POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

Executive Summary of the Thesis

# EEG Brain Functional Connectivity and Spectral Analysis for the Classification of Attention Deficit Hyperactivity Disorder Patients

Laurea Magistrale in Biomedical Engineering - Ingegneria Biomedica

**Authors:** Ludovica Gaspari, Eleonora Iascone

**Advisor:** Prof. Anna Maria Bianchi

**Co-advisors:** Alessandra Calcagno, Ing. , Stefania Coelli, Ph.D.

**Academic year:** 2020-2021

## 1. Introduction

Attention Deficit Hyperactivity Disorder (ADHD) is one of the most common neurodevelopmental disorders, affecting subject's normal cognitive and behavioral functioning [5]. It is highly predominant in childhood, affecting around 5% of children worldwide, however, if not treated, ADHD could persist into adulthood. Individuals with ADHD may have problems focusing on a single task for an extended period, sitting still for long or controlling impulsive behaviors. According to the kinds of symptoms, three different subtypes, namely Inattentive, Hyperactive/Impulsive and Combined types, can be identified.

Functional cognitive tests and observations of the behavior represent useful tools for the diagnostic process; nevertheless, there are several difficulties affecting the diagnosis, such as the subjectivity of ADHD symptoms and diagnostic classification scales filled in by parents or teachers. Indeed, over the last years, Electroencephalography (EEG) has been widely used in the diagnosis of ADHD, since this disorder is largely related to neurophysiological impairments.

This work of thesis aims at analyzing the EEG signals of children affected by ADHD, acquired during an attentional task, employing both Spectral and Functional Connectivity analyses, in order to find EEG-based biomarkers to be used as diagnostic tools for the disorder. The results obtained by these two types of analyses were then used as features for a Machine Learning-based classification, aimed at distinguishing the ADHD children from healthy ones.

## 2. Electroencephalography applied to ADHD

EEG is a technique to measure the electrical potentials reflecting the human brain electrical activity, characterized by an excellent temporal resolution (ms) and a poor spatial resolution. Thanks to its non-invasiveness, it is one of the most used tools to diagnose neurological diseases.

EEG waveforms are a mixture of several different frequency bands, generally subdivided in Delta (1-4Hz), Theta (4-8 Hz), Alpha (8-13

Hz), Beta 1 (13-22 Hz), Beta 2 (22-30 Hz) and Gamma (>30 Hz) waves. A Spectral Analysis could be conducted on the mentioned frequency bands, computing the *EEG Power Spectral Density (PSD)*, which reflects the distribution of the energy in the different frequency ranges. In order to study how functionally specialized brain areas (Frontal, Temporal, Parietal and Occipital) interact among each other, Brain Connectivity could be analyzed, to understand abnormal brain functions underlying various neurological disorders, by mapping the brain as a complex network [4].

According to the findings in literature, ADHD children were found to be characterized by elevated low frequency activity (Theta, located mainly in Frontal and Central regions) and decreased high frequencies (Alpha and Beta) with respect to age-matched control subjects during resting state. Nonetheless, the Spectral Analysis alone has not reached a definitive conclusion in finding an effective biomarker capable of discriminating the pathology from the healthy state, therefore, this type of analysis has often been coupled with Connectivity Analysis. Previous works applying Functional Connectivity in ADHD have underlined hypoactivity in brain functional networks involved in executive control and attentional processing, such as Fronto-Parietal and Ventral-Attentional networks; in addition, a reduced inhibition of the Default Mode Network, which is characterized by a higher level of activation during resting state, has been noticed during attentional tasks in ADHD.
Few studies have carried out a discrimination of the ADHD children from the control ones using Machine Learning techniques and most of them have focused their attention on the resting state condition, employing fMRI data or EEG Event-Related Potentials (ERP).
In conclusion, most of the studies have analyzed the resting-state condition, however, it is important to evaluate a neural network's ability to change from a passive to an active condition (cognitive task), reason why a temporal dynamic approach was adopted in the present work.

# 3. Materials and Methods

## 3.1. Study Design

Twenty-eight children (age 7 to 17 years) were enrolled in the study: 16 subjects (mean age 12.3, 3 females) affected by ADHD and 12 healthy controls (mean age 11.4, 5 females). As a first step of the experimental protocol, the EEG signals were recorded in an eyes-open resting state of two minutes, during which the participant looked at a screen. Afterwards, the subject performed a computerized Continuous Performance Test (CPT), during which the participants were instructed to press the left mouse button with their right index finger as soon as a letter different from X was presented (GO or target stimulus) and not to press the button when the X appeared on the screen (NoGO or No-target stimulus). The participants' behavioral performance was assessed using task-related scores, such as Commission and Omission Errors (CE and OE), Hit Response Time (HRT), HRT Standard Deviation (HRT SD) and CPT Variability.

## 3.2. Data Acquisition and Preprocessing

A continuous EEG was recorded using 62 surface electrodes, positioned according to the 10/20 International System, plus two electrodes used for collecting the electrooculogram (EOG) signals. FCz and AFz were the recording reference and the ground electrodes, respectively. Furthermore, after an appropriate low-pass hardware filter, EEG recordings were digitalized with a sampling rate equal to 500 Hz. A band-pass zero-phase FIR filter between 1 and 70 Hz was applied to the data together with a Notch filter at 50 Hz. After that, the signals were downsampled at 250 Hz. In order to remove ocular and muscular artefacts, the *Independent Component Analysis (ICA)* was performed on the signals. Subsequently, the two EOG electrodes and noisy channels (Tp9, Tp10 and Iz) were removed. As a last pre-processing step, a re-referencing of the EEG traces was performed with a *Laplacian filter*, in order to reduce the volume conduction artifact, that is the spreading of multiple brain sources at the scalp [3].

### 3.3. Spectral Analysis

The Welch's periodogram was computed for each EEG channel, in baseline and task condition, after normalization by subtracting the mean and dividing by its standard deviation. In order to extract, for each frequency band of interest, the relative *Power Spectral Density (PSD)*, which represents the percentage of power in one band relative to the total power of the signal [6], the integral of the PSD was computed in the mentioned frequency ranges. To evaluate the evolution of the brain activity during the task, changes in power ($\Delta$P%) relative to the baseline were considered with a time resolution of 10 seconds. Analyses were then conducted considering the whole CPT task (14 minutes) divided into 1-minute or 10-second windows, depending on analysis needs, in order to maintain the temporal information.

### 3.4. Connectivity Analysis

The *Imaginary Part of the Coherency (ImCoh)* was chosen as Functional Connectivity measure to estimate the magnitude of information flow between EEG paired signals. Unlike the real part of the Coherency, the imaginary part is insensitive to artifactual interactions caused by volume conduction [3]. Complex coherency was firstly computed based on Welch's method, obtaining a value every 10 seconds. Then, the absolute value of ImCoh was extracted in Alpha, Beta 1 and Theta frequency bands. The *Surrogation Test with Phase Randomization method* was applied in order to define a statistical threshold for assessing a 95% of significance of the coupling between the time series. Adjacency matrices 59x59 (where 59 is the number of considered electrodes) were obtained applying the estimated threshold, resulting in weighted graphs with only significant connections. From each adjacency matrix (i.e. graph), computed every 10 seconds, graph-based indices (*Degree, Strength, Betweenness Centrality, Global* and *Local Efficiency, Shortest Path Length* and *Clustering Coefficient*) were derived. Except for the Global Efficiency, which is a measure of the overall network, a region-wise analysis was performed considering the main four cerebral areas (Frontal, Central, Temporal and Parieto-Occipital) for all the mentioned indices. A statistical analysis employing *Wilcoxon test* for independent samples was adopted to determine which of the graph-based indices would have been more useful to discriminate between the two groups.

### 3.5. Machine Learning

Exploiting the Spectral and Connectivity information obtained from the CPT, a model aimed at the discrimination of ADHD and control groups was built, employing the *Support Vector Machine (SVM)* as non-linear classifier. The observations were the 10-seconds segments belonging to CPT (2184 samples), while the $\Delta$P% and the variations of graph indices with respect to the baseline, for each band and area, were selected as features (93 explanatory variables). The binary target (0/1) corresponds to CONTROL and ADHD, respectively. The dataset was subdivided in three parts, maintaining for each set the original balance in the target class and choosing randomly the corresponding observations. The *Test* set was created leaving out 8% (175 samples) of the entire dataset, while the *Training* and *Validation* sets were formed by 75% and 25% respectively of the remaining data. Subsequently, the Standardization was performed on the samples to make the features normally distributed. Three Feature Selection methods, *Analysis of Variance (ANOVA), Maximum Relevance Minimum Redundancy (MRMR)* and *Principal Component Analysis (PCA)*, were tested and compared to extract the most relevant features, before the application of the SVM classifier.

## 4. Results

The analysis of Performance Indices revealed that ADHD scores were characterized by higher median values in all the measures considered and, except for OE, by a higher variability with respect to healthy controls. The Wilcoxon test for independent samples performed on these indices has not highlighted significant differences, except for the HRT SD. A greater HRT SD is believed to reflect inefficient information flow during the execution of a cognitive task [2], so, variable responses might be due to inefficient or disrupting information processing.

### 4.1. Spectral Analysis

Median Scalp Maps of the two groups in each frequency band were compared, resulting in interesting findings in Theta and Alpha bands.

In *Theta*, from an overall point of view, the control group shows a greater sparse activation, as it can be seen in *Figure 1*. In literature, an increase in Theta $\Delta$P% is mainly appreciated in the Fronto-Parietal area, as could be observed partly in the obtained results. In Sauseng et al. and Gevins et al. studies the increased Theta activity in Frontal region was associated with a higher level of cognitive task demand.

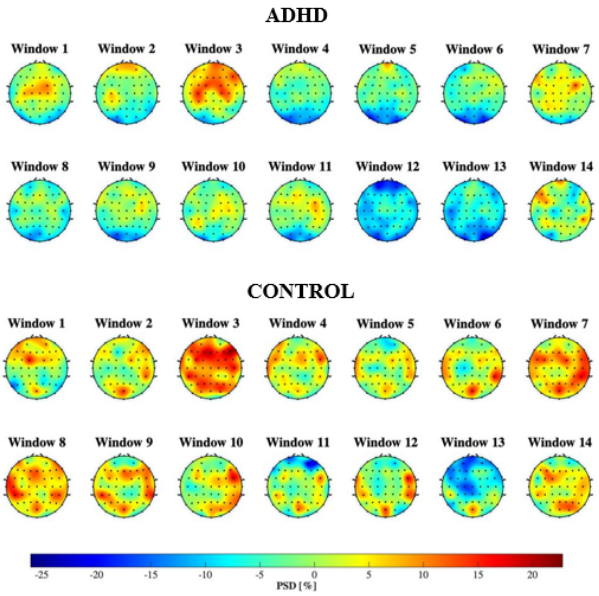



Figure 1: Theta median Topographical Scalp Maps, ADHD (top) and CONTROLS (bottom).

In *Alpha* band in the Parieto-Occipital zone a synchronization of these rhythms is present in both groups, but more accentuated in ADHDs. Generally, Alpha Parieto-Occipital rhythms are mitigated (desynchronization) in response to visual stimuli [1]; hence, their increase during the task, detected in the present work mainly in ADHDs, could be interpreted as difficulty for ADHD children in attending to and processing visual stimuli as efficiently as their healthy peers. In addition, ADHD group shows a greater desynchronization in both Central hemispheres with respect to Controls. From the literature, Alpha rhythms localized in the Central region (sensory-motor area), also called Mu-Rhythms [1], can be inhibited by the corresponding hand movement, while the muscle relaxation enhances them. Indeed, the higher desynchronization (the two blue spots visible in *Figure 2*) in ADHDs might be due to a greater strain of ADHD children to remain calm during the task, since hyperactivity is one of ADHD main symptoms. However, this hypothesis could not be strengthened, due to the lack of ADHD subtypes (inattentive and hyperactive).

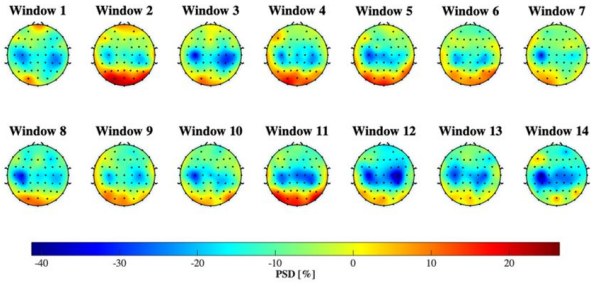



Figure 2: Alpha median Topographical Scalp Maps, ADHD

Because of the interesting behavior shown by ADHDs in Alpha on the motor cortex, the correlation between Alpha $\Delta$P% belonging to each 1-minute window and performance indices was computed: this was done in order to explore whether the supposed increased motor activity during cognitive task influenced the subject performance. Differently from what was expected, significant correlations ($p<0.05$) were found only for the Control group, revealing a negative correlation with CE and a positive correlation with HRT in all the windows. Hence, if Alpha activity increases in this area, it might signify that the controls tend to move less during the task, probably concentrating more and, consequently, committing less errors, spending also more time thinking about the response, not being impulsive.

### 4.2. Connectivity Analysis

From the graph analysis, a clear distinction between the two groups is visible during the execution of the CPT, as shown in *Figure 3*, while a similar behavior between them is present in the eyes-open resting state, with no tendency of forming clusters.

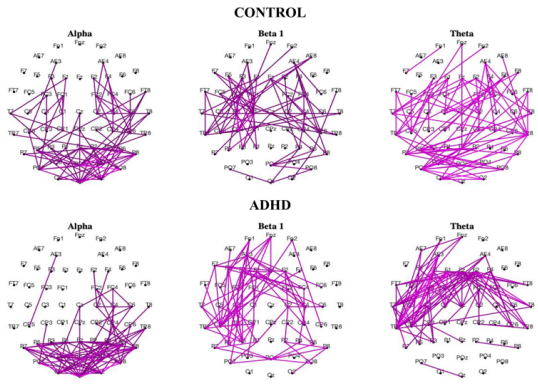



Figure 3: CPT graphs

Indeed, during CPT, in Controls there is an information flow among the Parieto-Occipital and Central brain regions in Alpha band, while ADHD patients are characterized by an information flow mostly limited to the only Parieto-Occipital area, indicating a higher network segregation. The dense connections visible in the Occipital lobe, for both groups, might be due to the visual task, since this lobe is responsible for receiving and integrating visual information.

By analyzing graph-based indices, *Temporal* area, as well as *Theta* band, were considered of particular interest. Indeed, ADHDs are characterized by higher values of Betweenness Centrality in Temporal area with respect to Controls and shorter paths between the Central and Temporal regions in Theta band. The major involvement of the Temporal area in ADHDs was hypothesized to be due to the adoption of verbally mediated strategies (silently reading letter to self or reminding self-instructions), employed to improve the performance. This is consistent with the role of Temporal area in verbal working memory. Furthermore, greater values of Global Efficiency were found in Theta band in the control group, suggesting the presence of an elevated integration of the brain during the task. The relevance of Theta band in cognitive tasks' performance was assessed in Martin Santiago et al. and Toth et al., hypothesizing that a higher integration may reflect how well the information flow is integrated before the response [2]. In addition, ADHDs have shown longer path lengths between Frontal and Parieto-Occipital areas with respect to Controls in Theta band, suggesting an inefficient information integration between these two areas, involved in attentional and visual processing respectively.

### 4.3. Machine Learning

As already mentioned, three methods of feature selection, ANOVA, MRMR and PCA, were chosen and compared. A positive outcome is that all the features selected by ANOVA method were also chosen by the MRMR; all of the attributes coincide with the graph indices resulted significant from the preliminary statistical test. Furthermore, most of the selected features were found in Theta band, coherently with the results of both Spectral and Connectivity analyses. In conclusion, the performances in terms of F1 score on the Test set are higher than 0.8 regardless of the feature selection method employed. The classifier reached better results in classifying the ADHD group with respect to the control one, probably due to the higher number of samples belonging to the ADHD class.

## 5. Conclusions

The present master thesis work presented an analysis of EEG activity of ADHD and control children during an attentional task, exploring Spectral characteristics and Brain Connectivity through a temporal dynamic approach, with the aim of discovering potential biomarkers to distinguish the pathology.

Contrasting with the literature and differently from what it was expected, no particularly interesting findings in the fluctuations of Beta 1 $\Delta$P% were discovered in the Frontal region. In literature, higher values of Frontal Beta power have been observed in the control group compared to the ADHD one. Since Frontal lobe is shown to be related to cognitive functions, a poor activity in this area could underline inability to focus on a task. However, a discriminating characteristic of functional connectivity found in Frontal region and Beta 1 band was the difference in Local Efficiency and Clustering Coefficient for the two groups: ADHDs are characterized by decreased values of these indices, underlying a low ability in exchanging information within the Frontal

region. In addition, ADHDs show longer path lengths between Frontal and Parieto-Occipital areas with respect to Controls, suggesting an inefficient information integration between them. Even if interesting results were not found in Beta 1 band in Frontal region through the spectral analysis, ANOVA and MRMR selected $\Delta$P% located in this band and area, together with some graph indices as features for the classification model.

The most relevant findings were related to Theta band and Temporal area, noticeable mainly through the Functional Connectivity and Graph Theory analyses. The major involvement of the Temporal area in Theta activity for ADHDs was hypothesized to be due to the adoption of verbally mediated strategies during the execution of the task. Among the graph-based indices, Betweenness Centrality, Global Efficiency and Shortest Path Length showed interesting results. Betweenness Centrality presented higher values in this area in ADHDs, which show also more efficient connections between the Central and Temporal regions. Regarding Global Efficiency, greater values were found in Theta band in the healthy patients, suggesting the presence of an elevated integration of the brain during the task; this is supported by findings in literature which associate higher values of this index in Theta to more stable results in the performance, deducted from the significant association of Global Efficiency with the HRT SD [2]. Indeed, a greater HRT SD is believed to reflect inefficient information flow during the execution of a cognitive task, so variable responses might be due to inefficient or disrupting information processing [2]. In support of this hypothesis, in the present work the HRT SD was the only performance index showing a significant difference between the two groups, with higher values found in the ADHD subjects.

Even though a significant biomarker has not been identified, the starting hypothesis, that a Spectral Analysis alone is not enough to reliably discriminate the disorder under study, was confirmed. Indeed, it was observed that functional brain connectivity is altered in children with ADHD, underlined by significantly different behavior of network indices compared to healthy subjects, especially in Theta band and Temporal area.

### 5.1. Limitations and Future Developments

The lack of performance data for each given stimulus, such as the participant's response and the corresponding reaction time, represented one of the main limitations associated to the current research. These parameters could have been used to understand how the brain activity change according to given or not given responses. Other aspects relative to the dataset that should be taken into account are the limited number of participants in the study, the large range of participants' age (7-17 years) and the lack of clinical identification of ADHD subtypes (inattentive, hyperactive/impulsive and combination types). A future study could be conducting an age-related analysis considering clusters of individuals with a more uniform age range and exploiting their clinical information in order to find differences in brain activity among the three ADHD subtypes.

In conclusion, the search for functional neural correlates of ADHD, and consequently for potential biomarkers of the disorder, is crucial in the pursuit of its prevention, early detection and more effective treatment. Combination of the ImCoh and graph theory methods, together with the dynamic temporal analysis, carried on in the present study, would be a very useful and novel approach for exploring underlying mechanisms of the brain and for diagnosing neurological disorders in the future.

## References

[1] *Quantitative EEG, Event-Related Potentials and Neurotherapy*. Elsevier, 2009.

[2] Keitaro Machida, Michael Murias, and Katherine A. Johnson. Electrophysiological Correlates of Response Time Variability During a Sustained Attention Task. *Frontiers in Human Neuroscience*, 13:363, October 2019.

[3] Guido Nolte, Ou Bai, Lewis Wheaton, Zoltan Mari, Sherry Vorbach, and Mark Hallett. Identifying true brain interaction from EEG data using the imaginary part of coherency. *Clinical Neurophysiology*, 115(10):2292–2307, October 2004.

[4] Mikail Rubinov and Olaf Sporns. Complex network

measures of brain connectivity: Uses and interpretations. *NeuroImage*, 52(3):1059–1069, September 2010.

[5] Kapil Sayal, Vibhore Prasad, David Daley, Tamsin Ford, and David Coghill. ADHD in children and young people: Prevalence, care pathways, and service provision. *The Lancet Psychiatry*, 5(2):175–186, February 2018.

[6] Jun-Sang Sunwoo, Kwang Su Cha, and Ki-Young Jung. Computational electroencephalography analysis for characterizing brain networks. *Annals of Clinical Neurophysiology*, 22(2):82–91, October 2020.

POLITECNICO DI MILANO
*School of Industrial and Information Engineering*
Master of Science in Biomedical Engineering


# EEG Brain Functional Connectivity And Spectral Analysis For The Classification of Attention Deficit Hyperactivity Disorder Patients

**Supervisor**: Prof. Anna Maria Bianchi
**Co-Supervisors**: Alessandra Calcagno, Ing.
Stefania Coelli, Ph.D.

**Authors**:
Ludovica Gaspari, 927456
Eleonora Iascone, 928485

Academic Year 2020-2021

# Abstract

Attention-Deficit/Hyperactivity Disorder (ADHD) is one of the most common neurodevelopmental disorders in childhood, affecting subject's normal cognitive and behavioral functioning. Since the clinical diagnosis may be affected by the subjectivity of ADHD symptoms and classification scales, Electroencephalography (EEG) has been widely used to support the ADHD diagnosis. This work of thesis aims at analyzing the EEG signals of children affected by ADHD, acquired during an attentional task, in order to find EEG-based biomarkers to be used as diagnostic tools for the disorder. Spectral features, Functional Connectivity, computed through the Imaginary part of the Coherency, and Graph Analysis are the methods employed in the present work: the study is conducted in the main frequency bands (Alpha, Beta 1 and Theta) and cerebral regions, taking into account the temporal dynamics of the entire task. The results obtained by these analyses were used as features for a Machine Learning-based classification, aimed at distinguishing the ADHD children from the healthy peers. The most relevant results from Spectral and Connectivity analyses were noticed in Theta band and Temporal area, as also confirmed by the feature selection methods employed in the Machine Learning. Despite the scarcity of the dataset and the variability in the patients' age, good performances have been reached in the Machine Learning-based classification, with a F1-score higher than 0.8. Even though a significant biomarker has not been identified, the starting hypothesis, that a Spectral Analysis alone is not enough to reliably discriminate the disorder under study, was confirmed.

**Keywords**: ADHD, EEG, Functional Connectivity, Theta, Temporal area, Imaginary part of the Coherency, graph indices

# Abstract in lingua italiana

Il Disturbo da Deficit di Attenzione/Iperattività (ADHD) è uno dei disturbi del neurosviluppo più comuni nell'infanzia, che colpisce il normale funzionamento cognitivo e comportamentale del soggetto. Dal momento che diversi fattori influenzano la diagnosi, ad esempio la soggettività dei sintomi e delle scale di classificazione, l'Elettroencefalografia (EEG) è uno dei metodi maggiormente impiegati nella diagnosi del disturbo. Questo lavoro di tesi ha come obiettivo l'analisi di segnali EEG di bambini affetti da ADHD, acquisiti durante un test di valutazione dell'attenzione sostenuta visiva, al fine di trovare dei biomarker da utilizzare come strumenti diagnostici per il disturbo. A tale scopo, tenendo conto della dinamica temporale dell'intero test, sono state analizzate le caratteristiche spettrali, la connettività funzionale, calcolata mediante la parte immaginaria della coerenza, e gli indici di grafo nelle principali bande di frequenza (Alfa, Beta 1 e Theta) e regioni cerebrali. I risultati ottenuti da queste analisi sono stati utilizzati come attributi volti a distinguere i bambini affetti da ADHD dai controlli sani, mediante un algoritmo di Machine Learning. Le maggiori differenze tra i due gruppi, ottenute dall'Analisi Spettrale e dalla Connettività Funzionale, sono state notate in banda Theta e nell'area temporale, come confermato anche dai metodi di selezione degli attributi impiegati nel Machine Learning. Malgrado i pochi dati a disposizione e la variabilità dell'età dei pazienti, è stata raggiunta una buona performance (F1 score maggiore di 0.8). Nonostante non sia stato identificato un biomarker significativo legato all'EEG, è stata confermata l'ipotesi che la sola analisi spettrale non è sufficiente a discriminare in maniera affidabile tale disturbo.

**Parole chiave**: ADHD, EEG, Connettività Funzionale, Theta, lobo Temporale, parte immaginaria della Coerenza, indici di grafo

# Contents

# Chapter 1

# Introduction

Attention-Deficit/Hyperactivity Disorder (ADHD) is one of the most common neurodevelopmental disorders in childhood, affecting around 5% of children worldwide [160]. Subjects affected by ADHD generally show an evident impairment in learning performance and in interacting with family members and peers; in addition, they often suffer of other psychiatric conditions, as anxiety and bipolar disorders [14]. ADHD is consistently more prevalent in males than females and highly predominant in children [161]; however, if not treated, ADHD could persist into adulthood, eventually representing a risk factor for drug addiction and criminality [160] [34] and probably implying difficulties with employment [160].

ADHD is characterized by symptoms of inattention, hyperactivity and impulsivity [33], whose occurrence substantially affects subjects' normal cognitive and behavioral functioning. Children with ADHD may have problems focusing on a single task for an extended period of time, sitting still for long or controlling impulsive behaviors, such as restlessness, fidgeting or excessive talking [23]. According to the types of symptoms, three different subtypes can be identified in the Diagnostic and Statistical Manual of Mental Disorders IV Text Revision (DSM IV-TR): Inattentive, Hyperactive/Impulsive and Combined types [73]. The combined type is when subjects show above-threshold levels of both symptoms categories.

Treatment of ADHD children, aimed at reducing ADHD symptoms [32], as medication with dopaminergic and noradrenergic activity, together with timely recognition of ADHD-type difficulties provide an opportunity to reduce long-term impairments [160]. Functional cognitive tests and observations of the behavior represent useful tools for the diagnostic process; however, there are several issues affecting the diagnosis, such as the difficulty of distinguishing attention

and behavioral symptoms due to other disorders or the subjectivity of ADHD symptoms and of diagnostic classification scales filled in by parents or teachers [73].

Indeed, over the last years, Electroencephalography (EEG) has been widely used in the diagnosis of ADHD, due to the fact that this disorder is largely related to neurophysiological impairments. In particular, most of the studies have been demonstrated that ADHD children show abnormal EEG activity compared to normal ones [130] [46].

This work of thesis aims at analyzing the EEG signals of children affected by ADHD, acquired during an attentional task, in order to find EEG-based biomarkers to be used as diagnostic tools for the disorder. The starting hypothesis was that a Spectral Analysis alone was not enough to reliably discriminate the disorder under study. However, a first exploration of EEG activity, in terms of Power Spectral Density in Alpha, Beta, Delta, Gamma and Theta frequencies, has been conducted, searching for evident differences between the two groups (ADHD and control subjects) along the whole task duration. Indeed, the attention task has been analyzed in a dynamic way, taking into account the temporal evolution of the brain activity. Subsequently, a Brain Connectivity Analysis was carried on, focusing the attention on the most used graph indices. The results obtained by these two types of exploring analyses were used as features for a Machine Learning-based classification, aimed at distinguishing the ADHD children from the healthy ones.

The thesis work is divided into the following chapters:

- **Chapter 1**: the aim of the thesis, together with a brief description of the disorder (ADHD), are introduced, explaining why EEG can be employed as a useful diagnostic tool.
- **Chapter 2**: theoretical background of the brain structure, EEG characteristics, brain connectivity and network theory are presented. The chapter also provides a review of the state of the art on the current use of spectral and connectivity analyses and machine learning techniques in ADHD diagnosis.
- **Chapter 3**: the methods and mathematical tools employed for EEG spectral and connectivity analyses are described in detail and a classification model for discriminating ADHD and control groups is also presented.
- **Chapter 4**: results of EEG spectral and connectivity analyses, together with

the performance reached by the different models built, are described and possible interpretations are given.

- **Chapter 5**: conclusions, limitations and possible future developments of the study are reported.

# Chapter 2

# Electroencefalography applied to ADHD

Before studying in deep what is Electroencephalography and how it could be applied for treating ADHD, an overview of the brain and its functional areas will be presented.

## 2.1 Brain Anatomy and Cerebral Cortex Functions

The **Central Nervous System (CNS)** is part of the nervous system and includes the *Brain* and the *Spinal Cord*. The brain is composed by billions of neurons which communicate among each other creating numerous connections and synapses [78]. As could be appreciated in *Figure 2.1*, three main parts of the brain could be identified [88]:

- The **Cerebrum** is composed by *Cerebral Cortex* (or *Gray Matter*), the outermost layer of the cerebrum containing neuronal cell bodies, and *White Matter*, mainly constituted of myelinated axons [88]. In addition, the Cerebrum is divided in left and right cerebral hemispheres.
- The **Cerebellum** is located in the lower part of the brain and its role is fundamental in the motor control [162].
- The **Brainstem** is the structure that links the Cerebrum to the Spinal Cord and Cerebellum [25].

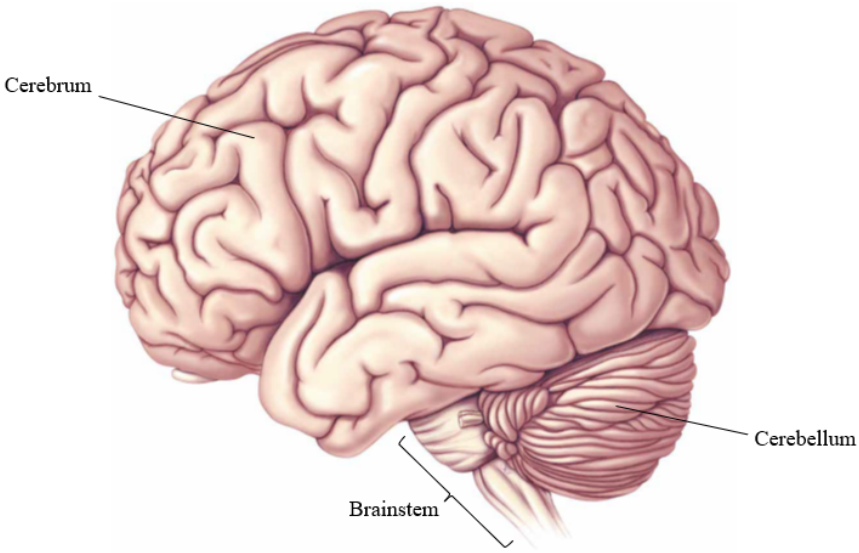



***Figure 2.1:*** *Lateral surface of the brain representing the three main parts of the brain: Cerebrum, Cerebellum and Brainstem. [27]*

The Cerebral Cortex is composed by four lobes *(Figure 2.2)*, each one characterized by a set of particular functions:

- The **Frontal Lobe** is located in the frontal part of the cerebral hemispheres and it is mainly involved in prospective memory (e.g. future planning) [137], speech and language production (Broca's area), voluntary movements control (Motor Cortex), personality, emotions and attention [88].
- The **Parietal Lobe** is positioned near the upper back portion of the skull, behind the Frontal lobe and superior to the Temporal lobe. Its role consists in receiving sensory information [30], sensorimotor planning, learning, language and spatial recognition [88]. Furthermore, this lobe has a key role in attention allocation [40].
- The **Temporal Lobe** is placed behind the Frontal lobe and it is generally associated with the processing of auditory information [139] and formation of visual memories, object recognition and in sequencing and organization of tasks.
- The **Occipital Lobe** is found in the posterior brain region and its main role is visual processing and interpretation [84].

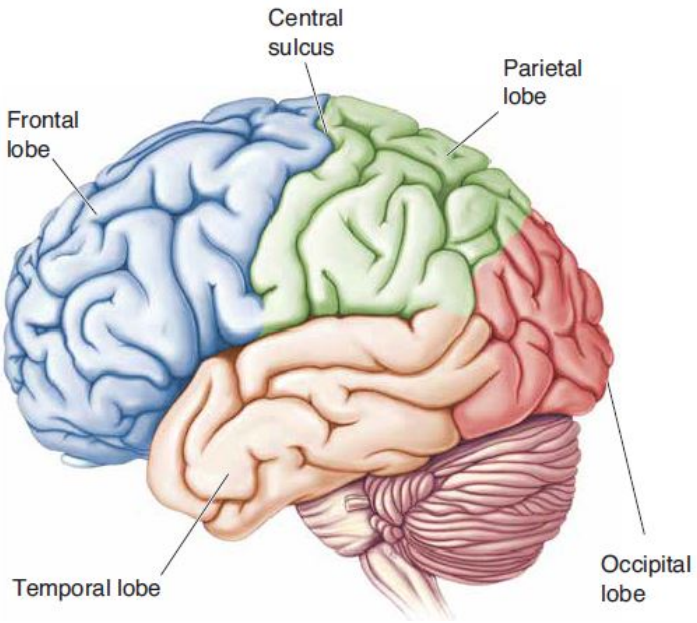



***Figure 2.2:*** *Representation of the main cerebral lobes: Frontal, Parietal, Temporal and Occipital. The Central sulcus separates the Frontal region from the Parietal one and the primary motor cortex from the primary somatosensory cortex [27].*

Several brain imaging studies have identified a distributed network of areas in the Frontal and Parietal lobes involved in sustained attention [151] [157]: the network formed by these two areas is known as *Frontal–Parietal Attention Network (FPAN)* and it has been studied in tasks involving sustained attention or at rest through functional connectivity, discovering a relation with attention performance [120] [166]. The FPAN supports executive control functions and ADHD affected subjects were found to exhibit deficits in this field [110]. For this reason, different studies have focused the attention on this network when dealing with ADHD.

Another network found to be impaired in ADHD is the *Ventral-Attention Network (VAN)*, which uses the Temporal-Parietal Junction (TPJ) and Ventral Frontal Cortex (VFC) to reorient attention to salient behaviorally relevant stimuli [66]. Furthermore, it was discovered by a recent meta-analysis that ADHD was characterized by an hypoactivation of this network, with respect to the control group, in tasks involving the modulation of attention in response to a target stimulus [56].

Finally, ADHD subjects were discovered to show a different behavior in the activation (and deactivation) of the resting state network (also referred as

task-negative network), called *Default Mode Network (DMN)*. DMN is active when the brain is at wakeful rest, while it is deactivated during some external goal-oriented tasks, such as visual attention or working memory tasks [45].

## 2.2 Electroencephalography

**Electroencephalography (EEG)** is a technique to measure the electrical potentials reflecting the human brain electrical activity [168] and it is one of the most useful tools to diagnose neurological diseases or to treat behavioral disturbances such as attention disorders [80].

EEG is a non-invasive technique, since the recordings are obtained through small electrodes placed in different locations on the scalp surface [168]. The signal acquired by each EEG sensor is a time course, whose amplitude represents the summed electrical activity, happening in a specific area of the cortex under the electrode surface [105], and a background activity coming from far neural sources. A typical EEG trace *(Figure 2.3)* is characterized by voltages, ranging from 0.5 to 100 µV in amplitude [180], in function of time [124].

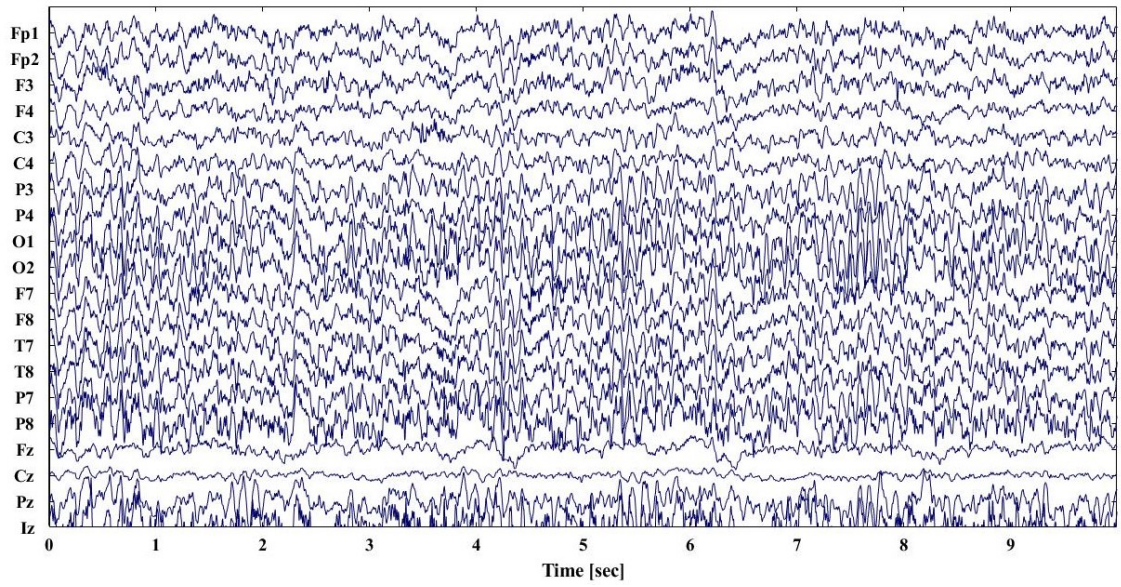



***Figure 2.3:*** *EEG trace of an ADHD patient during the attentional task: on the vertical axis the voltage in microvolt (µV) is displayed; on the x axis the time is shown in seconds (sec). The EEG signals were acquired through 64 sensors, but only the EEG traces of the first 19 electrodes (10-20 International System) are reported.*

The greatest advantage of EEG is the excellent temporal resolution (ms) [113], that allows to record complex patterns of neural activity, occurring within fractions of a second after the administration of a stimulus [180]; in addition, through EEG it is possible to capture the cognitive dynamics in the time frame in which cognition occurs [53].
On the other hand, it provides a poor spatial resolution, since the electrical currents recorded from the scalp do not always perfectly match with the specific underlying brain structure [113], but rather, one electrode can record a mixture of activities from many brain regions [53]. One of the main causes of this low spatial resolution is the volume conduction effect, i.e. the spreading of multiple brain sources at the scalp [37], that will be further explained in *Section 2.2.3*. However, through pre-processing methods, such as the Surface Laplacian [111], or by increasing the number of electrodes used for the acquisition [53], the spatial resolution can be enhanced.

Two different types of EEG analyses are employed in research [154]:

- *Qualitative (Clinical) EEG*: visual analysis of EEG patterns [59].
- *Quantitative EEG (QEEG)*: analysis that includes the study of frequency band powers, symmetry indices [59] and analysis of connectivity [112].

### 2.2.1 EEG Signal Generation

For a better understanding of EEG, in this section the mechanisms underlying the EEG signal generation will be explored.

As already mentioned, EEG is the recording of cerebral electrical potentials by sensors placed on the scalp [35]. Since the skull and the other layers constituting the human head provoke a great attenuation of the signals and most of the noise is produced either within the brain (internal noise) or over the scalp (external noise), the electrical activity generated by a single neuron is far too small to be detected by EEG [156]; therefore, only large groups of active neurons are able to generate a high enough potential detectable by scalp electrodes [156]. Indeed, the electrical activity acquired by EEG sensors is generated by groups of cerebral cortical neurons, located near the scalp where the electrodes are placed, with the same spatial orientation [124]. In particular, *Pyramidal Neurons*, represented in *Figure 2.4*, thanks to their unique perpendicular orientation of their long apical dendrites to the cortical surface, play a major role in the generation of the EEG signal [94].

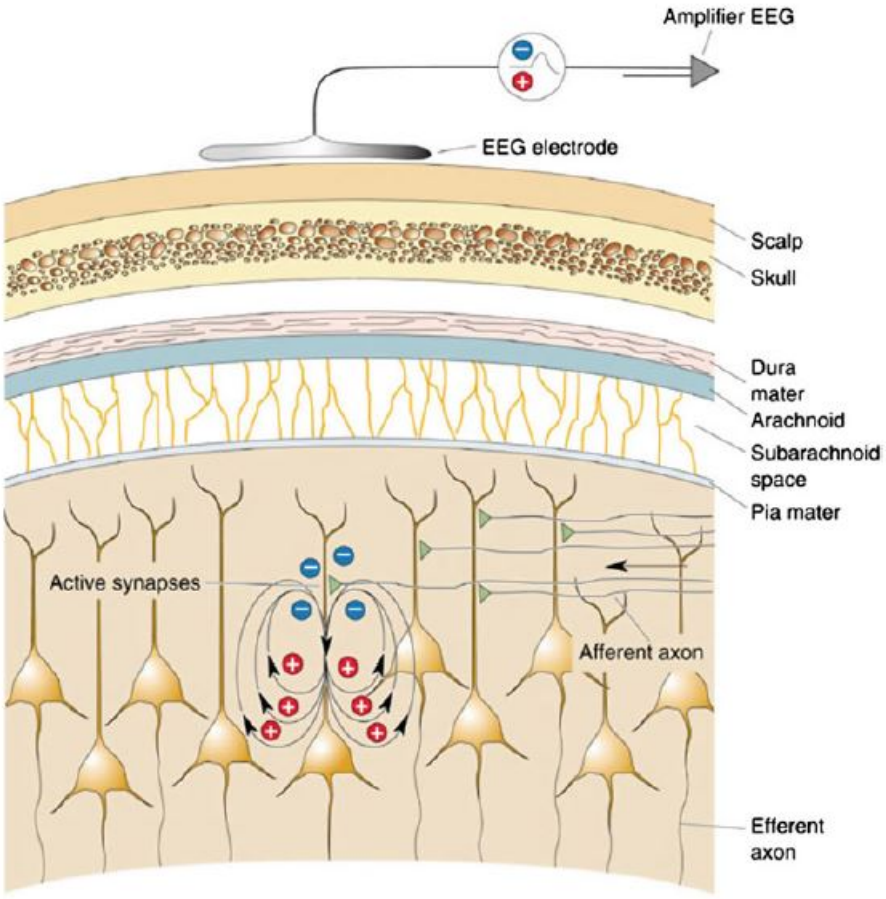



***Figure 2.4:*** *Generation of small electrical fields in Pyramidal Neurons [27].*

Neural electrical activity can be distinguished in *Action Potential (AP)* and *Postsynaptic Potential (PSP)*: when the PSP reaches the threshold conduction level in the postsynaptic neuron, the neuron starts firing, initiating the AP [168].
To be clear, EEG does not measure Action Potentials, but rather Postsynaptic Potentials [31]: indeed, AP are too short to sufficiently sum up, differently from PSP, which can last up to several tens of ms, being capable to produce potential changes that could be recorded extracellularly from the scalp [94]. To summarize, the principal contributors to the potentials measured by EEG are the summation of Excitatory and Inhibitory PSP of cortical Pyramidal Neurons [176].

### 2.2.2 EEG Waves and Spectral Analysis

EEG oscillations mirror the rhythmic fluctuations in the neurons excitability [176], with a frequency varying generally between 0.05 Hz and 500 Hz [41]. These oscillations could be described by three types of information [53]:

- *Frequency*: oscillations speed.
- *Power*: amount of energy contained in a frequency band.
- *Phase*: position along the sine wave at any given time point.

Frequency is one of the main criteria used to assess abnormalities in clinical EEGs [168]: indeed, EEG waveforms are a mixture of several different frequency bands, usually treated separately [113] through the signal decomposition, since they provide useful information about the subject state (e.g. wakefulness, sleep) [168] or the task demands [53].

Five major brain waves can be distinguished, characterized by different frequency bands and ordered from low to high frequencies [156] *(Figure 2.5)*:

- **Delta (0.5-4 Hz)**: these waves are mainly present during deep sleep and are generally associated with serious brain disorders and the waking state [168]. They are often found in infants and young children; furthermore, brain injuries, learning problems and severe ADHD are generally characterized by an increase in these rhythms [10].
- **Theta (4-8 Hz)**: Theta rhythms are involved in daydreaming, sleep and deep meditation [168]. These waves are seen to be prominent in ADHD, depression, hyperactivity, impulsivity and inattentiveness [10]. In general, in literature Theta band has been associated to sustained attention processes and working memory (central executive functions). Theta power increase is usually observed in the Fronto-Parietal area, due to the involvement of the Central Executive Network (CEN).
- **Alpha (8-13 Hz)**: Alpha waves appear mainly during eyes-closed condition or when the subject is in a relaxation state and are more visible in the Occipital lobe [168]. A reduced activity of Alpha power is caused by opening the eyes, anxiety, mental concentration or attention [156]. They have been also associated to the inability to focus [10]. In particular, recent studies interpreted these rhythms focusing on a distinction between Internally and Externally directed attention [118].
  As regard the External attention, when the patient is subjected to a visual

stimulation, Alpha power decreases over Occipital sites [118] [156]. Thus, it was hypothesized that the ability to suppress these rhythms is associated to the strength of attention to external stimuli [185]. The decrease is also observed during Sensorimotor tasks over Sensorimotor areas [147].
Contrarily, an increase in Alpha oscillations has been noticed during Internal tasks, as visual imagery or arithmetic operations [55]. However, the precise role of Alpha waves, together with their modulation provoked by the task, is still unclear [118].

- **Beta (13-30 Hz)**: Beta frequencies are usually varying symmetrically on both hemispheres in the Frontal Area and their arousal is more evident when the subject is actively engaged in mental tasks [168]. Beta is considered a brain waking rhythm, associated with active thinking, attention and focusing [156]. ADHD, depression and poor cognition could be characterized by a suppression of these oscillations [10]. Two different Beta rhythms could be identified: the first type, called *Rolandic Beta rhythms*, is located over the sensory-motor strip, while the *Frontal Beta rhythms* are found in the Frontal region [98]. A suppression (desynchronization) of Rolandic Beta rhythms could be observed during a GO/NOGO task, during the preparation of the movement, as well as during finger pressing; this desynchronization is then followed by rebound Beta synchronization [98]. Conversely to Rolandic rhythms, which are visible during motor-related tasks, Frontal Beta frequencies arise in cognitive tasks and are characterized by a synchronization, usually preceded by a desynchronization [98].

- **Gamma (>30 Hz)**: Gamma range comprehends frequencies above 30 Hz (mainly up to 45 Hz). These oscillations usually show very low amplitude values and their occurrence is rare, however, the detection of these frequencies could be useful to confirm the presence of certain brain diseases [156]. Cortical activity in this frequency range has been also linked to various cognitive processes, including attention [24] [150]: different studies [64] [74] [167] observed that an increase in Gamma power is related to arousal and attention.

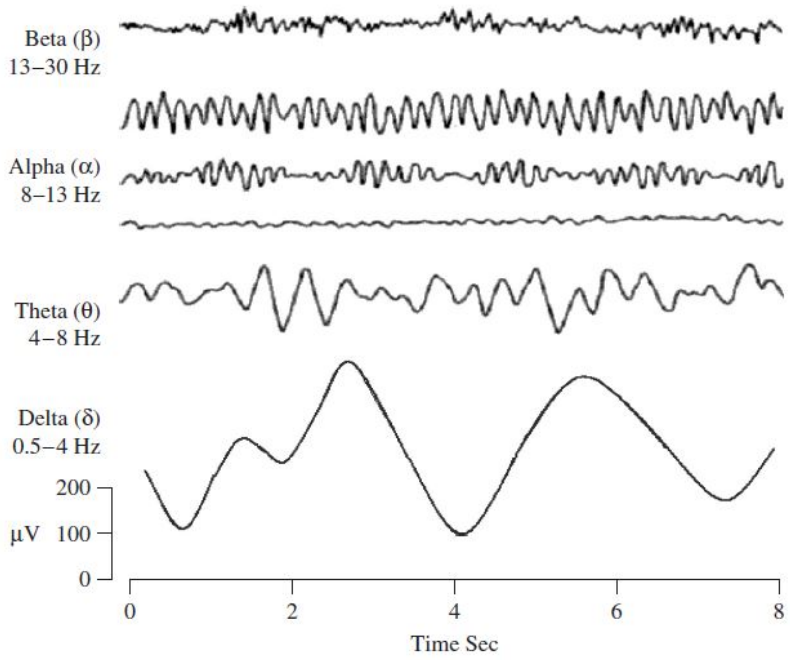



***Figure 2.5:*** *EEG Brain Waves: typical dominant brain frequencies, ordered from low to high frequencies [156].*

To perform an analysis of the mentioned frequency bands, **EEG Power Spectral Density (PSD)** is employed, since it reflects the distribution of the energy in the different frequency ranges. In particular, Spectral Analysis could be used to find associations between the oscillations within a certain band with a particular cognitive process or brain dysfunction [176]. The power associated to a frequency band can be expressed in absolute or relative terms; relative power represents the percentage of power in one band relative to all bands [92]. Calculating the relative EEG power with respect to the absolute one is advantageous since it helps to control for individual differences [154].

### 2.2.3 EEG Acquisition and Re-referencing

During an EEG acquisition protocol, it is important how to place the electrodes, since different lobes of the cerebral cortex are responsible for the processing of various types of activities [168]. The conventional electrode setting is the **International 10-20 System** [95], recommended by the International Federation of Societies for Electroencephalography and Clinical Neurophysiology [156]. The 10-20 System relies on the identification of anatomical landmarks, such as the Nasion, Inion and Preauricular points (electrodes connected to the left and right earlobes, A1 and A2), which are used to place the electrode at a fixed distance with respect to another electrode [82]. Specifically, the distances between neighbouring electrodes are either 10% or 20% of the total front-back (Nasion-Inion) or right-left (A1-A2) distance of the skull [168], as could be appreciated in *Figure 2.6*.
Extra electrodes could be added for the recording of Electrooculogram (EOG), Electrocardiogram (ECG) and Electromyogram (EMG) and exploited for the removal of artifacts [156].

The placement of the electrodes is named montage and at least two different types of montage could be identified [156]:

- *Bipolar (Differential) Montage*: each channel represents the difference between two neighbouring electrodes [168].
- *Referential Montage*: in this mode, generally one or two Reference Electrodes are considered, such as the vertex (Cz), linked-ears, one mastoid (TP10) or linked-mastoids, ipsilateral or controlateral ear or the tip of the nose [156]. Electrodes placed on the midline are often chosen, since they do not generate an amplification of the signal in one hemisphere with respect to the other [168].

Specifically, the referencing procedure consists in acquiring the values of the voltage recorded from each electrode with respect to a voltage value recorded in another place (reference electrode). So, even if there is a monopolar montage, it is acquired anyway a difference in potential [53].
It is needed to underline that the choice of a not relatively neutral reference might produce a topographic distortion [156]. Indeed, if the electric potential near the reference electrode is not neutral, the other electrode sites could contaminate the measurement, provoking a distortion of the temporal dynamic and power spectral analysis of the EEG recording [104]. However, if the data have been acquired with a given reference, a **re-referencing** is suggested to be applied with respect to another reference channel or channel combination [4].

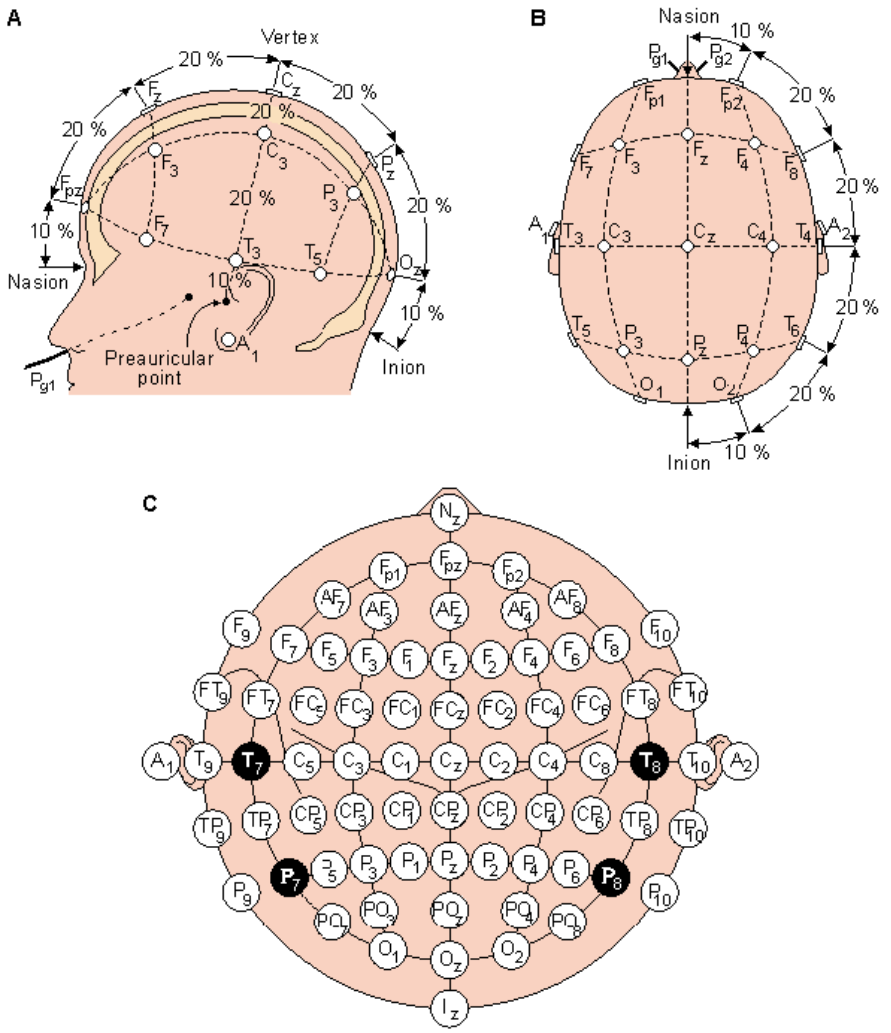



***Figure 2.6:*** *International 10-20 System: even numbers are referred to electrodes placed on the right hemisphere, whereas odd numbers to those placed on the left hemisphere. A) Sagittal view of the scalp; B) Top view of the scalp; C) 2D view of the electrode setup configuration [2].*

The **Common Average Referencing (CAR)** is generally applied to generate a more ideal electrode reference for EEG recordings. This technique is used to identify small signal sources in noisy recordings and has the advantage of being computationally simple [6]. CAR consists in subtracting the average potential of all the electrodes from each EEG channel, minimizing the contribution of uncorrelated noise sources [116]. However, it does not protect from the effect of

volume conduction, differently from the Surface Laplacian.
*Volume conduction* is an essential phenomenon to register the EEG activity and consists in the instantaneous propagation of the activity from the sources to the electrodes [37]. Recording sensors are not in direct contact with the cells generating the signal, thus this phenomenon is mainly due to the medium that separates the two [153]. However, it can cause the presence of an undesired high correlation between electrodes.

The **Surface Laplacian** spatial filter is a spatial band pass filter used to estimate the Current Source Density (CSD) at each electrode; after its application, high spatial frequency activity is preserved, while low spatial frequency activity is attenuated [53]. Its aim is to attenuate some of the volume conduction artifacts by transforming the source voltage recorded from the scalp into CSD, measured in $\mu V/cm^2$ [89]. Indeed, Laplacian Filter is capable to remove all the tangential components relative to the scalp, that result from the volume conduction, leaving only the perpedicular component. Since CSD is sensitive to local sources and insensitive to distant ones [13], the Laplacian filter is capable of minimizing volume conduction effects that may affect the results of the connectivity analysis [126]. Thus, Surface Laplacian is a recommended preprocessing step before performing a connectivity analysis. Moreover, it increases the SNR (Signal to Noise Ratio), allowing to obtain cleaner data, and improves topographical selectivity [53].

## 2.3 Brain Connectivity and Graph Theoretical Analysis

Information flow occurs at multiple scales in the brain, from the microscale among neurons to the macro scale among cortical regions. Magneto- and Electro-Encephalography (MEG, EEG) and functional Magnetic Resonance Imaging (fMRI) are some of the methods able to measure the information [187]. How information is processed, sent to, received by or shared between different brain regions is addressed in the research domain of Brain Connectivity [183].

**Brain Connectivity** and **Network Theory** have been widely used to study how functionally specialized brain areas interact among each other. Complex brain functions are the result of the connections of different brain regions in networks [152] and interactions could be described by means of different connectivity measures and indices based on Graph Theory [143]. Different types of Connectivity and the Graph Theory will be explained in detail in the following sections.

### 2.3.1 Brain Connectivity

Three types of Brain Connectivity can be defined:

- **Anatomical (or Structural) Connectivity**: physical connections between neighboring neurons or fiber tracks connecting neuron pools. The connections can be computed by estimating the brain's morphological (e.g. gray matter volume or cortical thickness) correlations in structural Magnetic Resonance Imaging (MRI) data, or by tracing the white matter projections linking cortical and sub-cortical regions through diffusion MRI (dMRI) data. [143].
- **Functional Connectivity**: statistical inter-dependency between two (or among more than two) brain regions, evaluated in terms of dependence between time series, acquired on different cerebral sites [44].
- **Effective Connectivity**: causal (or directed) influence of one brain area over another one [102].

The characteristics of brain's connectivity can be described by functional Segregation and Integration [39] [8]. *Segregation* identifies locally segregated processes, involving a cerebral area specialized in a specific task, as perceptual or motor processing [67]. Greater segregation means that many specialized different communities exist within a whole network [117]. *Integration* indicates the ability of specialized areas to interact among each other [152]. The higher the integration, the more the whole brain network works closely to process information [117].

The main steps followed to assess functional connectivity can be summarized in:

1. Computing the connectivity matrices through one of the measures explained in *Section 2.3.1.1*;
2. Applying a statistical threshold to eliminate weak and not significant functional connectivity values. The threshold can be retrieved, for example, through the Surrogate Method, a procedure useful to rule out dependency results that are not different to what would be expected by chance;
3. Characterizing the brain networks by using graph theoretical approaches, analyzed in *Section 2.3.2*.

There are two approaches to analyze synchronization among different areas, based on bivariate or multivariate measures. The bivariate or 'pair-wise' approach

considers the synchronization of each pair of EEG signals separately, while the second one takes into account more than two channels when estimating the interaction between pair of electrodes [85]. A bivariate measure has been chosen to assess the functional connectivity, since it was the most appropriate choice for the current application: indeed, pairwise methods allow to better manage a great number of nodes.

#### 2.3.1.1 Functional Connectivity Metrics

The literature provides a multitude of metrics to compute the Functional Connectivity [26] and each of these metrics has its own advantages and disadvantages. An overview of the distinctions between functional measures, considering a bivariate approach, is shown in *Figure 2.7*

.

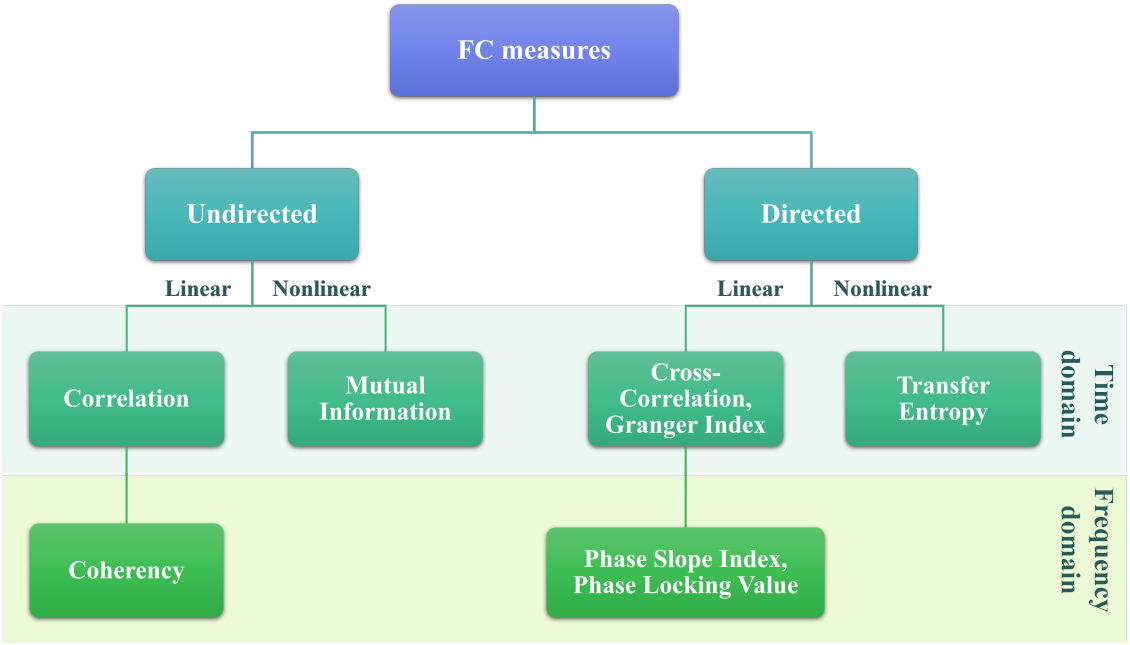



***Figure 2.7:*** *Functional Connectivity measures.*

A first distinction that could be detected is based on the presence, or absence, of the direction of the interaction. *Undirected Functional Connectivity* merely detects statistical dependence, without any information about the direction of influence, while the *Directed Functional Connectivity* indicates a causal influence, thus providing a direction (cause → effect). Within the first distinction, a second

subdivision can be made between Linear and Nonlinear measures. In general, these measures are relative to time or frequency domain and the calculation is either amplitude or phase-based.

*Linear Temporal Correlation* is a time domain measure and calculates the instantaneous linear relationship between two time series based on the amplitudes of the signals [36]. The *Cross-correlation* investigates the correlation between two time series that are shifted in time with respect to each other and this allows to retrieve the directionality of the interaction based on the time-lag between the activity [36], differently from the previous measure. The equivalent of the correlation in the frequency domain is the *Magnitude Squared Coherence*, which is the modulus of the Coherency [155]. The *Imaginary part of the Coherency* is another undirected functional connectivity measure, which has been used in the current work for the advantage of being less sensitive to volume conduction artifact.

Another way to analyze the interaction between two time series is to investigate how their phases are coupled and this can be done through phase synchronization measures, as *Phase Locking Value* or *Phase Lag Index*.

The *Mutual Information* and the *Transfer Entropy* are both based on probability functions and joint probability of the time series, and are considered nonlinear time domain measures [183]. While the Mutual Information assesses the undirected nonlinear relations between two signals, the Transfer Entropy indicates the directed nonlinear relations.

Another functional connectivity measure, exploiting the directionality in the time domain, is the *Granger Index*, which is based on the concept of Granger Causality. This concept expresses the ability to predict the future of one signal given the past information of another signal [191].

### 2.3.2 Graph Theoretical Analysis

Thanks to advanced neurophysiological and neuroimaging techniques, the brain can be mapped as a complex network at the macro-scale level. The concept of the brain as a complex network is fundamental to understand abnormal brain functions underlying various neurological disorders.

A **Network**, represented as a **Graph** G(N,K), is defined as a set of *nodes* N (i.e. electrodes) that represent brain regions interconnected by a set of *edges* K

(i.e. functional connectivity edges). The edges, also called *links*, are the entries of a NxN Adjacency Matrix, which represents the existence ($a_{ij}$) or strength ($w_{ij}$) of edges between each pair of nodes ($i$ and $j$) in the graph G. If this matrix is symmetric, the network is *Undirected*, otherwise it is *Directed*. In addition, a network can be classified as *Weighted* or *Unweighted (Binary)*, depending on whether the edges are assigned with different strengths (weights) [152]. The different kinds of graphs are displayed in *Figure 2.8*.

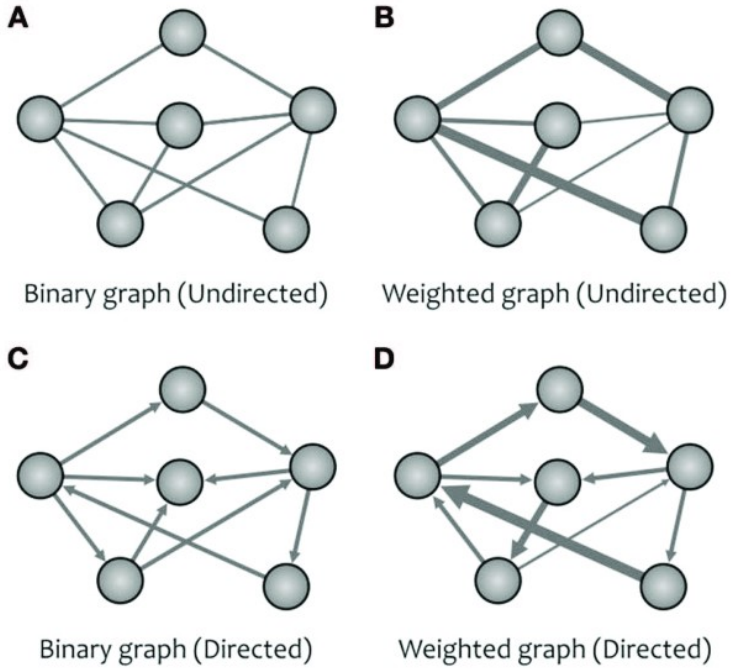



***Figure 2.8:*** *A network can be represented as a binary (A) or weighted (B) undirected graph. The directions of the causal effects among regions could be represented as arrows (C,D) [65].*

#### 2.3.2.1 Graph Indices

There are different graph theoretic indices, useful to quantify and characterize different aspects of a complex brain network, and they can be grouped into three main classes: Centrality, Integration and Segregation indices [73].
According to the different types of graph (binary/weighted, directed/undirected), there are different variants of the network indices; however, weighted and directed variants are typycally generalizations of binary undirected variants [152]. All the reported equations are relative to weighted and undirected graphs, which are those considered in the actual research.

*Centrality Indices*, as Degree, Strength or Betweenness Centrality, capture the relative importance of individual nodes within the network [165].

- **Betweenness Centrality** of a node $i$ is defined as the fraction of shortest paths between any pair of nodes that pass through the node $i$ [175] and quantifies how much a given node is in-between others.

$$BC_i = \frac{1}{N(N-1)} \sum_{h \neq i \neq j} \frac{\rho_{hj}(i)}{\rho_{hj}} \tag{2.1}$$

  $\rho_{hj}$ is the number of shortest paths between $h$ and $j$ and $\rho_{hj}(i)$ is the number of shortest paths between $h$ and $j$ passing through $i$.
  *Equation 2.1* is the same for both binary and weighted networks, provided that path lengths are computed on respective binary or weighted paths.
  Nodes with high Betweenness Centrality are highly connected to other nodes and work as hubs within the network [165].

- **Degree** of a node $i$ is computed as the number of links (equivalent to the number of neighbors) connected to that node and quantifies the node centrality [165].

$$D_i = \sum_{j \in N} a_{ij} \tag{2.2}$$

  $a_{ij}$ is the link connecting the node $i$ to the node $j$.

- **Strength**, the weighted variant of the degree, is defined as the sum of the weights of the links connected to the node [165].

$$S_i^w = \sum_{j \in N} w_{ij} \tag{2.3}$$

$w_{ij}$ is the weight of the link between the nodes $i$ and $j$.

*Integration indices*, such as Shortest Path Length and Global Efficiency, estimate the brain ability to rapidly combine specialized information of distributed brain regions and are based on the concept of *path* [60]. Paths in functional networks represent sequences of statistical associations between nodes and do not assume an information flow on anatomical links [152].

- **Shortest Path Length** is the shortest distance between node $i$ and node $j$.

$$SP_{ij}^{w} = \sum_{w_{uv} \in g_{i \leftrightarrow j}} w_{uv} \tag{2.4}$$

  $g_{i \leftrightarrow j}$ is the shortest path between the nodes $i$ and $j$.
  The *Characteristic Path Length* is defined as the average shortest path length in the network [152].

- **Global Efficiency** of the network G is the inverse of the average shortest path length and reflects how closely each node is connected [117].

$$GE^{w} = \frac{1}{n} \sum_{i \in N} E_i = \frac{\sum_{j \in N, j \neq i} (SP_{ij}^{w})^{-1}}{n-1} \tag{2.5}$$

  $E_i$ is the efficiency of node $i$ and ${SP_{ij}}^{w}$ is the weighted shortest path length between nodes $i$ and $j$.

*Segregation indices*, such as Clustering Coefficient and Local Efficiency, quantify the presence of densely interconnected groups of nodes, called clusters or modules [152].

- **Clustering Coefficient** of the node $i$ is the fraction of triangles (characterized by nodes as vertices and edges as sides) around that individual node and measures the robustness information transmission performance of the network [165].

$$C_i^{w} = \sum_{i \in N} \frac{2t_i^{w}}{D_i(D_i - 1)} \tag{2.6}$$

  $t_i$ is the number of triangles around a node $i$ and $D_i$ is the degree of node $i$. High clustering coefficient means that neighbours, to which the node is connected, are highly connected with each other [165].

- **Local Efficiency** of the node $i$ is the global efficiency computed on node neighborhood [152] and it measures the ability of each node to exchange information.

$$LE_i^w = GE^w(G_i) \tag{2.7}$$

GE($G_i$) is the global efficiency of $G_i$ which is the sub-graph composed of the neighbors of node $i$ [188].

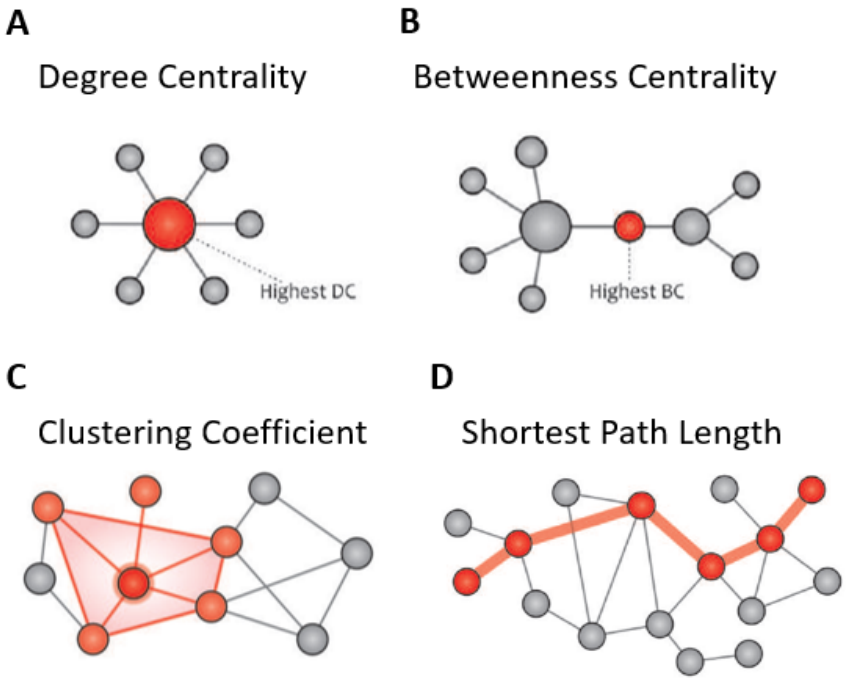



***Figure 2.9:*** *A) The Degree Centrality is the number of node neighbors; B) The Betweenness Centrality is a measure of the role of a node in acting as a bridge between different clusters; C) The Clustering Coefficient measures how much neighbors of a node are interconnected; D) The Characteristic Path Length measures the potential for the transmission of information, determined as the average Shortest Path Length across all pairs of nodes [65].*

## 2.4 State of the Art: EEG in ADHD

Electroencephalography-based measures are useful and non-invasive indices of Central Nervous System (CNS) function and are often used to assess both the resting-state and task-related dynamic features underlying neurophysiologic aspects of ADHD [154]. In particular, QEEG measures have been found to be useful as a clinical tool to diagnose ADHD, improving the accuracy and validity of the traditional behavioral cognitive tests employed for the diagnosis of the pathology [73].

As already mentioned in *Section 2.2*, a common method employed to obtain interesting information about this pathology is to study the power of the EEG signal in selected frequency bands, expressed in absolute or relative terms. Moreover, the ratio between the power in different frequency bands is another frequently used spectral measure in EEG analysis. Referring to ADHD, the *Theta/Beta Ratio (TBR)* is the most commonly employed one [73]: at rest, Theta oscillations are associated with drowsy and unfocused states, while Beta oscillations with cortical excitation [105].

According to the findings in literature, ADHD children were found to be characterized by elevated low frequency activity (absolute and relative Theta [158] [87] [52], located mainly in Frontal and Central regions, and increased posterior Delta [122] [52]) during resting state [154]. On the contrary, concentrating on high frequencies, most of the findings agree on a decrease in Alpha and Beta activities in ADHD with respect to age-matched control subjects [61] [43], although literature on the effect of ADHD in these bands is somehow inconsistent [73].

To date, an increased Theta band power and increased TBR have been considered the most robust and reproducible psychophysiological findings in ADHD [154]. However, elevated Theta power observed in ADHD might be a nonspecific biomarker of cortical dysfunction, since it is characteristic also of other disorders [73]. In addition, TBR reliability has not been replicated in several recent studies. In Arns et al. meta-analysis [17], a great heterogeneity among studies was observed, suggesting that the overall effects sizes may be an overestimation and that TBR did not have the reliability to support diagnosis of ADHD [154]. Furthermore, age and developmental course have been found to influence both Theta power and TBR more than the presence of ADHD [92].

Nonetheless, the Spectral Analysis alone has not reached a definitive conclusion in finding an effective biomarker capable of discriminating the pathology from the healthy state. As a consequence, this type of analysis has often been coupled with Connectivity Analysis, focusing the attention more on how efficiently the cortical areas communicate among each other.
The increasing investigation of brain connectivity analysis in ADHD studies, using the temporal resolution of EEG, allows to better understand transient changes in connectivity during cognitive tasks. Indeed, the connectivity retrieved by fMRI may not fully capture synchronization between faster brain oscillations, typically of fast-changing processes involving cognitive functions [127]. However, EEG connectivity metrics might be contaminated by volume conduction artifacts and most studies on ADHD have been characterized by this limitation, significantly affecting connectivity differences between the case (ADHD) and control group [138] [141]. For this reason, further investigation, by applying methods that make connectivity metrics uncontaminated, has been chosen for the current work.

Previous works applying Functional Connectivity in ADHD [189] [12] have underlined atypical patterns in brain functional networks involved in executive control, such as Fronto-Parietal network (also known as the Executive Control Circuit), attentional processing (Ventral-Attentional network) and resting states, as the Default-Mode network (DMN) [56] [182]. A reduced inhibition of the DMN, which is characterized by a higher level of activation during resting state and a lower one when subjects are engaged in cognitive tasks [38], has been underlined during attentional tasks in ADHD; at the same time, a hypoactivity in the task-positive networks (Fronto-Parietal and Ventral-Attentional networks) [164] has been noticed in ADHDs. One of the most common findings is that, in Frontal and Parietal areas, ADHD patients show elevated coherence in Theta band [73] and reduced coherence in Alpha band during the task [22].

Other studies [186] [146] have focused the attention on graph theoretical methods to understand which aspects of brain connectivity are disrupted in individuals with ADHD. A particularly studied graph-based index is the Global Efficiency, even if the results obtained and the interpretations disagree among the different studies. For example, Furlong et al.[69] suggested that increased Global Efficiency in ADHD individuals reflect disruption of information transfer across the brain, affecting cognitive functions; in addition, the study has found that this increase is associated with elevated ADHD symptoms severity [68]. On the other hand, decreased Global Efficiency has been detected in other studies [189] [28] and related to a loss of long-range communication due to structural abnormalities in ADHD.

To our knowledge, few studies have carried out a discrimination of the ADHD children from the control ones using Machine Learning techniques and most of them have focused their attention on the resting state condition [91] [107]. Tenev et al. study [179] is one of the few which considered the task (CPT and emotional continuous performance test), together with the eyes-open and eyes-closed conditions, to extract the relevant features for the classification. However, differently from the actual work, this study was provided with all the information relative to the different trials (e.g. performance indices, inter stimulus interval) and the starting dataset was larger. Furthermore, they did not take into account connectivity measures, but only the spectral ones. Ekhlasi et al. [62] exploited the brain connectivity for the classification, but, differently from the present study, they did not take into account the graph metrics as feature vectors, using instead the vectorized Connectivity Matrices as inputs of an Artificial Neural Network (ANN). Other studies [136] [134] [170] discriminated the two classes on the basis of Event Related Potentials (ERP) data and Smith et al. concluded that ERP information could be of considerable diagnostic utility and might be implemented in clinical practice as an additional diagnostic tool. Deep Learning techniques were not only employed for the classification of the pathology against physiology, but have been also utilized for the discrimination of the different ADHD subtypes, using both fMRI data [99] and EEG signals [11] during resting state.

In conclusion, most of the studies found in literature has analyzed ADHD mainly during Resting State condition, while only a few studies have focused their attention on Spectral and Connectivity analysis related to the task condition. However, it is important to evaluate a neural network's ability to change from a passive to an active condition (cognitive task). Studies based on the active condition have mainly analyzed the attentional task considering each segment of the EEG signal without taking into account the temporal dynamics [9] or investigating ERP [106] [127] evoked after the stimulus administration.

# Chapter 3

# Materials and Methods

## 3.1 Study Design

This section will focus on how the study was conducted, introducing a description of the study population and the acquisition protocol. Then, the EEG signals preprocessing will be also briefly described.

### 3.1.1 Participants

Twenty-eight subjects of age 7 to 17 years were enrolled in the study: 16 subjects (mean age 12.3, 3 females) affected by ADHD and 12 healthy controls (mean age 11.4, 5 females). Participants selection and data acquisition took place at the Unità Operativa di Neuropsichiatria dell'Infanzia e dell'Adolescenza dell'Azienda Socio-Sanitaria Territoriale Ospedali Santi Paolo e Carlo, in Milan.
General good health conditions, established by general and neurological exam, and the absence of cognitive disability or neurological pathology were required to all the participants as inclusion criteria. The absence of ADHD diagnosis based on DSM-IV criteria was needed for controls.
Regarding the exclusion criteria, the participants were not enrolled in the study if they presented a QI $< 80$, psychiatric and neurological disability, epilepsy, epileptiform abnormality in previous EEG exams and psychopharmacological drug assumption [13].

### 3.1.2 Protocol

During the experimental protocol, participants were seated in a room and asked to look at a computer screen, while EEG was performed.
As a first step, a baseline of two minutes with eyes-open, during which the participant looked at a meaningless figure, was acquired.

Afterwards, the subject performed a computerized **Continuous Performance Test (CPT)**, also called the Conners' "not-X" CPT, during which the letters of the English alphabet were presented in a randomized order on the screen. During the task, the participants were instructed to press the left mouse button with their right index finger as soon as a letter different from X was presented (GO or target stimulus) and not to press the button when the X appeared on the screen (NoGO or No-target stimulus).

The CPT is currently the most popular objective cognitive task used to support a clinical diagnosis of ADHD. This task generally allows to evaluate selective attention, sustained attention and impulsivity [169].
The participant's behavioral performance was assessed by the use of task-related scores, which can be associated to different aspects of attention, such as inattentiveness, impulsivity, sustained attention and vigilance [90]. A more detailed description of the Performance Indices will be provided in *Section 4.1.*

### 3.1.3 Data acquisition and Preprocessing

A continuous EEG was recorded using 62 surface electrodes, positioned according to the 10/20 International System, plus two electrodes used for collecting the Electrooculogram (EOG) signals. The positioning of the electrodes is displayed in *Figure 3.1.* In particular, FCz and AFz were the recording reference and the ground electrodes, respectively. Furthermore, after an appropriate low-pass hardware filter, EEG recordings were digitalized with a sampling rate equal to 500 Hz.

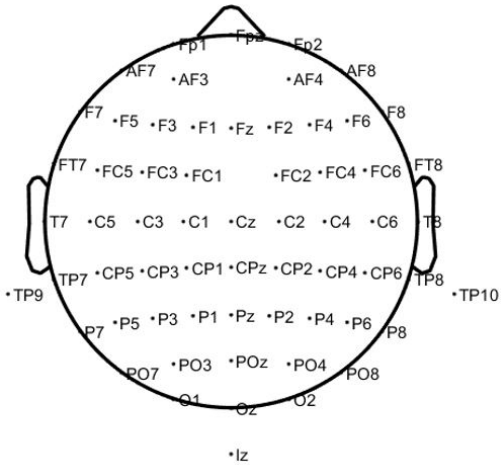



***Figure 3.1:*** *62 of 64 electrode locations, positioned according to the 10-20 International System.*

For the processing of the EEG signals, the MATLAB-based (The MathWorks, Inc., Natick, Massachusetts, United States) open source toolbox EEGLAB was used [58]. All the pre-processing steps are listed in *Figure 3.2*.

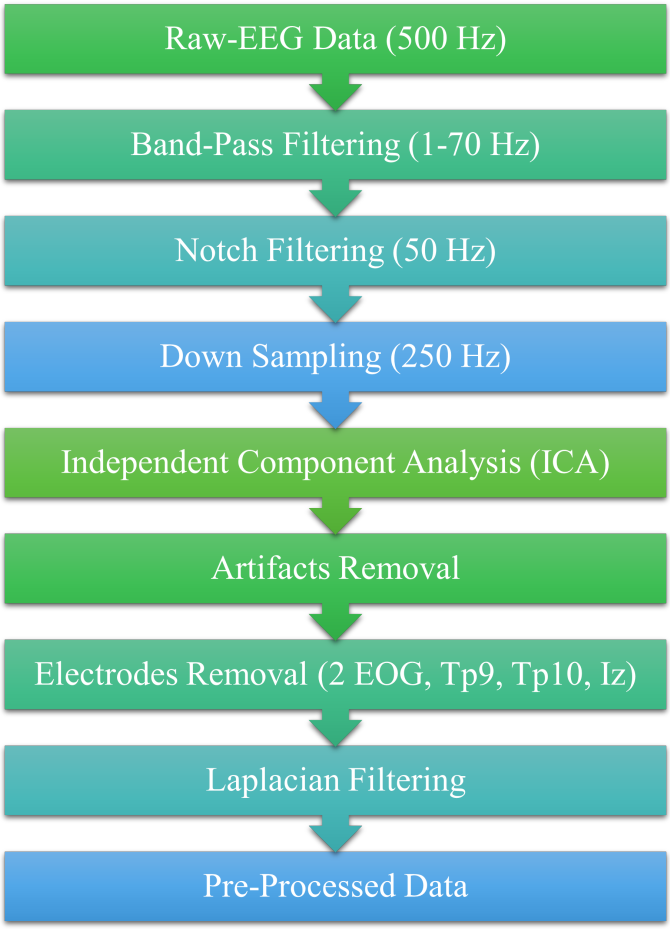
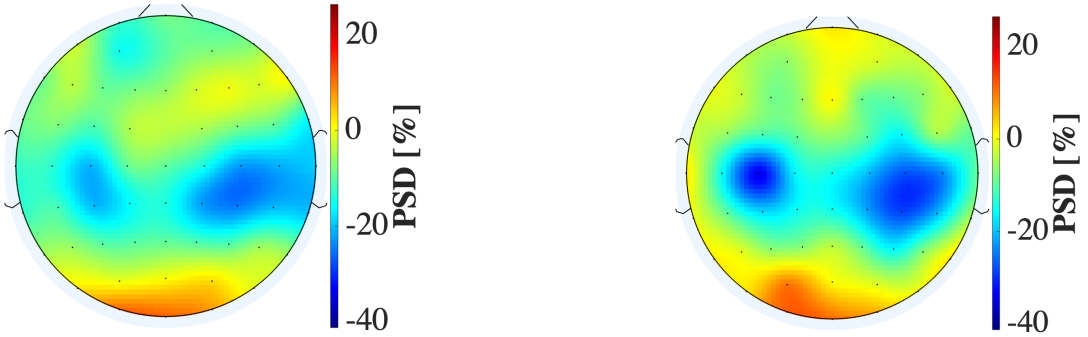



***Figure 3.2:*** *Pre-processing steps performed on the raw EEG signals.*

Firstly, a band-pass zero-phase FIR filter between 1 and 70 Hz was applied to the data in order to remove the low-frequency baseline drift and the noise at higher frequency, while introducing as little distortion as possible in the band of primary interest. A Notch filter at 50 Hz was also applied to attenuate the power line noise. After that, the signals were downsampled at 250 Hz to reduce data autocorrelation and memory occupation. In order to remove ocular and muscular artefacts, the *Independent Component Analysis (ICA)* was performed to separate

the data into statistically independent and spatially localized sources. Independent components associated to ocular artifacts (blinks and lateral movements), muscular noise and other not-brain related sources were visually identified and removed by an experienced operator. Subsequently, the two EOG electrodes were removed, together with Tp9, Tp10 and Iz EEG channels, since their signals were very noisy or the electrodes were not connected.

As a last pre-processing step, a *Re-referencing* of the EEG traces was performed. Two kinds of re-referencing methods were considered for application and compared *(Figure 3.3)*. Firstly, *Common Average Referencing (CAR)* was applied. Since it does not protect from the effects of volume conduction, so the activity tends to be less localized, *Laplacian Filtering* was performed on the signals, by using an adaptation of the *Current Source Density (CSD)* toolbox. Hence, all the analyses in the current research were conducted on Laplacian-filtered EEG data.




***(a) CAR*** ***(b) CSD***

***Figure 3.3:*** *Comparison between Common Average Referencing (CAR) and Current Source Density (CSD). The illustrated Scalp Maps, relative to the third 1-minute window of CPT, represent the Alpha Power Spectral Density (PSD) with respect to the baseline, computed as the median of the ADHD patients.*

## 3.2 Performance Indices Analysis

This first analysis was conducted in order to check if the Performance Indices alone were able to discriminate between pathology and physiology.
As mentioned in the previous section, the CPT performance is evaluated through the scores obtained analyzing different behavioral Performance Indices. Specifically, the indices considered in the present study are:

- *Omission Error (OE)*: failure to press the key when the critical signal (target) appears. High omission rates indicate that the subject did not respond to the target stimuli because of a specific reason, such as difficulty in focusing. Indeed, OE is considered a measure of inattention [169].

- *Commission Error (CE)*: a response is given when a non-target is shown [21]. Depending on the Hit Response Time (HRT), high commission rates may indicate either inattention or impulsivity. If high CE scores correspond to slow reaction times (high HRT), the subject was probably inattentive to the type of stimulus presented and, therefore, responded at a high non-target rate. If, on the other hand, high CE scores are combined with fast reaction times, the subject likely failed to control his impulsivity in responding to non-targets. In this case, high commission rates would reflect impulsivity rather than inattention [169].

- *CPT Hit Response Time (HRT)* : average speed of correct responses; it is the latency between the target onset and the participant's response [21]. An atypically slow HRT may indicate inattention (especially when commission rates are high). Alternatively, a very fast HRT, when combined with high CE scores, may indicate impulsiveness.

- *CPT HRT Standard Deviation (HRT SD)*: measure of response speed consistency. A high HRT SD may be indicative of inattention, suggesting that the subject was less engaged and that the stimuli were processed less efficiently during the task [54].

- *CPT Variability*: takes into account the task variability. High variability values indicate that subject attention and processing efficiency vary throughout the test [54].

Since the performance values of two ADHD subjects were missing, these two participants were excluded from behavioral analysis. All the indices were provided by the clinical team performing the acquisition as global t-scored values.

To understand which of these behavioral indices were significant to distinguish the two groups (i.e ADHD patients and the Controls), a Wilcoxon test for independent samples was carried out within each group for each performance index.
The *Wilcoxon test* has been chosen instead of t-test, since the latter hypothesizes a reasonably normal distribution for the population [19]. From the observation of the boxplots *(Figure 3.4)*, it can be deduced, indeed, that the assumption of a normal distribution cannot be done for the current performance indices. The test is based on a null hypothesis, for which the two groups have the same distribution with the same median. Rejecting the null hypothesis means that the p-value is less than the significance level, set to 0.05 in the actual research: a p-value near 0 shows high ability of the feature in discriminating the groups, while when it is close to 1 indicates high similarity of distributions [12]. It is worth mentioning that only one value per patient was acquired for each behavioral parameter, corresponding to the whole CPT (global parameters).

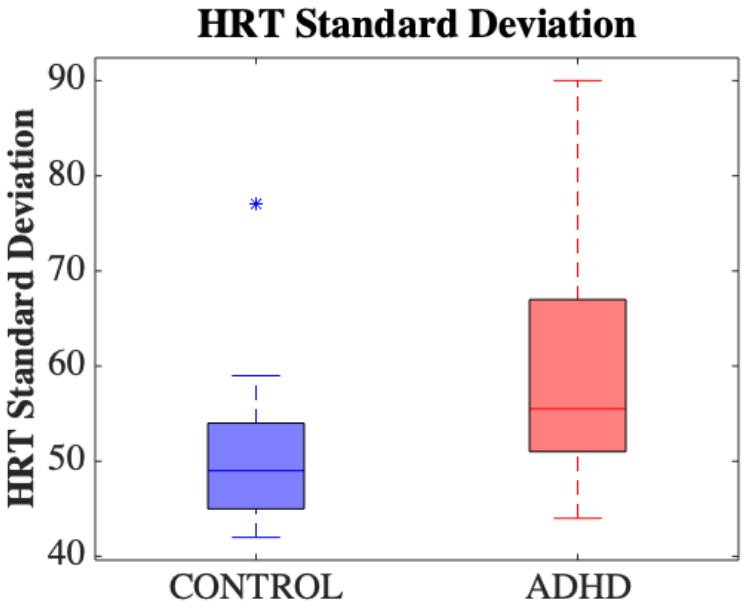



***Figure 3.4:*** *Boxplot of Control (blue) and ADHD (red) distributions of Hit Response Time Standard Deviation (HRT SD). * represents an outlier.*

## 3.3 EEG analysis

This section deals with the EEG signals analysis methods applied in this work:

- **Spectral Analysis**
- **Functional Connectivity Analysis**

### 3.3.1 Spectral Analysis

In this section all the steps concerning the time-frequency analysis computed on the signals will be explained. The aim of the analysis was to understand if significant differences could be observed between ADHD patients and Controls. This type of analysis is mainly conducted considering the variation of the CPT spectral power with respect to the resting state period.

#### 3.3.1.1 Power Spectral Density

Firstly, the Power Spectral Density (PSD) was computed through a function based on *Welch's Periodogram*. The Welch's method is a non-parametric approach used to estimate the PSD. It can be considered an improvement of the standard Bartlett's method, since it is more robust to noise and reduces the variance of the individual PSD values by applying an explicit tapering window to the signal [190]. This technique consists in segmenting the time signal into successive blocks, obtaining the periodogram for each block, and finally averaging the periodograms [171]. The estimation of the PSD is given by *Equation 3.1*:

$$\hat{S}_x(f_k) = \frac{1}{K} \sum_{m=0}^{K-1} P_{x_m,M}(f_k) \tag{3.1}$$

where $P_{x_m,M}$ is the periodogram of the m-th block, with m=0:K-1.

In this study, the relative Power Spectral Density was estimated every 10 seconds of the signal, without overlapping. The choice of the segment length (10 seconds) was made taking into account that the CPT duration is quite long, and a very high temporal resolution was not required in this analysis.
Within the 10-seconds segment the spectrum was computed with Hamming tapering of 1 second and a chosen overlap of 50% (125 samples of overlap, since the sampling frequency was 250 Hz), obtaining a frequency resolution of 1 Hz. The final spectral density was obtained by averaging the 19 periodograms of the 10-seconds segment. In order to get a single power value for each frequency band

of interest, the integral of the PSD was computed in the subsequent frequency ranges:

a. *Delta*: 1-4 Hz

b. *Theta*: 4-8 Hz

c. *Alpha*: 8-13 Hz

d. *Beta 1*:13-22 Hz

e. *Beta 2*: 22-30 Hz

f. *Gamma*: 30-45 Hz

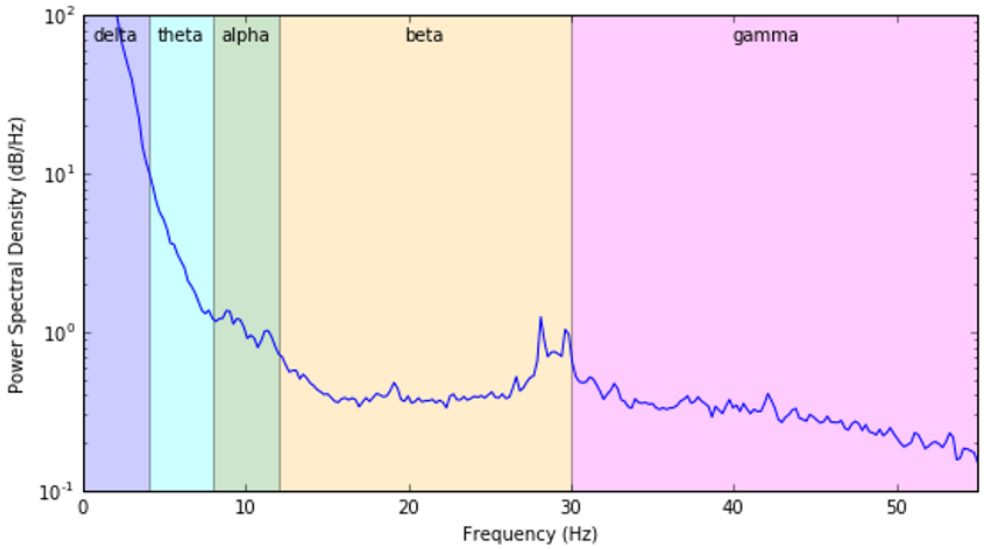



***Figure 3.5:*** *Power Spectral Density [dB/Hz] of EEG recording in function of frequency [Hz] [1].*

The normalization of the power is computed within the function in order to reduce the dependence of the extracted power from the amplitude of the signal itself. This procedure allows to obtain a more robust comparison between signals coming from different cortical areas and across subjects [140].

#### 3.3.1.2 Power variation evaluation

When exploring a cognitive task, it is important to choose an appropriate baseline to observe the changes in brain activity. Both eyes-closed and eyes-open conditions are suitable baseline choices, but they show different characteristics of the brain activities. In the current study, the eyes-open condition was considered as baseline, since it is more appropriate when dealing with tasks that involve visual processing. Thus, it is assured that the EEG changes are associated to the task considered [135]. To evaluate the changes in the brain activity during the task with respect to the baseline, the *Percentage Change* in the band-specific power value was computed. The results are interpreted as changes in power relative to the power of the baseline period, according to the *Equation 3.2*.

$$\Delta P\% = 100 * \frac{P_{activity_{tf}} - \overline{P_{baseline_f}}}{\overline{P_{baseline_f}}} \tag{3.2}$$

The bar over the baseline indicates the mean across the baseline time points; t and f are time and frequency indicators [53].
In the present work, the baseline value was subtracted to each 10-seconds power of CPT. Thus, the percentage variation of the spectral densities, relative to each frequency band and to each patient, was obtained.

#### 3.3.1.3 Topographical representation: Scalp Maps

Most of the studies on ADHD analyze the CPT (generally lasting few minutes) by condensing the results in a single value, loosing the temporal information. In the current study, since the CPT duration was of 14 minutes, it was decided to study the Spectral Density dynamically, choosing temporal windows of 1 minute each. The value for each window was obtained by averaging the PSD values of the six 10-seconds segments, mentioned previously. This procedure was conducted for every participant of the two groups, considering separately the frequency bands mentioned before, with the aim of detecting some differences among control and ADHD patients. For a meaningful visualization of the described quantification of the brain power activations, using the 'topoplot' EEGLAB function, the topographical representation of the power on the scalp maps was obtained. An example of the scalp map is shown in *Figure 3.6*:

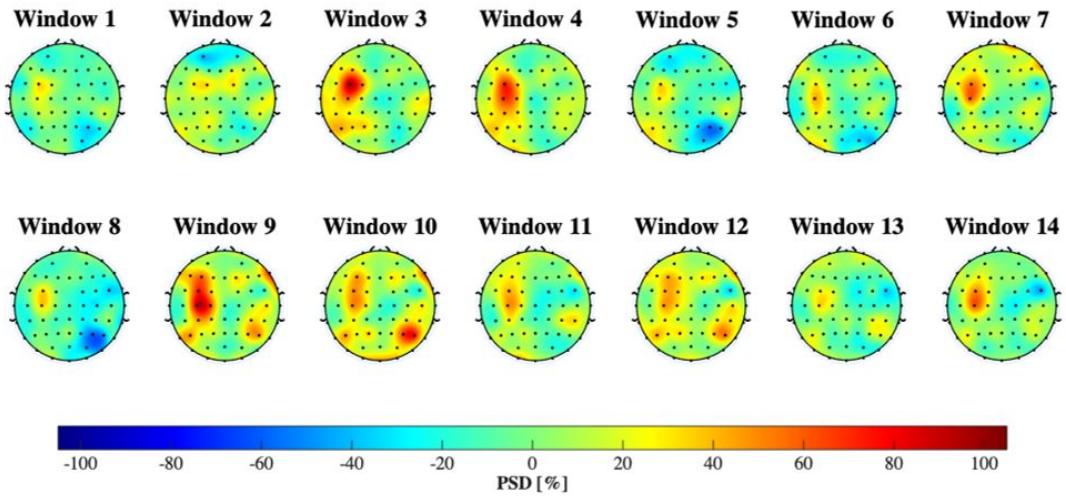



***Figure 3.6:*** *Scalp Map: Patient 07, Control, Beta1 Band. The minimum and the maximum values of the colorbar were chosen according to the minimum and maximum values of the overall population (considering both ADHD and control patients), for each frequency band. Indeed, in order to perform a correct comparison between all subjects, the color scale is recommended to be equal. The unit of measure is the percentage variation of PSD.*

In order to represent and compare the power activations for both the control and the ADHD group, median scalp maps were obtained. The median was preferred to the mean since it is more robust to outliers. The computation was done for every electrode, considering the two groups separately.

#### 3.3.1.4 Cortical Areas-specific Analysis

Another type of analysis conducted concerns the median temporal trend of the power percentage variation considering 4 different areas, each one divided in the two hemispheres, in order to investigate the laterality and regionality of the brain activity. The following areas were thus considered *(Figure 3.7)*:

- *Left Frontal*: Fp1, F1, F3, F5, F7, AF3, AF7
- *Right Frontal*: Fp2, F2, F4, F6, F8, AF4, AF8
- *Left Parieto-Occipital*: P1, P3, P5, P7, PO3, PO7, O1
- *Right Parieto-Occipital*: P2, P4, P6, P8, PO4, PO8, O2
- *Left Temporal*: T7, FC5, CP5, C5, FT7, TP7

- *Right Temporal*: T8, FC6, CP6, C6, FT8, TP8
- *Left Central*: C3, FC1, CP1, C1, FC3, CP3
- *Right Central*: C4, FC2, CP2, C2, FC4, CP4

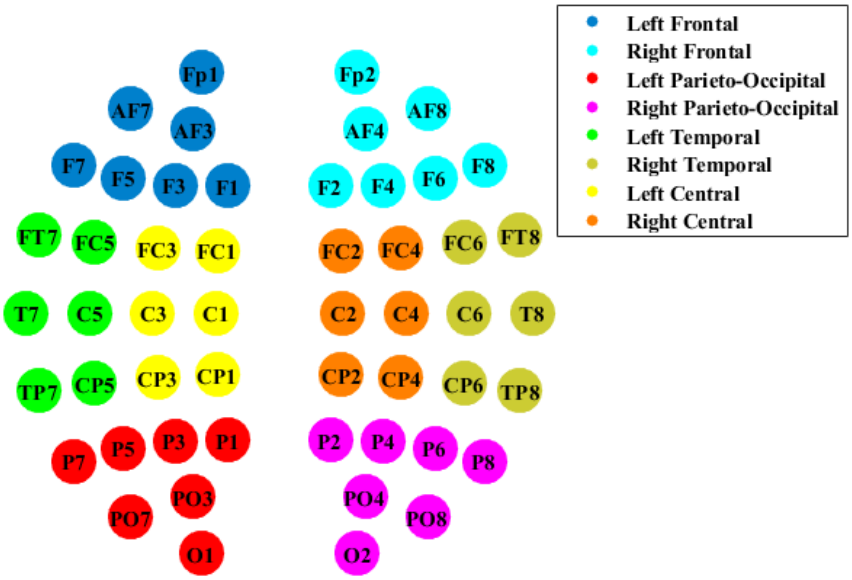



***Figure 3.7:*** *Subdivision of EEG electrodes according to the different brain areas (Frontal, Temporal, Central and Parieto-Occipital), divided into Right and Left hemispheres.*

For each area the median value of $\Delta$P% among the relative electrodes was computed for each subject. This type of analysis is useful to achieve an easier interpretation of the results instead of an investigation conducted on the electrodes considered singularly. Unlike the scalp maps, the temporal trend was visualized in only Alpha, Beta 1 and Theta bands: indeed, a great number of studies [132] [131] [23] [113] [172] [130] have noticed high levels of Theta and/or decreased levels of Alpha and Beta, which might be associated to the ADHD [142]. Moreover, the trend was displayed considering 10-seconds and not 1-minute windows, because, differently from the previous one, this type of analysis lends itself better to this temporal scale, providing more interpretable results. In order to visualize and compare the temporal trend for the ADHD and control group, the median value among the patients was computed for the two different groups. In *Figure 3.8*, the median temporal trend relative to Theta in the Left Parieto-Occipital region is displayed.

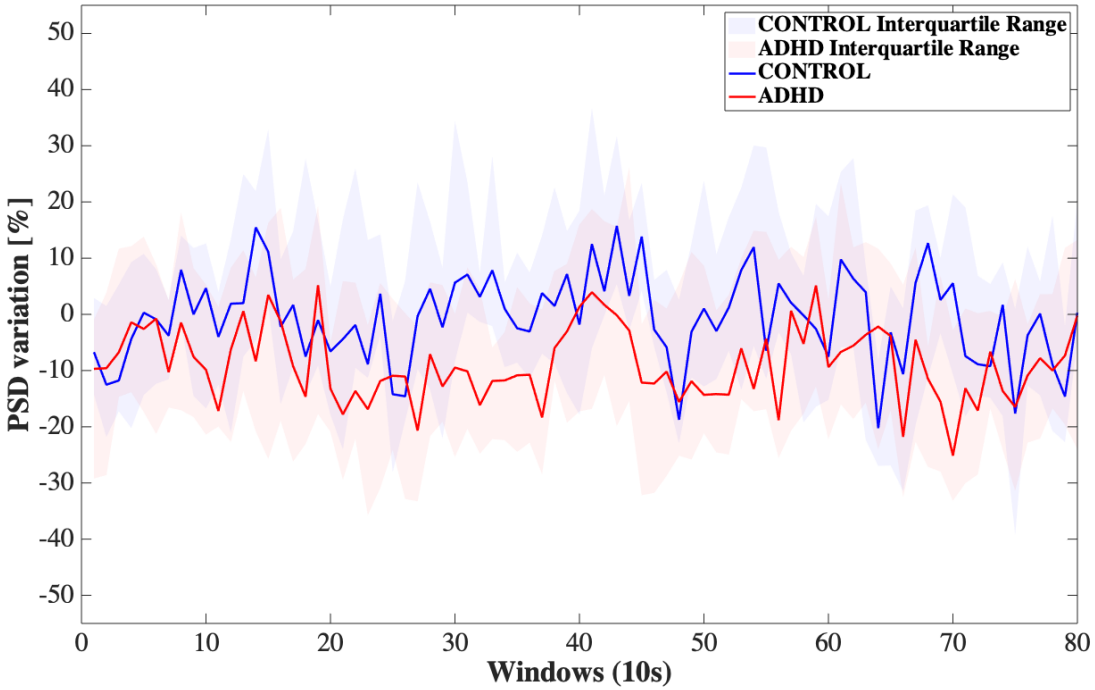



***Figure 3.8:*** *Median Temporal Trend of the Power Spectral Density (PSD) variation with respect to the baseline in Theta (4-8 Hz) band, in Left Parieto-Occipital Cerebral Region. The interquartile range is plotted around the median temporal trends relative to the Controls (blue) and ADHDs (red).*

#### 3.3.1.5 Correlation Analysis between EEG and Performance Indices

Based on the results obtained in *Section 3.3.1.3*, an investigation of the correlation between EEG-based spectral parameters and performance indices was carried on. A first correlation was performed between a global EEG index, computed per subject as the median value of $\Delta$P% along the 13 minutes of the task in Alpha band in Right and Left Central regions, and global performance indices. Successively, a second analysis was conducted taking into account the temporal evolution, correlating each one-minute value of $\Delta$P% with the global performance indices.

The initial task duration corresponded to 14 minutes, but in this analysis and in all the subsequent analyses, the last minute was discarded since, at the end of the acquisition, the patient performance is generally considered not to be so reliable due to fatigue.

The decision about the Alpha band and the Central area was made after having noticed an interesting behavior in the $\Delta$P% of the ADHD group, as it will be explained in *Section 4.2*. Both hemispheres were analyzed separately, but the corresponding electrodes were chosen differently from the previous analysis, subsequently to the observation of the channels involved in the diversified behavior between groups:

- *Left Central*: C3, FC1, CP1, FC5, CP5 C1, FC3, CP3, C5
- *Right Central*: C4, FC2, CP2, FC6, CP6, C2, FC4, CP4, C6

Before proceeding further with the computation of the correlation, outliers were removed from both the variables distributions under analysis, through the method of quartiles. Indeed, one definition of outlier is any data point that lies below the first quartile and above the third quartile more than 1.5 Interquartile Range (IQR), which is the difference between the values of the third and first quartiles [163].

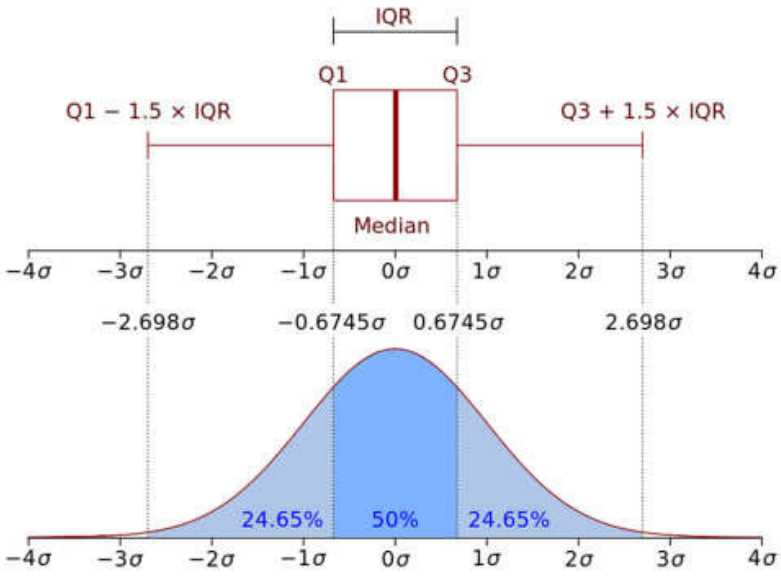



***Figure 3.9:*** *Representation of First Quartile (Q1), Third Quartile (Q3) and Interquartile Range (IQR) of a normal distribution [145]*

.

Since the numerosity of the population in the two groups was too small to suppose a normal distribution [133] and normality and linearity of variables are fundamental

requirements to apply Pearson's correlation, the *Spearman's correlation* was employed. Spearman's correlation is a non-parametric method, usually adopted when the assumption of a bivariate normal distribution is not sustainable [18]. Furthermore, this approach is more robust with respect to outliers than Pearson's correlation coefficient [174].

Since the Spectral Analysis alone was not sufficient to find significant biomarkers capable of discriminating the pathology, a Connectivity Analysis was carried on.

### 3.3.2 Functional Connectivity Analysis

In this section, the mathematical tools and methods employed for Connectivity Analysis of the EEG signals will be explained. A brief overview of all the steps performed in the Connectivity Analysis is presented in *Figure 3.10*.

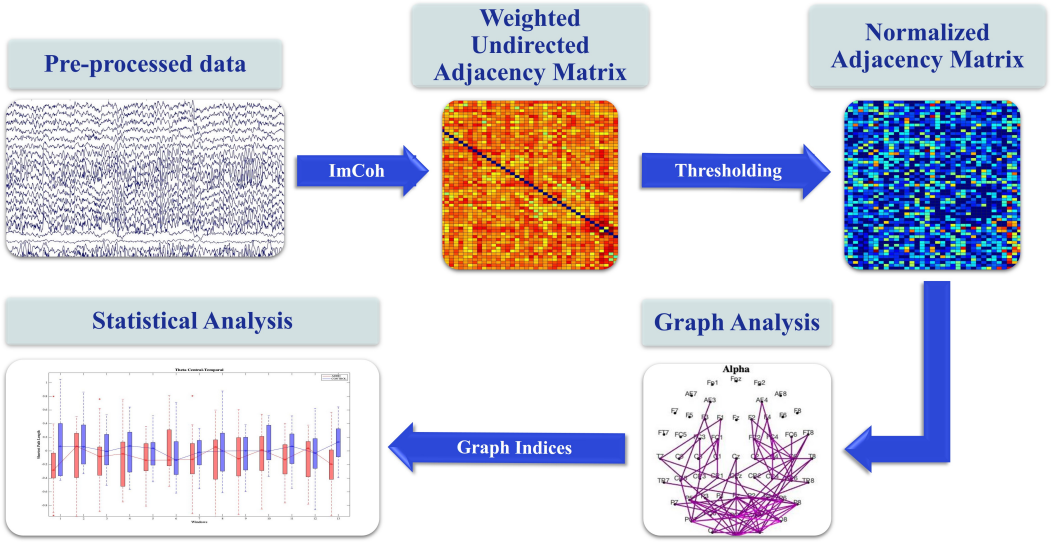



***Figure 3.10:*** *Connectivity Workflow: the Connectivity Analysis starts from pre-processed data, from which the Weighted Undirected Adjacency Matrices were obtained through the Imaginary Part of the Coherency (ImCoh). Subsequently, a thresholding was applied in order to obtain only the significant connections ($p<0.05$). From the Graph obtained from the Adjacency Matrix graph-based measures are derived. A statistical analysis is then performed on these indices to identify the most significant ones to distinguish the two groups.*

A bivariate analysis was performed, analyzing pairs of electrodes separately, and conducted on the two acquisition phases, i.e. eyes-open resting state and CPT. The eyes-open condition was considered in order to study the 13 minutes of CPT variation with respect to a baseline.

#### 3.3.2.1 Connectivity Measure: Imaginary Part of Coherency

The *Imaginary Part of the Coherency (ImCoh)* was chosen as Functional Connectivity measure to estimate magnitude of information flow between the bivariate EEG signals and to represent a Functional Connectivity Network (FCN) estimation at the sensors level.
Unlike the real part of the coherency, the imaginary part is insensitive to artifactual interactions caused by volume conduction [138]. This is due to the fact that ImCoh is only sensitive to synchronizations of two processes which are time-lagged to each other [138]: indeed, it captures the non-instantaneous connectivity between brain activities from EEG channels that are phase-lagged (i.e. delay-based). On the other hand, volume conduction affects multiple scalp channels with near-zero phase delays. The result is that, when two signals at the same frequency have identical phase values, possibly due to volume conduction artifacts, ImCoh is zero [144].

The coherency between two EEG signals, which could be considered a generalization of the correlation in the frequency domain, is a measure of the linear relationship between the two at the same frequency [97].
Coherency *(Equation 3.3)* is defined as the cross-spectrum between the two signals, divided by the square root of the product of each signal power spectrum at the same frequency [97].

$$C_{ij}(f) = \frac{S_{ij}(f)}{\sqrt{S_{ii}(f) * S_{jj}(f)}} \tag{3.3}$$

If $X_i$ and $X_j$ are the discrete Fourier transforms of the respective time series, the Cross Spectrum is defined as an average over a sufficiently large number of windows (K) of the product of the two discrete Fourier transforms [138].

$$S_{ij}(f) = \frac{1}{K} \sum_{k=1}^{K} X_{i,k}(f) * \overline{X_{j,k}(f)} \tag{3.4}$$

Complex coherency was firstly computed based on Welch's method, the same method used for PSD computation, obtaining a value every 10 seconds. The complex coherency is computed every 1 second within the 10 seconds window and, unlike to the PSD computation, the overlap was set to zero and Hanning window was chosen. Then, the absolute value of ImCoh was extracted. In order to obtain

ImCoh in the specific frequency band, Alpha, Beta1 and Theta, values of ImCoh, corresponding to frequency bins of that specific band, were averaged.

#### 3.3.2.2 Testing the significance of the connectivity: Surrogation Test

The Surrogation Test with Phase Randomization method [148] was applied in order to define a statistical threshold for assessing the significance of the coupling between the time series. This method consists in generating new time series from the original ones: the amplitudes are maintained while the phase relationship is destroyed, by randomly shuffling the phase in the frequency domain and then transforming them back to the time domain [119].
The number of surrogates has been chosen equal to 50. The imaginary part of the coherency in Alpha, Beta1 and Theta frequency bands was computed for the 50 surrogates, through the same method explained in *Section3.3.2.1* and the threshold was extracted as the percentile corresponding to 95% of significance.

**Adjacency matrices** were obtained applying the estimated threshold: the results from the original connectivity values below the threshold were set to zero, while the others were maintained, resulting in a weighted graph with only significant connections.
The adjacency matrices 59x59 (where 59 is the number of considered electrodes) were normalized, in order to make comparable all the values among each other.

For each subject and for the two conditions (baseline and CPT), weighted adjacency matrices were obtained for each frequency band (Alpha, Beta1 and Theta).

#### 3.3.2.3 Graph Analysis

The statistical significance of the above connectivity measure was evaluated using indices based on graph theory. From each adjacency matrix (i.e. graph), computed every 10 seconds, the following graph-based indices were derived using MATLAB functions provided in the Brain Connectivity Toolbox (BCT) [152], obtaining a value of the index every 10 seconds, for each band and for each acquisition phase:

- *Measures of centrality*: **Degree**, **Strength** and **Betweenness Centrality**;
- *Measures of integration*: **Shortest Path Length** and **Global Efficiency**;
- *Measures of segregation*: **Clustering Coefficient** and **Local Efficiency**;

To study the information flow between different regions of the brain, the region-wise analysis was performed, taking into account the main four cerebral areas shown in *Figure 3.11*:

- *Frontal*: Fp1,Fp2, F3, F4, F7, F8, Fz, F1, F2, AF3, AF4, F5, F6, AF7, AF8, Fpz;
- *Central*: C3, C4, , Cz, FC1, FC2, CP1, CP2, C1, C2, CPz, FC3, FC4, CP3, CP4;
- *Temporal*: T7, T8, FC5, FC6, CP5, CP6, C5, C6, FT7, FT8, TP7, TP8;
- *Parieto Occipital*: PO7, PO8, POz, Oz, P5, P6, PO3, PO4, P1, P2, Pz, P7, P8, P3, P4, O1, O2;

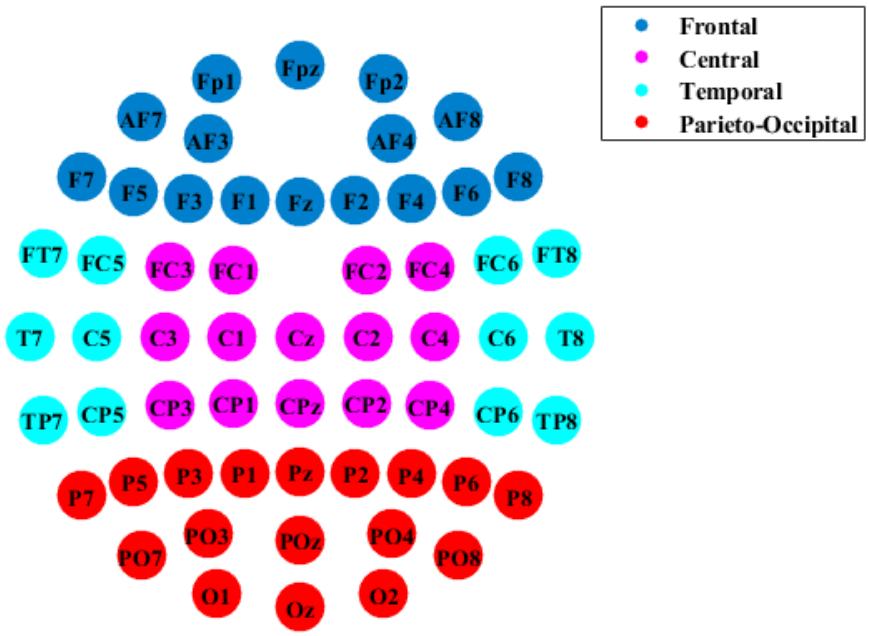



***Figure 3.11:*** *Subdivision of EEG electrodes according to the different brain areas (Frontal, Temporal, Central and Parieto-Occipital).*

The mentioned indices, except for Global Efficiency, which is a measure of the overall network, are measures relative to each electrode (i.e. node-based indices). Thus, in order to obtain a value of the index for area, the median of the values corresponding to each electrode of the specific area was computed.

Concerning the Shortest Path Length, since it is represented by a matrix of distances between all pairs of electrodes, the sub-matrices corresponding to all pairs of previously mentioned regions were extracted. In order to have a value for each pair, the sub-matrices median was considered, obtaining a characteristic path length for each pair of regions.

- **Statistical Analysis**

To evidence different connectivity patterns during the task between the two groups (i.e. healthy participants and patients with ADHD), it was deemed appropriate to subtract the baseline from the indices, as it was already done in the Spectral Analysis. The baseline was obtained averaging the 6 values of the indices along 1 minute of the resting state and then subtracting it from each 10-second segment value relative to the task.

A statistical analysis employing Wilcoxon test for independent samples was then adopted to determine which of the graph-based indices would have been more useful to discriminate between the two groups. The data entries for the test were the values of the indices in every 10-second segment of CPT (for a total of 78 segments per patient): this resulted in a vector 1x1248 (78x16) for the ADHD population and 1x936 (78x12) for the Controls, for every index, band and area under examination. Before performing the test, outliers, identified as data outside the 1.5 inter-quartile range, were removed.

Based on the results retrieved from the statistical analysis, only a subset of connectivity indices (pvalue $< 0.05$) was selected for a subsequent investigation of the temporal trend during CPT. This subset was further screened, since some of the indices did not carry interesting results or provided similar information, like Clustering Coefficient and Local Efficiency.

The decision to analyze the indices temporal trend was taken since no such type of ADHD study was found in literature and could probably reveal an interesting decreasing or increasing trend in the graph measures. It could be considered a novel approach of applying graph theoretical measures, useful for investigating the brain activity in the two different groups.

A 1-minute temporal scale was employed and the values per minute were again obtained through the average of the 10-second values.

## 3.4 Machine Learning

In this section, a model for discriminating ADHD and control groups will be described, exploiting only the information contained in the CPT. Indeed, many works have tried to discriminate the pathology considering only the resting state, while almost no one has performed classification using the CPT information. The aim of merging Graph theoretical and Spectral approaches with Machine Learning techniques was to provide a further way to investigate the ADHD impairment, as well as mapping predictions to a single individual case.

The workflow for the Machine Learning-based Classification is illustrated in *Figure 3.12*.

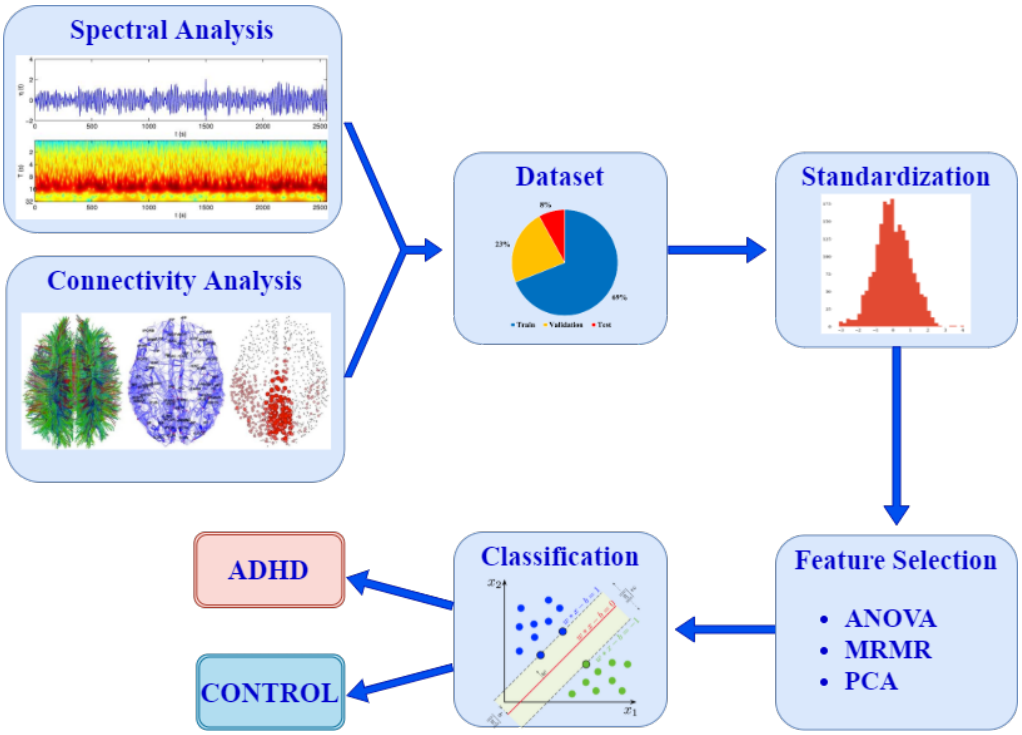



***Figure 3.12:*** *Machine Learning-based Classification Workflow. The initial features of the dataset were derived from the Spectral and Connectivity Analyses. The dataset was splitted in Train, Validation and Test sets and a Standardization was applied to normalize the data. Feature selection methods (ANOVA, MRMR and PCA) were used to extract the most relevant attributes and compared. A non-linear classifier was employed for discriminating the two groups (ADHD and CONTROL).*

For the implementation, Python language has been used and the code has been written and executed on Colaboratory, a product from Google Research. The algorithm was applied on the dataset, which has been previously prepared by standardizing the data and choosing a subset of features in order to reduce complexity and dimensionality.

### 3.4.1 Data Preparation

The starting dataset is composed as follows:

1. The *Observations* are the 10-second segments belonging to CPT, for a total of 2184 samples.

2. The *Features* are the variations of Graph Indices and of Power Spectral Densities with respect to the baseline, for each band (Alpha, Beta 1 and Theta) and area (Frontal, Central, Temporal and Parieto-Occipital), reaching 93 explanatory variables. As explained previously, the extraction of the features for each area was done by computing the median of the electrodes' values of that specific area.

3. The binary *Target* (0/1) corresponds to CONTROL and ADHD, respectively.

As can be seen in *Figure 3.14*, the dataset did not show a great imbalance in the data, considering the target, indeed it was not deemed necessary to apply any method to balance it.

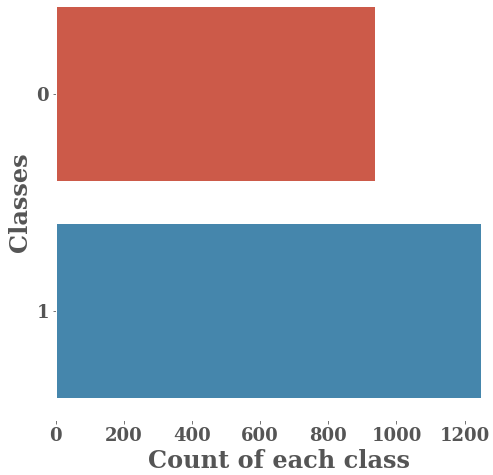



***Figure 3.13:*** *Counting of Dataset Observations: 936 belonging to Controls (0) and 1248 belonging to ADHD (1).*

Before proceeding with the split in Train, Validation and Test sets, a preliminary step was to remove the features relative to the Degree Index, since the information provided was nearly equal to that carried by the variables related to the Strength Index. The reason to preserve the Strength with respect to the Degree was that the first one carries the additional information of the weight of the link.

The dataset was then subdivided in three parts *(Figure 3.14)*, maintaining for each set the original balance in the target class:

- **Test Set**: 8% of the entire dataset (175 samples) was randomly chosen and used as Test Set to assess the classification performance when new data are considered.
- **Training Set**: 75% of the remaining dataset (1506 samples) was randomly selected to create the Training Set, whose utility is to train the classifier.
- **Validation Set**: the other 25% of the data (503 observations) was used as Validation Set, employed to predict the responses for the observations through the fitted model [86].

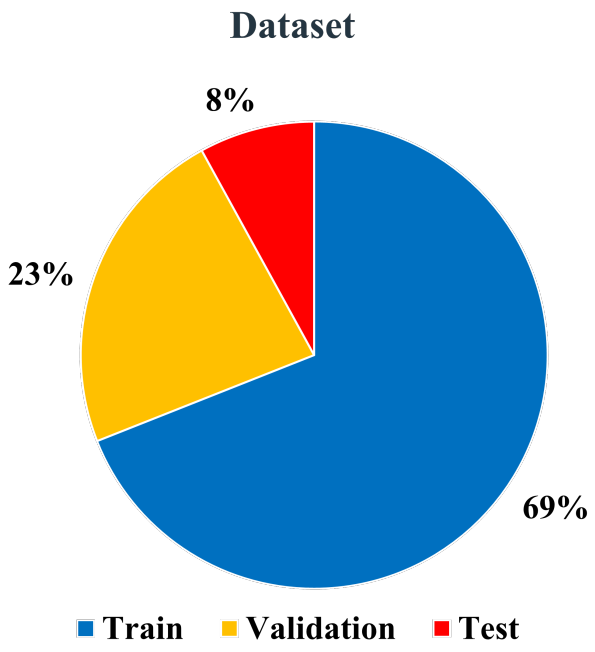
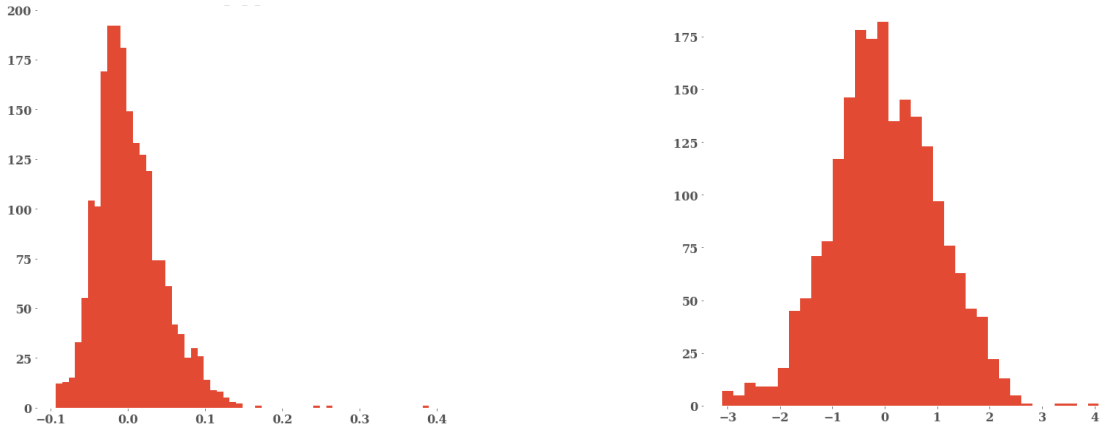



***Figure 3.14:*** *Dataset split in Train (69%), Validation (23%) and Test (8%) sets.*

Subsequently, **Standardization** was performed on the Training, merged with Validation set, and Test set separately, when using Analysis of Variance and Maximum Relevance Minimum Redundancy feature selection methods; while, when adopting Principal Components Analysis, the standardization was executed on the entire dataset.

Standardization is useful to make the features normally distributed: indeed, a normal-like distribution (zero-mean and unitary variance) is required to perform many statistical methods. The *Power Transformer* standardization was selected and implemented in Python environment, to make the data more Gaussian-like. An example of the effect of standardization on the Betweenness Centrality index in Alpha band and Central area is reported in *Figure 3.15*.




***(a)*** *No Standardization* ***(b)*** *Standardization: Power Transformer*

***Figure 3.15:*** *Histograms of Clustering Coefficient: Alpha Central.*

#### 3.4.1.1 Feature Selection

*Feature Selection* is an essential step when dealing with a classification problem, especially in datasets with many features. The aim is to reduce the original feature set to a smaller one, preserving the relevant information for predicting the response, while discarding the redundant one [50]. Furthermore, feature selection methods are useful to speed up the training of the model, lower the complexity and increase the generalization potential of the classifier [51]. Different methods have been tried in the present study, that are summarized in *Figure 3.16*.

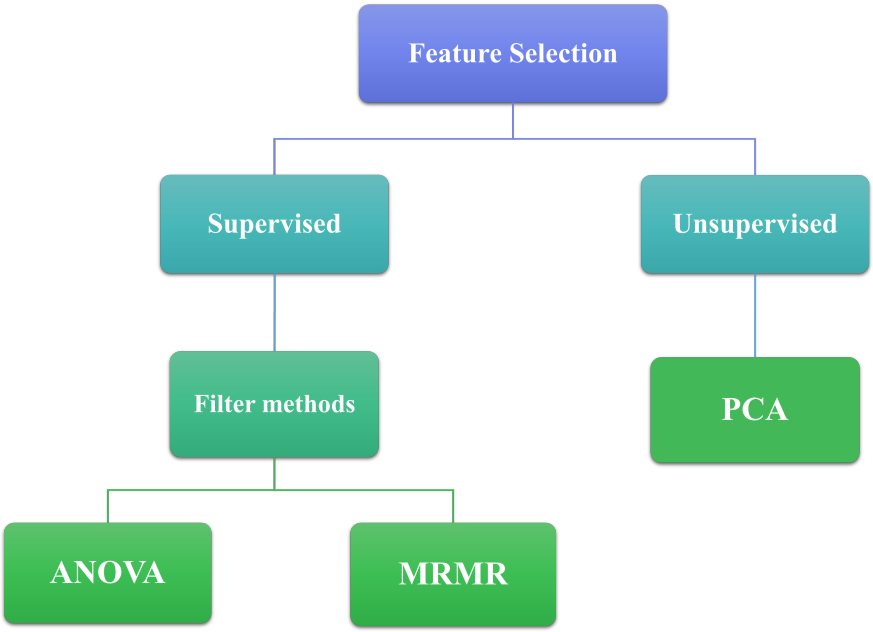



***Figure 3.16:*** *Scheme of the tested Feature Selection Methods.*

A first distinction in Feature Selection is between Supervised and Unsupervised methods: the difference is in the usage of the target variable to select the features. Indeed, while *Unsupervised Methods* ignore the target information, *Supervised Methods* include the knowledge of the classes in the selection [100].

*Filter Methods* are a particular type of supervised approaches, which act independently from the chosen classifier and filter out features using heuristics or the characteristics of the given data [76].

**Analysis of Variance** and **Maximum Relevance Minimum Redundancy** are two examples of these methods, which have been applied in the current study. Regarding Unsupervised feature selection, the attention was focused on **Principal Component Analysis**.

Indeed, these particular statistical approaches are recommended when dealing with a starting dataset composed by numerical features and a categorical target variable, as in this case. It is important to underline that Supervised methods need to be applied only to the Training dataset, while Unsupervised ones could consider the whole dataset, since the target information is not taken into account.

- **Analysis of Variance (ANOVA)**

**ANOVA** is a parametric statistical test that demonstrates how strongly a feature can distinguish a class from another class by a value (p-value), which varies between 0 and 1. To fulfil this goal, an *F-statistic*, or *F-test*, is employed to test the equality of the means of the groups, by comparing variance among groups relative to variance within groups [103]. The more the variance between the groups, the more they differ from each other, as can be seen in *Figure 3.17*. The results of this test are used to see which features can be removed from the dataset, according to their independence from the target variable [101].


***(a) Not significantly different groups*** ***(b) Significantly different groups***

***Figure 3.17:*** *As the variance between the groups increases, the population distributions differ from each other [109].*

In the current study, an inbuilt Python function namely SelectKBest, which provides a score for each feature, based on the F-test, was used. The higher the score, the higher the prediction power of the variable. Considering the variables in a descending order according to the scores, it was noted that there were many features characterized by a score around 0 and few ones with a score over 15. Hence, a threshold of 15 was chosen as a criterion to feature selection, picking only variables that overcome this score, resulting in 16 features.
Before applying ANOVA, as mentioned before, a data Standardization was performed in order to obtain Gaussian variables, since Gaussianity is a fundamental requirement for the F-test.

One drawback of ANOVA is that it does not take into account the presence of the correlation among variables: this means that two or more variables could carry redundant information, not improving the prediction. Therefore, a Correlation Analysis with *Pearson's correlation coefficient* was performed after

having applied ANOVA, with the consequent removal of the redundant variables. The choice of which features to remove was done according to the score provided by ANOVA test, dropping out other 2 features.

- **Maximum Relevance Minimum Redundancy (MRMR)**

The **MRMR** algorithm is part of a special group of filter-based feature selection approaches, whose aim is to select highly predictive (relevance), but at the same time uncorrelated (redundancy) features [149].

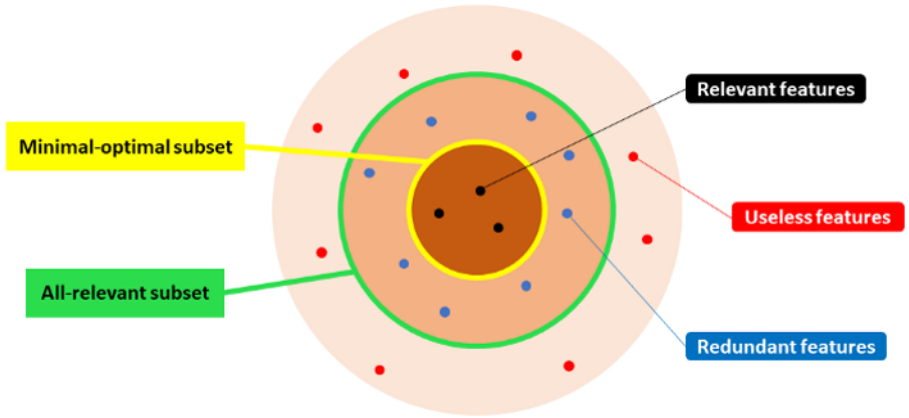



***Figure 3.18:*** *Minimal-optimal methods, such as the MRMR, aim at identifying a small set of features that have the maximum possible predictive power, finding a minimal-optimal subset of attributes [125].*

In the current study, a variant of this method, the **F-test Correlation Quotient (FCQ)**, was chosen, since it is the most suitable for continuous features. Indeed, the relevance is computed through the F-statistics, while the redundancy by using the Pearson's correlation coefficient, as shown in *Equation 3.5*:

$$score_{FCQ}(f_i) = \frac{F(f_i, target)}{\frac{1}{|S|}\sum_{f_s \in S} \rho(f_s, f_i)} \tag{3.5}$$

$\rho(f_s,f_i)$ is the Pearson correlation, $F(f_i, target)$ is the F-statistic and S the selected feature set [192].

Similarly to the ANOVA, this algorithm could be applied thanks to the normalization procedure. However, in this case it was necessary to set a priori the number of the desired features, which was initially chosen according to the ANOVA result and, then, adjusted for several trials. After this procedure, the optimal number of selected features resulted to be 20.

- **Principal Component Analysis (PCA)**

The last feature selection algorithm analyzed was **Principal Component Analysis (PCA)**. Differently from the previous cases, PCA is an unsupervised method, so it could be applied to the whole dataset, since the target is never considered to choose the features. This technique performs the dimensionality reduction of a set of variables by finding linear combinations among the features, called *principal components*, that successively have maximum variance for the data, and then recasting the data along the orthogonal component axes [63].


*(a) 2D* *(b) 3D*

***Figure 3.19:*** *Representation of Principal Components space [3].*

Therefore, the aim is to reduce the number of features of the dataset, while preserving as much information as possible.

In the present investigation, the first 28 principal components have been picked, explaining 90% of the total variance, since, from the subsequent ones, the cumulative percentage of explained variance increase was very small.

### 3.4.2 Classification

After selecting the relevant attributes, different classifiers were tested, the data sets were feed to the classifiers and the models were built.

The final choice fell on the **Support Vector Machines (SVM)**, a non-linear, non-parametric method for supervised learning, suitable for binary classification. SVMs can be useful when the data are not regularly distributed or have an unknown distribution. Moreover, they offer a good generalization when choosing appropriately the parameters, being robust even if the training dataset has some bias [20]. SVM is based on a non-probabilistic approach which uses spatial and geometric properties instead of using the probability distribution of data points for the discrimination [179]. Indeed, SVMs work by separating the points belonging to different classes with a hyperplane (decision boundary), or a set of hyperplanes, having the largest distance to the nearest data point of any class [79].

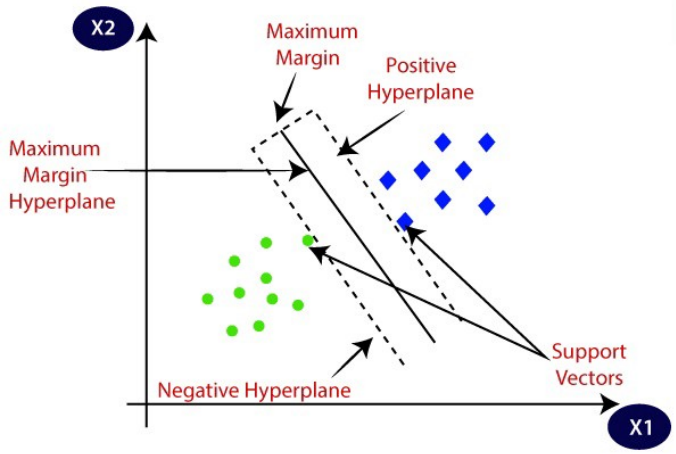



***Figure 3.20:*** *SVM separating hyperplane [5].*

*Kernel functions (K)* are nonlinear functions which are useful when the data are not linearly separable, because they map the data points to another higher dimensional space in which these points can be linearly separable [179]. Considering a higher dimensional space might cause overfitting (poor model generalization), which indicates that the model achieves a high accuracy with Training Set but no good performances with unseen examples [77]. Thus, this model would be very sensitive to noise and even very small changes in data point values may change the classification results. To overcome this issue, a soft margin is introduced for

allowing some examples to be misclassified [77].

For this reason, the *Regularization Hyperparameter C* is introduced to find a trade-off between misclassification of training examples against simplicity of the decision boundary. When a high value of C is chosen, the aim is to classify all the training examples correctly, while low values are used to make the decision surface smooth. Another important parameter is the *Kernel Coefficient* $\gamma$, which regulates the influence of a single training example: the larger this value, the lower the model generalization [72].

To find the most appropriate parameters for the classifier (K, $\gamma$ and C), a *Gridsearch* was implemented in Python, with *F1* as score. F1 *(Equation 3.6)*, also called F-score, is a measure of accuracy that takes into account both Precision *(Equation 3.7)* and Recall *(Equation 3.8)*, by doing a harmonic average [75] [173]:

$$F - measure = \frac{(\beta^2 + 1) * precision * recall}{\beta^2 * precision + recall} \tag{3.6}$$

$$precision = \frac{tp}{tp + fp} \tag{3.7}$$

$$recall = \frac{tp}{tp + fn} \tag{3.8}$$

*tp* are the true positives, *fp* the false positives and *fn* the false negatives. When dealing with a binary classification, a false positive is an error which incorrectly indicates the presence of a condition or a disease when the disease is not present; on the contrary, a false negative occurs when the test indicates the absence of the condition when it is present. A true positive occurs when the model correctly predict the positive class, while a true negative when the negative class is correctly predicted [184].

Regarding the values of K, C and $\gamma$ to put in the Grid Search, they have been picked according to the most common values used in literature [83]:

- K = 'linear', 'poly', 'sigmoid', 'rbf'
- C = $2^{-5}$, $2^{-3}$, ..., $2^{15}$
- $\gamma$= $2^{-15}$, $2^{-13}$, ..., $2^{3}$

In addition, to generalize the chosen classifier model to an independent data set which is different from the one used for training the model [16], a 10-Fold Cross-Validation was implemented in the Grid Search. Cross-validation is a technique that splits the training data into $n$ smaller sets: it uses *n-1* of the folds as training data, while the remaining part of the data is used for validating the resulting model. This procedure is followed for each of the $n$ folds, so that each of the folds is used exactly once as a testing set. The final estimation of the cross validation is obtained by averaging the results computed in each iteration. [83].

# Chapter 4

# Results and Discussion

## 4.1 Performance Indices Analysis

As introduced in *Chapter 3*, the first preliminary step performed in the current work was the investigation of the selected Performance Indices. The detection of differences in the distributions of the two populations (ADHD and Controls) was assessed by conducting the Wilcoxon test for independent samples and analyzing their relative boxplots, shown in *Figure 4.1*.

As can be noticed from the boxplots, ADHD children's scores were characterized by higher median values in all the measures considered and, except for Omission Error, by a higher variability with respect to healthy controls. This is coherent with the literature, since ADHD patients tend to make more Commission and Omission errors than normal controls and take more time to respond [114]. Nevertheless, the Wilcoxon test performed on these indices has not highlighted significant differences, except for the HRT SD, suggesting that these measures were not sufficient to clearly discriminate the pathology.

A greater HRT SD is believed to reflect inefficient information flow during the execution of a cognitive task [117], so, variable responses might be due to inefficient or disrupting information processing. Connectivity patterns of the brain could be useful to measure the efficiency of this information processing.

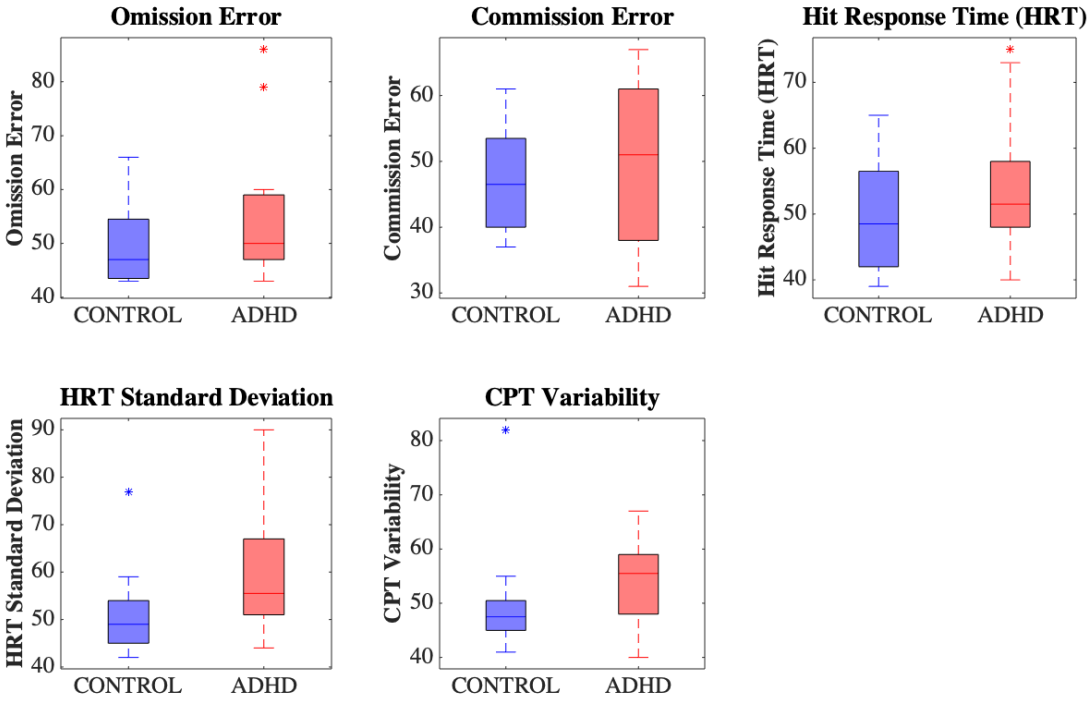



***Figure 4.1:*** *Boxplots of Control (blue) and ADHD (red) distributions of Performance Indices: Omission Error (OE), Commission Error (CE), Hit Response Time (HRT), HRT Standard Deviation (HRT SD) and CPT Variability; * represents an outlier.*

## 4.2 EEG Analysis

In this section the results of Spectral and Functional Connectivity analyses will be shown and commented.

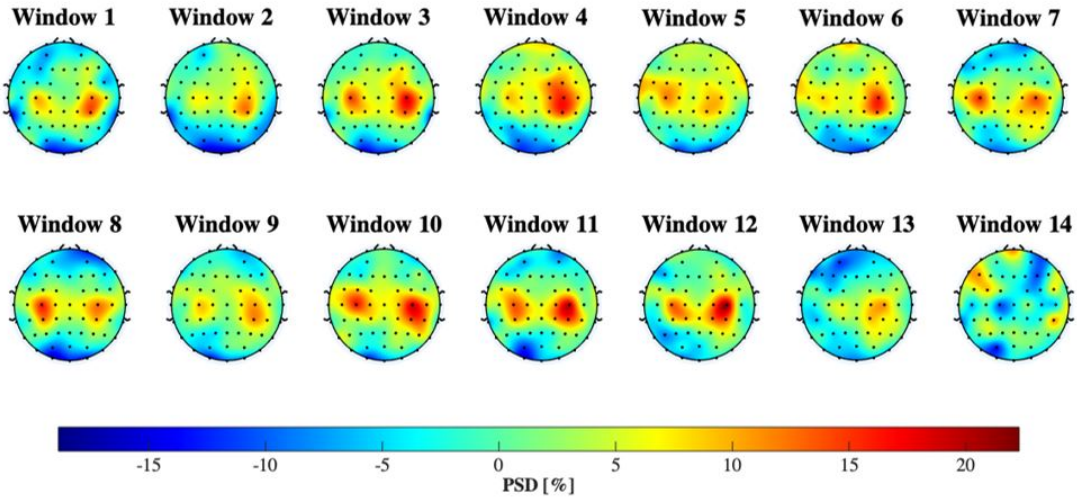
### 4.2.1 Spectral Analysis

As concerns the scalp maps, the images shown are relative to 1-minute windows, while the temporal trend considers different brain areas and hemispheres and takes into account 10-second windows for a better time resolution, as already discussed. Single subject scalp maps analysis reveals large heterogeneity in EEG spectral power characteristics within both groups. In order to have a more global vision of the difference between ADHD and healthy subjects, the median plots of the two groups, representing $\Delta$P%, are shown in the different frequency bands. In addition, the median scalp maps of resting state are displayed, as a tool to justify some results obtained.

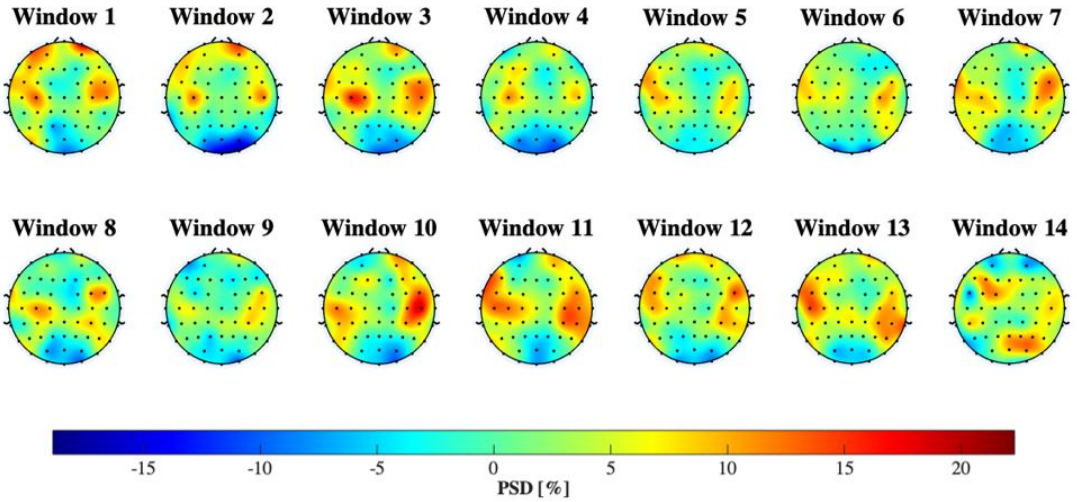
- **Delta (1-4 Hz)**

In median **Scalp Maps** of ADHD and Control groups, referred to the brain activity during CPT in the Delta frequency range and displayed in *Figure 4.2*, two spots of activation (i.e. synchronization or higher power with respect to the resting state) are clearly visible in both groups, but more evident in the right Central hemisphere in the ADHD group. Higher values of Delta power in ADHD group during task on Centro-Parietal regions were also found in Chen et al. studies [49] [48]. In the control group the synchronization seems to be located mainly in the Temporal zone.

Regarding the Fixation, the majority of findings in literature agrees on the presence of higher power in Delta rhythms in ADHD children with respect to their peers with typical development. Some of these studies proposed that an increase in Delta power reflects the maturation lag in ADHD children [23] [123]. Contrasting with the literature, the median Delta power activity along 1-minute baseline, illustrated in *Figure 4.3*, seems more synchronized in Central areas for the Controls but more spread over the whole scalp in the ADHD group. In fact, a lower activation in the Controls with respect to the ADHD group is found in Fronto-Temporal area.

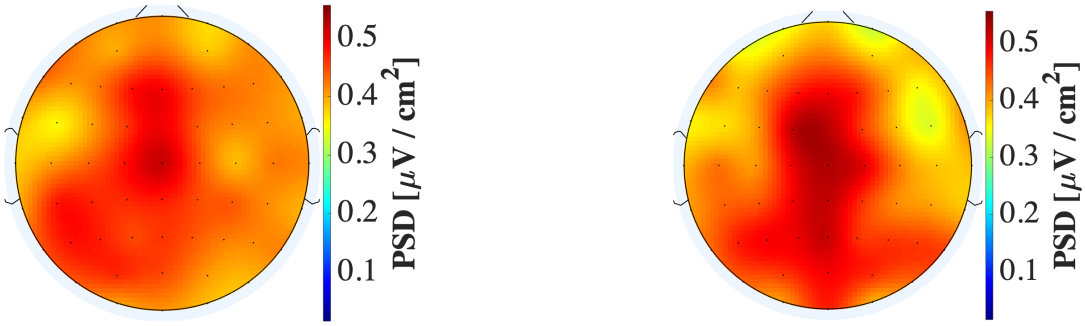



*(a) ADHD*




*(b) CONTROL*

***Figure 4.2:*** *Median Scalp Maps of ADHDs (a) and Controls (b) in Delta band (1-4 Hz), relative to each 1-minute window of CPT: the Power Spectral Density percentage variation (PSD %) with respect to the baseline is displayed. The minimum and the maximum values of the colorbar were chosen according to the minimum and maximum values of the median of the ADHD and CONTROL population during CPT.*



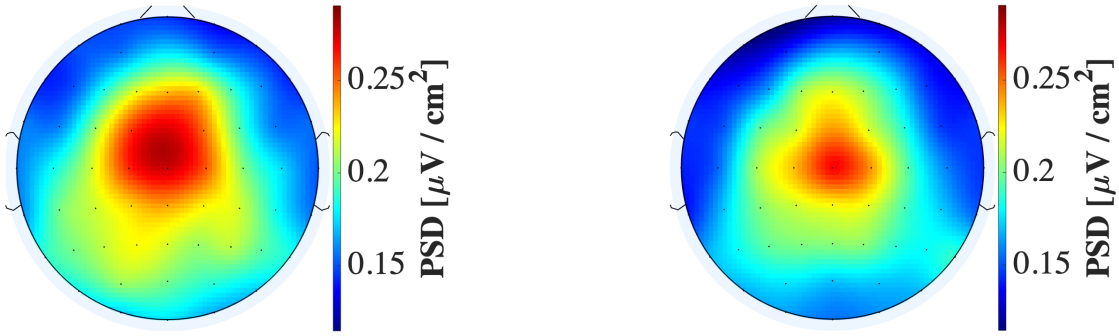

***(a) ADHD*** ***(b) CONTROL***

***Figure 4.3:*** *Median Scalp Maps of ADHDs (a) and Controls (b) in Delta (1-4 Hz) band, relative to 1-minute of eyes-open resting state: the Power Spectral Density (PSD) is displayed with a unit of measure equal to $\mu V/cm^2$. The minimum and the maximum values of the colorbar were chosen according to the minimum and maximum values of the median of the ADHD and CONTROL population during the baseline.*

No additional interesting results were found from the higher time resolution display of temporal trends.

- **Theta (4-8 Hz)**

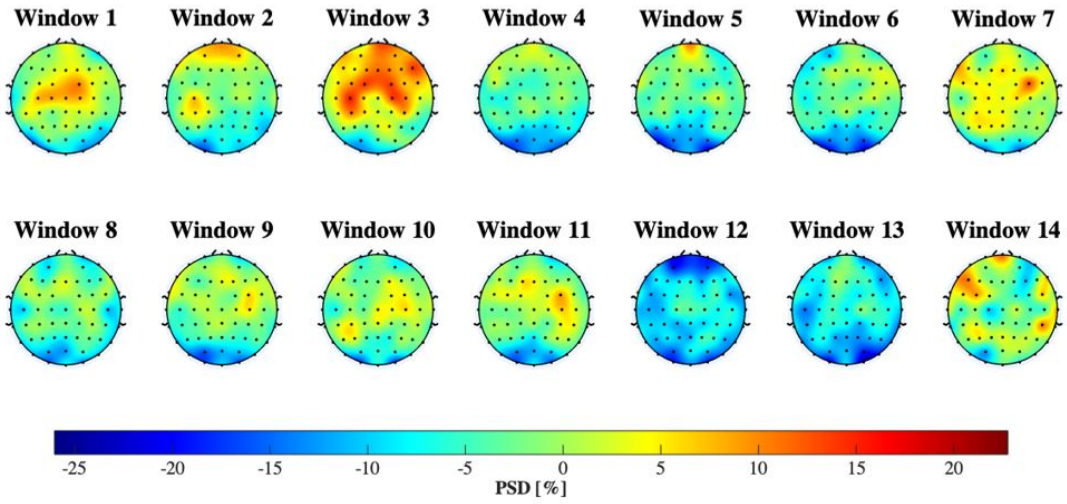
As depicted in *Figure 4.4*, showing the median **Scalp Maps** during eyes-open resting condition in Theta band, ADHD patients are characterized by a higher Theta power located in the Fronto-Central lobe with respect to the control group. Abnormally increased Theta power during resting condition is one of the most robust electroencephalographic anomalies in ADHD, mainly in Frontal brain regions, and this has been linked to a decrease in attention [81] and to drowsy and unfocused states [105].



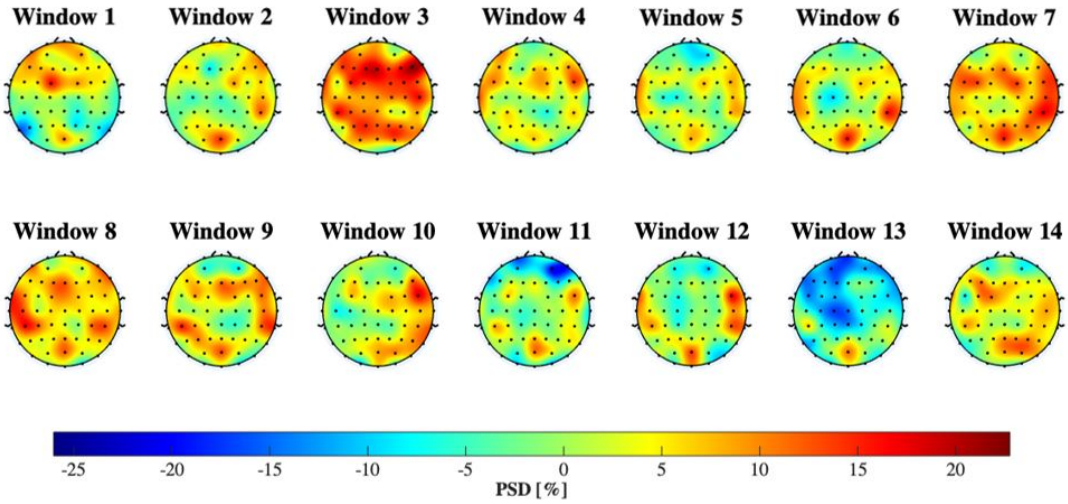

***(a) ADHD*** ***(b) CONTROL***

***Figure 4.4:*** *Median Scalp Maps of ADHDs (a) and Controls (b) in Theta band (4-8 Hz), relative to 1-minute of eyes-open resting state: the Power Spectral Density (PSD) is displayed with a unit of measure equal to $\mu V/cm^2$. The minimum and the maximum values of the colorbar were chosen according to the minimum and maximum values of the median of the ADHD and CONTROL population during the baseline.*

Focusing on CPT, as we can see in *Figure 4.5*, some differences between ADHD and control children could be observed. From an overall point of view, the control group shows a greater sparse Theta $\Delta$P% in some windows. Focusing on the single windows, in both groups, in window 3, there is an evident increase in Theta power. In general, an increase in Theta power is appreciated in the Fronto-Parietal area, as could be observed partly in these results. In Sauseng et al. [159] this increased Frontal Theta activity was associated with a higher level of cognitive task demand, as also observed by Gevins et al. [70].
In ADHD scalp map there is a desynchronization of Theta rhythm in the Parieto-Occipital lobe, not present in the median plot of healthy subjects.




*(a) ADHD*

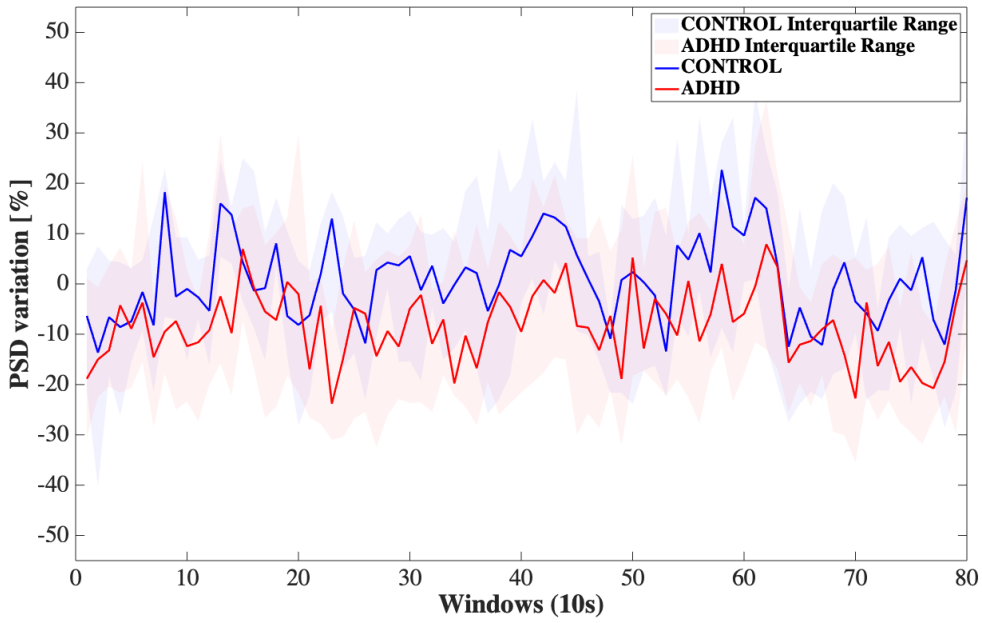



*(b) CONTROL*

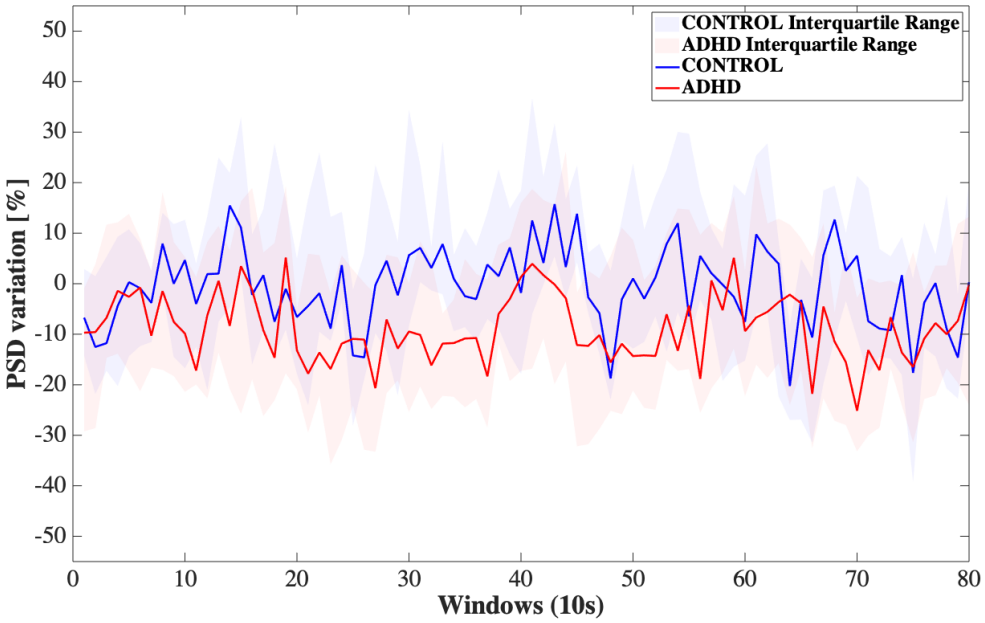
***Figure 4.5:*** *Median Scalp Maps of ADHDs (a) and Controls (b) in Theta band (4-8 Hz), relative to each 1-minute window of CPT: the Power Spectral Density percentage variation (PSD %) with respect to the baseline is displayed. The minimum and the maximum values of the colorbar were chosen according to the minimum and maximum values of the median of the ADHD and CONTROL population during CPT.*

Notable results in Theta band median **Temporal Trend** have been found in the Right and Left Parieto-Occipital zone, as shown in *Figure 4.6*, where a greater difference in values could be appreciated in both lobes, with higher values exhibited by healthy subjects. This is coherent with what shown in the scalp maps, illustrated in *Figure 4.5*.

A similar outcome is obtained in the Temporal cerebral regions, in *Figure 4.7*, where higher values relative to Controls are shown.

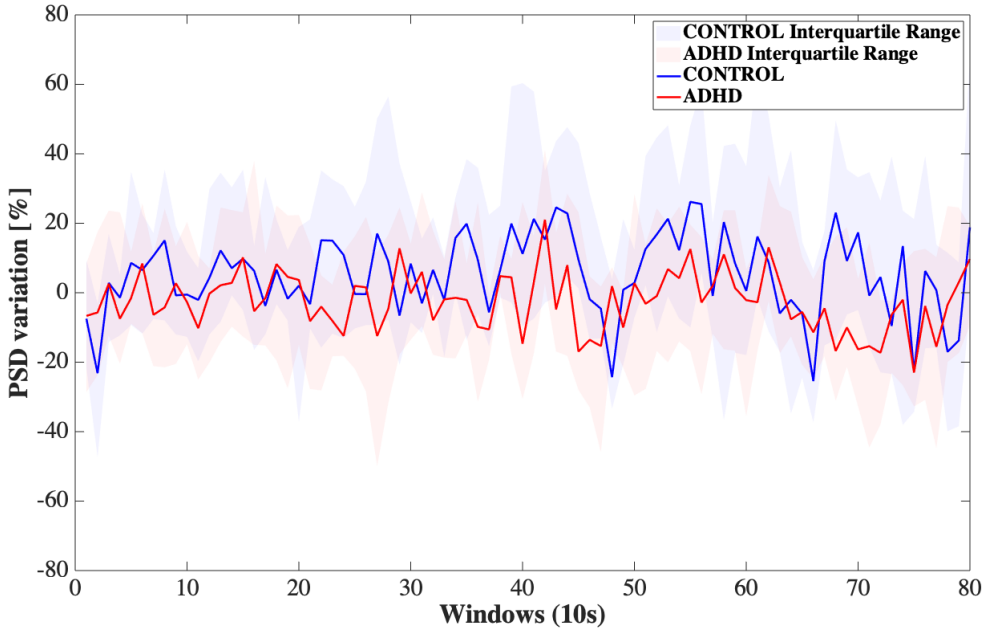



*(a) Right Parieto-Occipital*

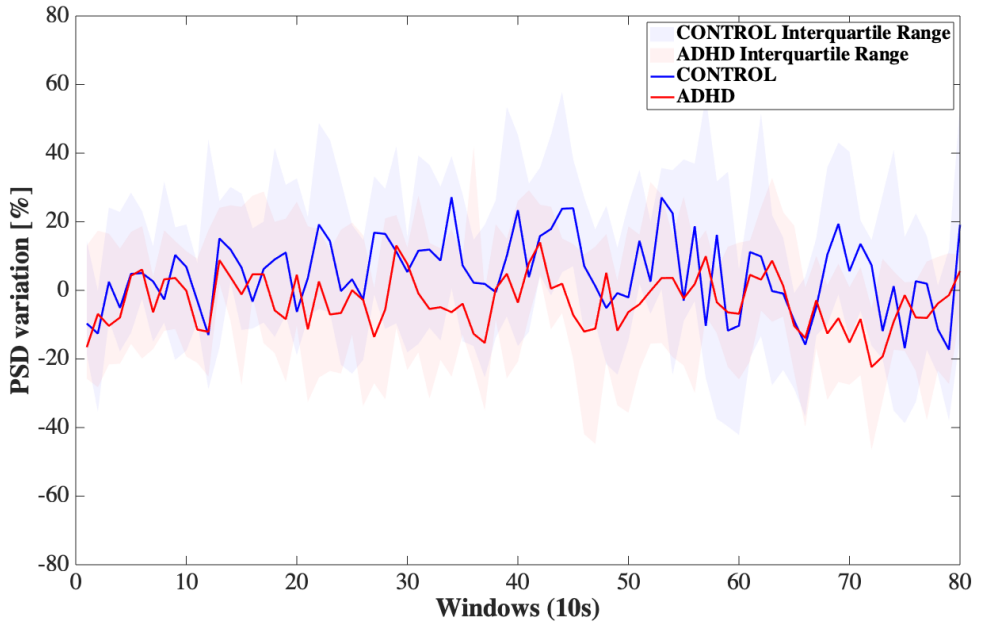



*(b) Left Parieto-Occipital*

***Figure 4.6:*** *Median Temporal Trend of the Power Spectral Density (PSD) variation with respect to the baseline in Theta (4-8 Hz) band, in Right (a) and Left (b) Parieto-Occipital area. The interquartile range is plotted around the median temporal trends relative to the Controls (blue) and ADHDs (red).*




*(a) Right Temporal*




*(b) Left Temporal*

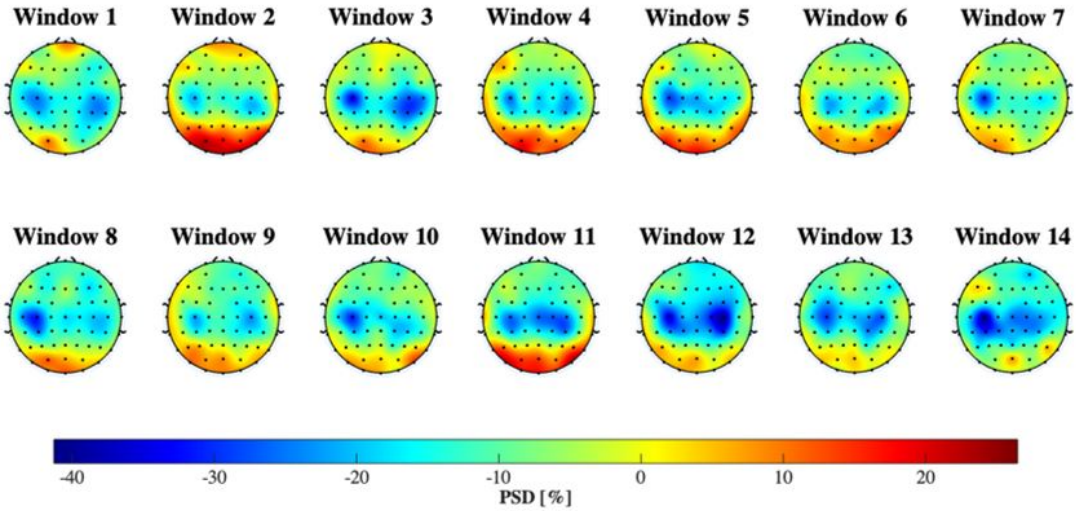
***Figure 4.7:*** *Median Temporal Trend of the Power Spectral Density (PSD) variation with respect to the baseline in Theta (4-8 Hz) band, in Right (a) and Left (b) Temporal area. The interquartile range is plotted around the median temporal trends relative to the Controls (blue) and ADHDs (red).*

- **Alpha (8-13 Hz)**

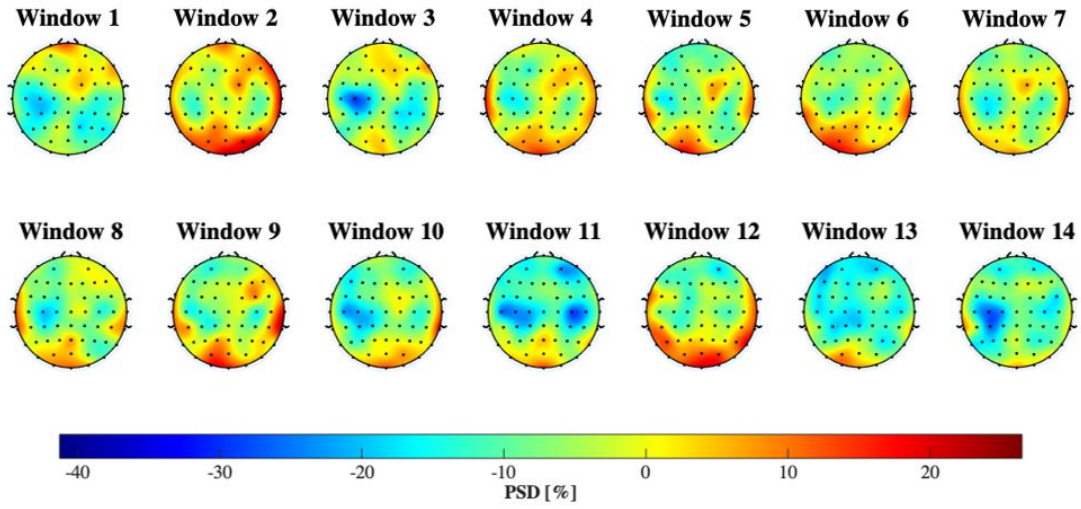
In *Figure 4.8* the median **Scalp Maps** of CPT in Alpha band are shown. From this representation it seems that ADHD subjects are characterized by a higher desynchronization of $\Delta$P%, more evident in the Central area, with respect to the Controls. Regarding the eyes-open resting state, shown in *Figure 4.9*, the baseline of ADHD is characterized by a higher synchronization in Central and Parieto-Occipital regions; this may indicate that the ADHD participants hardly suppress Alpha activity during eyes-open baseline compared to Controls. Klimesh [96] hypothesized that an increased Alpha power during resting state was related to impaired cognitive performance both in ADHD and control children.

Focusing on different cerebral areas, at least two types of Alpha rhythmicity can be identified: Posterior (Parieto-Occipital) Alpha rhythms and Sensory-motor (Central) Alpha rhythms, also called Mu-Rhythms [7].

*Posterior Alpha rhythms* are recorded over Parieto-Occipital areas and these rhythms are usually more accentuated (synchronization) when the subject has the eyes closed, while they are mitigated (desynchronization) in response to visual stimuli [7]. In the current work, the power of Parieto-Occipital rhythms is higher in the task condition when compared with the eyes-open condition (reaching at maximum 20% of PSD variation) for both groups in most of the analyzed 1-minute windows. In particular, this Parieto-Occipital synchronization is more evident in the ADHD group.
A possible explanation could be the difficulty for ADHD children in attending to and processing visual stimuli as efficiently as children without ADHD. In support of this interpretation, in the study conducted by Lenartowicz et al. [106] it was observed that the decrease in Alpha is attenuated in tasks involving visual selective attention, especially in ADHD inattentive type.
However, the role of Parieto-Occipital oscillations and how the different components of a task (processing of external stimuli, internal elaboration, and task demand) interact to affect Alpha power are still unclear [118].

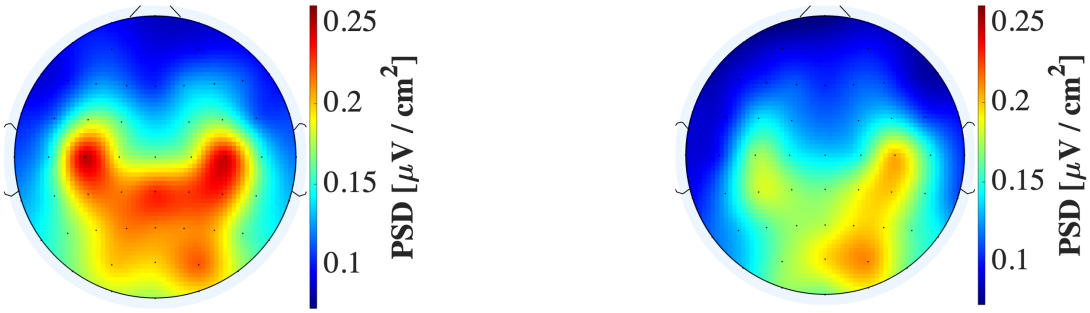



*(a) ADHD*




*(b) CONTROL*

***Figure 4.8:*** *Median Scalp Maps of ADHDs (a) and Controls (b) in Alpha band (8-13 Hz), relative to each 1-minute window of CPT: the Power Spectral Density percentage variation (PSD %) with respect to the baseline is displayed. The minimum and the maximum values of the colorbar were chosen according to the minimum and maximum values of the median of the ADHD and CONTROL population during CPT.*




***(a) ADHD*** ***(b) CONTROL***

***Figure 4.9:*** *Median Scalp Maps of ADHDs (a) and Controls (b) in Alpha band (8-13 Hz), relative to 1-minute of eyes-open resting state: the Power Spectral Density (PSD) is displayed with a unit of measure equal to $\mu V/cm^2$. The minimum and the maximum values of the colorbar were chosen according to the minimum and maximum values of the median of the ADHD and CONTROL population during the baseline.*

*Mu-Rhythms* are a particular type of Alpha waves, localized on the sensory-motor strip of the cortex (for this reason they are also called sensory-motor Alpha rhythms). These rhythms can be inhibited by the corresponding hand movement, while the muscle relaxation enhances them [7].
In the ADHD group, the desynchronization of these rhythms is more evident with respect to the Controls. Since the task is performed with the right hand, the suppression of Mu-Rhythms is more accentuated in contralateral (left) hemisphere. However, two blue spots can be observed in the central area in both hemispheres, since an ipsilateral desynchronization happens immediately after the controlateral one [177]. The reason why the lateralization could not be appreciated is due to the low temporal resolution (1 minute).
Since hyperactivity is one of the ADHD main symptoms, the higher desynchronization of these rhythms could be interpreted as a higher difficulty in ADHD children to remain calm during the task. However, this hypothesis could not be strengthened, due to the lack of the ADHD subtypes (inattentive and hyperactive).

Regarding the median **Temporal Trend** of the Alpha $\Delta$P%, an interesting decrease during CPT can be seen in the Right Frontal hemisphere, shown in *Figure 4.10*. Moreover, the PSD values belonging to healthy subjects are slightly higher than those of ADHD ones. This difference in PSD values is further more evident in the Right Central area, as shown in *Figure 4.11*.

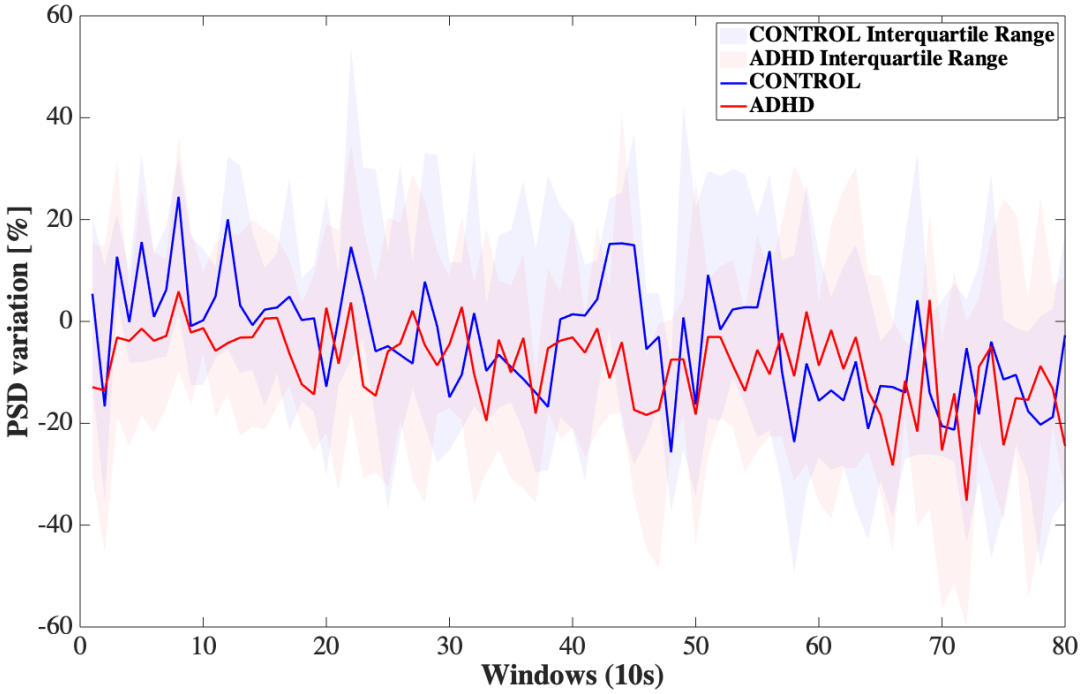



***Figure 4.10:*** *Median Temporal Trend of the Power Spectral Density (PSD) variation with respect to the baseline in Apha (8-13 Hz) band, in Right Frontal area. The interquartile range is plotted around the median temporal trends relative to the Controls (blue) and ADHDs (red).*

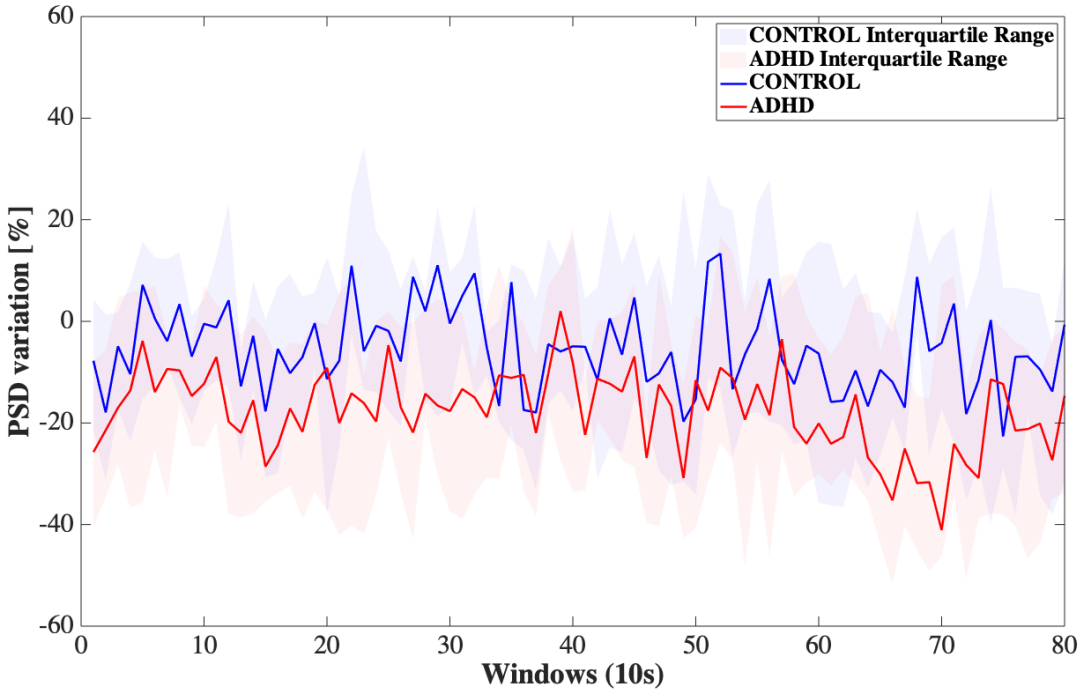
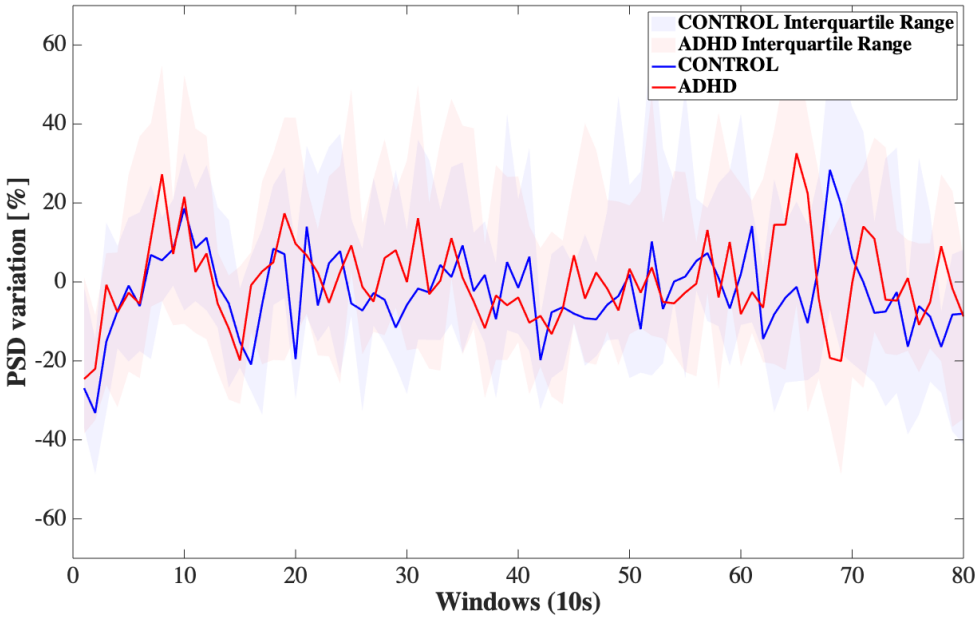
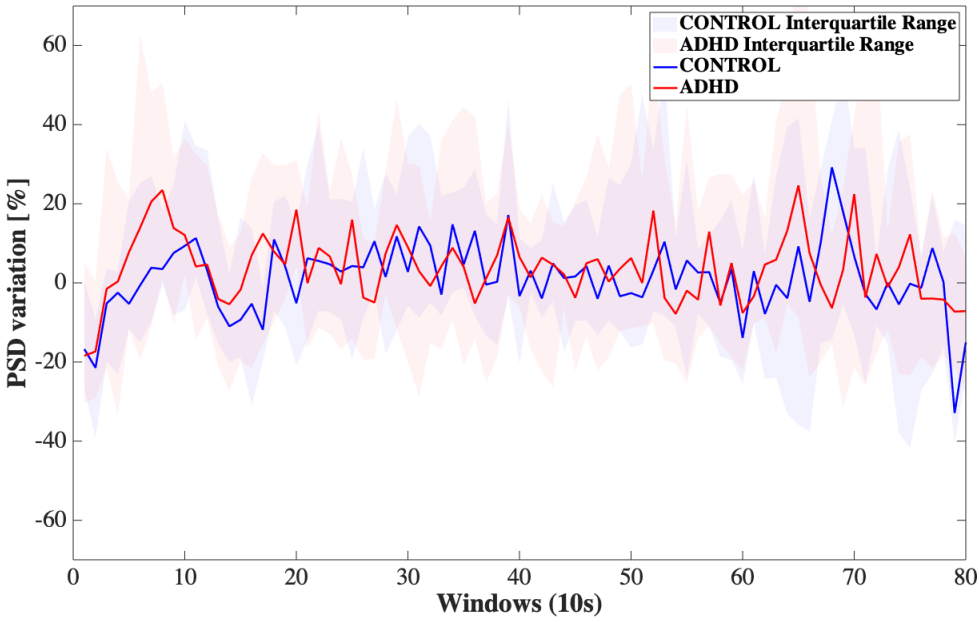



***Figure 4.11:*** *Median Temporal Trend of the Power Spectral Density (PSD) variation with respect to the baseline in Apha (8-13 Hz) band, in Right Central area. The interquartile range is plotted around the median temporal trends relative to the Controls (blue) and ADHDs (red).*

The last interesting result in Alpha band is relative to the Parieto-Occipital regions, illustrated in *Figure 4.12*, where, at the beginning of the task, a common ascending trend for both groups until the 10th window (100 sec) is observed, followed by a decrease and an assessment of the $\Delta$P%. The peak could be probably due to the start of the visual stimulus, while the assessment could be interpreted as a habituation to the task.




*(a) Right Parieto-Occipital*



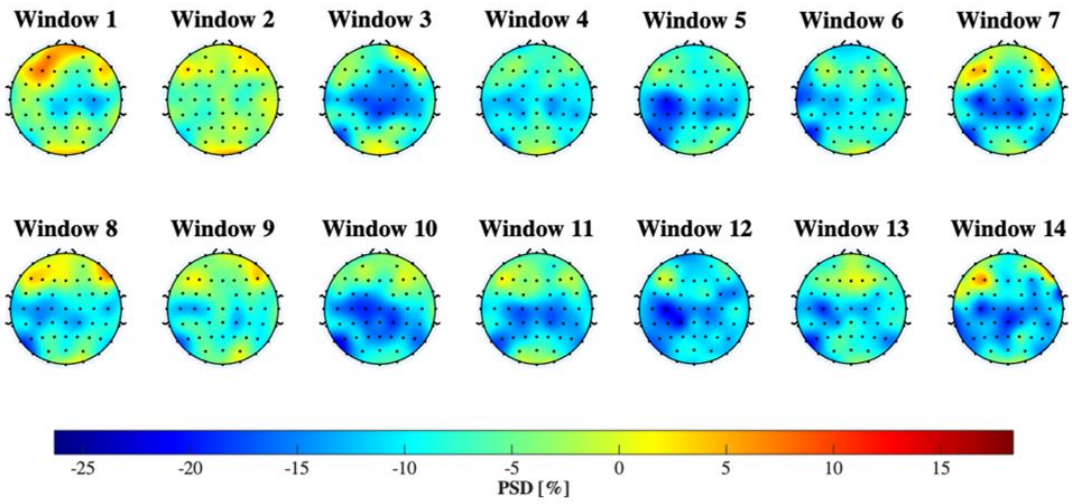

*(b) Left Parieto-Occipital*

***Figure 4.12:*** *Median Temporal Trend of the Power Spectral Density (PSD) variation with respect to the baseline in Apha (8-13 Hz) band, in Right (a) and Left (b) Parieto-Occipital area. The interquartile range is plotted around the median temporal trends relative to the Controls (blue) and ADHDs (red).*

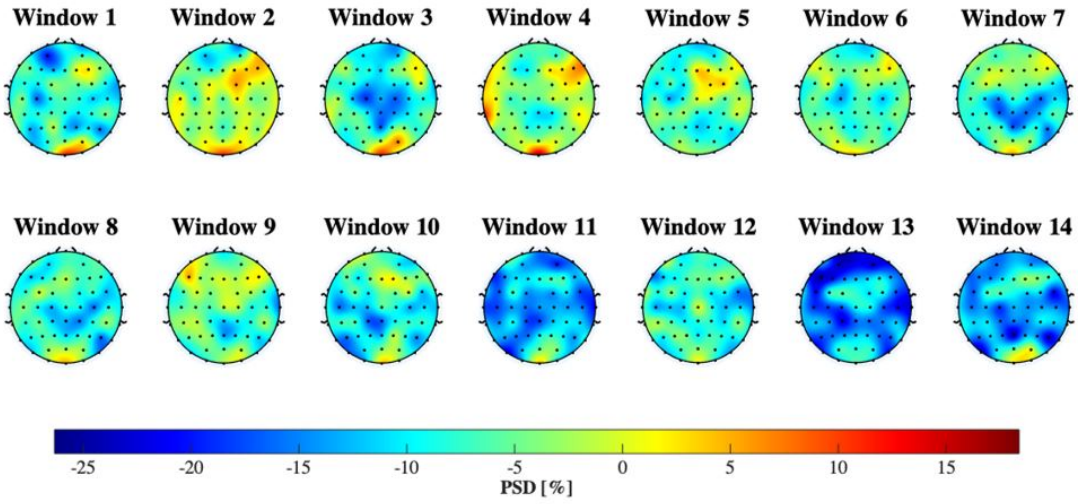
- **Beta 1 (13-22 Hz)**

In *Figure 4.13*, representing Beta 1 median **Scalp Maps** during CPT, a desynchronization in the central region could be noted for ADHD group, in almost all the windows. Furthermore, during the baseline condition, as we can see in *Figure 4.14*, where Beta 1 baseline activity is reported, Controls show a globally higher Beta 1 power, as expected [115], with a higher activation on the Occipital area with respect to ADHDs.

Surprisingly, during the task the control group does not show visible activation, as expected from the literature [108] [113], but only a slight synchronization (reaching at maximum 10% of PSD variation) in Frontal area in some of the windows. In addition, most of the last windows are characterized by a decrease in the Beta 1 power. Since this is an unexpected result, the interpretation on the basis of the literature is difficult because no previous works, to the best of our knowledge, have dynamically analyzed sustained attention tasks in a population of children. A possible naïve interpretation on the basis of our data could be that a loss of attention occurred for Controls at the end of the task. The fact that during CPT no synchronization of the Beta 1 power is observed might be due to a low task difficulty.

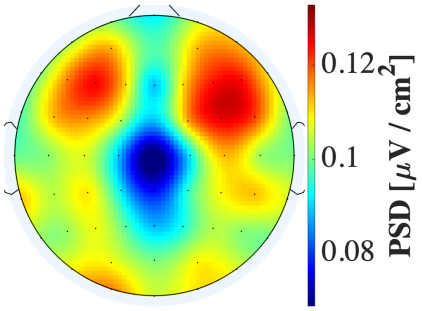



*(a) ADHD*

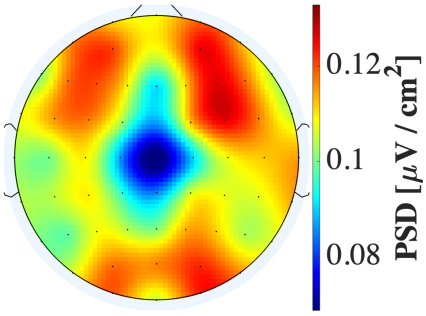



*(b) CONTROL*

***Figure 4.13:*** *Median Scalp Maps of ADHDs (a) and Controls (b) in Beta 1 band (13-22 Hz), relative to each 1-minute window of CPT: the Power Spectral Density percentage variation (PSD %) with respect to the baseline is displayed. The minimum and the maximum values of the colorbar were chosen according to the minimum and maximum values of the median of the ADHD and CONTROL population during CPT.*




*(a) ADHD*

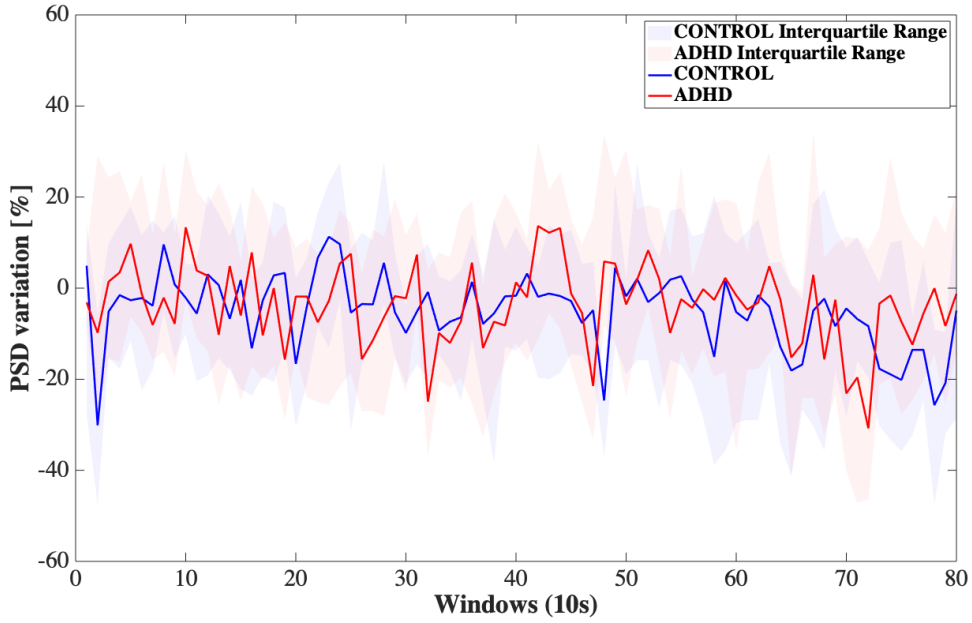



*(b) CONTROL*

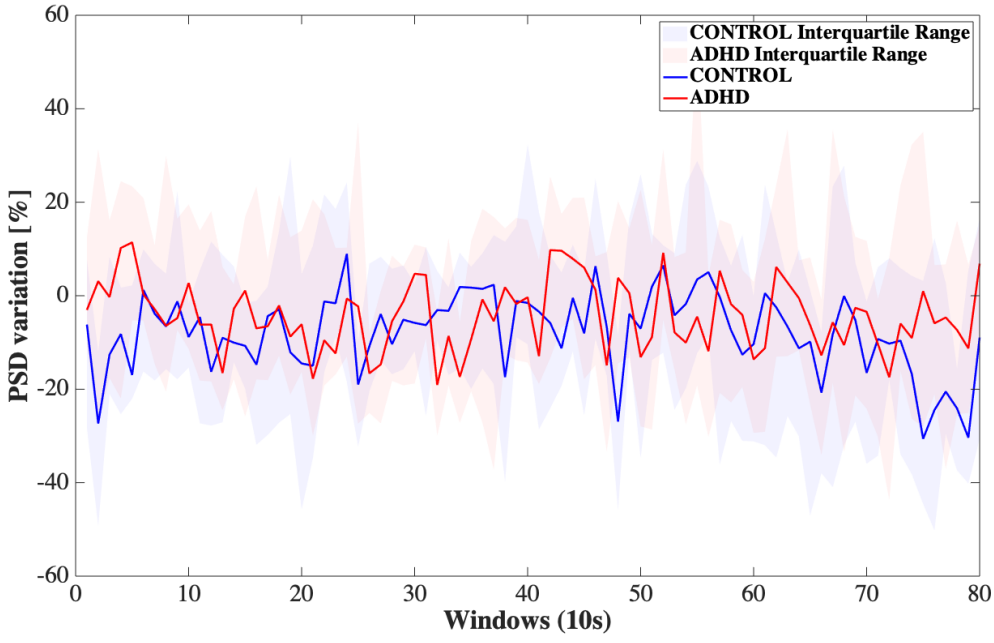
***Figure 4.14:*** *Median Scalp Maps of ADHDs (a) and Controls (b) in Beta 1 (13-22 Hz) band, relative to 1-minute of eyes-open resting state: the Power Spectral Density (PSD) is displayed with a unit of measure equal to $\mu V/cm^2$. The minimum and the maximum values of the colorbar were chosen according to the minimum and maximum values of the median of the ADHD and CONTROL population during the baseline.*

Observing the median **Temporal Trend** in the Frontal brain area, visualized in *Figure 4.15*, higher values of the control group compared with the ADHD one would be expected, since frontal lobes are supposed to be related to cognitive functions and a poor activity in this area could underline inability to focus on a task. On the contrary, either in the right and left frontal region there is no evident overall difference between the two classes.

Nevertheless, a difference is visible in the Left Parieto-Occipital region, as shown in *Figure 4.16*, where the control trend is slightly less negative compared to ADHD trend.




*(a) Right Frontal*




*(b) Left Frontal*

***Figure 4.15:*** *Median Temporal Trend of the Power Spectral Density (PSD) variation with respect to the baseline in Beta 1 (13-22 Hz) band, in Right (a) and Left (b) Frontal area. The interquartile range is plotted around the median temporal trends relative to the Controls (blue) and ADHDs (red).*

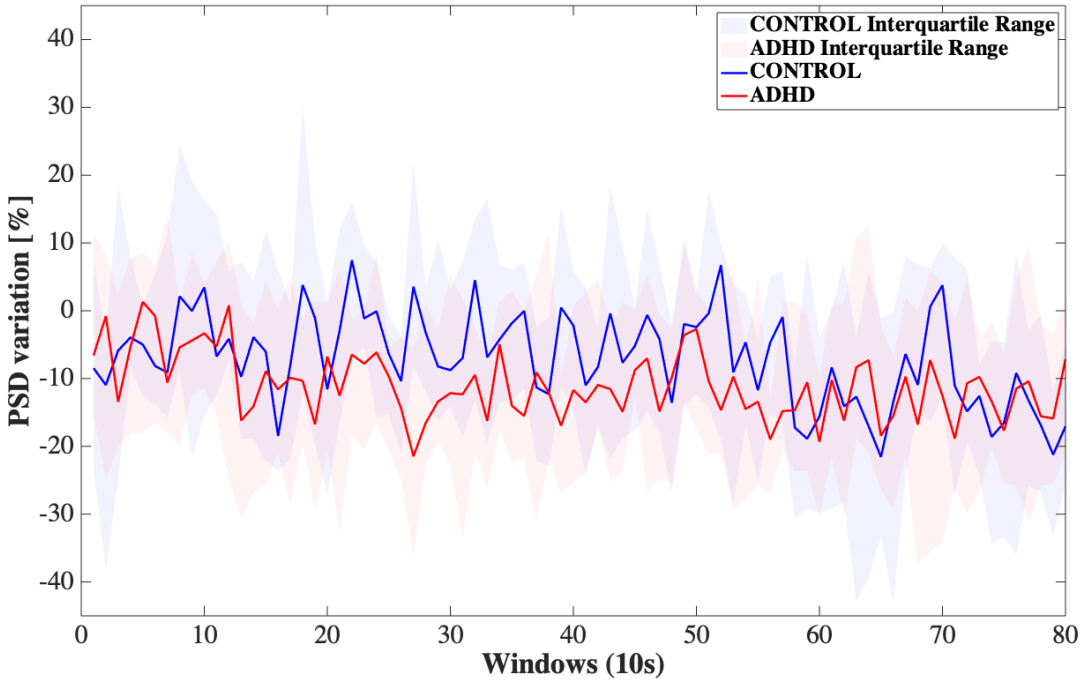


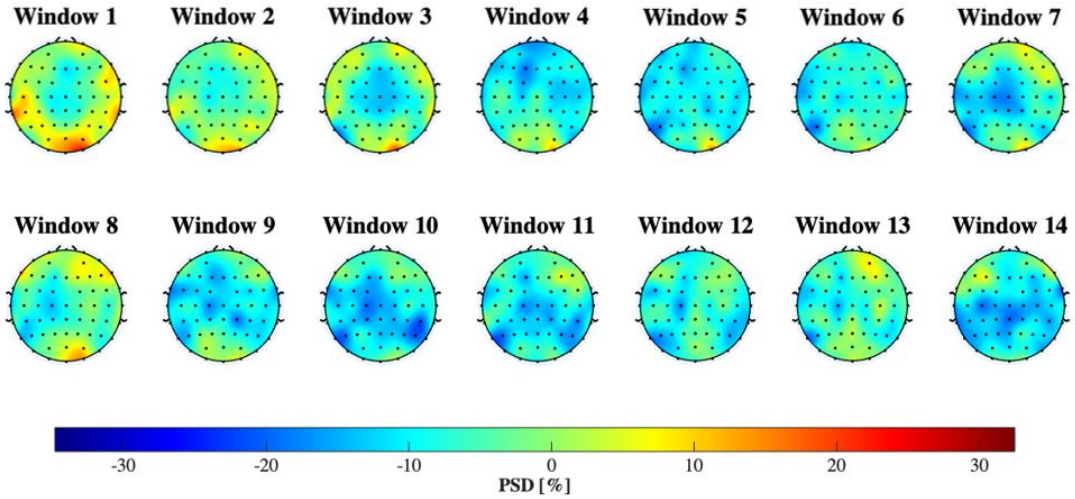

***Figure 4.16:*** *Median Temporal Trend of the Power Spectral Density (PSD) variation with respect to the baseline in Beta 1 (13-22 Hz) band, in Left Parieto-Occipital area. The interquartile range is plotted around the median temporal trends relative to the Controls (blue) and ADHDs (red).*

- **Beta 2 (22-30 Hz)**

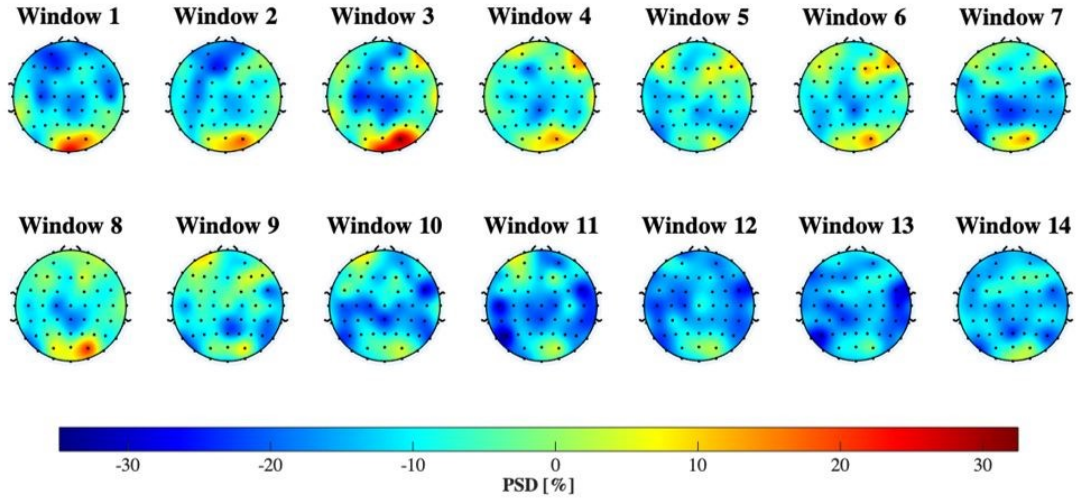
Exploring high Beta oscillations **Scalp Maps** of CPT, illustrated in *Figure 4.17*, a difference between ADHD children and their healthy peers can be observed in the Occipital area, where the activity seems increased in Controls. No significant results have been found in literature, since many studies do not split the Beta band in low and high frequencies, and most of them mainly investigate the low Beta band.

Observing the baseline activity in *Figure 4.18*, similar activations among the groups can be noticed, with a slightly higher synchronization in Controls in the Frontal area.

Since no noteworthy results were gained in the temporal trend analysis, they were not reported.

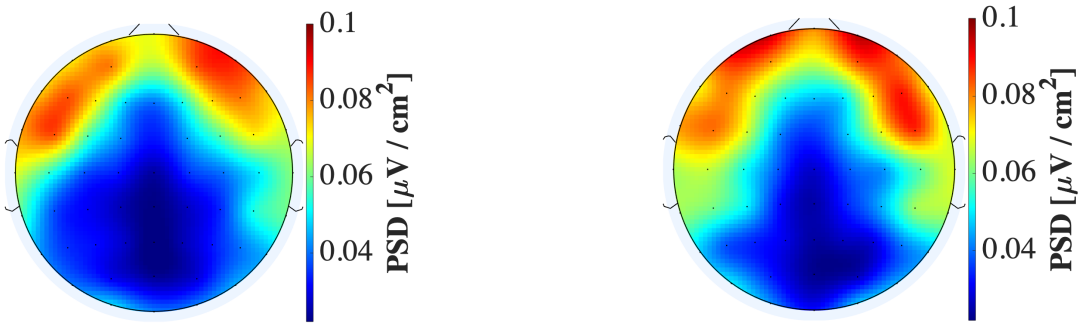



*(a) ADHD*




*(b) CONTROL*

***Figure 4.17:*** *Median Scalp Maps of ADHDs (a) and Controls (b) in Beta 2 band (22-30 Hz), relative to each 1-minute window of CPT: the Power Spectral Density percentage variation (PSD %) with respect to the baseline is displayed. The minimum and the maximum values of the colorbar were chosen according to the minimum and maximum values of the median of the ADHD and CONTROL population during CPT.*

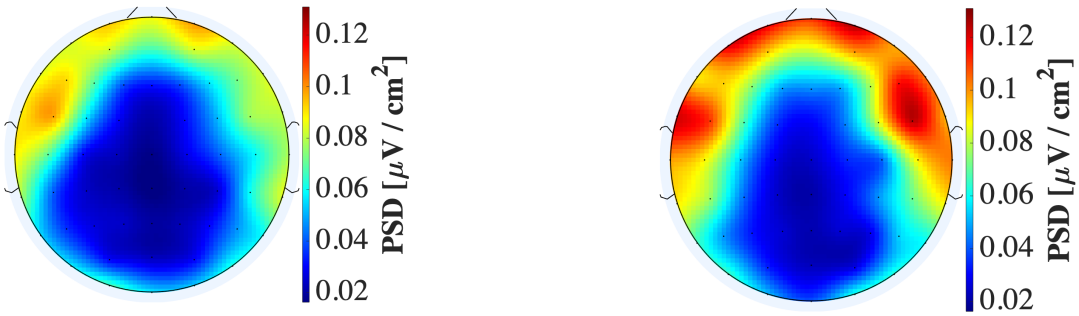



***(a) ADHD*** ***(b) CONTROL***

***Figure 4.18:*** *Median Scalp Maps of ADHDs (a) and Controls (b) in Beta 2 (22-30 Hz) band, relative to 1-minute of eyes-open resting state: the Power Spectral Density (PSD) is displayed with a unit of measure equal to $\mu V/cm^2$. The minimum and the maximum values of the colorbar were chosen according to the minimum and maximum values of the median of the ADHD and CONTROL population during the baseline.*

- **Gamma (30-45 Hz)**

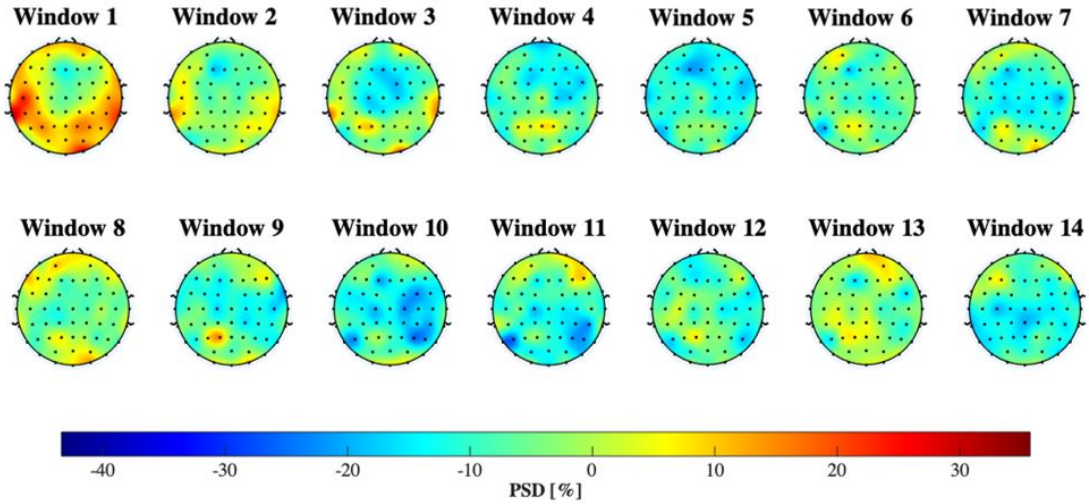
In *Figure 4.19*, Gamma median **Scalp Maps** of Controls during eyes-open condition show enhanced Gamma in the Fronto-Temporal region compared with the ADHD group. In Benasich et al. [29], reduced Gamma during eyes-open resting state and poor cognitive skills, attention and inhibition control were found to be directly linked. In general, the literature reports that enhanced Gamma activity, which is associated with greater attention, precedes correct responses in a target detection task. Thus, if Gamma activity is important in cognitive and attentional processing, reduced spontaneous Gamma activity in ADHD subjects with respect to control ones may imply deficits associated with the disorder.



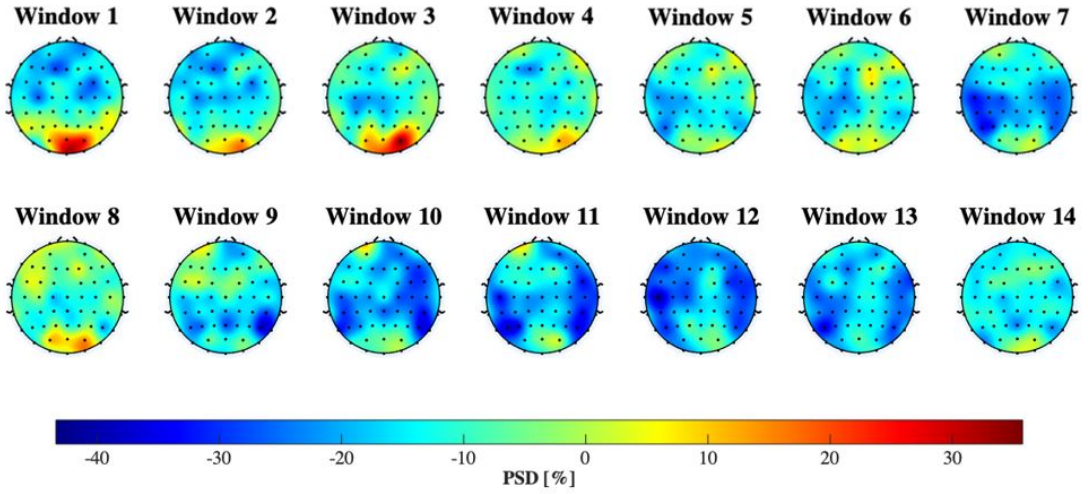

***(a) ADHD*** ***(b) CONTROL***

***Figure 4.19:*** *Median Scalp Maps of ADHDs (a) and Controls (b) in Gamma (30-45 Hz) band, relative to 1-minute of eyes-open resting state: the Power Spectral Density (PSD) is displayed with a unit of measure equal to $\mu V/cm^2$. The minimum and the maximum values of the colorbar were chosen according to the minimum and maximum values of the median of the ADHD and CONTROL population during the baseline.*

Shifting focus to CPT, *Figure 4.20* reveals a sparse desynchronization from window 7 to 13, except for window 8, in the control group.




*(a) ADHD*




*(b) CONTROL*

***Figure 4.20:*** *Median Scalp Maps of ADHDs (a) and Controls (b) in Gamma band (30-45 Hz), relative to each 1-minute window of CPT: the Power Spectral Density percentage variation (PSD %) with respect to the baseline is displayed. The minimum and the maximum values of the colorbar were chosen according to the minimum and maximum values of the median of the ADHD and CONTROL population during CPT.*

#### 4.2.1.1 Correlation between EEG and Performance indices

From the scalp maps results, a higher Alpha desynchronization in the motor cortex of ADHD subjects, shown in *Figure 4.8*, was underlined, probably due to hyperactivity. To explore whether the supposed increased motor activity during cognitive task influenced the subject performance, a median value of ΔP% in Alpha band in Right and Left Central regions was obtained per subject, in order to compute the correlation with the performance indices.

As in the previous analyses, both hemispheres were analyzed separately, but the corresponding electrodes were chosen differently:

- *Left Central*: C3, FC1, CP1, FC5, CP5 C1, FC3, CP3, C5
- *Right Central*: C4, FC2, CP2, FC6, CP6, C2, FC4, CP4, C6

The correlation was performed, after having removed the outliers, through the method mentioned before. No significant correlations were found either in the Right or Left Central areas for ADHD subjects, differently from what it was expected. On the contrary, significant correlations were found with the Commission Error and Hit Response Time for control group, as reported in the *Table 4.1*.

| **Alpha ΔP%** | **OE** | **CE** | **HRT** | **HRT SD** | **CPT Var** |
|---|---|---|---|---|---|
| **Right Central** | 0.070 | -0.607 * | 0.732 * | 0.283 | 0.278 |
| **Left Central** | 0.225 | -0.540 * | 0.563 * | 0.347 | 0.251 |

***Table 4.1:*** *Controls Correlation Table of Performance indices and Alpha median ΔP% in Right and Left Central lobes.*
** represents a significant correlation corresponding to a p-value < 0.05. Abbreviations → OE = Omission Error, CE = Commission Error, HRT = Hit Response Time, HRT SD = Standard Deviation of Hit Response Time, CPT Var = CPT Variability.*

In response to these results, a further analysis was carried on, by correlating the values of ΔP% relative to each one-minute window (for a total of 13 windows) with the behavioral indices. This decision was taken in order not to loose the temporal information, coherently with the previous analyses. The choice to correlate values relative to each window (Alpha EEG index) with values belonging to the whole task (performance indices) was obliged, since values of behavioral indices after every stimulus were not acquired.

Firstly, boxplots of ADHD and control ΔP%, for all the 13 windows in Right Central and Left Central regions, were inspected, in order to see if a different behavior was present in particular windows, with the aim of focusing the attention more on these windows when performing the correlation. Since the behavior within groups was mostly constant along the entire duration of CPT, all the windows were taken into account.

In confirmation of the previous global results, no significant correlations were found for ADHD subjects, while significant correlations in some windows were obtained for control group in the same indices mentioned above. The correlation tables of Control subjects in Right Central and Left Central brain areas are visualized in *Table 4.2* and in *Table 4.3*, respectively.

| **Alpha ΔP%** | **OE** | **CE** | **HRT** | **HRT SD** | **CPT Var** |
|---|---|---|---|---|---|
| **Window 1** | 0.490 | -0.530 | 0.661 * | 0.219 | 0.590 |
| **Window 2** | -0.243 | -0.305 | 0.574 | -0.137 | -0.037 |
| **Window 3** | 0.286 | -0.133 | 0.361 | 0.091 | 0.397 |
| **Window 4** | -0.504 | -0.425 | 0.294 | -0.119 | -0.219 |
| **Window 5** | -0.131 | -0.449 | 0.473 | -0.041 | 0.114 |
| **Window 6** | -0.208 | -0.442 | 0.396 | -0.087 | 0.174 |
| **Window 7** | 0.279 | -0.312 | 0.515 | 0.397 | 0.224 |
| **Window 8** | 0.162 | -0.740 * | 0.851 ** | 0.333 | 0.301 |
| **Window 9** | -0.205 | -0.470 | 0.055 | -0.299 | -0.109 |
| **Window 10** | -0.224 | -0.384 | 0.310 | -0.043 | 0.061 |
| **Window 11** | -0.050 | -0.292 | 0.456 | 0.463 | 0.456 |
| **Window 12** | -0.332 | -0.635 * | 0.364 | 0.155 | -0.183 |
| **Window 13** | -0.063 | -0.761 * | 0.683 * | 0.425 | 0.205 |

***Table 4.2:*** *Controls Correlation Table of Alpha Right Central ΔP% and Performance indices. * represents a significant correlation corresponding to a p-value < 0.05; ** corresponding to a p-value < $10^{-4}$. Abbreviations → OE = Omission Error, CE = Commission Error, HRT = Hit Response Time, HRT SD = Standard Deviation of Hit Response Time, CPT Var = CPT Variability.*

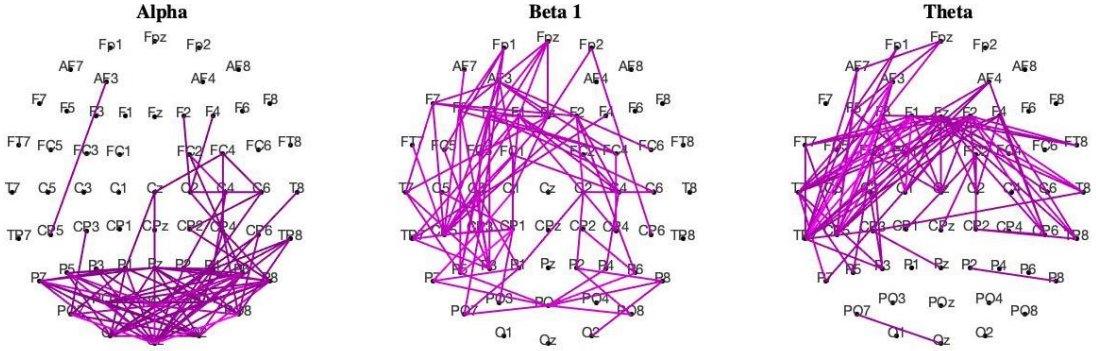
| Alpha ΔP% | OE | CE | HRT | HRT SD | CPT Var |
|---|---|---|---|---|---|
| **Window 1** | 0.342 | -0.561 | 0.844 ** | 0.215 | 0.292 |
| **Window 2** | 0.018 | -0.460 | 0.732 * | -0.018 | 0.009 |
| **Window 3** | 0.205 | -0.386 | 0.494 | 0.123 | 0.219 |
| **Window 4** | 0.120 | -0.340 | 0.389 | 0.384 | 0.429 |
| **Window 5** | 0.134 | -0.484 | 0.487 | 0.160 | 0.292 |
| **Window 6** | 0.208 | -0.614 * | 0.445 | 0.333 | 0.324 |
| **Window 7** | 0.081 | -0.449 | 0.389 | -0.023 | 0.027 |
| **Window 8** | 0.286 | -0.435 | 0.606 * | 0.416 | 0.416 |
| **Window 9** | 0.205 | -0.475 | 0.305 | -0.122 | 0.377 |
| **Window 10** | -0.105 | -0.228 | 0.150 | 0.018 | 0.292 |
| **Window 11** | 0.342 | -0.311 | 0.688 * | 0.555 | 0.419 |
| **Window 12** | -0.071 | -0.365 | 0.200 | 0.402 | 0.009 |
| **Window 13** | 0.335 | -0.593 * | 0.602 * | 0.416 | 0.283 |

***Table 4.3:*** *Controls Correlation Table of Alpha Left Central ΔP% and Performance indices. * represents a significant correlation corresponding to a p-value < 0.05; ** corresponding to a p-value < $10^{-4}$. Abbreviations → OE = Omission Error, CE = Commission Error, HRT = Hit Response Time, HRT SD = Standard Deviation of Hit Response Time, CPT Var = CPT Variability.*

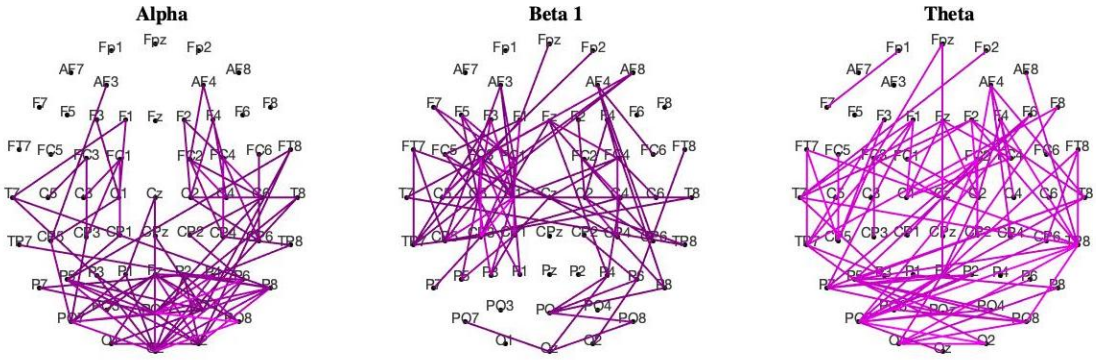
A negative correlation with CE and a positive correlation with HRT was noticed in all the windows in both the correlation tables of Controls. This means that an increasing in Alpha Central brain activity corresponds to a lower number of Commission Errors and a higher time spent to give a response. Higher Alpha values in sensory-motor area could signify that the Controls are more calm and tend to move less during the task, probably concentrating more and, consequently, committing less errors. This observation, together with the fact that ADHD children commit more CE, might reinforce the previous hypothesis that the desynchronization in Alpha, found in the Central area in ADHD children, could be due to hyperactivity. In addition, the higher HRT means that the patient spends more time thinking about the response, not being impulsive. Indeed, impulsiveness and hyperactivity are often found together in the ADHD-hyperactive type.
Finally, it is important to underline that the correlations resulted significant were found only in these two performance indices, with more significant windows in the Left Central lobe.

### 4.2.2 Functional Connectivity Analysis

*Figure 4.21* shows the median graphs of the entire duration of CPT, calculated in each band for patients with ADHD and Healthy Controls.




***(a) ADHD***




***(b) CONTROL***

***Figure 4.21:*** *Median CPT Connectivity Graphs, obtained through ImCoh, of ADHD (top) and control (below) subjects. Only the statistically significant edges (p-value $< 0.05$) are shown. The brighter the link color, the stronger the functional connectivity.*

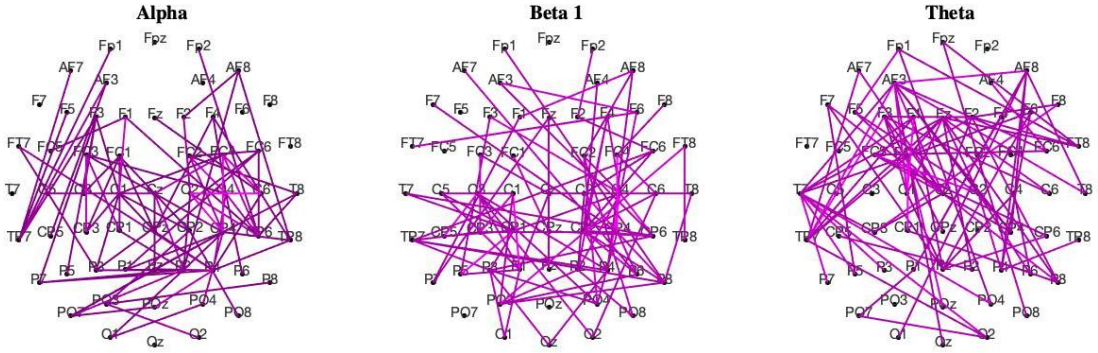
Regarding the **Alpha band**, in Controls there is an information flow among the Parieto-Occipital and Central brain regions, while ADHD patients are characterized by an information flow mostly limited to the only Parieto-Occipital area, indicating a higher network segregation. Segregation in the Occipital area in ADHD patients was also observed in Ansari et al. [15]. The dense connections visible in the

Occipital lobe, for both groups, might be due to the visual task, since the Occipital lobe is responsible for receiving and integrating visual information, also revealed by Li et al.'s study [47]. Hence, defects found in the connectivity in this area may reveal a brain function deficit during visual processing and may be due to a higher inattention of ADHD group to the visual stimulus [15].

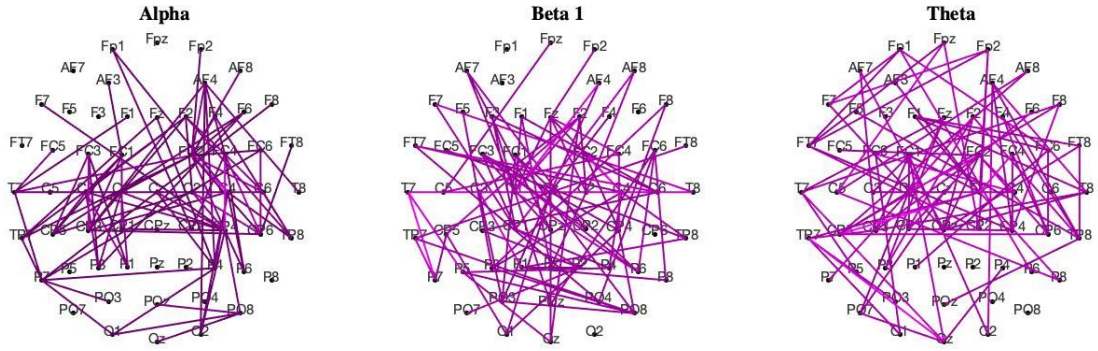
In **Theta band**, it can be noticed that communication in Controls occurs among the entire brain, except for the prefrontal region. On the contrary, the ADHD graph in Theta shows an accentuated communication between nodes of the Central and Temporal areas, more strongly in the left hemisphere.

This lateralization characterizing the ADHD patients, is also evident in **Beta 1**, in which functional connectivity patterns are present in the left hemisphere among all regions, except for the Occipital area. In many studies regarding rest, simple and complex perfomance task [46] [56] [57], this atypical lateralized brain activity of ADHD has been observed; in particularly, in Saleh M. H. Mohamed et al. research [129], it has been found to be linked to self-reported inattention symptoms.

A different situation could be retrieved by analyzing the median graphs of the two groups during resting state, shown in *Figure 4.22*. A clear distinction between the two groups in each band is not visible: in both groups there is no evident tendency of forming clusters, differently from the CPT.




*(a) ADHD*




*(b) CONTROL*

***Figure 4.22:*** *Median Connectivity Graphs during resting state, obtained through ImCoh, of ADHD (top) and CONTROL (below) subjects. Only the statistically significant edges (p-value $<$ 0.05) are shown. The brighter the link color, the stronger the functional connectivity.*

#### 4.2.2.1 Graph Analysis

Based on the statistical results, as already mentioned in *Chapter 3*, the significant graph indices ($p<10^{-4}$) were further screened, resulting in the measures listed in the *Table 4.4*. During the screening phase, it was paid attention to maintain at least a measure belonging to Centrality, Segregation and Integration indices. In fact, Segregation and Integration are particularly important since, as already proposed by Cao et al. [44], individuals with ADHD may have a disrupted balance of them.

The brain regions found of interest for Betweenness Centrality in Beta 1 and Theta, as well as for Local Efficiency and for Shortest Path Length in Theta, included nodes of the Frontal and Parieto-Occipital regions. Since the Fronto-Parietal network has been implicated in attentional and executive processes and it is thought to be impaired in ADHD [60], it is interesting that in these two areas the indices of the ADHD and healthy subjects show different behaviors, able to discriminate the two populations.

| **Indices** | **Alpha** | **Beta 1** | **Theta** |
|---|---|---|---|
| **Betweenness Centrality** | Central<br>Parieto-Occipital | Frontal<br>Central<br>Temporal | Frontal<br>Temporal<br>Parieto-Occipital |
| **Local Efficiency** | | Frontal<br>Central | Central<br>Parieto-Occipital |
| **Shortest Path Length** | Central-PO | | Frontal-PO<br>Central-Temporal |
| **Global Efficiency** | ✓ | | ✓ |

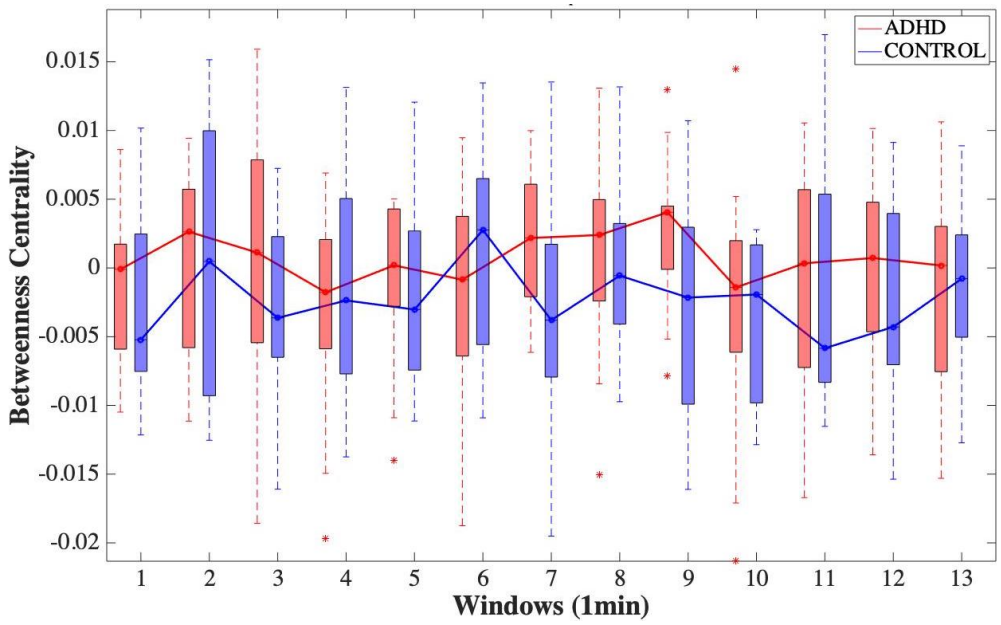
***Table 4.4:*** *Significant Graph Indices resulting from Wilcoxon test. Abbreviation → PO = Parieto-Occipital.*

Changes in the selected graph measures might reveal whether there are particular connectivity patterns with characteristics which dynamically change during the task; thus, a temporal trend investigation was conducted on these indices, resulting in 13 boxplots (one for every minute of CPT) per group. It should be recalled that the measures displayed are variations with respect to the baseline.

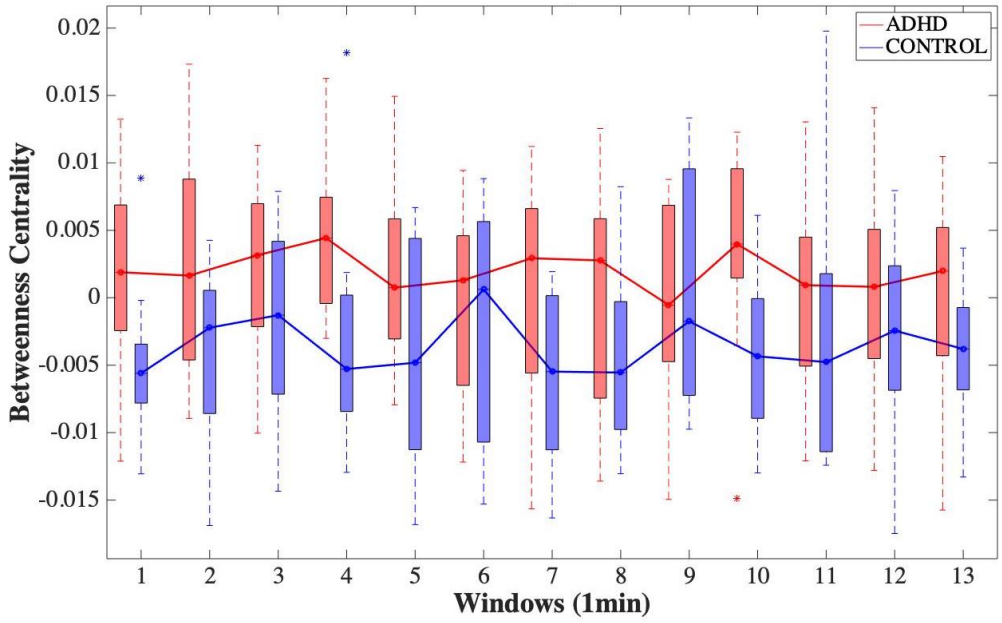
- **Betweenness Centrality**

Firstly, the attention will be focused on the **Temporal** region. In *Figure 4.23*, the boxplots and median temporal trends of Betweenness Centrality of the two groups in Beta 1 and Theta band, are shown. As can be noticed in both bands, ADHD median trend assumes higher values than controls one, and is characterized by a positive variation with respect to the baseline, while the control trend is characterized mostly by a negative variation.
It can be deducted that ADHD patients show an unexpected involvement of this area during the attentional task, also confirmed by the inspection of the connectivity graphs (*Figure 4.21*), where more and stronger connections are visible in this area, particularly in Beta 1. A similar observation was made by Tamm et al. [178], who hypothesized that this unusual activation (noticed, in their work, in the left middle and superior temporal gyri) was due to the fact that children in ADHD group may tend to adopt verbally mediated strategies for the task (i.e. silently reading letter to self or silently reminding self-instructions); in their case, this behavior was observed in 60% of ADHD children compared to 50% in the control group. The location of such activation is consistent with the role of Temporal area in verbal working memory [42] and may reflect an attempt to enhance task performance. Additionally, this region activation may be also associated with visual recognition of objects [93]. However, further investigation on Left Temporal cortex activation needs to be conducted.

Regarding the **Frontal** region, in Beta 1 band, illustrated in *Figure 4.24*, ADHD trend assumes higher values than the control one. The two trends do not show any variation with respect to the baseline in the first 2 minutes; in the following minutes the ADHD median trend assumes values over 0, while the Controls trend decreases with respect to the baseline. Differently from Beta 1 Frontal, the Theta Frontal trend of ADHD *(Figure 4.24)* is lower than control one and decreases with respect to the baseline.

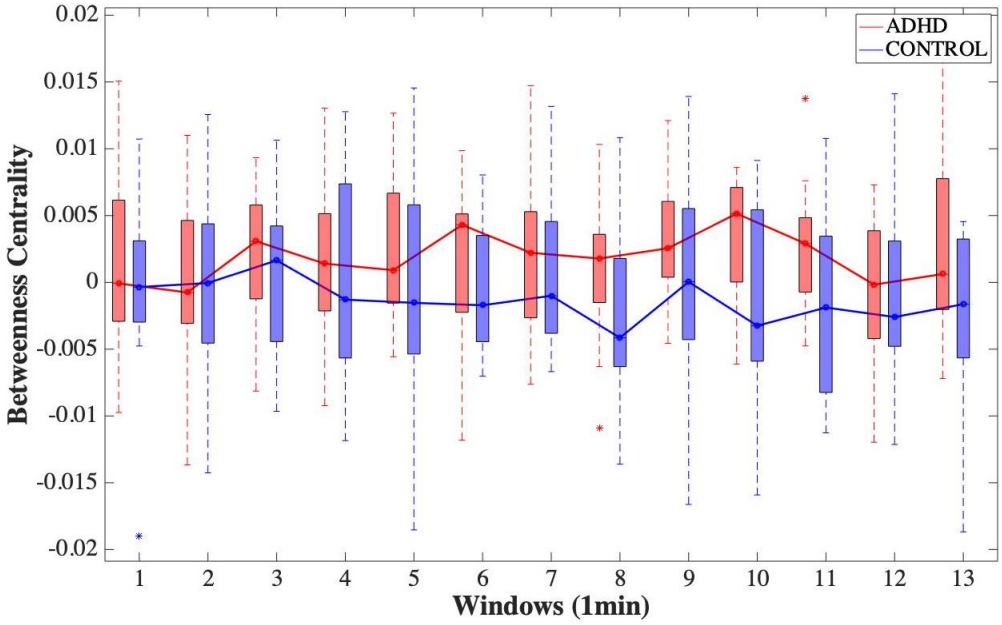



*(a) Beta 1*

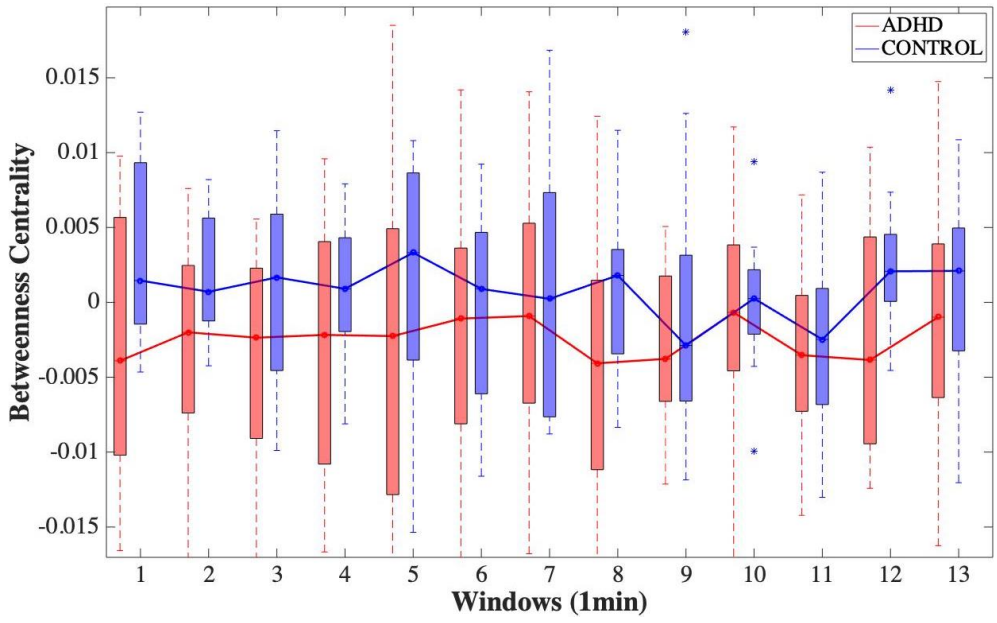



*(b) Theta*

***Figure 4.23:*** *Boxplots and Median Temporal Trends of Betweenness Centrality of Control (blue) and ADHD (red) group in Temporal cerebral region in Beta 1 (13-22 Hz) and Theta (4-8 Hz) band. * represents an outlier*




*(a) Beta 1*




*(b) Theta*

***Figure 4.24:*** *Boxplots and Median Temporal Trends of Betweenness Centrality of Control (blue) and ADHD (red) group in Frontal cerebral region in Beta 1 (13-22 Hz) and Theta (4-8 Hz) band. * represents an outlier.*

Shifting the attention on **Central and Parieto-Occipital** regions, the first area for Betweenness centrality resulted significant only in Beta 1 *(Figure 4.25)*, while the second one only in Theta *(Figure 4.26)*.

*Figure 4.25*, representing the Central area, shows a positive trend of Controls, opposite to ADHD one, which is characterized by a negative variation with respect to the baseline.

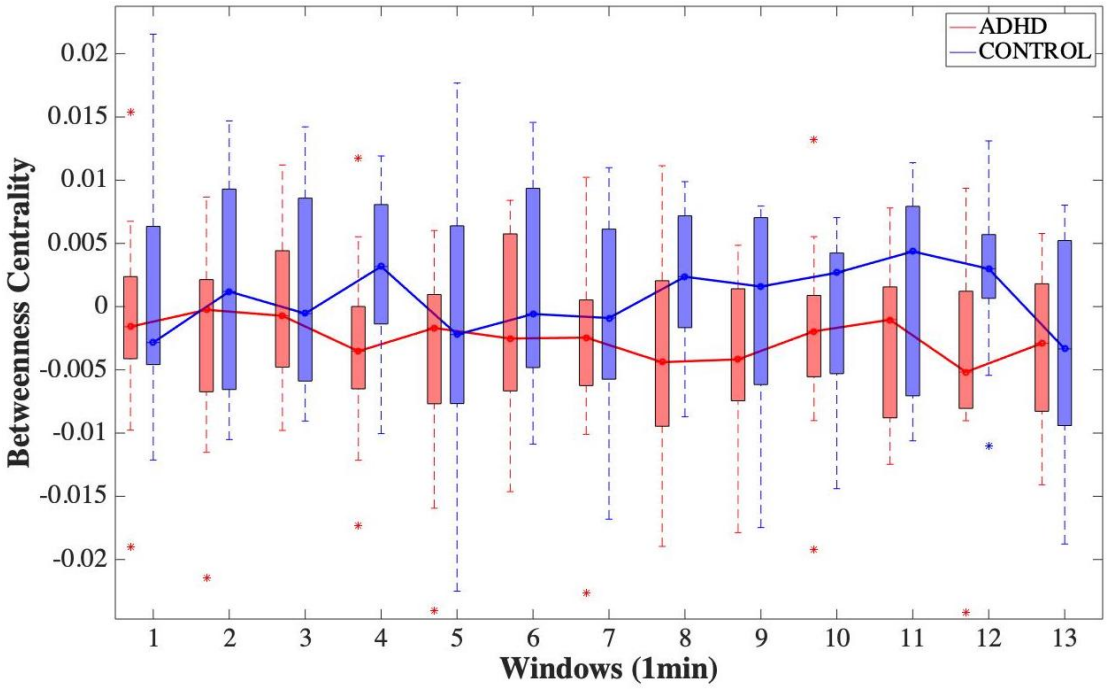



***Figure 4.25:*** *Boxplots and Median Temporal Trends of Betweenness Centrality of control (blue) and ADHD (red) group in Central cerebral region in Beta 1 band (13-22 Hz). * represents an outlier.*

*Figure 4.26* shows a positive median trend for Controls, higher with respect to the ADHD one, in Parieto-Occipital region.

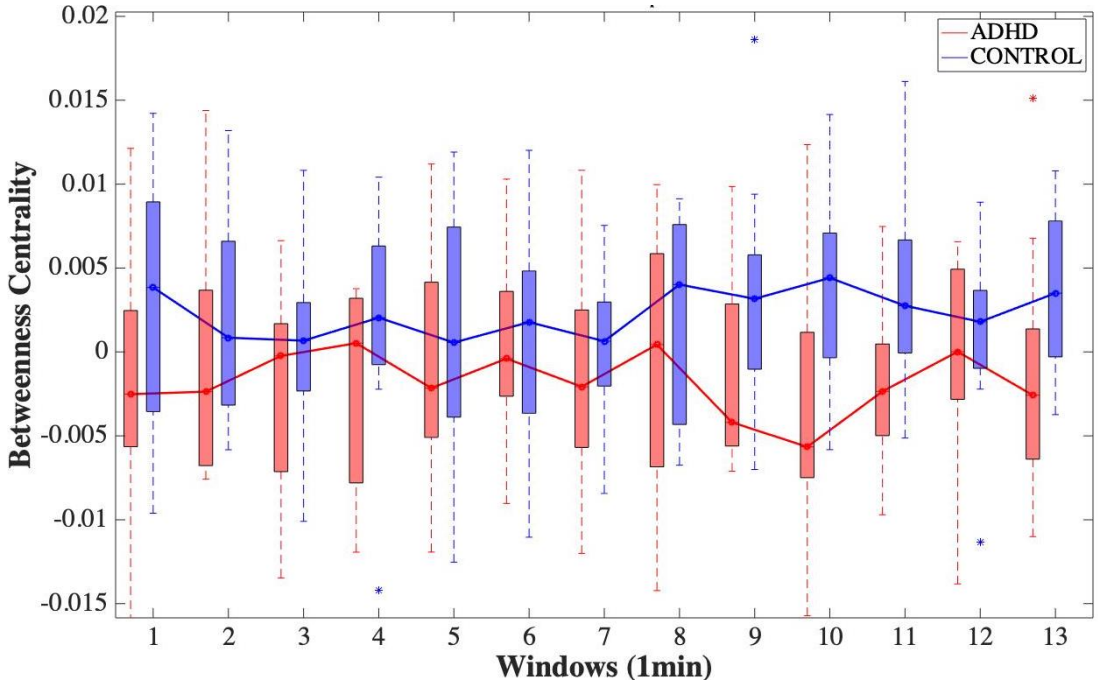



***Figure 4.26:*** *Boxplots and Median Temporal Trends of Betweenness Centrality of control (blue) and ADHD (red) group in Parieto-Occipital cerebral region in Theta (4-8 Hz) band. * represents an outlier.*

- **Global Efficiency**

The ADHD Global Efficiency median trends in Alpha *(Figure 4.27)* and Theta *(Figure 4.28)* display an opposite behavior, since the Alpha trend is characterized by higher values belonging to ADHD patients, while the Theta trend shows slightly lower values than the healthy one.

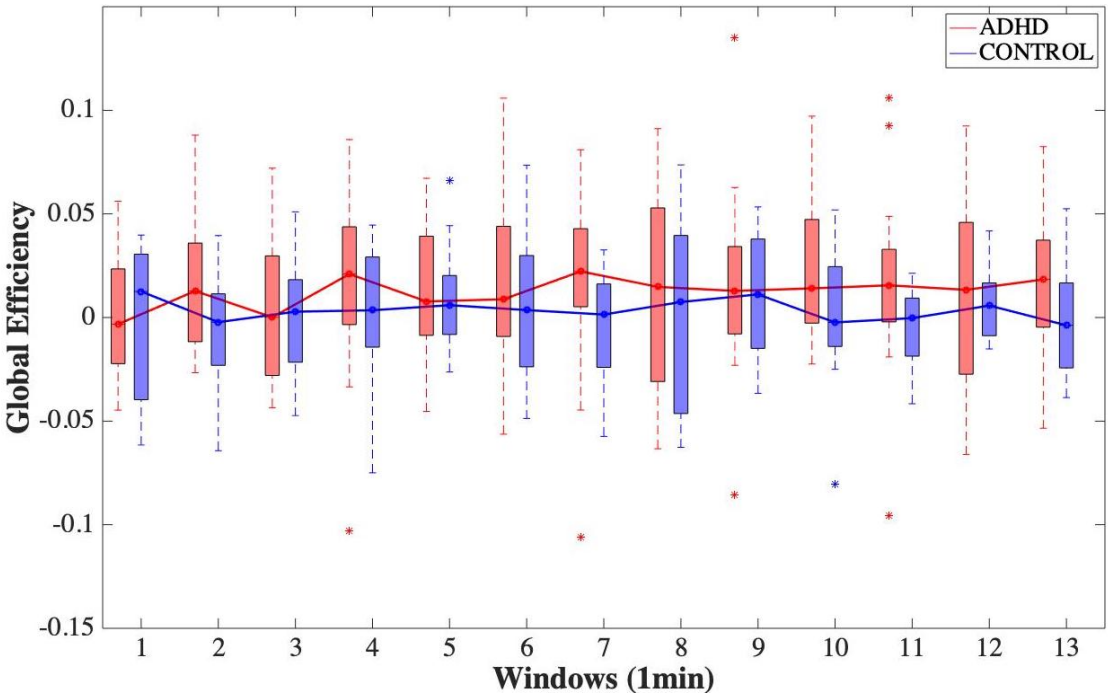



***Figure 4.27:*** *Boxplots and Median Temporal Trends of Global Efficiency of control (blue) and ADHD (red) group in Alpha (8-13 Hz) band. * represents an outlier.*

Focusing on Theta *(Figure 4.28)*, the results suggest that Controls are characterized by a greater integration of the brain during the task: this same phenomenon was observed in Machida et al. [117] and it was associated with a more stable performance, since Global Efficiency in Theta band was significantly associated with HRT SD. Indeed, the Theta band may play a crucial role in the ability of the brain to produce stable responses during a task: only in this band, the Global Efficiency was associated with HRT SD in the study conducted by Machida et al. Furthermore, the importance of Theta band in cognitive task performance was

reported in several studies [121][181]. It was also suggested by Sauseng et al. [159] that long range connectivity in this band might contribute in integrating sensory information into executive control and coordinating information from different brain cortical areas during a mental task to produce an output [128]. In conclusion, a higher integration observed in Theta band may reflect how well the information flow is integrated before the response [117].

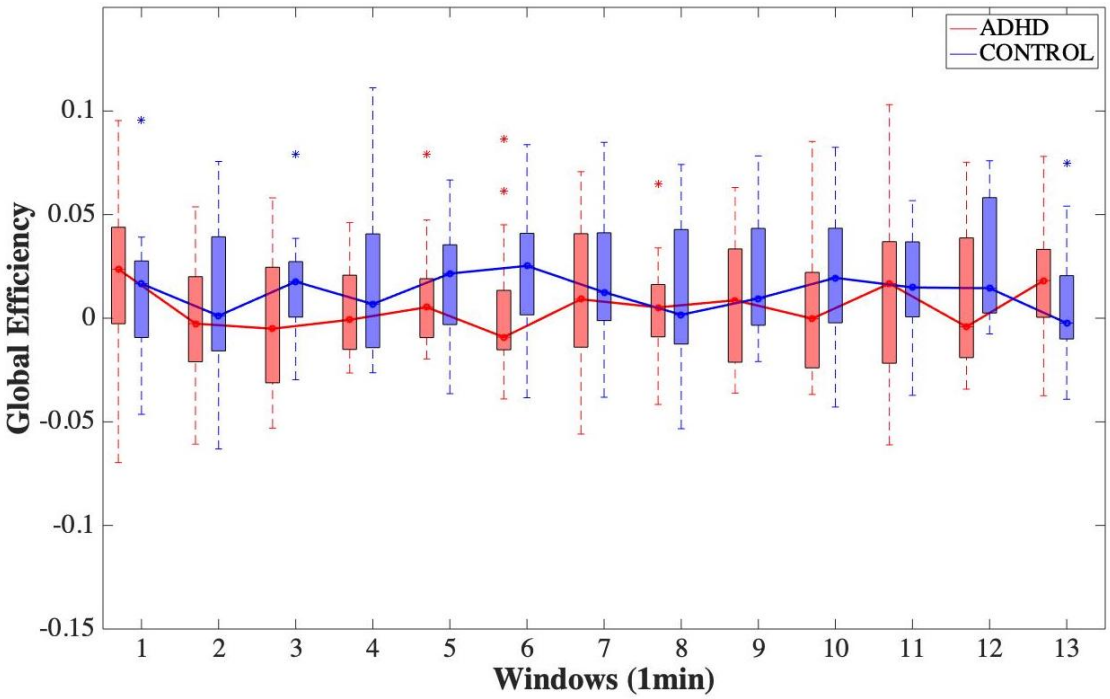



***Figure 4.28:*** *Boxplots and Median Temporal Trends of Global Efficiency of control (blue) and ADHD (red) group in Theta (13-22 Hz) band. * represents an outlier.*

- **Local Efficiency**

In *Figure 4.29*, boxplots of Beta 1 Local Efficiency in Frontal region for both groups are shown and a lower median trend of ADHDs along the 13 one-minute windows can be noticed. Comparing to the healthy children, ADHDs are characterized by a lower ability of exchanging information in Frontal region. The same situation has been noted for Clustering Coefficient in the same band and area. A decreased Clustering Coefficient in ADHDs indicates that local neighbourhoods are, one average, less fully interconnected than in typical developing Controls. This is consistent with a decreased Local Efficiency, since the decrease in number of links in a local neighbourhood diminishes the network efficiency in that region.

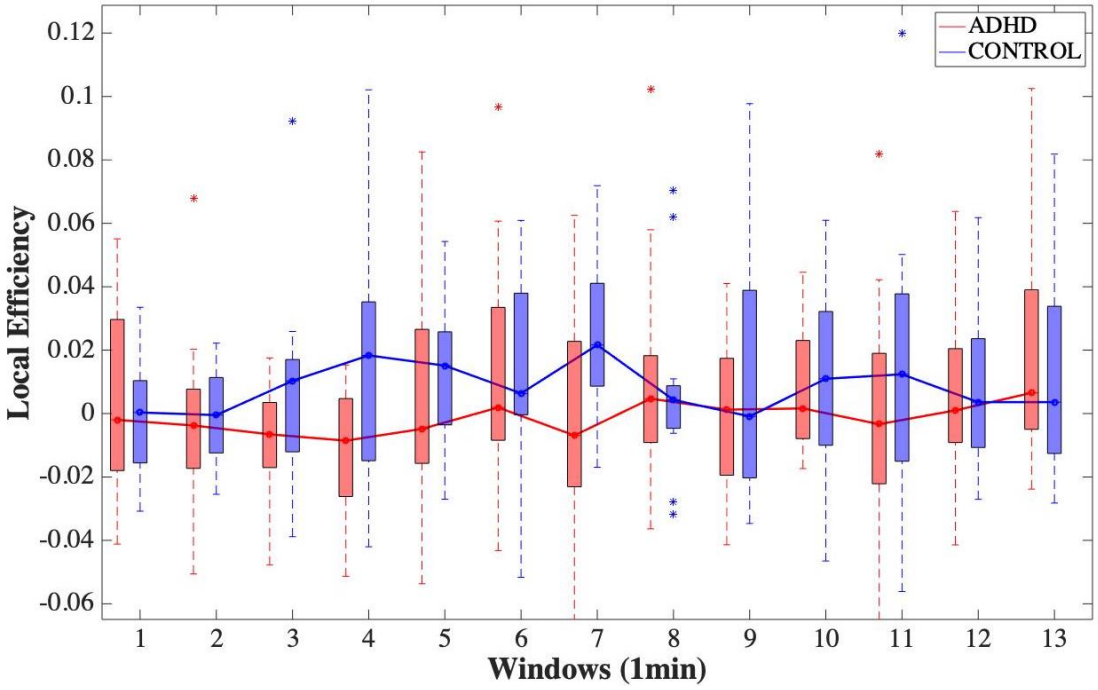



***Figure 4.29:*** *Boxplots and Median Temporal Trends of Local Efficiency of control (blue) and ADHD (red) group in Frontal cerebral region in Beta 1 (13-22 Hz) band. * represents an outlier.*

- **Shortest Path Length**

As regards the Shortest Path Length in Theta band, the connections linking the Frontal and the Parieto-Occipital areas, displayed in *Figure 4.30*, are longer (higher path lengths) in the ADHD group than in healthy patients; this result suggests that ADHD children may be characterized by an inefficient information integration [71], supported also by the results obtained for the Global Efficiency in Theta, that was found decreased in ADHDs.

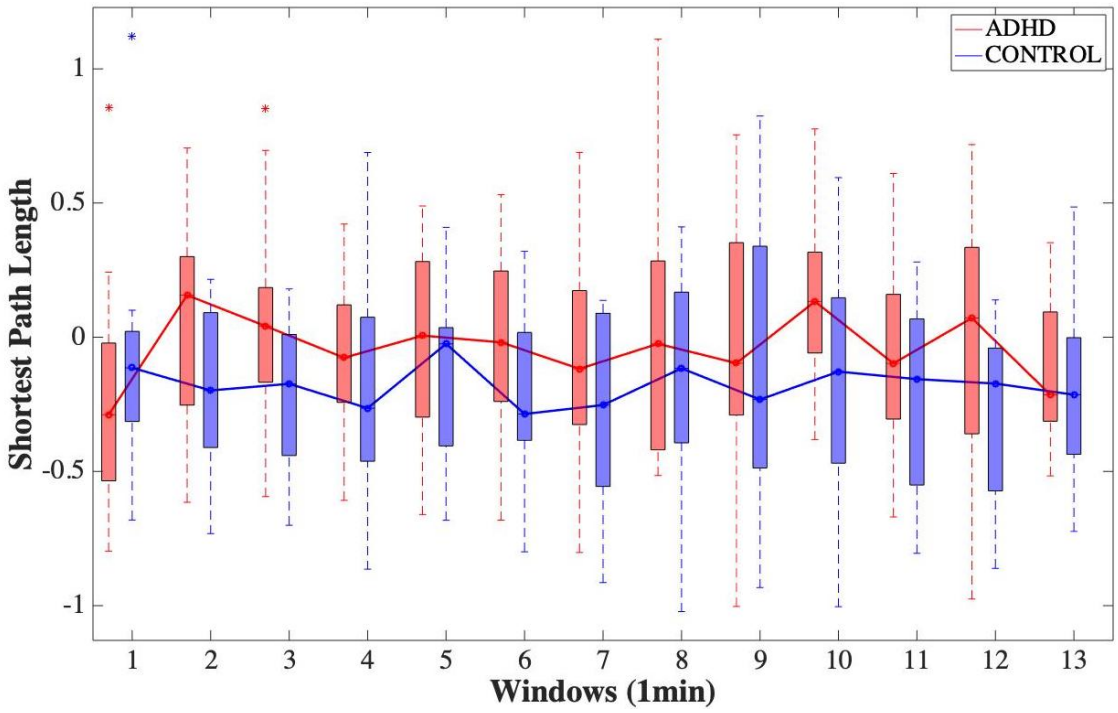



***Figure 4.30:*** *Boxplots and Median Temporal Trends of Shortest Path Length of control (blue) and ADHD (red) group between Frontal and Parieto-Occipital cerebral region in Theta (4-8 Hz) band. * represents an outlier.*

On the contrary, always considering Theta band, the paths connecting Central and Temporal areas are shorter in the ADHD patients with respect to controls *(Figure 4.31)*, confirming the denser communication between nodes belonging to these regions noticed in *Figure 4.21* and the higher Betweenness Centrality found in Theta Temporal, shown in *Figure 4.23*.

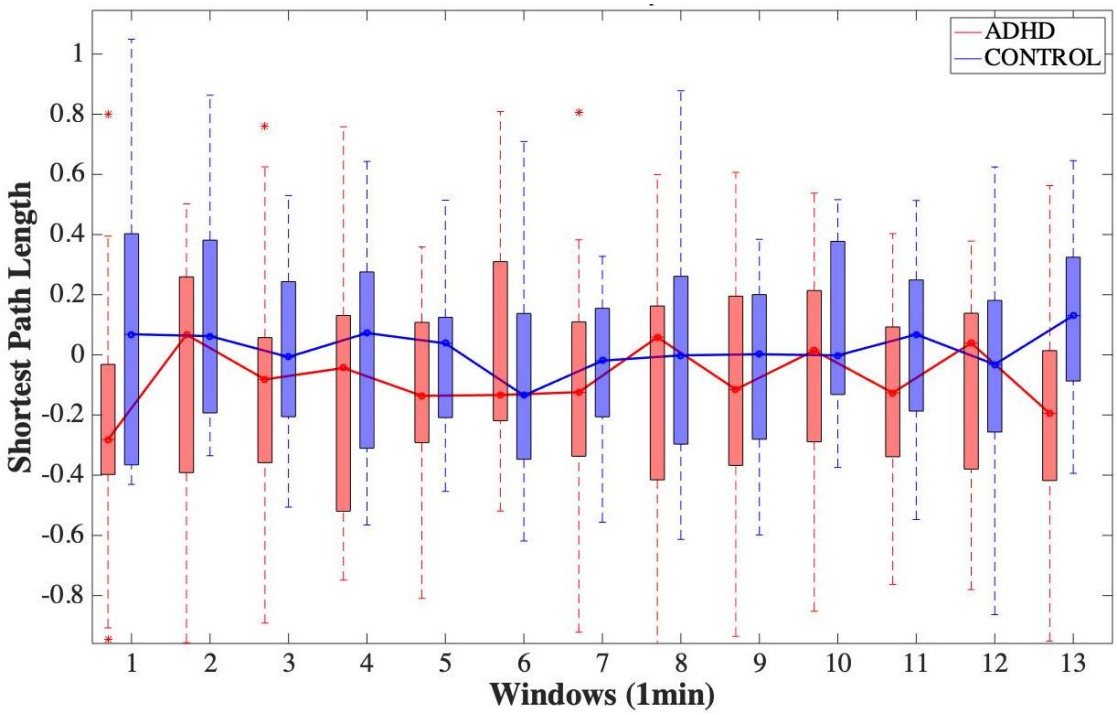



***Figure 4.31:*** *Boxplots and Median Temporal Trends of Shortest Path Length of control (blue) and ADHD (red) group between Central and Temporal cerebral region in Theta (4-8 Hz) band.* * *represents an outlier.*

## 4.3 Machine Learning

In this section, the features selected by using the three different methods (ANOVA, MRMR and PCA) and the performances reached by the Support Vector Machine classifier in distinguishing ADHD and control subjects will be presented.

- **ANOVA**

As described in *Section 3.4.1.1*, 16 out of 81 initial features, have been extracted with ANOVA method. After applying the Pearson's correlation, whose resulting matrix is shown in *Figure 4.32*, other two variables were excluded from the dataset, since they presented correlation values of 0.98 with other two features.

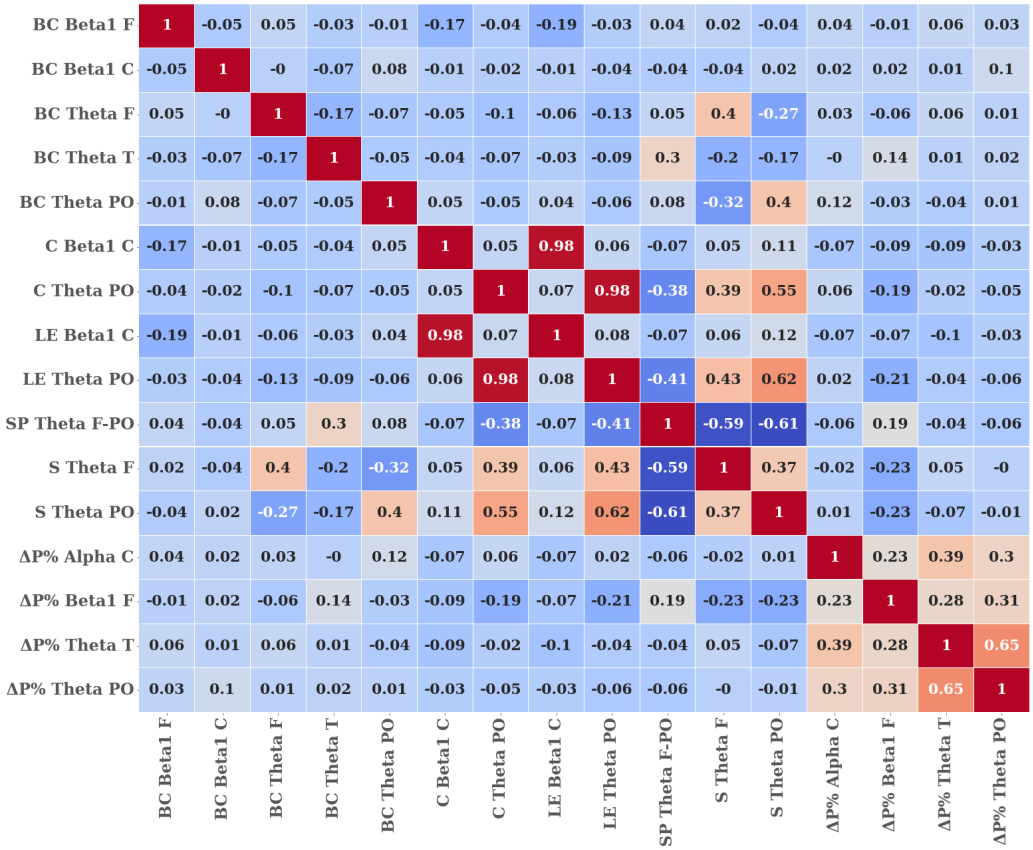



***Figure 4.32:*** *ANOVA Correlation Matrix. Graph Indices → LE: Local Efficiency, BC: Betweenness Centrality, C: Clustering Coefficient, S: Strength, SP: Shortest Path Length. Cerebral areas → C: Central, F: Frontal, T: Temporal, PO: Parieto-Occipital.*

Indeed, Local Efficiency in Beta 1 Central (LE Beta1 C) showed a correlation of 0.98 with the Clustering Coefficient in the same band and area (C Beta1 C); the same situation was observed for the Local Efficiency in Theta Parieto-Occipital (LE Theta PO) with the corresponding Clustering Coefficient (C Theta PO). This result could be easily explained by the fact that they are both segregation measures and, thus, they bring similar information. At the end, Local Efficiency in Theta Parieto-Occipital (LE Theta PO) and Clustering Coefficient in in Beta 1 Central (C Beta1 C) have been eliminated since, according to the F-score, they presented a lower predictive power towards the target than the others.

The final subset of 14 explanatory variables is displayed in *Table 4.5*.

| **Features** | **Alpha** | **Beta 1** | **Theta** |
|---|---|---|---|
| **Betweenness Centrality** | | Frontal<br>Central | Frontal<br>Temporal<br>Parieto-Occipital |
| **Clustering Coefficient** | | | Parieto-Occipital |
| **Local Efficiency** | | Central | |
| **Strength** | | | Frontal<br>Parieto-Occipital |
| **Shortest Path Length** | | | Frontal-PO |
| **ΔP [%]** | Central | Frontal | Temporal<br>Parieto-Occipital |

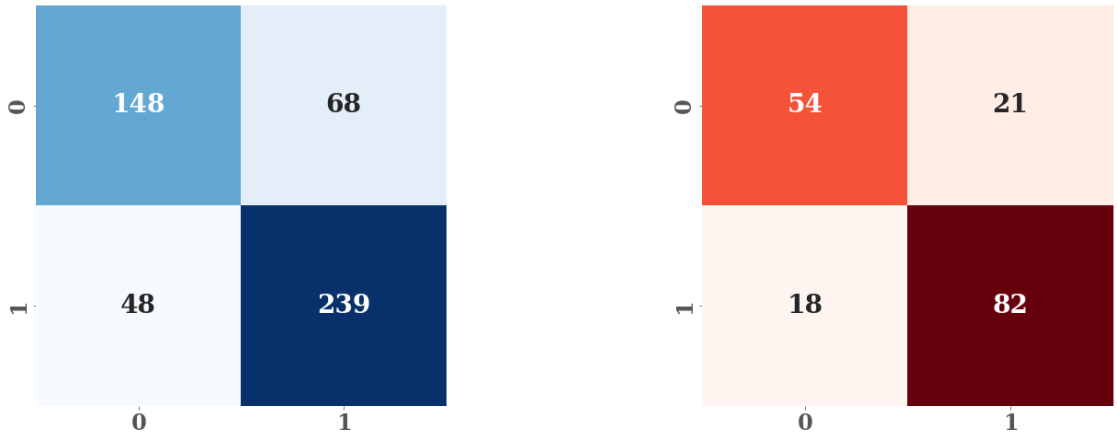
***Table 4.5:*** *Final selected features using ANOVA method.*

The final optimal classifier, obtained from the Grid Search and illustrated in the *Table 4.6*, is the one with the parameters that maximize the F1 score. The values relative to the F-score in the three data sets are also illustrated in *Table 4.6*.

| SVC(C=1, gamma=0.125, kernel='rbf') | | |
|---|---|---|
| **Train Set** | **Validation Set** | **Test Set** |
| F1: 0.809 | F1: 0.804 | F1: 0.807 |

***Table 4.6:*** *Final classifier and F1 scores obtained in the three sets adopting ANOVA method.*

The Confusion Matrices, obtained from validation and test phases, are displayed in *Figure 4.33*.




***(a)*** *Validation Set* ***(b)*** *Test Set*

***Figure 4.33:*** *ANOVA Confusion Matrices: Controls (0) and ADHD (1). The number of True Negative, False Negative, False Positive and True positive cases are displayed in the first, second, third and fourth quadrants respectively.*

- **MRMR**

The same procedure followed for ANOVA was also applied for MRMR technique. As already said, the difference with ANOVA consists in the fact of choosing a priori the number of desired features K (20).

Even if MRMR takes into account the redundancy among variables, it was considered advisable to make a further check on the correlation of the variables, as shown in *Figure 4.34*.

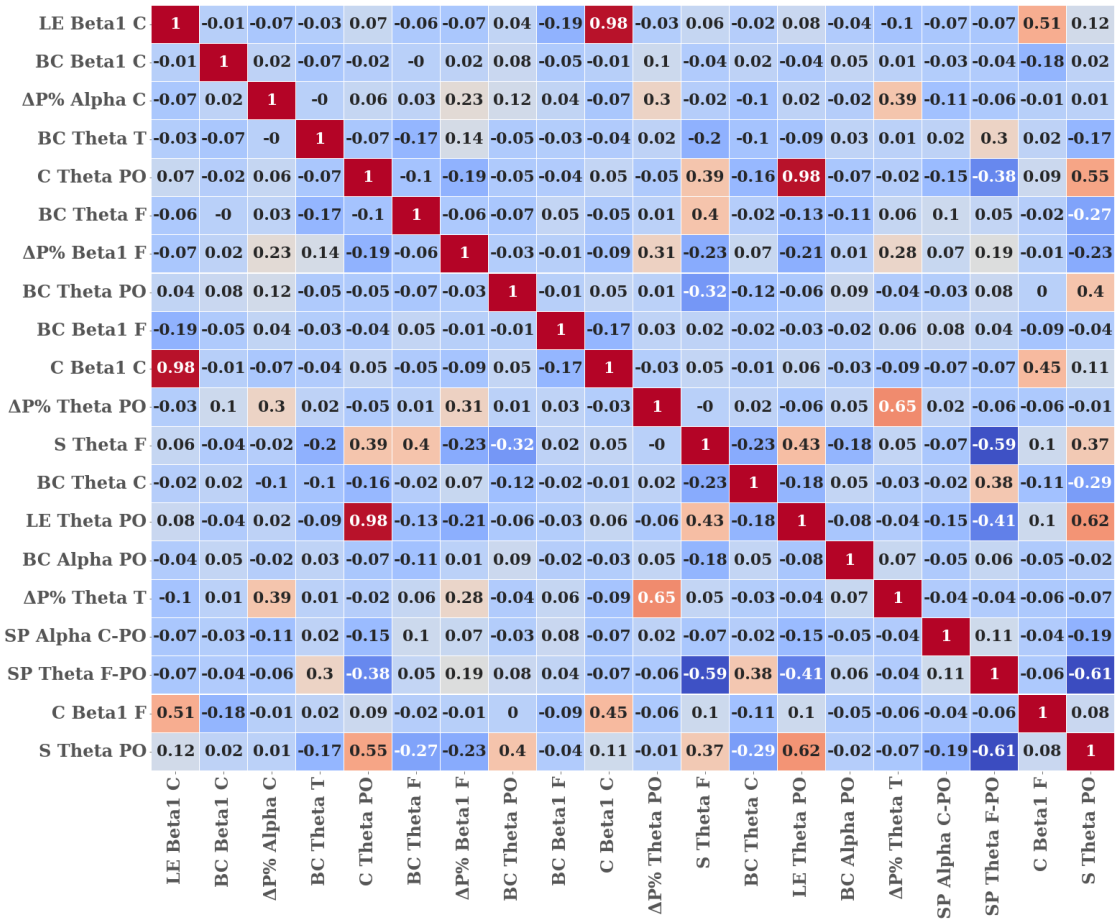



***Figure 4.34:*** *MRMR Correlation Matrix. Graph Indices → LE: Local Efficiency, BC: Betweenness Centrality, C: Clustering Coefficient, S: Strength, SP: Shortest Path Length. Cerebral areas → C: Central, F: Frontal, T: Temporal, PO: Parieto-Occipital.*

Since a great correlation (0.98) emerged between Clustering Coefficient and Local Efficiency in Beta 1 Central (C Beta1 C and LE Beta1 C) and Theta Parieto-Occipital (C Theta PO and LE Theta PO), the Clustering Coefficient in Beta 1 (C Beta1 C) and Local Efficiency in Theta Parieto-Occipital (LE Theta PO) were removed from the set because of the lower predictive power than the others (verified by ANOVA).

The 18 final selected attributes are displayed in *Table 4.7*.

| **Features** | **Alpha** | **Beta 1** | **Theta** |
|---|---|---|---|
| **Betweenness Centrality** | Parieto-Occipital | Frontal<br>Central | Frontal<br>Central<br>Temporal<br>Parieto-Occipital |
| **Clustering Coefficient** | | Frontal | Parieto-Occipital |
| **Local Efficiency** | | Central | |
| **Strength** | | | Frontal<br>Parieto-Occipital |
| **Shortest Path Length** | Central-PO | | Frontal-PO |
| **ΔP [%]** | Central | Frontal | Temporal<br>Parieto-Occipital |

***Table 4.7:*** *Final selected features using MRMR method.*

As can be noticed, most of the selected features are the same as the ones selected by ANOVA, except for the Betweenness Centrality, Modularity and Strength. In particular, BC in Alpha Parieto-Occipital and in Theta Central, and Modularity in Alpha Temporal were selected by MRMR, but not by ANOVA; on the contrary, Strength in Theta Parieto-Occipital was chosen only by ANOVA.
It is worth to underline that some of the attributes selected by both methods are those that showed interesting results in the Spectral and Connectivity analyses, as the Alpha ΔP% in Central area or the Theta Betweenness Centrality in Temporal area. In addition, it is positive that Beta 1 ΔP% in Frontal area and the SP between Frontal and Parieto-Occipital regions were chosen, since they are involved

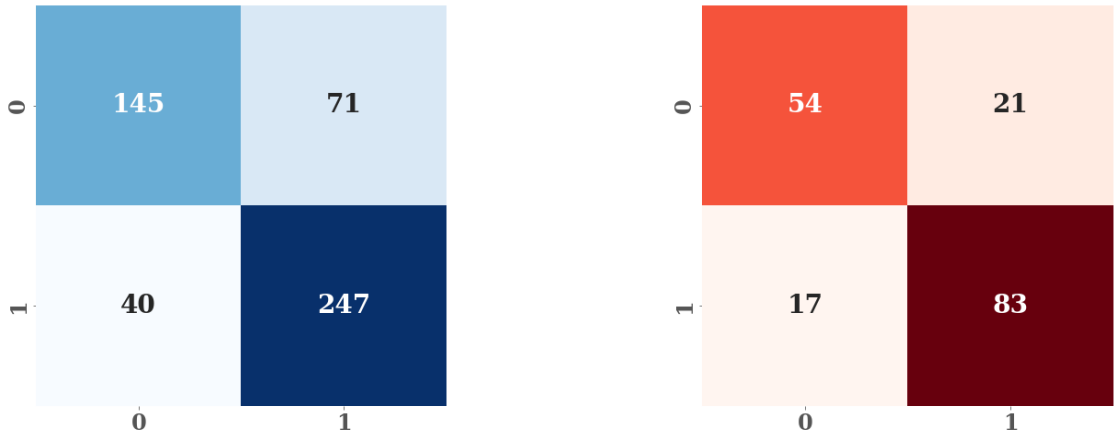
in Attentive tasks (Frontal Area) and Visual Processing (Parieto-Occipital).

Finally, the optimal classifier, which is the same as the one chosen by the GridSearch using the features picked by ANOVA, is reported in the *Table 4.8*, with the relative F1 scores.

| SVC(C=1, gamma=0.125, kernel='rbf') | | |
|---|---|---|
| **Train Set** | **Validation Set** | **Test Set** |
| F1: 0.821 | F1: 0.816 | F1:0.813 |

***Table 4.8:*** *Final classifier and F1 scores obtained in the three sets adopting MRMR method.*

Confusion matrices obtained on Validation and Test set are shown in *Figure 4.35*.




***(a)*** *Validation Set* ***(b)*** *Test Set*

***Figure 4.35:*** *MRMR Confusion Matrices: Controls (0) and ADHD (1). The number of True Negative, False Negative, False Positive and True positive cases are displayed in the first, second, third and fourth quadrants respectively.*

- **PCA**

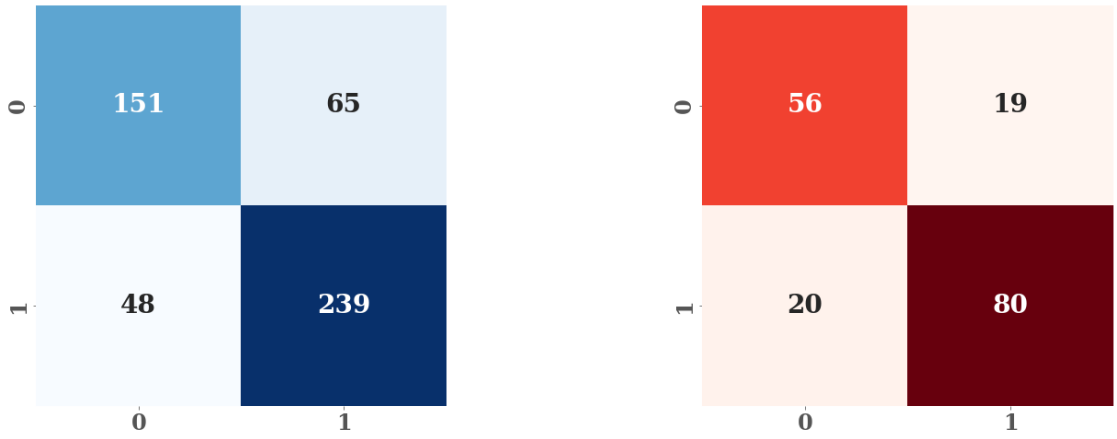
As mentioned before in *Section 3.4.1.1*, 28 principal components were considered out of 81.

The dataset was then subdivided in Train, Validation and Test sets and the GridSearch was applied to the Train, resulting in the optimal classifier in *Table 4.9*:

| SVC(C=2, gamma=0.03125, kernel='rbf') | | |
|---|---|---|
| **Train Set** | **Validation Set** | **Test Set** |
| F1: 0.810 | F1: 0.808 | F1: 0.804 |

***Table 4.9:*** *Final classifier and F1 scores obtained in the three sets adopting PCA method.*

With the chosen classifier, the subsequent Confusion Matrices in *Figure 4.36* were obtained:




***(a)*** *Validation Set* ***(b)*** *Test Set*

***Figure 4.36:*** *PCA Confusion Matrices: Controls (0) and ADHD (1). The number of True Negative, False Negative, False Positive and True positive cases are displayed in the first, second, third and fourth quadrants respectively.*

In conclusion, it can be noticed that the classification performance, for all the three different models, achieved good results, with no big difference among them.

However, the best performance was achieved by the model trained on the subset of features selected by PCA, but the computational time was greater than the one of the other two methods.

# Chapter 5

# Conclusions

The present master thesis work presented an analysis of EEG activity of ADHD and control children during an attentional task, exploring Spectral characteristics and the Brain Connectivity, with the aim of discovering potential biomarkers to distinguish the pathology. A temporal dynamic approach, considering the task entire duration, was adopted to investigate abnormal brain activity in ADHD children.

The signal processing, including the data cleaning, filtering, resampling and artifacts rejection, has not been performed in the current study, but it was already done in a previous work. On the contrary, the re-referencing procedure through Laplacian filter, as well as the removal of noisy electrodes, was executed in this work in order to reduce the remaining artifacts, as the volume conduction effect, which may affect the connectivity analysis. For the same reason, a functional connectivity measure insensitive to this artifact, as the ImCoh, was chosen.

Contrasting with the literature and differently from what it was expected, no particularly interesting findings in the fluctuations of Beta 1 power with respect to baseline were discovered in the Frontal region. A plausible explanation could be that higher frequency bands suffer more from the adoption of a continuous spectral analysis with respect to an event-related analysis. In literature, higher values of Frontal Beta power have been observed in the control group compared with the ADHD one [108] [113]. Since Frontal lobes are shown to be related to cognitive functions, a poor activity in this area could underline inability to focus on a task. However, interesting differences among groups were noticed in Alpha power in both hemispheres of Central region, finding two desynchronized spots in both groups, but more evident in the ADHDs. This higher desynchronization of Alpha frequencies was interpreted as a greater difficulty of ADHD children

in remaining calm when performing the task and it may be associated with the hyperactivity symptom. Results in the correlations of Alpha $\Delta$P% with CE and HRT (negative and positive respectively) could strengthen the hypothesis that the calmer is the patient (higher Alpha), the less is the number of Commission Errors committed and the longer is the time spent thinking about the response, not being impulsive. However, this hypothesis could not be confirmed due to the absence of symptomatology (inattentiveness and hyperactivity) of the patients.

Theta band and Temporal area were considered of particular interest, mainly in the Functional Connectivity and Graph Theory analyses, since ADHD group showed an unusual behavior in this band and area during the CPT. The major involvement of the Temporal area in ADHDs was hypothesized to be due to the adoption of verbally mediated strategies (silently reading letter to self or reminding self-instructions) employed to improve the performance: this is consistent with the role of Temporal area in verbal working memory [42].

Focusing the attention on graph-based indices, the Betweenness Centrality presented higher values in ADHDs in the Temporal area and the paths connecting the Central region with the Temporal one were shorter than those belonging to Controls. Regarding Global Efficiency, greater values were found in Theta band in the control group, suggesting the presence of an elevated integration of the brain during the task. The relevance of Theta band in cognitive tasks performance was assessed in several studies, suggesting that a higher integration may reflect how well the information flow is integrated before the response. Furthermore, in literature, higher values of Global Efficiency in Theta band have been associated to more stable results in the performance, deducted from the significant association of Global Efficiency with the HRT SD [117]. Indeed, a greater HRT SD is believed to reflect inefficient information flow during the execution of a cognitive task, so variable responses might be due to inefficient or disrupting information processing [117]. In support of this hypothesis, in the present work the HRT SD was the only performance index which showed significant differences between the two groups, with higher values found in the ADHD subjects.

The most discriminating characteristic found in Frontal region and Beta 1 band in connectivity analysis was the difference in Local Efficiency and Clustering Coefficient for the two groups: ADHDs are characterized by a decreased Local Efficiency and Clustering Coefficient, underlying a low ability of exchanging information within the Frontal region. In addition, ADHDs show longer path lengths between Frontal and Parieto-Occipital areas with respect of Controls,

suggesting an inefficient information integration between these two areas.

Even if significant results were not found in Beta 1 band in Frontal region through the spectral analysis, ANOVA and MRMR selected $\Delta$P% located in this band and area, together with some graph indices, as features for the classification model. A positive outcome is that all the features selected by ANOVA method were also chosen by the MRMR and all of them coincide with the graph indices resulted significant from the Wilcoxon test; furthermore, most of the selected features were found in Theta band, coherently with the results of both Spectral and Connectivity analyses. Finally, the performances in terms of F1 score on the test set reached results higher than 0.8, regardless of the feature selection method employed, demonstrating robustness of the algorithm.

One of the main limitations associated to the current research was the lack of performance data for each given stimulus, such as the participant's response and the corresponding reaction time. Indeed, these parameters could have been used to understand how the brain activity change according to given or not given responses, allowing an Event-Related Potentials analysis. Other aspects relative to the dataset that should be taken into account are the limited number of participants in the study, the large range of participants' age (7-17 years) and the lack of clinical identification of ADHD subtypes (inattentive, hyperactive/impulsive and combination types). Hence, a possible improvement of this analysis could consist in considering a different dataset, populated with larger groups of ADHD and matched control subjects and provided of stimulus-related behavioral features. Furthermore, the classical EEG bands in children could be adapted, since children show different brain mechanisms and the adoption of the main standard frequency bands could not provide accurate results [96]. Moreover, a future study could be an age-related analysis considering clusters of individuals with a more uniform age range, exploiting their clinical information in order to find differences in brain activity among the three ADHD subtypes.

It is important to highlight that, even if good performances were reached in the Machine Learning algorithm, this classification has not a real clinical validity, since it is performed on more than one temporal window belonging to the same subject (and consequently correlated among each other), not classifying the single patient. Hence, the decision of creating the test set with randomically chosen windows was taken due to the fact that selecting all the windows belonging to a single subject could have introduced correlations among the data, distorting the results.

Since most of the studies on ADHD were conducted analyzing the resting state or considering the CPT in a non-dynamic way, there have been some difficulties in finding works similar to the current one in order to make useful comparisons between the findings and to reach a consistent result.

In conclusion, the search for functional neural correlates of ADHD, and consequently for potential biomarkers of the disorder, is crucial in the pursuit of its prevention, early detection and more effective treatment. Combination of the ImCoh and graph theory methods, together with the dynamic temporal analysis, carried on in the present study, would be a very useful and novel approach for exploring underlying mechanisms of the brain and for diagnosing neurologic disorder in the future.

# Bibliography


[1] EEG – still breathing. https://still-breathing.net/tag/eeg/.

[2] ElectrodePositions1020.PNG (604×700). https://www.bci2000.org/mediawiki/images/1/15/ElectrodePositions1020.PNG.

[3] En:pca [Analysis of community ecology data in R]. https://www.davidzeleny.net/anadat-r/doku.php/en:pca.

[4] Re-referencing. https://eeglab.org/tutorials/ConceptsGuide/rereferencing_background.html.

[5] Support Vector Machine (SVM) Algorithm - Javatpoint. https://www.javatpoint.com/machine-learning-support-vector-machine-algorithm.

[6] *Clinical Neurophysiology*. Elsevier, Amsterdam Paris, 2003.

[7] *Quantitative EEG, Event-Related Potentials and Neurotherapy*. Elsevier, 2009.

[8] Structure and function of complex brain networks. *Dialogues in Clinical Neuroscience*, 15(3):247–262, September 2013.

[9] Ali Kareem Abbas, Ghasem Azemi, Sajad Amiri, Samin Ravanshadi, and Amir Omidvarnia. Effective connectivity in brain networks estimated using EEG signals is altered in children with ADHD. *Computers in Biology and Medicine*, 134:104515, July 2021.

[10] Priyanka A. Abhang, Bharti W. Gawali, and Suresh C. Mehrotra. Technical Aspects of Brain Rhythms and Speech Parameters. In *Introduction to EEG- and Speech-Based Emotion Recognition*, pages 51–79. Elsevier, 2016.

[11] Amirmasoud Ahmadi, Mehrdad Kashefi, Hassan Shahrokhi, and Mohammad Ali Nazari. Computer aided diagnosis system using deep

convolutional neural networks for ADHD subtypes. *Biomedical Signal Processing and Control*, 63:102227, January 2021.

[12] Mehran Ahmadlou, Hojjat Adeli, and Amir Adeli. Graph Theoretical Analysis of Organization of Functional Brain Networks in ADHD. *Clinical EEG and Neuroscience*, 43(1):5–13, January 2012.

[13] Allwin Alex, S. Coelli, A. M. Bianchi, L. Ponzini, E. Buzzi, and M. P. Canevini. EEG analysis of brain activity in attention deficit hyperactivity disorder during an attention task. In *2017 IEEE 3rd International Forum on Research and Technologies for Society and Industry (RTSI)*, pages 1–4, Modena, Italy, September 2017. IEEE.

[14] Michael J. Aminoff. *Encyclopedia of the Neurological Sciences*. Academic Press, Amsterdam, 2. ed edition, 2014.

[15] Sheida Ansari Nasab, Shirin Panahi, Farnaz Ghassemi, Sajad Jafari, Karthikeyan Rajagopal, Dibakar Ghosh, and Matjaž Perc. Functional neuronal networks reveal emotional processing differences in children with ADHD. *Cognitive Neurodynamics*, July 2021.

[16] Sylvain Arlot and Alain Celisse. A survey of cross-validation procedures for model selection. *Statistics Surveys*, 4(none), January 2010.

[17] Martijn Arns, C. Keith Conners, and Helena C. Kraemer. A Decade of EEG Theta/Beta Ratio Research in ADHD: A Meta-Analysis. *Journal of Attention Disorders*, 17(5):374–383, July 2013.

[18] R Artusi, P Verderio, and E Marubini. Bravais-Pearson and Spearman Correlation Coefficients: Meaning, Test of Hypothesis and Confidence Interval. page 4.

[19] Deborah Ashby. Practical statistics for medical research. Douglas G. Altman, Chapman and Hall, London, 1991. No. of pages: 611. Price: £32.00. *Statistics in Medicine*, 10(10):1635–1636, October 1991.

[20] Laura Auria and R. A. Moro. Support Vector Machines (SVM) as a Technique for Solvency Analysis. *SSRN Electronic Journal*, 2008.

[21] Joan C Ballard. Assessing Attention: Comparison of Response-Inhibition and Traditional Continuous Performance TestsÃ. page 20.

[22] R Barry. EEG coherence in attention-deficit/hyperactivity disorder: A comparative study of two DSM-IV types. *Clinical Neurophysiology*, 113(4):579–585, April 2002.

[23] Robert J Barry, Adam R Clarke, and Stuart J Johnstone. A review of electrophysiology in attention-deficit/hyperactivity disorder: I. Qualitative and quantitative electroencephalography. *Clinical Neurophysiology*, 114(2):171–183, February 2003.

[24] E Başar, C Başar-Eroğlu, S Karakaş, and M Schürmann. Brain oscillations in perception and memory. *International Journal of Psychophysiology*, 35(2-3):95–124, March 2000.

[25] Hayden Basinger and Jeffery P. Hogg. Neuroanatomy, Brainstem. In *StatPearls*. StatPearls Publishing, Treasure Island (FL), 2021.

[26] André M. Bastos and Jan-Mathijs Schoffelen. A Tutorial Review of Functional Connectivity Analysis Methods and Their Interpretational Pitfalls. *Frontiers in Systems Neuroscience*, 9, January 2016.

[27] Mark F. Bear, Barry W. Connors, and Michael A. Paradiso. *Neuroscience: Exploring the Brain*. Wolters Kluwer, Philadelphia, fourth edition edition, 2016.

[28] Richard Beare, Chris Adamson, Mark A. Bellgrove, Veronika Vilgis, Alasdair Vance, Marc L. Seal, and Timothy J. Silk. Altered structural connectivity in ADHD: A network based analysis. *Brain Imaging and Behavior*, 11(3):846–858, June 2017.

[29] April A. Benasich, Zhenkun Gou, Naseem Choudhury, and Kenneth D. Harris. Early cognitive and language skills are linked to resting frontal gamma power across the first 3 years. *Behavioural Brain Research*, 195(2):215–222, December 2008.

[30] Giovanni Berlucchi and Giuseppe Vallar. Chapter 1 - The history of the neurophysiology and neurology of the parietal lobe. In Giuseppe Vallar and H. Branch Coslett, editors, *Handbook of Clinical Neurology*, volume 151 of *The Parietal Lobe*, pages 3–30. Elsevier, January 2018.

[31] Andrea Biasiucci, Benedetta Franceschiello, and Micah M. Murray. Electroencephalography. *Current Biology*, 29(3):R80–R85, February 2019.

[32] Joseph Biederman. Attention-Deficit/Hyperactivity Disorder: A Selective Overview. page 6.

[33] Joseph Biederman. Attention-Deficit/Hyperactivity Disorder: A Selective Overview. *Biological Psychiatry*, 57(11):1215–1220, June 2005.

[34] Joseph Biederman, Carter R. Petty, K. Yvonne Woodworth, Alexandra Lomedico, Laran L. Hyder, and Stephen V. Faraone. Adult Outcome of Attention-Deficit/Hyperactivity Disorder: A Controlled 16-Year Follow-Up Study. *The Journal of Clinical Psychiatry*, 73(07):941–950, July 2012.

[35] C D Binnie and P F Prior. Electroencephalography. *Journal of Neurology, Neurosurgery & Psychiatry*, 57(11):1308–1319, November 1994.

[36] J. D. Bonita, L. C. C. Ambolode, B. M. Rosenberg, C. J. Cellucci, T. A. A. Watanabe, P. E. Rapp, and A. M. Albano. Time domain measures of inter-channel EEG correlations: A comparison of linear, nonparametric and nonlinear measures. *Cognitive Neurodynamics*, 8(1):1–15, February 2014.

[37] Clemens Brunner, Martin Billinger, Martin Seeber, Timothy R. Mullen, and Scott Makeig. Volume Conduction Influences Scalp-Based Connectivity Estimates. *Frontiers in Computational Neuroscience*, 10, 2016.

[38] Randy L. Buckner, Jessica R. Andrews-Hanna, and Daniel L. Schacter. *The Brain's Default Network*: *Anatomy, Function, and Relevance to Disease*. *Annals of the New York Academy of Sciences*, 1124(1):1–38, March 2008.

[39] Ed Bullmore and Olaf Sporns. Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, 10(3):186–198, March 2009.

[40] George Bush. Attention-Deficit/Hyperactivity Disorder and Attention Networks. *Neuropsychopharmacology*, 35(1):278–300, January 2010.

[41] György Buzsáki and Andreas Draguhn. Neuronal Oscillations in Cortical Networks. *Science*, 304(5679):1926–1929, June 2004.

[42] Roberto Cabeza, Florin Dolcos, Reiko Graham, and Lars Nyberg. Similarities and Differences in the Neural Correlates of Episodic Memory Retrieval and Working Memory. *NeuroImage*, 16(2):317–330, June 2002.

[43] Enoch Callaway. Hyperactive Children's Event-Related Potentials Fail to Support Underarousal and Maturational-Lag Theories. *Archives of General Psychiatry*, 40(11):1243, November 1983.

[44] Miao Cao, Ni Shu, Qingjiu Cao, Yufeng Wang, and Yong He. Imaging Functional and Structural Brain Connectomics in

Attention-Deficit/Hyperactivity Disorder. *Molecular Neurobiology*, 50(3):1111–1123, December 2014.

[45] F. Xavier Castellanos and Erika Proal. Large-Scale Brain Systems in ADHD: Beyond the Prefrontal-Striatal Model. *Trends in Cognitive Sciences*, 16(1):17–26, January 2012.

[46] Robert J. Chabot and Gordon Serfontein. Quantitative electroencephalographic profiles of children with attention deficit disorder. *Biological Psychiatry*, 40(10):951–963, November 1996.

[47] Chunli Chen, Huan Yang, Yasong Du, Guangzhi Zhai, Hesheng Xiong, Dezhong Yao, Peng Xu, Jianhua Gong, Gang Yin, and Fali Li. Altered Functional Connectivity in Children with ADHD Revealed by Scalp EEG: An ERP Study. *Neural Plasticity*, 2021:1–9, May 2021.

[48] I-Chun Chen, Chih-Hao Chang, Yang Chang, Dar-Shong Lin, Cheng-Hsiu Lin, and Li-Wei Ko. Neural Dynamics for Facilitating ADHD Diagnosis in Preschoolers: Central and Parietal Delta Synchronization in the Kiddie Continuous Performance Test. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 29:1524–1533, 2021.

[49] I-Chun Chen, Chia-Ling Chen, Chih-Hao Chang, Zuo-Cian Fan, Yang Chang, Dar-Shong Lin, Cheng-Hsiu Lin, and Li-Wei Ko. Task-Related Brain Oscillation and Coherence Analysis Between High and Low Cognitive Proficiencies in Preschoolers With Attention Deficit Hyperactivity Disorder. Preprint, In Review, December 2020.

[50] Rung-Ching Chen, Christine Dewi, Su-Wen Huang, and Rezzy Eko Caraka. Selecting critical features for data classification based on machine learning methods. *Journal of Big Data*, 7(1):52, December 2020.

[51] Xue-wen Chen and Michael Wasikowski. FAST: A roc-based feature selection metric for small samples and imbalanced data classification problems. In *Proceeding of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD 08*, page 124, Las Vegas, Nevada, USA, 2008. ACM Press.

[52] Adam R. Clarke, Robert J. Barry, Rory McCarthy, and Mark Selikowitz. Electroencephalogram differences in two subtypes of Attention-Deficit/Hyperactivity Disorder. *Psychophysiology*, 38(2):212–221, March 2001.

[53] Mike X Cohen. *Analyzing Neural Time Series Data: Theory and Practice*. The MIT Press, 2014.

[54] C Keith Conners. Continuous Performance Test 3rd Edition. page 9, 2014.

[55] Nicholas R. Cooper, Adrian P. Burgess, Rodney J. Croft, and John H. Gruzelier. Investigating evoked and induced electroencephalogram activity in task-related alpha power increases during an internally directed attention task:. *NeuroReport*, 17(2):205–208, February 2006.

[56] Samuele Cortese, Clare Kelly, Camille Chabernaud, Erika Proal, Adriana Di Martino, Michael P. Milham, and F. Xavier Castellanos. Toward Systems Neuroscience of ADHD: A Meta-Analysis of 55 fMRI Studies. *American Journal of Psychiatry*, 169(10):1038–1055, October 2012.

[57] Ana Cubillo, Rozmin Halari, Anna Smith, Eric Taylor, and Katya Rubia. A review of fronto-striatal and fronto-cortical brain abnormalities in children and adults with Attention Deficit Hyperactivity Disorder (ADHD) and new evidence for dysfunction in adults with ADHD during motivation and attention. *Cortex*, 48(2):194–215, February 2012.

[58] Arnaud Delorme and Scott Makeig. EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1):9–21, March 2004.

[59] Jakob I. Doerrfuss, Tayfun Kilic, Michael Ahmadi, Martin Holtkamp, and Joachim E. Weber. Quantitative and Qualitative EEG as a Prediction Tool for Outcome and Complications in Acute Stroke Patients. *Clinical EEG and Neuroscience*, 51(2):121–129, March 2020.

[60] Anderson dos Santos Siqueira, Claudinei Eduardo Biazoli Junior, William Edgar Comfort, Luis Augusto Rohde, and João Ricardo Sato. Abnormal Functional Resting-State Networks in ADHD: Graph Theory and Pattern Recognition Analysis of fMRI Data. *BioMed Research International*, 2014:1–10, 2014.

[61] R. A. Dykman, P. J. Holcomb, D. M. Oglesby, and P. T. Ackerman. Electrocortical frequencies in hyperactive, learning-disabled, mixed, and normal children. *Biological Psychiatry*, 17(6):675–685, June 1982.

[62] Ali Ekhlasi, Ali Motie Nasrabadi, and Mohammadreza Mohammadi. Classification of the Children with ADHD and Healthy Children Based on

the Directed Phase Transfer Entropy of EEG Signals. *Frontiers in Biomedical Technologies*, June 2021.

[63] Brian Everitt and David C. Howell, editors. *Encyclopedia of Statistics in Behavioral Science*. John Wiley & Sons, Hoboken, N.J, 2005.

[64] J. Fan, J. Byrne, M. S. Worden, K. G. Guise, B. D. McCandliss, J. Fossella, and M. I. Posner. The Relation of Brain Oscillations to Attentional Networks. *Journal of Neuroscience*, 27(23):6197–6206, June 2007.

[65] Farzad V. Farahani, Waldemar Karwowski, and Nichole R. Lighthall. Application of Graph Theory for Identifying Connectivity Patterns in Human Brain Networks: A Systematic Review. *Frontiers in Neuroscience*, 13:585, June 2019.

[66] Michael D. Fox, Maurizio Corbetta, Abraham Z. Snyder, Justin L. Vincent, and Marcus E. Raichle. Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems. *Proceedings of the National Academy of Sciences*, 103(26):10046–10051, June 2006.

[67] Karl J. Friston. Functional and Effective Connectivity: A Review. *Brain Connectivity*, 1(1):13–36, January 2011.

[68] Sarah Furlong. Resting-state EEG Connectivity in Young Children with ADHD. page 18.

[69] Sarah Furlong, Jessica R. Cohen, Joseph Hopfinger, Jenna Snyder, Madeline M. Robertson, and Margaret A. Sheridan. Resting-state EEG Connectivity in Young Children with ADHD. *Journal of Clinical Child & Adolescent Psychology*, pages 1–17, August 2020.

[70] A Gevins, M E Smith, L McEvoy, and D Yu. High-resolution EEG mapping of cortical activation related to working memory: Effects of task difficulty, type of processing, and practice. *Cerebral Cortex*, 7(4):374–385, June 1997.

[71] Amir Hossein Ghaderi, Cognitive Neuroscience Laboratory, Department of Psychology, University of Tabriz, Tabriz, Iran., Mohammad Ali Nazari, Cognitive Neuroscience Laboratory, Department of Psychology, University of Tabriz, Tabriz, Iran., Hassan Shahrokhi, Research Center of Psychiatry and Behavioral Sciences, Tabriz University of Medical Sciences, Tabriz, Iran., Amir Hossein Darooneh, and Department of Physics, Faculty of Sciences, University of Zanjan, Zanjan, Iran. Functional Brain Connectivity Differences Between Different ADHD Presentations: Impaired Functional Segregation in

ADHD-Combined Presentation but not in ADHD-Inattentive Presentation. *Basic and Clinical Neuroscience Journal*, 8(4):267–278, July 2017.

[72] Aditi Goel and Saurabh Kr. Srivastava. Role of Kernel Parameters in Performance Evaluation of SVM. In *2016 Second International Conference on Computational Intelligence & Communication Technology (CICT)*, pages 166–169, Ghaziabad, India, February 2016. IEEE.

[73] Julian Gonzalez, Guzmán Alba, Ernesto Pereda, Soledad Mañas, Almudena González, and Leopoldo Méndez. Electroencephalography signatures of attention-deficit/hyperactivity disorder: Clinical utility. *Neuropsychiatric Disease and Treatment*, page 2755, October 2015.

[74] Sara L. Gonzalez Andino, Cristoph M. Michel, Gregor Thut, Theodor Landis, and Rolando Grave de Peralta. Prediction of response speed by anticipatory high-frequency (gamma band) oscillations in the human brain. *Human Brain Mapping*, 24(1):50–58, January 2005.

[75] Cyril Goutte and Eric Gaussier. A Probabilistic Interpretation of Precision, Recall and F-Score, with Implication for Evaluation. In David Hutchison, Takeo Kanade, Josef Kittler, Jon M. Kleinberg, Friedemann Mattern, John C. Mitchell, Moni Naor, Oscar Nierstrasz, C. Pandu Rangan, Bernhard Steffen, Madhu Sudan, Demetri Terzopoulos, Dough Tygar, Moshe Y. Vardi, Gerhard Weikum, David E. Losada, and Juan M. Fernández-Luna, editors, *Advances in Information Retrieval*, volume 3408, pages 345–359. Springer Berlin Heidelberg, Berlin, Heidelberg, 2005.

[76] Lilli Haar, Katharina Anding, Konstantin Trambitckii, and Gunther Notni. Comparison between Supervised and Unsupervised Feature Selection Methods:. In *Proceedings of the 8th International Conference on Pattern Recognition Applications and Methods*, pages 582–589, Prague, Czech Republic, 2019. SCITEPRESS - Science and Technology Publications.

[77] Henry Han and Xiaoqian Jiang. Overcome Support Vector Machine Diagnosis Overfitting. *Cancer Informatics*, 13s1:CIN.S13875, January 2014.

[78] P. Hartmann, A. Ramseier, F. Gudat, M. J. Mihatsch, and W. Polasek. [Normal weight of the brain in adults in relation to age, sex, body height and weight]. *Der Pathologe*, 15(3):165–170, June 1994.

[79] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning*. Springer Series in Statistics. Springer New York, New York, NY, 2009.

[80] Neep Hazarika, Jean Zhu Chen, Ah Chung Tsoi, and Alex Sergejew. Classification of EEG signals using the wavelet transform. *Signal Processing*, 59(1):61–72, May 1997.

[81] Daniel F. Hermens, Eleonore X.C. Soei, Simon D. Clarke, Michael R. Kohn, Evian Gordon, and Leanne M. Williams. Resting EEG theta activity predicts cognitive performance in attention-deficit hyperactivity disorder. *Pediatric Neurology*, 32(4):248–256, April 2005.

[82] Uwe Herwig, Peyman Satrapi, and Carlos Schönfeldt-Lecuona. Using the International 10-20 EEG System for Positioning of Transcranial Magnetic Stimulation. *Brain Topography*, 16(2):95–99, 24.

[83] Chih-Wei Hsu, Chih-Chung Chang, and Chih-Jen Lin. A Practical Guide to Support Vector Classification. page 16.

[84] Trevor Huff, Navid Mahabadi, and Prasanna Tadi. Neuroanatomy, Visual Cortex. In *StatPearls*. StatPearls Publishing, Treasure Island (FL), 2021.

[85] Mahdi Jalili, Elham Barzegaran, and Maria G. Knyazeva. Synchronization of EEG: Bivariate and Multivariate Measures. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 22(2):212–221, March 2014.

[86] Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. *An Introduction to Statistical Learning*, volume 103 of *Springer Texts in Statistics*. Springer New York, New York, NY, 2013.

[87] Troy Janzen, Ken Graap, Stephan Stephanson, Wilma Marshall, and George Fitzsimmons. Differences in baseline EEG measures for ADD and Normally Achieving preadolescent males. *Biofeedback and Self-Regulation*, 20(1):65–82, March 1995.

[88] Khalid H. Jawabri and Sandeep Sharma. Physiology, Cerebral Cortex Functions. In *StatPearls*. StatPearls Publishing, Treasure Island (FL), 2021.

[89] Jürgen Kayser and Craig E. Tenke. On the benefits of using surface Laplacian (current source density) methodology in electrophysiology. *International Journal of Psychophysiology*, 97(3):171–173, September 2015.

[90] C. Keith Conners, Gill Sitarenios, and Lindsay E. Ayearst. Conners' Continuous Performance Test Third Edition. In Jeffrey S. Kreutzer, John DeLuca, and Bruce Caplan, editors, *Encyclopedia of Clinical Neuropsychology*, pages 929–933. Springer International Publishing, Cham, 2018.

[91] Vikas Khullar, Karuna Salgotra, Harjit Pal Singh, and Davinder Pal Sharma. Deep Learning-Based Binary Classification of ADHD Using Resting State MR Images. *Augmented Human Research*, 6(1):5, December 2021.

[92] Hanni Kiiski, Marc Bennett, Laura M. Rueda-Delgado, Francesca R. Farina, Rachel Knight, Rory Boyle, Darren Roddy, Katie Grogan, Jessica Bramham, Clare Kelly, and Robert Whelan. EEG spectral power, but not theta/beta ratio, is a neuromarker for adult ADHD. *European Journal of Neuroscience*, 51(10):2095–2109, May 2020.

[93] Jae-Jin Kim, Benedicto Crespo-Facorro, Nancy C. Andreasen, Daniel S. O'Leary, Baiquan Zhang, Gregory Harris, and Vincent A. Magnotta. An MRI-Based Parcellation Method for the Temporal Lobe. *NeuroImage*, 11(4):271–288, April 2000.

[94] Timo Kirschstein and Rüdiger Köhling. What is the Source of the EEG? *Clinical EEG and Neuroscience*, 40(3):146–149, July 2009.

[95] G. H. Klem, H. O. Lüders, H. H. Jasper, and C. Elger. The ten-twenty electrode system of the International Federation. The International Federation of Clinical Neurophysiology. *Electroencephalography and Clinical Neurophysiology. Supplement*, 52:3–6, 1999.

[96] Wolfgang Klimesch. EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis. *Brain Research Reviews*, 29(2):169–195, April 1999.

[97] Mark A Kramer. An Introduction to Field Analysis Techniques: The Power Spectrum and Coherence. page 8, 2013.

[98] Juri D. Kropotov. Chapter 3 - Beta Rhythms. In Juri D. Kropotov, editor, *Quantitative EEG, Event-Related Potentials and Neurotherapy*, pages 59–76. Academic Press, San Diego, January 2009.

[99] Deping Kuang, Xiaojiao Guo, Xiu An, Yilu Zhao, and Lianghua He. Discrimination of ADHD Based on fMRI Data with Deep Belief Network. In David Hutchison, Takeo Kanade, Josef Kittler, Jon M. Kleinberg, Alfred Kobsa, Friedemann Mattern, John C. Mitchell, Moni Naor, Oscar Nierstrasz, C. Pandu Rangan, Bernhard Steffen, Demetri Terzopoulos, Doug Tygar, Gerhard Weikum, De-Shuang Huang, Kyungsook Han, and Michael Gromiha, editors, *Intelligent Computing in Bioinformatics*, volume 8590, pages 225–232. Springer International Publishing, Cham, 2014.

[100] Max Kuhn and Kjell Johnson. *Applied Predictive Modeling*. Springer, New York, 2013.

[101] Max Kuhn and Kjell Johnson. *Feature Engineering and Selection: A Practical Approach for Predictive Models*. Chapman & Hall/CRC Data Science Series. CRC Press, Taylor & Francis Group, Boca Raton London New York, 2020.

[102] E. W. Lang, A. M. Tomé, I. R. Keck, J. M. Górriz-Sáez, and C. G. Puntonet. Brain Connectivity Analysis: A Short Survey. *Computational Intelligence and Neuroscience*, 2012:1–21, 2012.

[103] Martin G. Larson. Analysis of Variance. *Circulation*, 117(1):115–121, January 2008.

[104] Xu Lei and Keren Liao. Understanding the Influences of EEG Reference: A Large-Scale Brain Network Perspective. *Frontiers in Neuroscience*, 11:205, 2017.

[105] Agatha Lenartowicz and Sandra K. Loo. Use of EEG to Diagnose ADHD. *Current Psychiatry Reports*, 16(11):498, November 2014.

[106] Agatha Lenartowicz, Ali Mazaheri, Ole Jensen, and Sandra K Loo. ABERRANT MODULATION OF BRAIN OSCILLATORY ACTIVITY AND ATTENTIONAL IMPAIRMENT IN ADHD. page 24, 2019.

[107] Sheng-Fu Liang, Tsung-Hao Hsieh, Pin-Tzu Chen, Ming-Long Wu, Chun-Chia Kung, Chun-Yu Lin, and Fu-Zen Shaw. Differentiation between resting-state fMRI data from ADHD and normal subjects: Based on functional connectivity and machine learning. In *2012 International Conference on Fuzzy Theory and Its Applications (iFUZZY2012)*, pages 294–298, Taichung, Taiwan, November 2012. IEEE.

[108] Yu-Chi Liao, Nai-Wen Guo, Bei-Yi Su, Shin-Jaw Chen, Hsing-Fang Tsai, and Kuan-Ying Lee. Frontal Beta Activity in the Meta-Intention of Children With Attention Deficit Hyperactivity Disorder. *Clinical EEG and Neuroscience*, 52(2):136–143, April 2021.

[109] Gabriel Romero Liguori and Luiz Felipe Pinho Moreira. Operating with Data - Statistics for the Cardiovascular Surgeon: Part III. Comparing Groups. *Brazilian Journal of Cardiovascular Surgery*, 33(6), 2018.

[110] Hsiang-Yuan Lin, Wen-Yih Isaac Tseng, Meng-Chuan Lai, Kayako Matsuo, and Susan Shur-Fen Gau. Altered Resting-State Frontoparietal Control

Network in Children with Attention-Deficit/Hyperactivity Disorder. *Journal of the International Neuropsychological Society*, 21(4):271–284, April 2015.

[111] Xiang Liu, Oleksandr Makeyev, and Walter Besio. Improved Spatial Resolution of Electroencephalogram Using Tripolar Concentric Ring Electrode Sensors. *Journal of Sensors*, 2020:1–9, June 2020.

[112] Livia Livint Popa, Hanna Dragos, Cristina Pantelemon, Olivia Verisezan Rosu, and Stefan Strilciuc. The Role of Quantitative EEG in the Diagnosis of Neuropsychiatric Disorders. *Journal of Medicine and Life*, 13(1):8–15, 2020.

[113] Sandra K. Loo and Russell A. Barkley. Clinical Utility of EEG in Attention Deficit Hyperactivity Disorder. *Applied Neuropsychology*, 12(2):64–76, June 2005.

[114] Bruno J. Losier, Patrick J. McGrath, and Raymond M. Klein. Error Patterns on the Continuous Performance Test in Non-Medicated and Medicated Samples of Children With and Without ADHD: A Meta-Analytic Review. *Journal of Child Psychology and Psychiatry*, 37(8):971–987, November 1996.

[115] Joel F. Lubar. Discourse on the development of EEG diagnostics and biofeedback for attention-deficit/hyperactivity disorders. *Biofeedback and Self-Regulation*, 16(3):201–225, September 1991.

[116] Kip A. Ludwig, Rachel M. Miriani, Nicholas B. Langhals, Michael D. Joseph, David J. Anderson, and Daryl R. Kipke. Using a Common Average Reference to Improve Cortical Neuron Recordings From Microelectrode Arrays. *Journal of Neurophysiology*, 101(3):1679–1689, March 2009.

[117] Keitaro Machida, Michael Murias, and Katherine A. Johnson. Electrophysiological Correlates of Response Time Variability During a Sustained Attention Task. *Frontiers in Human Neuroscience*, 13:363, October 2019.

[118] Elisa Magosso, Francesca De Crescenzio, Giulia Ricci, Sergio Piastra, and Mauro Ursino. EEG Alpha Power Is Modulated by Attentional Changes during Cognitive Tasks and Virtual Reality Immersion. *Computational Intelligence and Neuroscience*, 2019:1–18, June 2019.

[119] Thomas Maiwald, Enno Mammen, Swagata Nandi, and Jens Timmer. Surrogate Data — A Qualitative and Quantitative Analysis. In Rainer Dahlhaus, Jürgen Kurths, Peter Maass, and Jens Timmer, editors,

*Mathematical Methods in Signal Processing and Digital Image Analysis*, pages 41–74. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.

[120] Sebastian Markett, Martin Reuter, Christian Montag, Gesine Voigt, Bernd Lachmann, Sarah Rudorf, Christian E. Elger, and Bernd Weber. Assessing the function of the fronto-parietal attention network: Insights from resting-state fMRI and the attentional network test: Assessing the Function of the Fronto-Parietal Attention Network. *Human Brain Mapping*, 35(4):1700–1709, April 2014.

[121] Oscar Martin-Santiago, Javier Gomez-Pilar, Alba Lubeiro, Marta Ayuso, Jesús Poza, Roberto Hornero, Myriam Fernandez, Sonia Ruiz de Azua, Cesar Valcarcel, and Vicente Molina. Modulation of brain network parameters associated with subclinical psychotic symptoms. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 66:54–62, April 2016.

[122] Milos Matoušek, Peder Rasmussen, and Christopher Gillberg. EEG Frequency Analysis in Children with So-Called Minimal Brain Dysfunction and Related Disorders. In B. Saletu, C. Perris, and D. Kemali, editors, *Advances in Biological Psychiatry*, volume 15, pages 102–108. S. Karger AG, April 1985.

[123] Masato Matsuura, Yoshiro Okubo, Michio Toru, Takuya Kojima, Yu He, Yucun Shen, and Chung Kyoon Lee. A cross-national EEG study of children with emotional and behavioral problems: A WHO collaborative study in the Western Pacific region. *Biological Psychiatry*, 34(1-2):59–65, July 1993.

[124] Mayo Clinic College of Medicine, Rochester, Minnesota, Erik St. Louis, Lauren Frey, University of Colorado, Denver, Colorado, Jeffrey Britton, Mayo Clinic College of Medicine, Rochester, Minnesota, Jennifer Hopp, University of Maryland, Baltimore, Maryland, Pearce Korb, University of Colorado, Denver, Colorado, Mohamad Koubeissi, George Washington University, Washington, District of Columbia, William Lievens, University of Alabama, Birmingham, Alabama, Elia Pestana-Knight, and Cleveland Clinic Foundation, Cleveland, Ohio. *Electroencephalography (EEG): An Introductory Text and Atlas of Normal and Abnormal Findings in Adults, Children, and Infants*. American Epilepsy Society, 2016.

[125] Samuele Mazzanti. "MRMR" Explained Exactly How You Wished Someone Explained to You. https://towardsdatascience.com/mrmr-explained-exactly-how-you-wished-someone-explained-to-you-9cf4ed27458b, April 2021.

[126] Christoph M. Michel and Micah M. Murray. Towards the utilization of EEG as a brain imaging tool. *NeuroImage*, 61(2):371–385, June 2012.

[127] Giorgia Michelini, Joseph Jurgiel, Ioannis Bakolis, Celeste H. M. Cheung, Philip Asherson, Sandra K. Loo, Jonna Kuntsi, and Iman Mohammad-Rezazadeh. Atypical functional connectivity in adolescents and adults with persistent and remitted ADHD during a cognitive control task. *Translational Psychiatry*, 9(1):137, December 2019.

[128] Hiroaki Mizuhara, Li-Qun Wang, Koichiro Kobayashi, and Yoko Yamaguchi. A long-range cortical network emerging with theta oscillation in a mental task:. *NeuroReport*, 15(8):1233–1238, June 2004.

[129] Saleh M. H. Mohamed, Norbert A. Börger, Reint H. Geuze, and Jaap J. van der Meere. Brain lateralization and self-reported symptoms of ADHD in a population sample of adults: A dimensional approach. *Frontiers in Psychology*, 6, September 2015.

[130] Vincent J. Monastra. *Unlocking the Potential of Patients with ADHD: A Model for Clinical Practice.* American Psychological Association, Washington, 2008.

[131] Vincent J. Monastra, Joel F. Lubar, and Michael Linden. The development of a quantitative electroencephalographic scanning process for attention deficit–hyperactivity disorder: Reliability and validity studies. *Neuropsychology*, 15(1):136–144, 2001.

[132] Vincent J. Monastra, Joel F. Lubar, Michael Linden, Peter VanDeusen, George Green, William Wing, Arthur Phillips, and T. Nick Fenger. Assessing attention deficit hyperactivity disorder via quantitative electroencephalography: An initial validation study. *Neuropsychology*, 13(3):424–433, July 1999.

[133] Md. Moniruzzaman Moni and Muhammad Shuaib. A Comparison of Power of Normality Tests: Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors, Anderson-Darling and Jarque-Bera Tests. 2015.

[134] Andreas Mueller, Gian Candrian, Juri D Kropotov, Valery A Ponomarev, and Gian-Marco Baschera. Classification of ADHD patients on the basis of independent ERP components using a machine learning system. *Nonlinear Biomedical Physics*, 4(S1):S1, June 2010.

[135] Mohammad Ali Nazari, Fabrice Wallois, Ardalan Aarabi, and Patrick Berquin. Dynamic changes in quantitative electroencephalogram during continuous performance test in children with attention-deficit/hyperactivity disorder. *International Journal of Psychophysiology*, 81(3):230–236, September 2011.

[136] Adeleh Dehghani Nazhvani, Reza Boostani, Somayeh Afrasiabi, and Khadijeh Sadatnezhad. Classification of ADHD and BMD patients using visual evoked potential. *Clinical Neurology and Neurosurgery*, 115(11):2329–2335, November 2013.

[137] Kerryn Neulinger, Joanne Oram, Helen Tinson, John O'Gorman, and David H. K. Shum. Prospective memory and frontal lobe function. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition*, 23(2):171–183, 2016.

[138] Guido Nolte, Ou Bai, Lewis Wheaton, Zoltan Mari, Sherry Vorbach, and Mark Hallett. Identifying true brain interaction from EEG data using the imaginary part of coherency. *Clinical Neurophysiology*, 115(10):2292–2307, October 2004.

[139] Kirill V. Nourski. Auditory processing in the human cortex: An intracranial electrophysiology perspective. *Laryngoscope Investigative Otolaryngology*, 2(4):147–156, August 2017.

[140] Pablo Núñez, Jesús Poza, Carlos Gómez, Saúl J. Ruiz-Gómez, Víctor Rodríguez-González, Miguel Ángel Tola-Arribas, Mónica Cano, and Roberto Hornero. Analysis of Electroencephalographic Dynamic Functional Connectivity in Alzheimer's Disease. In Lenka Lhotska, Lucie Sukupova, Igor Lacković, and Geoffrey S. Ibbott, editors, *World Congress on Medical Physics and Biomedical Engineering 2018*, volume 68/2, pages 165–168. Springer Singapore, Singapore, 2019.

[141] Paul L. Nunez, Ramesh Srinivasan, Andrew F. Westdorp, Ranjith S. Wijesinghe, Don M. Tucker, Richard B. Silberstein, and Peter J. Cadusch. EEG coherency. *Electroencephalography and Clinical Neurophysiology*, 103(5):499–515, November 1997.

[142] Geir Ogrim, Juri Kropotov, and Knut Hestad. The quantitative EEG theta/beta ratio in attention deficit/hyperactivity disorder and normal controls: Sensitivity, specificity, and behavioral correlates. *Psychiatry Research*, 198(3):482–488, August 2012.

[143] Elzbieta Olejarczyk and Wojciech Jernajczyk. Graph-based analysis of brain connectivity in schizophrenia. *PLOS ONE*, 12(11):e0188629, November 2017.

[144] Julie Onton and Scott Makeig. Information-based modeling of event-related brain dynamics. In *Progress in Brain Research*, volume 159, pages 99–120. Elsevier, 2006.

[145] Husein Perez and Joseph H. M. Tah. Improving the Accuracy of Convolutional Neural Networks by Identifying and Removing Outlier Images in Datasets Using t-SNE. *Mathematics*, 8(5):662, April 2020.

[146] Steven E. Petersen and Olaf Sporns. Brain Networks and Cognitive Architectures. *Neuron*, 88(1):207–219, October 2015.

[147] G. Pfurtscheller, A. Stancák, and Ch. Neuper. Event-related synchronization (ERS) in the alpha band — an electrophysiological correlate of cortical idling: A review. *International Journal of Psychophysiology*, 24(1-2):39–46, November 1996.

[148] Dean Prichard and James Theiler. Generating surrogate data for time series with several simultaneously measured variables. *Physical Review Letters*, 73(7):951–954, August 1994.

[149] Milos Radovic, Mohamed Ghalwash, Nenad Filipovic, and Zoran Obradovic. Minimum redundancy maximum relevance feature selection approach for temporal gene expression data. *BMC Bioinformatics*, 18(1):9, December 2017.

[150] Supratim Ray, Ernst Niebur, Steven S. Hsiao, Alon Sinai, and Nathan E. Crone. High-frequency gamma activity (80–150Hz) is increased in human cortex during selective attention. *Clinical Neurophysiology*, 119(1):116–133, January 2008.

[151] Andrew F. Rossi, Luiz Pessoa, Robert Desimone, and Leslie G. Ungerleider. The prefrontal cortex and the executive control of attention. *Experimental brain research. Experimentelle Hirnforschung. Experimentation cerebrale*, 192(3):489–497, January 2009.

[152] Mikail Rubinov and Olaf Sporns. Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage*, 52(3):1059–1069, September 2010.

[153] Seward B. Rutkove. Introduction to Volume Conduction. In Andrew S. Blum and Seward B. Rutkove, editors, *The Clinical Neurophysiology Primer*, pages 43–53. Humana Press, Totowa, NJ, 2007.

[154] Jacqueline F. Saad, Michael R. Kohn, Simon Clarke, Jim Lagopoulos, and Daniel F. Hermens. Is the Theta/Beta EEG Marker for ADHD Inherently Flawed? *Journal of Attention Disorders*, 22(9):815–826, July 2018.

[155] T H Sander, A Bock, S Leistner, A Kühn, and L Trahms. Coherence and imaginary part of coherency identifies cortico-muscular and cortico-thalamic coupling. In *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, pages 1714–1717, Buenos Aires, August 2010. IEEE.

[156] Saeid Sanei and Jonathan Chambers. EEG Signal Processing. page 313.

[157] Martin Sarter, Ben Givens, and John P Bruno. The cognitive neuroscience of sustained attention: Where top-down meets bottom-up. *Brain Research Reviews*, 35(2):146–160, April 2001.

[158] James H. Satterfield, Dennis P. Cantwell, Leonard I. Lesser, and Robert L. Podosin. Physiological Studies of the Hyperkinetic Child: I. *American Journal of Psychiatry*, 128(11):1418–1424, May 1972.

[159] P. Sauseng, J. Hoppe, W. Klimesch, C. Gerloff, and F. C. Hummel. Dissociation of sustained attention from central executive functions: Local activity and interregional connectivity in the theta range. *European Journal of Neuroscience*, 25(2):587–593, January 2007.

[160] Kapil Sayal, Vibhore Prasad, David Daley, Tamsin Ford, and David Coghill. ADHD in children and young people: Prevalence, care pathways, and service provision. *The Lancet Psychiatry*, 5(2):175–186, February 2018.

[161] Thomas E. Schlaepfer and Charles B. Nemeroff, editors. *Neurobiology of Psychiatric Disorders*. Number 3rd ser., v. 105 in Handbook of Clinical Neurology. Elsevier, Edinburgh, 2012.

[162] Jeremy D. Schmahmann and David Caplan. Cognition, emotion and the cerebellum. *Brain*, 129(2):290–292, February 2006.

[163] Neil C Schwertman, Margaret Ann Owens, and Robiah Adnan. A simple more general boxplot method for identifying outliers. *Computational Statistics & Data Analysis*, 47(1):165–174, August 2004.

[164] Richard B. Silberstein, Andrew Pipingas, Maree Farrow, Florence Levy, Con K. Stough, and David A. Camfield. Brain functional connectivity abnormalities in attention-deficit hyperactivity disorder. *Brain and Behavior*, 6(12), December 2016.

[165] Timothy J. Silk, Charles B. Malpas, Richard Beare, Daryl Efron, Vicki Anderson, Philip Hazell, Brad Jongeling, Jan M. Nicholson, and Emma Sciberras. A network analysis approach to ADHD symptoms: More than the sum of its parts. *PLOS ONE*, 14(1):e0211053, January 2019.

[166] Timothy J. Silk, Alasdair Vance, Nicole Rinehart, John L. Bradshaw, and Ross Cunnington. Dysfunction in the Fronto-Parietal Network in Attention Deficit Hyperactivity Disorder (ADHD): An fMRI Study. *Brain Imaging and Behavior*, 2(2):123–131, June 2008.

[167] W. Singer. Neuronal synchrony: A versatile code for the definition of relations? *Neuron*, 24(1):49–65, 111–125, September 1999.

[168] Siuly Siuly, Yan Li, and Yanchun Zhang. *Electroencephalogram (EEG) and Its Background*, pages 3–21. Springer International Publishing, Cham, 2016.

[169] Ortal Slobodin. The Utility of the CPT in the Diagnosis of ADHD in Individuals with Substance Abuse: A Systematic Review. *European Addiction Research*, 26(Suppl. 4-5):283–294, 2020.

[170] Janette L. Smith, Stuart J. Johnstone, and Robert J. Barry. Aiding diagnosis of attention-deficit/hyperactivity disorder and its subtypes: Discriminant function analysis of event-related potential data: Discriminant function analysis of ERP data in AD/HD subtypes. *Journal of Child Psychology and Psychiatry*, 44(7):1067–1075, October 2003.

[171] Julius O. Smith. *Spectral Audio Signal Processing.* Stanford University, CCRMA, Stanford, Calif, 2011.

[172] Steven M. Snyder and James R. Hall. A Meta-analysis of Quantitative EEG Power Associated With Attention-Deficit Hyperactivity Disorder:. *Journal of Clinical Neurophysiology*, 23(5):441–456, October 2006.

[173] Marina Sokolova, Nathalie Japkowicz, and Stan Szpakowicz. Beyond Accuracy, F-Score and ROC: A Family of Discriminant Measures for Performance Evaluation. In David Hutchison, Takeo Kanade, Josef Kittler, Jon M. Kleinberg, Friedemann Mattern, John C. Mitchell, Moni Naor, Oscar Nierstrasz, C. Pandu Rangan, Bernhard Steffen, Madhu Sudan, Demetri Terzopoulos, Dough Tygar, Moshe Y. Vardi, Gerhard Weikum, Abdul Sattar, and Byeong-ho Kang, editors, *AI 2006: Advances in Artificial Intelligence*, volume 4304, pages 1015–1021. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006.

[174] C. Spearman. The Proof and Measurement of Association between Two Things. *The American Journal of Psychology*, 15(1):72, January 1904.

[175] Olaf Sporns, Christopher J. Honey, and Rolf Kötter. Identification and Classification of Hubs in Brain Networks. *PLoS ONE*, 2(10):e1049, October 2007.

[176] Jun-Sang Sunwoo, Kwang Su Cha, and Ki-Young Jung. Computational electroencephalography analysis for characterizing brain networks. *Annals of Clinical Neurophysiology*, 22(2):82–91, October 2020.

[177] Giulia Tacchino, Marta Gandolla, Stefania Coelli, Riccardo Barbieri, Alessandra Pedrocchi, and Anna M. Bianchi. EEG Analysis During Active and Assisted Repetitive Movements: Evidence for Differences in Neural Engagement. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 25(6):761–771, June 2017.

[178] Leanne Tamm, Vinod Menon, Jessica Ringel, and Allan L. Reiss. Event-Related fMRI Evidence of Frontotemporal Involvement in Aberrant Response Inhibition and Task Switching in Attention-Deficit/Hyperactivity Disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, 43(11):1430–1440, November 2004.

[179] Aleksandar Tenev, Silvana Markovska-Simoska, Ljupco Kocarev, Jordan Pop-Jordanov, Andreas Müller, and Gian Candrian. Machine learning approach for classification of ADHD adults. *International Journal of Psychophysiology*, 93(1):162–166, July 2014.

[180] M Teplan. FUNDAMENTALS OF EEG MEASUREMENT. *MEASUREMENT SCIENCE REVIEW*, 2:12, 2002.

[181] Brigitta Tóth, Roland Boha, Márton Pósfai, Zsófia Anna Gaál, Anikó Kónya, Cornelis Jan Stam, and Márk Molnár. EEG synchronization characteristics of functional connectivity and complex network properties of memory maintenance in the delta and theta frequency bands. *International Journal of Psychophysiology*, 83(3):399–402, March 2012.

[182] Lucina Q. Uddin, A.M. Clare Kelly, Bharat B. Biswal, Daniel S. Margulies, Zarrar Shehzad, David Shaw, Manely Ghaffari, John Rotrosen, Lenard A. Adler, F. Xavier Castellanos, and Michael P. Milham. Network homogeneity reveals decreased integrity of default-mode network in ADHD. *Journal of Neuroscience Methods*, 169(1):249–254, March 2008.

[183] Pieter van Mierlo, Margarita Papadopoulou, Evelien Carrette, Paul Boon, Stefaan Vandenberghe, Kristl Vonck, and Daniele Marinazzo. Functional brain connectivity from EEG in epilepsy: Seizure prediction and epileptogenic focus localization. *Progress in Neurobiology*, 121:19–35, October 2014.

[184] Don van Ravenzwaaij and John P. A. Ioannidis. True and false positive rates for different criteria of evaluating statistical evidence from clinical trials. *BMC Medical Research Methodology*, 19(1):218, December 2019.

[185] Simo Vanni, Antti Revonsuo, and Riitta Hari. Modulation of the Parieto-Occipital Alpha Rhythm during Object Detection. *The Journal of Neuroscience*, 17(18):7141–7147, September 1997.

[186] Petra E. Vértes and Edward T. Bullmore. Annual Research Review: Growth connectomics – the organization and reorganization of brain networks during normal and abnormal development. *Journal of Child Psychology and Psychiatry*, 56(3):299–320, March 2015.

[187] Huifang E. Wang, Christian G. BÃ©nar, Pascale P. Quilichini, Karl J. Friston, Viktor K. Jirsa, and Christophe Bernard. A systematic framework for functional connectivity measures. *Frontiers in Neuroscience*, 8, December 2014.

[188] Jinhui Wang, Xinian Zuo, and Yong He. Graph-Based Network Analysis of Resting-State Functional MRI. *Frontiers in Systems Neuroscience*, 4:16, June 2010.

[189] Liang Wang, Chaozhe Zhu, Yong He, Yufeng Zang, QingJiu Cao, Han Zhang, Qiuhai Zhong, and Yufeng Wang. Altered small-world brain functional networks in children with attention-deficit/hyperactivity disorder. *Human Brain Mapping*, 30(2):638–649, February 2009.

[190] P. Welch. The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Transactions on Audio and Electroacoustics*, 15(2):70–73, June 1967.

[191] Timothy O. West, David M. Halliday, Steven L. Bressler, Simon F. Farmer, and Vladimir Litvak. Measuring directed functional connectivity using non-parametric directionality analysis: Validation and comparison with non-parametric Granger Causality. *NeuroImage*, 218:116796, September 2020.

[192] Zhenyu Zhao, Radhika Anand, and Mallory Wang. Maximum Relevance and Minimum Redundancy Feature Selection Methods for a Marketing Machine Learning Platform. *arXiv:1908.05376 [cs, stat]*, August 2019.

# List of Figures

# List of Tables

# List of Acronyms

**ADHD** Attention-Deficit/Hyperactivity Disorder

**EEG** Electroencephalography

**DSM IV-TR** Diagnostic and Statistical Manual of Mental Disorders IV Text Revision

**CNS** Central Nervous System

**FPAN** Frontal–Parietal Attention Network

**VAN** Ventral-Attention Network

**TPJ** Temporal-Parietal Junction

**VFC** Ventral Frontal Cortex

**DMN** Default Mode Network

**QEEG** Quantitative EEG

**AP** Action Potential

**PSP** Postsynaptic Potential

**CEN** Central Executive Network

**PSD** Power Spectral Density

**CPT** Continuous Performance Test

**EOG** Electrooculogram

**ECG** Electrocardiogram

**EMG** Electromyogram

**MEG** Magneto-Electroencephalography

**fMRI** functional Magnetic Resonance Imaging

**SNR** Signal to Noise Ratio

**TBR** Theta/Beta Ratio

**ANN** Artificial Neural Network

**ERP** Event Related Potentials

**ICA** Independent Component Analysis

**CAR** Common Average Referencing

**CSD** Current Source Density

**OE** Omission Error

**CE** Commission Error

**HRT** Hit Response Time

**HRT SD** Hit Response Time Standard Deviation

**IQR** Interquartile Range

**ImCoh** Imaginary Part of the Coherency

**FCN** Functional Connectivity Network

**BCT** Brain Connectivity Toolbox

**ANOVA** Analysis of Variance

**MRMR** Maximum Relevance Minimum Redundancy

**PCA** Principal Component Analysis

**FCQ** F-test Correlation Quotient

**SVM** Support Vector Machines