrality mapping for analyzing connectivity patterns in fMRI data of the human brain". In: *PloS one* 5.4, e10232.

Lotte, F. (2014). "A tutorial on EEG signal-processing techniques for mental-state recognition in brain–computer interfaces". In: *Guide to brain-computer music interfacing*, pp. 133–161.

Lotte, F., L. Bougrain, and M. Clerc (2015). *Electroencephalography (EEG)-based brain-computer interfaces*.

Lotte, F. et al. (2007). "A review of classification algorithms for EEG-based brain–computer interfaces". In: *Journal of neural engineering* 4.2, R1.

Lotte, F. et al. (2018). "A review of classification algorithms for EEG-based brain–computer interfaces: a 10 year update". In: *Journal of neural engineering* 15.3, p. 031005.

Lotze, M. and U. Halsband (2006). "Motor imagery". In: *Journal of Physiology-paris* 99.4-6, pp. 386–395.

Lu, H. et al. (2010). "Regularized common spatial pattern with aggregation for EEG classification in small-sample setting". In: *IEEE transactions on Biomedical Engineering* 57.12, pp. 2936–2946.

Luppino, G and G Rizzolatti (2000). "The organization of the frontal motor cortex". In: *Physiology* 15.5, pp. 219–224.

Ma, A. and R. J. Mondragón (2015). "Rich-cores in networks". In: *PloS one* 10.3, e0119678.

Mahjoory, K. et al. (2017). "Consistency of EEG source localization and connectivity estimates". In: *Neuroimage* 152, pp. 590–601.

Makhtar, S. N. et al. (2020). "Improved functional connectivity network estimation for brain networks using multivariate partial coherence". In: *Journal of neural engineering* 17.2, p. 026013.

Marinazzo, D. et al. (2011). "Nonlinear connectivity by Granger causality". In: *Neuroimage* 58.2, pp. 330–338.

Markett, S. et al. (2018). "Working memory capacity and the functional connectome-insights from resting-state fMRI and voxelwise centrality mapping". In: *Brain imaging and behavior* 12.1, pp. 238–246.

Marrelec, G. et al. (2006). "Partial correlation for functional brain interactivity investigation in functional MRI". In: *Neuroimage* 32.1, pp. 228–237.

Márton, L. et al. (2014). "Detrended fluctuation analysis of EEG signals". In: *Procedia Technology* 12, pp. 125–132.

Maslov, S. and K. Sneppen (2002). "Specificity and stability in topology of protein networks". In: *Science* 296.5569, pp. 910–913.

Masoudi-Nejad, A., F. Schreiber, and Z. R. M. Kashani (2012). "Building blocks of biological networks: a review on major network motif discovery algorithms". In: *IET systems biology* 6.5, pp. 164–174.

McDougle, S. D., R. B. Ivry, and J. A. Taylor (2016). "Taking aim at the cognitive side of learning in sensorimotor adaptation tasks". In: *Trends in cognitive sciences* 20.7, pp. 535–544.

McFarland, D. J. et al. (2000). "Mu and beta rhythm topographies during motor imagery and actual movements". In: *Brain topography* 12.3, pp. 177–186.

McSharry, P. E., L. A. Smith, and L. Tarassenko (2003). "Prediction of epileptic seizures: are nonlinear methods relevant?" In: *Nature medicine* 9.3, pp. 241–242.

Mesulam, M.-M. (2000). *Principles of behavioral and cognitive neurology*. Oxford University Press.

Meunier, D. et al. (2009). "Age-related changes in modular organization of human brain functional networks". In: *Neuroimage* 44.3, pp. 715–723.

Michel, C. M. and D. Brunet (2019). "EEG source imaging: a practical review of the analysis steps". In: *Frontiers in neurology* 10, p. 325.

Michel, C. M. et al. (2004). "EEG source imaging". In: *Clinical neurophysiology* 115.10, pp. 2195–2222.

Miller, K. J. et al. (2010). "Cortical activity during motor execution, motor imagery, and imagery-based online feedback". In: *Proceedings of the National Academy of Sciences* 107.9, pp. 4430–4435.

Milo, R. et al. (2002). "Network motifs: simple building blocks of complex networks". In: *Science* 298.5594, pp. 824–827.

Mokienko, O. A. et al. (2013). "Increased motor cortex excitability during motor imagery in brain-computer interface trained subjects". In: *Frontiers in computational neuroscience* 7, p. 168.

Monti, R. P. et al. (2017). "Real-time estimation of dynamic functional connectivity networks". In: *Human brain mapping* 38.1, pp. 202–220.

Mottaz, A. et al. (2018). "Modulating functional connectivity after stroke with neurofeedback: Effect on motor deficits in a controlled cross-over study". In: *NeuroImage: Clinical* 20, pp. 336–346.

Munzert, J., B. Lorey, and K. Zentgraf (2009). "Cognitive motor processes: the role of motor imagery in the study of motor representations". In: *Brain research reviews* 60.2, pp. 306–326.

Naito, E. et al. (2002). "Internally simulated movement sensations during motor imagery activate cortical motor areas and the cerebellum". In: *Journal of Neuroscience* 22.9, pp. 3683–3691.

Netoff, T. I. et al. (2006). "Detecting coupling in the presence of noise and nonlinearity". In: *Handbook of Time Series Analysis: Recent Theoretical Developments and Applications*, pp. 265–282.

Neuper, C. (1999). "Motor imagery and ERD". In: *Handbook of electroencephalography and clinical neurophysiology* 6, pp. 303–322.

Neuper, C. and G. Pfurtscheller (2001). "Event-related dynamics of cortical rhythms: frequency-specific features and functional correlates". In: *International journal of psychophysiology* 43.1, pp. 41–58.

Neuper, C., M. Wörtz, and G. Pfurtscheller (2006). "ERD/ERS patterns reflecting sensorimotor activation and deactivation". In: *Progress in brain research* 159, pp. 211–222.

Newman, M (2006). "Modularity and community structure in networks". In: *Proceedings of the national academy of sciences* 103.23, pp. 8577–8582.

Newman, M. (2012). "Networks: An introduction. 2010: Oxford university press". In: *Artif. Life* 18, pp. 241–242.

Newman, M. E. (2004). "Analysis of weighted networks". In: *Physical review E* 70.5, p. 056131.

Niedermeyer, E. and F. L. da Silva (2005). *Electroencephalography: basic principles, clinical applications, and related fields*. Lippincott Williams & Wilkins.

Nolte, G. et al. (2004). "Identifying true brain interaction from EEG data using the imaginary part of coherency". In: *Clinical neurophysiology* 115.10, pp. 2292–2307.

Nunez, P. L., R. Srinivasan, et al. (2006). *Electric fields of the brain: the neurophysics of EEG*. Oxford University Press, USA.

Nunez, P. L. et al. (1997). "EEG coherency: I: statistics, reference electrode, volume conduction, Laplacians, cortical imaging, and interpretation at multiple scales". In: *Electroencephalography and clinical neurophysiology* 103.5, pp. 499–515.

Obando, C. and F. De Vico Fallani (2017). "A statistical model for brain networks inferred from large-scale electrophysiological signals". In: *Journal of The Royal Society Interface* 14.128, p. 20160940.

Obando, C. et al. (2019). "Temporal connection signatures of human brain networks after stroke". In: *arXiv preprint arXiv:1907.10009*.

Onnela, J.-P. et al. (2005). "Intensity and coherence of motifs in weighted complex networks". In: *Physical Review E* 71.6, p. 065103.

Orsborn, A. L. and B. Pesaran (2017). "Parsing learning in networks using brain–machine interfaces". In: *Current opinion in neurobiology* 46, pp. 76–83.

Ozdemir, A., E. M. Bernat, and S. Aviyente (2017). "Recursive tensor subspace tracking for dynamic brain network analysis". In: *IEEE Transactions on Signal and Information Processing over Networks* 3.4, pp. 669–682.

Paluš, M. (1996). "Nonlinearity in normal human EEG: cycles, temporal asymmetry, nonstationarity and randomness, not chaos". In: *Biological cybernetics* 75.5, pp. 389–396.

Palva, S. and J. M. Palva (2007). "New vistas for α-frequency band oscillations". In: *Trends in neurosciences* 30.4, pp. 150–158.

Pereda, E., R. Q. Quiroga, and J. Bhattacharya (2005). "Nonlinear multivariate analysis of neurophysiological signals". In: *Progress in neurobiology* 77.1-2, pp. 1–37.

Pfurtscheller, G. and A Aranibar (1977). "Event-related cortical desynchronization detected by power measurements of scalp EEG". In: *Electroencephalography and clinical neurophysiology* 42.6, pp. 817–826.

Pfurtscheller, G. and F. L. Da Silva (1999). "Event-related EEG/MEG synchronization and desynchronization: basic principles". In: *Clinical neurophysiology* 110.11, pp. 1842–1857.

Pfurtscheller, G. and C. Neuper (2001). "Motor imagery and direct brain-computer communication". In: *Proceedings of the IEEE* 89.7, pp. 1123–1134.

Pfurtscheller, G. and C. Neuper (1997). "Motor imagery activates primary sensorimotor area in humans". In: *Neuroscience letters* 239.2-3, pp. 65–68.

Pfurtscheller, G. et al. (1997). "EEG-based discrimination between imagination of right and left hand movement". In: *Electroencephalography and clinical Neurophysiology* 103.6, pp. 642–651.

Pfurtscheller, G. et al. (2006). "Mu rhythm (de) synchronization and EEG single-trial classification of different motor imagery tasks". In: *NeuroImage* 31.1, pp. 153–159.

Pichiorri, F et al. (2011). "Sensorimotor rhythm-based brain–computer interface training: the impact on motor cortical responsiveness". In: *Journal of neural engineering* 8.2, p. 025020.

Pichiorri, F. et al. (2015). "Brain–computer interface boosts motor imagery practice during stroke recovery". In: *Annals of neurology* 77.5, pp. 851–865.

Pikovsky, A., M. Rosenblum, J. Kurths, et al. (2001). "A universal concept in nonlinear sciences". In: *Self* 2, p. 3.

Porro, C. A. et al. (1996). "Primary motor and sensory cortex activation during motor performance and motor imagery: a functional magnetic resonance imaging study". In: *Journal of Neuroscience* 16.23, pp. 7688–7698.

Raichle, M. E. et al. (2001). "A default mode of brain function". In: *Proceedings of the National Academy of Sciences* 98.2, pp. 676–682.

Ramoser, H, J Muller-Gerking, and G Pfurtscheller (1999). "Designing optimal spatial filters for single-trial EEG classification in a movement task". In: *IEEE Trans. Rehab. Eng* 8, pp. 441–446.

Ringo, J. L. (1991). "Neuronal interconnection as a function of brain size". In: *Brain, behavior and evolution* 38.1, pp. 1–6.

Rizzolatti, G. and G. Luppino (2001). "The cortical motor system". In: *Neuron* 31.6, pp. 889–901.

Roberts, J. A. et al. (2017). "Consistency-based thresholding of the human connectome". In: *NeuroImage* 145, pp. 118–129.

Rodrigues, P. L. C. et al. (2017). "Dimensionality Reduction for BCI classification using Riemannian geometry". In: *BCI 2017-7th International Brain-Computer Interface Conference*.

Romero, D., V. N. Ioannidis, and G. B. Giannakis (2017). "Kernel-based reconstruction of space-time functions on dynamic graphs". In: *IEEE Journal of Selected Topics in Signal Processing* 11.6, pp. 856–869.

Rosenberg, J. et al. (1989). "The Fourier approach to the identification of functional coupling between neuronal spike trains". In: *Progress in biophysics and molecular biology* 53.1, pp. 1–31.

Roth, M. et al. (1996). "Possible involvement of primary motor cortex in mentally simulated movement: a functional magnetic resonance imaging study." In: *Neuroreport* 7.7, pp. 1280–1284.

Rubinov, M and O Sporns (2011). "Weight-conserving characterization of complex functional brain networks". In: *Neuroimage* 56.4, pp. 2068–2079.

Rubinov, M. and O. Sporns (2010). "Complex network measures of brain connectivity: uses and interpretations". In: *Neuroimage* 52.3, pp. 1059–1069.

Sabaté, M., B. González, and M. Rodríguez (2004). "Brain lateralization of motor imagery: motor planning asymmetry as a cause of movement lateralization". In: *Neuropsychologia* 42.8, pp. 1041–1049.

Saha, S. and M. Baumert (2020). "Intra-and inter-subject variability in EEG-based sensorimotor brain computer interface: a review". In: *Frontiers in computational neuroscience* 13, p. 87.

Sakkalis, V. and M. Zervakis (2009). *Linear and nonlinear synchronization analysis and visualization during altered states of consciousness*. INTECH Open Access Publisher.

Salvador, R. et al. (2005). "Neurophysiological architecture of functional magnetic resonance images of human brain". In: *Cerebral cortex* 15.9, pp. 1332–1342.

Sanei, S. and J. A. Chambers (2013). *EEG signal processing*. John Wiley & Sons.

Santoso, S et al. (1997). "Time-series analysis of nonstationary plasma fluctuations using wavelet transforms". In: *Review of scientific instruments* 68.1, pp. 898–901.

Schirrmeister, R. T. et al. (2017). "Deep learning with convolutional neural networks for EEG decoding and visualization". In: *Human brain mapping* 38.11, pp. 5391–5420.

Schreiber, T. (2000). "Measuring information transfer". In: *Physical review letters* 85.2, p. 461.

Seger, C. A. and E. K. Miller (2010). "Category learning in the brain". In: *Annual review of neuroscience* 33, p. 203.

Seger, C. A. (1994). "Implicit learning." In: *Psychological bulletin* 115.2, p. 163.

Serrano, M. Á., M. Boguná, and A. Vespignani (2009). "Extracting the multiscale backbone of complex weighted networks". In: *Proceedings of the national academy of sciences* 106.16, pp. 6483–6488.

Sharma, N., J.-C. Baron, and J. B. Rowe (2009). "Motor imagery after stroke: relating outcome to motor network connectivity". In: *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society* 66.5, pp. 604–616.

Sharma, N., V. M. Pomeroy, and J.-C. Baron (2006). "Motor imagery: a backdoor to the motor system after stroke?" In: *Stroke* 37.7, pp. 1941–1952.

Sheng, J. et al. (2020). "Characteristics and variability of functional brain networks". In: *Neuroscience Letters* 729, p. 134954.

Sherbondy, A. J., M. C. Rowe, and D. C. Alexander (2010). "MicroTrack: an algorithm for concurrent projectome and microstructure estimation". In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, pp. 183–190.

Shine, J. M. et al. (2015). "Estimation of dynamic functional connectivity using Multiplication of Temporal Derivatives". In: *NeuroImage* 122, pp. 399–407.

Sitaram, R. et al. (2017). "Closed-loop brain training: the science of neurofeedback". In: *Nature Reviews Neuroscience* 18.2, pp. 86–100.

Skidmore, F et al. (2011). "Connectivity brain networks based on wavelet correlation analysis in Parkinson fMRI data". In: *Neuroscience letters* 499.1, pp. 47–51.

Slepian, D. (1978). "Prolate spheroidal wave functions, Fourier analysis, and uncertainty—V: The discrete case". In: *Bell System Technical Journal* 57.5, pp. 1371–1430.

Smith, R. E. et al. (2013). "SIFT: Spherical-deconvolution informed filtering of tractograms". In: *Neuroimage* 67, pp. 298–312.

Solodkin, A. et al. (2004). "Fine modulation in network activation during motor execution and motor imagery". In: *Cerebral cortex* 14.11, pp. 1246–1255.

Sporns, O. and R. F. Betzel (2016). "Modular brain networks". In: *Annual review of psychology* 67, p. 613.

Sporns, O., R. Kötter, and K. J. Friston (2004). "Motifs in brain networks". In: *PLoS biology* 2.11, e369.

Stam, C. J. (2014). "Modern network science of neurological disorders". In: *Nature Reviews Neuroscience* 15.10, pp. 683–695.

Stam, C. J. and J. C. Reijneveld (2007). "Graph theoretical analysis of complex networks in the brain". In: *Nonlinear biomedical physics* 1.1, pp. 1–19.

Stam, C. J. and B. Van Dijk (2002). "Synchronization likelihood: an unbiased measure of generalized synchronization in multivariate data sets". In: *Physica D: Nonlinear Phenomena* 163.3-4, pp. 236–251.

Stanley, M. L. et al. (2015). "Changes in brain network efficiency and working memory performance in aging". In: *PLoS One* 10.4, e0123950.

Stefano Filho, C. A., R. Attux, and G. Castellano (2018). "Can graph metrics be used for EEG-BCIs based on hand motor imagery?" In: *Biomedical Signal Processing and Control* 40, pp. 359–365.

Stephan, K. et al. (1995). "Functional anatomy of the mental representation of upper extremity movements in healthy subjects". In: *Journal of neurophysiology* 73.1, pp. 373–386.

Stinear, C. M., M. K. Fleming, and W. D. Byblow (2006). "Lateralization of unimanual and bimanual motor imagery". In: *Brain research* 1095.1, pp. 139–147.

Stiso, J. et al. (2020). "Learning in brain-computer interface control evidenced by joint decomposition of brain and behavior". In: *Journal of neural engineering* 17.4, p. 046018.

Tang, J. et al. (2010). "Small-world behavior in time-varying graphs". In: *Physical Review E* 81.5, p. 055101.

Tangermann, M. et al. (2012). "Review of the BCI competition IV". In: *Frontiers in neuroscience*, p. 55.

Taubert, M. et al. (2011). "Long-term effects of motor training on resting-state networks and underlying brain structure". In: *Neuroimage* 57.4, pp. 1492–1498.

Telesford, Q. K. et al. (2011). "The ubiquity of small-world networks". In: *Brain connectivity* 1.5, pp. 367–375.

Termenon, M. et al. (2016). "The "hub disruption index," a reliable index sensitive to the brain networks reorganization. a study of the contralesional hemisphere in stroke". In: *Frontiers in computational neuroscience*, p. 84.

Tewarie, P. et al. (2015). "The minimum spanning tree: an unbiased method for brain network analysis". In: *Neuroimage* 104, pp. 177–188.

Theiler, J. and P. E. Rapp (1996). "Re-examination of the evidence for low-dimensional, nonlinear structure in the human electroencephalogram". In: *Electroencephalography and clinical Neurophysiology* 98.3, pp. 213–222.

Thompson, M. C. (2019). "Critiquing the concept of BCI illiteracy". In: *Science and engineering ethics* 25.4, pp. 1217–1233.

Thomson, D. J. (1982). "Spectrum estimation and harmonic analysis". In: *Proceedings of the IEEE* 70.9, pp. 1055–1096.

Tumminello, M. et al. (2005). "A tool for filtering information in complex systems". In: *Proceedings of the National Academy of Sciences* 102.30, pp. 10421–10426.

Uribe, L. F. S. et al. (2019). "A correntropy-based classifier for motor imagery brain-computer interfaces". In: *Biomedical Physics & Engineering Express* 5.6, p. 065026.

Valencia, M. et al. (2008). "Dynamic small-world behavior in functional brain networks unveiled by an event-related networks approach". In: *Physical Review E* 77.5, p. 050905.

Valencia, M. et al. (2009). "Complex modular structure of large-scale brain networks". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 19.2, p. 023119.

Van Den Heuvel, M. P. and O. Sporns (2011). "Rich-club organization of the human connectome". In: *Journal of Neuroscience* 31.44, pp. 15775–15786.

Van Den Heuvel, M. P. et al. (2013). "Abnormal rich club organization and functional brain dynamics in schizophrenia". In: *JAMA psychiatry* 70.8, pp. 783–792.

Van Wijk, B. C., C. J. Stam, and A. Daffertshofer (2010). "Comparing brain networks of different size and connectivity density using graph theory". In: *PloS one* 5.10, e13701.

Vanderah, T. and D. Gould (2020). *Nolte's The Human Brain E-Book: An Introduction to its Functional Anatomy*. Elsevier Health Sciences.

Vansteensel, M. J. et al. (2016). "Fully implanted brain–computer interface in a locked-in patient with ALS". In: *New England Journal of Medicine* 375.21, pp. 2060–2066.

Vespignani, A. (2018). *Twenty years of network science*.

Vicente, R. et al. (2008). "Dynamical relaying can yield zero time lag neuronal synchrony despite long conduction delays". In: *Proceedings of the National Academy of Sciences* 105.44, pp. 17157–17162.

Vico Fallani, F. de et al. (2012). "Redundancy in functional brain connectivity from EEG recordings". In: *International Journal of Bifurcation and Chaos* 22.07, p. 1250158.

Vidal, J. J. (1973). "Toward direct brain-computer communication". In.

Vinck, M. et al. (2011). "An improved index of phase-synchronization for electrophysiological data in the presence of volume-conduction, noise and sample-size bias". In: *Neuroimage* 55.4, pp. 1548–1565.

Walz, J. M. et al. (2013). "Simultaneous EEG-fMRI reveals temporal evolution of coupling between supramodal cortical attention networks and the brainstem". In: *Journal of Neuroscience* 33.49, pp. 19212–19222.

Wander, J. D. et al. (2013). "Distributed cortical adaptation during learning of a brain–computer interface task". In: *Proceedings of the National Academy of Sciences* 110.26, pp. 10818–10823.

Wang, L. et al. (2009). "Altered small-world brain functional networks in children with attention-deficit/hyperactivity disorder". In: *Human brain mapping* 30.2, pp. 638–649.

Watts, D. J. and S. H. Strogatz (1998). "Collective dynamics of 'small-world'networks". In: *nature* 393.6684, pp. 440–442.

Westlake, K. P. and S. S. Nagarajan (2011). "Functional connectivity in relation to motor performance and recovery after stroke". In: *Frontiers in systems neuroscience* 5, p. 8.

Wibral, M. et al. (2013). "Measuring information-transfer delays". In: *PloS one* 8.2, e55809.

Wig, G. S. (2017). "Segregated systems of human brain networks". In: *Trends in cognitive sciences* 21.12, pp. 981–996.

Wilson, J. A. et al. (2009). "Using an EEG-based brain-computer interface for virtual cursor movement with BCI2000". In: *JoVE (Journal of Visualized Experiments)* 29, e1319.

Winterhalder, M et al. (2004). "Nonlinear dynamics in EEG from epileptic patients: Is it possible to predict seizures?" In: *AIP Conference Proceedings*. Vol. 742. 1. American Institute of Physics, pp. 216–221.

Winterhalder, M., B. Schelter, and J. Timmer (2006). *Handbook of Time Series Analysis: Recent Theoretical Developments and Applications*. John Wiley & Sons.

Wolpaw, J. R. and D. J. McFarland (2004). "Control of a two-dimensional movement signal by a noninvasive brain-computer interface in humans". In: *Proceedings of the national academy of sciences* 101.51, pp. 17849–17854.

Wolpaw, J. R. et al. (2002). "Brain–computer interfaces for communication and control". In: *Clinical neurophysiology* 113.6, pp. 767–791.

Wu, Z. et al. (2020). "A comprehensive survey on graph neural networks". In: *IEEE transactions on neural networks and learning systems* 32.1, pp. 4–24.

Xu, J., M. Grosse-Wentrup, and V. Jayaram (2020). "Tangent space spatial filters for interpretable and efficient Riemannian classification". In: *Journal of Neural Engineering* 17.2, p. 026043.

Xu, L et al. (2014). "Motor execution and motor imagery: a comparison of functional connectivity patterns based on graph theory". In: *Neuroscience* 261, pp. 184–194.

Yger, F., M. Berar, and F. Lotte (2016). "Riemannian approaches in brain-computer interfaces: a review". In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 25.10, pp. 1753–1762.

Yi, W. et al. (2014). "Evaluation of EEG oscillatory patterns and cognitive process during simple and compound limb motor imagery". In: *PloS one* 9.12, e114853.

Zalesky, A. et al. (2014). "Time-resolved resting-state brain networks". In: *Proceedings of the National Academy of Sciences* 111.28, pp. 10341–10346.

Zhang, H. et al. (2012). "Improved recognition of error related potentials through the use of brain connectivity features". In: *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. Ieee, pp. 6740–6743.

Zhang, R. et al. (2019). "Using brain network features to increase the classification accuracy of MI-BCI inefficiency subject". In: *Ieee Access* 7, pp. 74490–74499.

Zhou, B. et al. (2016). "A fully automated trial selection method for optimization of motor imagery based brain-computer interface". In: *PloS one* 11.9, e0162657.