UNIVERSITY OF SOUTHAMPTON

Faculty of Engineering and Physical Sciences
School of Electronics and Computer Science

Complexity and Synchronisation Analysis of Electroencephalogram Signals for early Prediction of Neurodevelopmental Disorders

by

Dalal Mohammed Bakheet

ORCiD: 0000-0003-2227-6801

A thesis for the degree of
Doctor of Philosophy

July 2022

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING AND PHYSICAL SCIENCES
School of Electronics and Computer Science

Doctor of Philosophy

COMPLEXITY AND SYNCHRONISATION ANALYSIS OF ELECTROENCEPHALOGRAM SIGNALS FOR EARLY PREDICTION OF NEURODEVELOPMENTAL DISORDERS

by Dalal Mohammed Bakheet

Neurodevelopmental disorders (NDDs) are a group of neurological disorders emerging during early development and impacting higher-level brain function. Early identification of NDDs raises the possibility of improving outcomes. However, NDDs are still clinically detected through a subjective evaluation approach that lacks biological evidence and requires an extended period to identify the disorder. This thesis aims to identify biomarkers from the electroencephalogram (EEG) for abnormal brain dynamics and utilise these biomarkers for NDDs prediction. Nonlinear time series analysis methods, mainly complexity (entropies) and synchronisation (node degree of the phase-lag index) features are investigated for this purpose due to their ability to reflect brain dynamics. A machine learning framework that combines advanced nonlinear EEG processing methods for NDDs identification has been developed and evaluated. The proposed framework was first used with a dataset of children with autism spectrum disorder (ASD) and controls to validate its classification efficacy. This exploration resulted in high performance, suggesting complexity features as biomarkers for ASD prediction. The framework was then used to explore at-birth EEG characteristics of infants with Hypoxic-Ischemic Encephalopathy (HIE) who later developed cerebral palsy (CP). High performance has been reached, and the proposed features were reported as potential biomarkers for early CP prediction. A regression model has also been developed to explore the correlation between the EEG characteristics of the HIE infants and their two-year cognitive scores. The results suggest the proposed features as biomarkers for early cognitive function prediction. Finally, the framework was used with a dataset of individuals with a major depressive disorder to validate its ability to predict their depression severity. Compared to the state-of-the-art research, an acceptable regression performance has been reached, and the proposed features have been suggested as biomarkers for depression severity prediction. This work lays the foundation for evidence-based decision-making applications for early prediction of CP and cognitive outcomes of infants with HIE, paving the way for establishing tailored intervention programs at an appropriate point during development to improve the outcomes.

Contents

List of Figures

List of Tables

Declaration of Authorship

I, Dalal Mohammed Bakheet, declare that this thesis and the work presented in it is my own and has been generated by me as the result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Parts of this work have been published as: (Bakheet and Maharatna, 2021; Bakheet et al., 2021; Alotaibi et al., 2022)

Signed:..... Date:.....

Acknowledgements

All praise and thanks to Almighty Allah (Subhanahu Wa Taalaa) for His greatness and for giving me the strength and patience to struggle through during the last years of my degree. It is because of His blessing that it has been possible for me to complete this work.

I would like to take this opportunity to thank my country, represented by the Royal Embassy of Saudi Arabia Cultural Bureau in London, for providing me with this scholarship and their continuous support during the previous academic years.

Special thanks and appreciation are due to Prof. Koushik Maharatna for his constant support and guidance. His continuous encouragement, enthusiasm, and faith in me were the motivation to do my best. I hope this effort will come up to his expectations.

I would also like to take the opportunity to acknowledge Dr. Brigitte Vollmer and Dr. Konn, Daniel, for their co-operation in describing and discussing the provided dataset and their insight on the work done.

Next, I would like to thank my lovely friends, Noura Alotaibi, Kholoud Alghamdi, Muna Altherwi, Ghadah Alzaidy and Nouf Alothman, who supported me throughout my time here in Southampton and being with me in the difficult and sweet moments. A special appreciation is extended to my friend Noura Alotaibi, who is also my colleague in the research group and co-author in several manuscripts. Her shared knowledge and deepest discussions strongly contributed to the completion of this work. Thank you for always being there for me.

Deepest gratitude goes to my beloved parents for their support; without their prayers, I would not have reached this far. Words fail to express my appreciation to my husband, Dr. Hatem Bakheet, whose dedication, support and love have taken the load off my shoulders. My daughters Safana, Salma and Sara, and my son Abdulmalik deserve special mention for their love and support. From the bottom of my heart, Thank you my sweetest family.

To my parents, family, friends and mentors . . .

Definitions and Abbreviations

ACC	Accuracy
AD	Alzheimer's disease
$ADHD$	Attention-deficit/hyperactivity disorder
$ANOVA$	Analysis of variance
AUC	Area under the curve
$aEEG$	Amplitude-integrated electroencephalogram
$ApEn$	Approximate entropy
ASD	Autism spectrum disorder
$BH-FDR$	Benjamini-Hochberg false discovery rate
$BSITD-III$	Bayley scales of infant and toddler development III
$cEEG$	Conventional grading electroencephalogram
CP	Cerebral palsy
DA	Discriminant analysis
$DSM-5$	Diagnostic and statistical manual of mental, fifth edition
$DSM-IV-TR$	Diagnostic and statistical manual of mental disorders, 4th edition, text revision
EEG	Electroencephalography
EMD	Empirical mode decomposition
$EEMD$	Ensemble empirical mode decomposition
ERP	Event-related potential
FE	Facial expression
FFT	Fast Fourier transform
FIR	Finite Impulse Response
$fMRI$	functional magnetic resonance imaging
$HCoh$	Intrahemispheric coherence
HT	Hilbert transform
HIE	Hypoxic-ischemic encephalopathy
ICA	Independent components analysis
$ICD-10$	International classification of diseases, tenth edition
$ICoh$	Interhemispheric coherence
IMF	Intrinsic mode function
(kNN)	k-Nearest neighbors
$LOOCV$	Leave-one-out cross-validation

MAE	Mean absolute error
Max	Maximum
MCI	Mild cognitive impairment
MEG	Magnetoencephalopham
MF	Mean frequency
Min	Minimum
MDD	Major depressive disorder
MRI	Magnetic resonance imaging
$MEMD$	Multivariate empirical mode decomposition
MSE	Multiscale entropy
ND	Node degree
$NDDs$	Neurodevelopmental disorders
$NA-MEMD$	Noise-assisted multivariate empirical mode decomposition
PEn	Permutation entropy
PLI	Phase lag index
PLV	Phase-locking value
PS	Phase synchronisation
PSD	Power spectral density
$qEEG$	Quantitative electroencephalogram
R	Pearson's correlation coefficient
ROI	Region of interest
$RMSE$	Root mean square error
$RUSBoost$	Random undersampling boost
ROC	Receiver operating characteristic
$SampEn$	Sample entropy
$sEMD$	Standard empirical mode decomposition
$STFT$	Short-time Fourier transform
$SMOTEBoost$	Synthetic minority oversampling boost
SNR	Signal to Noise Ratio
$SpEn$	Spectral entropy
Std	Standard deviation
SVM	Support vector machines
TD	Typically developing
TNR	True negative rate
TPR	True positive rate
UHS	University hospital of Southampton
VC	Volume conduction
$WPPSI-III$	Wechsler preschool and primary scale of intelligence III
WT	Wavelet transform

Chapter 1

Introduction

1.1 Motivation

Neurodevelopmental disorders (NDDs) are a group of disorders linked to the development of the neurological system and brain function impairment. NDDs include different disorders such as autism spectrum disorder (ASD), attention-deficit/hyperactivity disorder (ADHD), intellectual disabilities, learning disabilities and cerebral palsy (CP). These disorders are often complex and heterogeneous, and they reveal abnormalities from cellular to behavioral levels. Thus, individuals with NDDs may suffer from difficulties in motor, cognition, language and learning skills (Bowman and Varcin, 2018).

More than 19,000 UK households were interviewed for the 2019/20 family resources survey (GOV.UK, 2021). Based on this survey, children were much more likely affected by social/behavioral and learning impairments than adults. Around 45% of disabled children suffered from a social/behavioral impairment, and 35% reported a learning impairment. These disorders have a severe socio-economic impact on individuals, and in turn, families and society. The study carried out by the London School of Economics and Political Science estimated that the UK spends around £32 billion per year only for managing autism (Iemmi et al., 2017). Early prediction of neurodevelopmental impairments and person-centric intervention could help in preventing severe disabilities, maximising life chances, promoting opportunities for social participation, and improving outcomes (Cioni et al., 2016).

A systematic Cochrane review followed with preterm infants, who are at high risk for NDDs, provided evidence that early intervention during infancy positively affects cognitive and motor outcomes with the cognitive benefits continuing to pre-school age (Spittle et al., 2012). The plasticity phenomenon associated with young brains is the reason behind the efficiency of early intervention. It helps modify structural and functional brain connectivity in response to experiences and environmental stimuli (Berardi

et al., 2000; Cioni et al., 2016). Early intervention is based, in turn, on early NDDs prediction.

NDDs often originate during the intrauterine development and age of birth, then continue throughout life. These disorders often result from different factors that include low birth weight, preterm delivery, early chemical exposure, and deprivation of oxygen during birth processes (leading to hypoxic-ischemic encephalopathy (HIE) brain dysfunction). All of these factors are converging on pathways serving various aspects of brain development (De Felice et al., 2015). Thus, although some behaviors and symptoms of NDDs may only be visible later during individual development, they are likely to have roots in atypical brain structure and function that can be traced back to the early months of life.

Nevertheless, most NDDs are still clinically detected through self-reported tests and behavioral assessments that often rely on parent-rated symptoms (Falkmer et al., 2013). This process is affected by different factors such as uncertainty in answering examination questions, the ability of caregivers to detect and report the disorder's symptoms accurately, and subjective opinions of clinicians in evaluating a condition. Moreover, most NDDs symptoms usually do not appear until at least the second year of the child's life, which delays the disorder identification. Consequently, the critical developmental window for early intervention may be missed (Bosl et al., 2018; Bowman and Varcin, 2018). The critical period of brain plasticity is defined as the window in which the development of brain functional characteristics is strongly dependent and shaped by environmental stimuli and experience (Cioni et al., 2016). These limitations of the behavioral-based evaluation approach stipulate the need for quantitive and objective measures that can be used to predict different NDDs. If further developed in clinical applications, these measures can aid early and accurate prediction of such disorders in the future.

Recently, neuroimaging and neurophysiological techniques, such as functional magnetic resonance imaging (fMRI), magnetoencephalopham (MEG) and electroencephalography (EEG), have been used as promising tools to study brain dynamics that can reflect the abnormalities associated with NDDs. These methods can help to detect impairments in brain functions of individuals even during their early days of life (Bowman and Varcin, 2018).

EEG is considered in this research due to its properties, including portability and comparative ease of use, making it the preferred technique for children and infants. Moreover, the high temporal resolution of the EEG signals results in an accurate estimation of brain dynamics, depending largely on studying neural activation over time (Bowman and Varcin, 2018).

EEG signals usually produce a huge amount of complex and multidimensional data, considered as a vital source for predicting, diagnosing, and planning therapy of NDDs

(Siuly and Zhang, 2016). Although the investigation of such data can be viewed as a valuable tool to identify NDDs, it requires signal analysis methods that provide reliable biological markers (biomarkers) for good decision-making in disorder prediction. Visual inspection of such neurological data may be done by experts/neurologists to detect signals' abnormalities. However, they cannot interpret the recorded data optimally as their human eye-brain system cannot process this vast amount of noisy and high dimensional information (Siuly and Zhang, 2016).

Signal analysis plays an important role in understanding the brain dynamics by mapping the recorded signals into quantitive measures that could reveal abnormal brain function. Various methods have been proposed for quantifying the EEG signals to reflect different NDDs, but developing effective and accurate techniques that account for the critical characteristics of EEG has always been challenging. Nonstationarity is the first behavior of EEG where the signal characteristics change over time (Sanei and Chambers, 2007). Time-frequency domain methods are considered the most appropriate approaches to consider such a property of the signals (Sanei and Chambers, 2007). On the other hand, the EEG signal is considered an output of a nonlinear system, where the changes in the brain metabolism due to physiological and biological phenomena or the activity of the brain itself make the brain (as a system) nonlinear (Sanei and Chambers, 2007). Studies suggested that the nonlinear structure could be found in 60 –70% of the signals, even in newborns (Stam, 2005).

Although the analysis of such a system is very complicated, the advances in the theory of nonlinear dynamical systems allow the study of pattern formation in the complex neuronal networks of the brain (Stam, 2005). The nonlinear time series analysis of EEG includes two primary approaches; one characterises the nonlinear behavior within the EEG signal (such as the entropy measures), while the second quantifies the nonlinear synchronisation between signals from different brain regions (such as the phase synchronisation measures). For simplicity, the EEG complexity and EEG synchronisation are used to refer to the first and second analysis methods, respectively. Despite these two approaches quantifying different characteristics of EEG, they depend in part on each other in which the complexity of a single signal does, to some extent, reflect the interaction with other signals. Likewise, the synchronisation between signals depends on the complexity of each signal (Stam, 2005).

Comparing time-frequency nonlinear characteristics of EEGs between typically developing (TD) and individuals with specific NDDs could aid in classifying these populations and thus assist in their clinical prediction. Nevertheless, investigation of such EEG features for the early prediction of most NDDs is generally limited. In addition, most of the works done in the literature in this field computed the nonlinear complexity and synchronisation measures from the time-frequency domain using traditional

methods such as wavelet transform, which will be described in Section 2.3.3. Such time-frequency analysis methods require a predefined basis (usually according to the traditional brain waves discussed in Section 2.2.1) to decompose the signals. However, the well-known variability between subjects in neural oscillations of interest could lead to missing potentially meaningful brain dynamics (Sweeney-Reed and Nasuto, 2009; Saby and Marshall, 2012). Empirical mode decomposition (EMD)-based methods, which will be discussed in Section 2.3.4, are considered a good candidate to address this problem as they decompose the signals adaptively, and thus, all potentially meaningful brain dynamics are included in the analysis. Therefore, employing such methods to develop an adaptive analytical tool that measures nonlinear characteristics of EEG could better reveal atypical brain functions associated with different NDDs and consequently may improve their prediction accuracy.

1.2 Aims and Objectives

This thesis is motivated by the fact that early identification can reduce the impacts of NDDs. It is also inspired by the limited applications of the adaptive time-frequency nonlinear EEG analysis to this field in general. Thus, the hypothesis that will be tested is to what extent utilising EEG complexity and synchronisation features computed from the EMD-based domain could help in identifying quantitative biomarkers for abnormal brain dynamics and whether these biomarkers could assist in the early prediction of NDDs.

It should be pointed out here that NDDs are a heterogeneous set of disorders, and consequently, different biomarkers might apply to different conditions. This thesis is mainly concerned with at-risk infants (with HIE) who developed CP and/or cognitive function deficits at a later age. Thus, the focus is to explore the efficacy of the proposed method to early predict the outcomes in this population. The ability of the proposed approach to reflect abnormal brain activity is further validated on available datasets of two other populations (Children with ASD and individuals with major depression disorder (MDD)).

Detailed thesis objectives are listed in the following:

1. To explore whether the EMD-based nonlinear complexity measures (particularly the sample entropy (SampEn)) of the task-related EEG signals can serve as biomarkers for ASD prediction. This investigation has been carried out on a benchmark dataset of children with ASD and their neurotypical peers (aged between 6 to 13 years old) recorded previously under age-appropriate emotional facial expression tasks (Apicella et al., 2012). Although employing such a dataset does not comply with the main aim of early identification of NDDs, as a proven environment, it

could verify the utilisation of the proposed methodology of computing the nonlinear measures from the EMD-based domain to explore some biomarkers that are associated with the brain abnormality.

2. To investigate whether the EMD-based nonlinear complexity measures (SampEn, permutation entropy (PEn), and spectral entropy (SpEn)) of resting-state EEG signals recorded from at-risk infants can be considered as early features to predict their CP outcome. This exploration has been conducted on a dataset of infants with HIE treated with hypothermia at the University Hospital of Southampton (UHS) and followed up under a long-term clinical follow-up program at the UHS. Following a neurological examination at 24 months of age by a pediatric neurologist, the neuromotor function was assessed, and the diagnosis of CP was made according to the criteria of the Surveillance of CP in Europe Working Group (Cans, 2000). The choice of this dataset was guided by its suitability to answer the overall question of exploring quantitive biomarkers for early NDDs prediction.

3. To conduct further analysis on resting-state EEG signals recorded from at-risk neonates (the same dataset used in objective 2) employing the EMD-based nonlinear synchronisation measures (particularly the node degree parameters derived from the phase-lag index (ND of PLI)) and evaluate whether such features could improve the performance of the CP prediction.

4. To investigate the potential of the EMD-based nonlinear complexity and synchronisation measures (SampEn, PEn, SpEn, and ND of PLI) of the signals recorded from at-risk infants to serve as biomarkers for predicting their cognitive function outcome. This exploration has been carried out on the similar dataset used in objectives 2 and 3, considering the neuropsychological outcomes of the infants evaluated at 24 months through the Bayley scales of infant and toddler development III (BSITD-III) (Weiss et al., 2010).

5. To validate the efficacy of the method proposed in objective 4 (mainly using EMD-based SampEn, PEn, SpEn, and ND of PLI measures) to identify potential biomarkers for depression severity prediction. This investigation has been conducted on a resting-state EEG dataset of individuals with MDD (aged between 18 to 25 years old), and their Beck Depression Inventory (BDI) scores (Beck et al., 1996; Cavanagh and Allen, 2017). The choice of this dataset was guided by its availability and suitability to verify the utilisation of the proposed method for cognitive scores prediction.

1.3 Challenges

The primary challenges towards fulfilling the research objectives can be summarised in the following:

- ***Multivariate signal analysis:*** EEG datasets are inherently high dimensional signals recorded from a group of electrodes placed on the scalp over different brain regions. For example, the dataset of the HIE infants who developed CP used in this thesis includes 26 subjects, each having EEG signals of approximately two minutes recorded from 19 channels. It is typically useful to consider potential biomarkers from these multidimensional signals for proper disorder prediction. However, analysing such signals using multivariate EMD-based methods is not an easy process and is of high computational complexity. This limitation is discussed further in Sections 4.7 and 5.8.

- ***Limited and imbalanced number of data:*** The main challenge of this study was the limited and imbalanced number of samples in the datasets, particularly the dataset of the HIE infants. The data was initially collected from 30 neonates born at term equivalent age (38-42 weeks of gestation) with HIE. Out of 30 infants under prospective monitoring, 26 have completed the neurological examination at 24 months. Consequently, only six infants developed CP while 20 had normal neuromotor development. Furthermore, only 20 children completed the cognitive function assessment at 24 months, and the distribution of their cognitive scores was skewed such that the majority of the children developed a normal cognitive function.

  As a prospective study, the long-term monitoring procedure needs to be carried out again to add more data. Fortunately, UHS has recently standardised a protocol for conducting such a prospective study on all newborns with HIE. Thus, a much larger trial can be undertaken in the future. Moreover, since this is a common problem in clinical studies, the machine learning community has developed different methods and algorithms to handle such limited and imbalanced data problems.

  In addition, we have conducted further investigations to validate the ability of the proposed framework to reveal atypical brain activity on larger and balanced datasets of individuals with MDD and ASD, respectively.

1.4 Contributions

This thesis is concerned with individuals with NDDs and is motivated by the fact that early identification of these disorders can improve the outcome and reduce the socio-economic impacts. Several studies have been conducted to investigate the efficacy of EEG analysis to provide quantitative biomarkers to characterise abnormal brain activities. However, the investigation of nonlinear EEG characteristics for early prediction of NDDs is generally limited. In addition, most of the works done in this field accounted for the nonstationarity property of EEG by computing the nonlinear measures using

traditional time-frequency analysis methods (such as wavelet transform), which could lead to missing potentially meaningful brain dynamics.

In view of this, the primary contributions of this research work can be outlined as follows:

1. Estimates the complexity (SampEn, PEn, and SpEn) and synchronisation (ND of PLI) features from the EMD-based domain to consider the nonlinear and non-stationary characteristics of the signals. EMD-based methods further account for the variations between subjects in neural oscillations of interest where they decompose the EEG adaptively and do not rely on filter cut-offs considering all dynamics involved in the signals.

2. Identifies the frequency ranges of each component (decomposed using EMD-based methods) and finds their corresponding traditional brain waves (i.e., delta, theta, alpha, beta and gamma). Such a method helps in understanding brain abnormalities associated with various brain disorders, comparing the findings with state-of-the-art results, and identifying biomarkers in a way that could be useful to clinicians.

3. Shows that the nonlinear complexity features (SampEn) of the alpha and theta brain waves and the linear features (maximum, minimum, and standard deviation) of the delta-band components can be considered biomarkers of impaired facial expression processing in autistic children, aiding in their prediction.

4. Shows that the nonlinear complexity measures (SampEn, PEn, and SpEn) of resting-state EEGs at the alpha-band can serve as potential biomarkers of early abnormal brain function and can be used to discriminate at-risk infants (with HIE) who developed CP by two years from those with the normal neuromotor outcome.

5. Reports relationships between the nonlinear complexity (SampEn, PEn, and SpEn) and synchronisation (ND of PLI) features of EEG signals of at-risk infants who developed CP and those with the normal neuromotor outcome at beta-, alpha-, and delta-band, suggesting these features as biomarkers for the early prediction of CP.

6. Reports significant correlations between PEn, SpEn and ND of PLI features (calculated at the left hemisphere from alpha- and delta-band) of at-birth EEG (recorded from infants with neonatal HIE) and cognitive scores, suggesting these features as potential biomarkers for early cognitive function prediction.

7. Confirms the findings of (Mohammadi and Moradi, 2021) which found that complexity and synchronisation features could be used as biomarkers for depression severity prediction. Particularly, the current study shows significant correlations between SpEn and ND of PLI (calculated from beta- and delta-band) and BDI scores, proving the efficacy of these features for depression severity prediction.

1.5 Outline of the thesis

The rest of the thesis is organised as follows: Chapter 2 describes various existing techniques for EEG signal analysis together with a deeper discussion of the related state-of-the-art studies toward NDDs prediction. Chapter 3 provides an overall process of the proposed framework. Chapter 4 gives detailed methodology and results of investigating the ability of the EEG complexity, recorded under a facial emotional expression task, to improve the performance of ASD prediction.

Chapter 5 explains the experimental results, which were carried out to explore the efficiency of EEG complexity to classify the signals of infants who developed CP from those with the normal neuromotor outcome. Chapter 6 provides findings of employing the synchronisation features for CP prediction as well as joint analysis of the signals' complexity and synchronisation.

Chapter 7 describes the correlation analysis carried out to investigate relationships between the nonlinear features extracted from infants' EEG and their corresponding cognitive scores evaluated at 24 months, while Chapter 8 provides a similar investigation conducted on a much larger dataset of MDD individuals to validate Chapter 7's methodology. The conclusion, along with the highlights of future works that can follow this research, are detailed in Chapter 9.

1.6 List of publications

1. Bakheet, D. and Maharatna, K. (2021), Linear and nonlinear analysis of intrinsic mode function after facial stimuli presentation in children with autism spectrum disorder, Computers in Biology and Medicine 133(2), p104376.

2. Bakheet, D., Alotaibi, N., Konn, D., Vollmer, B. and Maharatna, K. (2021), Prediction of cerebral palsy in newborns with hypoxic-ischemic encephalopathy using multivariate eeg analysis and machine learning, IEEE Access pp. 11.

3. Alotaibi, N., Bakheet, D., Konn, D., Vollmer, B. and Maharatna, K. (2022), Cognitive outcome prediction in infants with neonatal hypoxic-ischemic encephalopathy based on functional connectivity and complexity of the electroencephalography signal, Frontiers in Human Neuroscience 15.

Chapter 2

Background and Literature Review

This chapter provides the theoretical basis of brain signal processing and its application in NDDs prediction, highlighting recent developments and knowledge gaps in the field. Section 2.1 gives an overview of the brain signals and the clinical tools used to record them, focusing on the EEGs —the brain signals used in this research. Different neurophysiological properties of EEG used in the NDDs prediction are then introduced in Section 2.2. Advanced signal processing methods used to extract informative biomarkers for NDDs are reviewed in Section 2.3.

Nonlinear time series analysis of EEG signals and its applications to NDDs prediction are discussed in Section 2.4 and Section 2.5. Section 2.6 discusses the characteristics of ASD, CP and cognitive function impairments and reviews state-of-the-art EEG analysis methods carried out for their prediction.

2.1 Neurological Signals

The neural activity of the human brain begins at a very early stage, between the 17th and 23rd weeks of prenatal development, and represents the status of brain function. This fact motivates the application of advanced brain signal processing methods to reflect different neurological conditions (Sanei and Chambers, 2007). Thus, understanding the neuronal activities, as well as the mechanisms underlying the generation of signals along with their recording techniques, are essential for NDDs prediction. The following sections aid in understanding the biology behind the human brain and the clinical tools that help study its neurological function.

2.1.1 Signal Acquisition Techniques

The brain functions and the physiology behind it are usually investigated by applying several signal acquisition techniques. These techniques range from invasive measurements of brain activity, such as electrocorticogram and microelectrode arrays, to noninvasive recording such as EEG (Cohen, 2014). Some of these devices measure the electrical component of the electromagnetic field of the brain, such as EEG and electrocorticogram. Other techniques consider the magnetic component of the brain's electromagnetic field, such as MEG (Cohen, 2014). Other acquisition devices estimate the hemodynamic response to neuronal activity, such as fMRI and Near-infrared spectroscopy. These devices do not directly measure the neuronal activation but instead rely on the level of oxygenated and deoxygenated hemoglobin, which in turn reflects the neural activity in different brain regions (Sanei and Chambers, 2007).

Choosing the right neurological acquisition tool depends on the brain's impairment that needs to be identified (Siuly and Zhang, 2016). For example, EEG is often used to identify brain-related disorders such as autism, stroke, epilepsy, and alcoholism-related disorders. On the other hand, to detect neurological conditions such as tumours, blood clots, and the location of strokes, a computerised tomography scan is applied to monitor the cross-sectional images of the brain. Magnetic resonance imaging (MRI) tests help physicians identify and monitor blood flow problems and degenerative disorders such as multiple sclerosis and brain injury from trauma. fMRI measures small changes in blood flow as a person does tasks. Thus, it is suitable to investigate the brain in action (i.e., while moving or speaking) and to detect areas of the brain that become active and pinpoint how long they stay active (Siuly and Zhang, 2016).

fMRI has several advantages supporting its application to investigate brain abnormalities associated with different NDDs. These advantages include its noninvasive acquisition nature with an excellent spatial resolution, where structures of the size of a few millimeters can be localised with the fMRI. This improved resolution permits examining correlated brain activity across different brain regions and provides a more accurate interpretation of the brain dynamics associated with distinct NDDs. However, fMRI suffers from some disadvantages, which make it unsuitable for the identification of NDDs. The most critical limitation is its relatively low temporal resolution compared to the methods that directly measure the electrical brain activity, such as EEG. This low resolution affects detecting brain dynamics in the same time frame at which cognition occurs. Moreover, fMRI devices are technically expensive and demanding, making them difficult to be used with infants (Sanei and Chambers, 2007).

Nevertheless, fMRI has been widely used in the literature to reveal brain malfunctions associated with NDDs (even with infants). While the review provided in this thesis is biased towards EEG-based NDDs-detection technique, fMRI review can be found in different articles such as (Ernst et al., 2015). The rest of the thesis focuses on the EEG

technology and its uses for NDDs identification, as it is the recording technique used in this research. EEG was the tool of choice in this thesis due to its advantages discussed in the next section.

2.1.2 Electroencephalography (EEG)

Before describing the recording mechanism and the analysis aspects of EEG, it is necessary to briefly review the neurons' structure and function to understand how the signals measure information dynamics of the brain. Neurons consist of a cell body, an axon, and dendrites (Figure 2.1). A cell body has a single nucleus and mainly produces electrical discharges. Neurons produce electrical impulses when triggered by a stimulus, which can be any activity inside or outside the body that evokes a physical or psychological response. The axon serves as an output channel to transmit electrical impulses, and it is connected via junctions (called synapses) to the dendrites of other neurons (Sanei and Chambers, 2007).

Communication between neurons is based on action potentials that travel along the axon of cells and release neurotransmitters (which are specialised chemicals) when arriving at a synapse. In turn, the neurotransmitters trigger a flow of ions across the cell membranes in the direction governed by the membrane potential. The flow of positive ions of sodium, Na+, potassium, K+, calcium, Ca++, and chlorine, Cl−, across the cell membrane leads to a change in membrane potential, i.e., changing the potential difference between the soma (body of a neuron) and apical dendrites, which branch from neurons. If the membrane potential reaches a threshold of approximately $60 - 70$ μV with negative polarity, an action potential is triggered, and information is transmitted through the axon to other nerve cells (Sanei and Chambers, 2007).

An EEG signal is a measurement of currents that flow during synaptic excitations of the dendrites of pyramidal neurons located in the cerebral cortex. However, the signal is attenuated due to the different layers inside a human head, including the brain, skull, and scalp. As a result, only a large population of active neurons can generate enough potential that is measurable by the electrodes on the scalp, which gives a coarse view of neural activity. EEG signals are assumed to be generated by distinct neuronal populations formed by the interconnection of the individual neurons (Sanei and Chambers, 2007).

EEG signals can be recorded invasively or noninvasively, measuring the electrical components of the electromagnetic field of the brain that are generated by neuronal activity. Thus, electro- (referring to the registration of brain electrical activities), encephalo- (referring to emitting the signals from the head), and gram (or graphy), which means drawing or writing, were combined to form the term EEG (Sanei and Chambers, 2007).

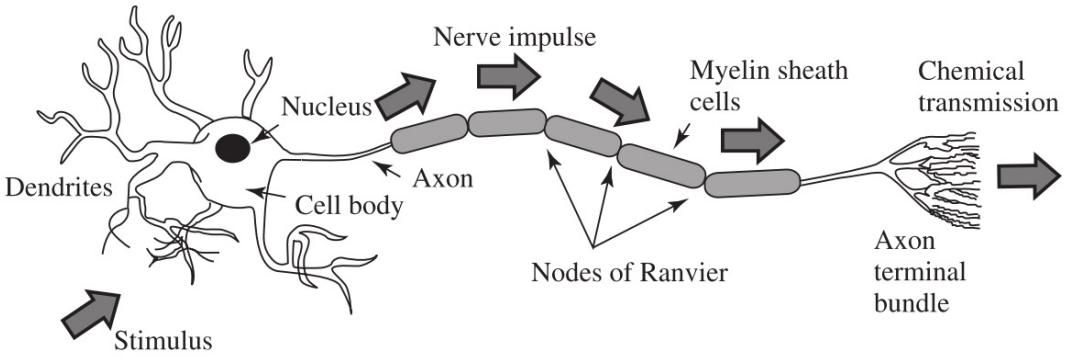


FIGURE 2.1: Structure of a neuron (reproduced from (Sanei and Chambers, 2007)).

The discoverer of human EEG signals was Hans Berger, who began his studies in 1920 (Sanei and Chambers, 2007). Since its discovery, the use of EEG has resulted in the continuous development of clinical studies, recognition, and diagnosis of many physiological and neurological abnormalities of the brain.

Noninvasive EEG is often recorded at many brain locations simultaneously using one electrode at each position. These electrodes are stuck to the scalp with a conductive gel improving contact impedance between the skin and the electrodes. Amplifiers (one for each channel) are often used to transform the signals from analog to digital (to be fed into a computer) (Sanei and Chambers, 2007). For the application of a larger number of electrodes, an electrode cap is often used. The distance between adjacent electrodes is usually in the range of one to a few centimetres, and available EEG caps can record up to 256 channels. However, 128/256 channels produce high-density EEG recordings that require a very high computational complexity analysis. Typically, and primarily in the clinic, a much lower number of channels is used.

Figure 2.2 represents conventional electrode positioning (also called 10/20) for the 19 and 75 electrodes recommended by the American EEG society. The International 10/20 system is usually used for naming and defining electrode positions. In this system, the electrode location is linked to the underlying area of the cerebral cortex. The system considers some constant distances (i.e., the total front-back or right-left distance of the skull). Then, it uses 10% or 20% of the specified interval as the distances between adjacent electrodes. Each electrode is named using a letter to mark the lobe (C: Central; F: Frontal; T: Temporal; P: Parietal; O: Occipital) and a number to represent the hemisphere location. Odd numbers refer to electrode positions on the left hemisphere, while even numbers on the right region. The letter z (zero) is used to refer to electrodes located on the midline (Sanei and Chambers, 2007).

EEG recordings have numerous advantages that can help in revealing abnormal brain function of individuals with NDDs. First of all, EEG offers millisecond order temporal resolution of neural dynamics (Bowman and Varcin, 2018). This high resolution allows capturing cognitive dynamics in the same time frame in which cognition occurs. For

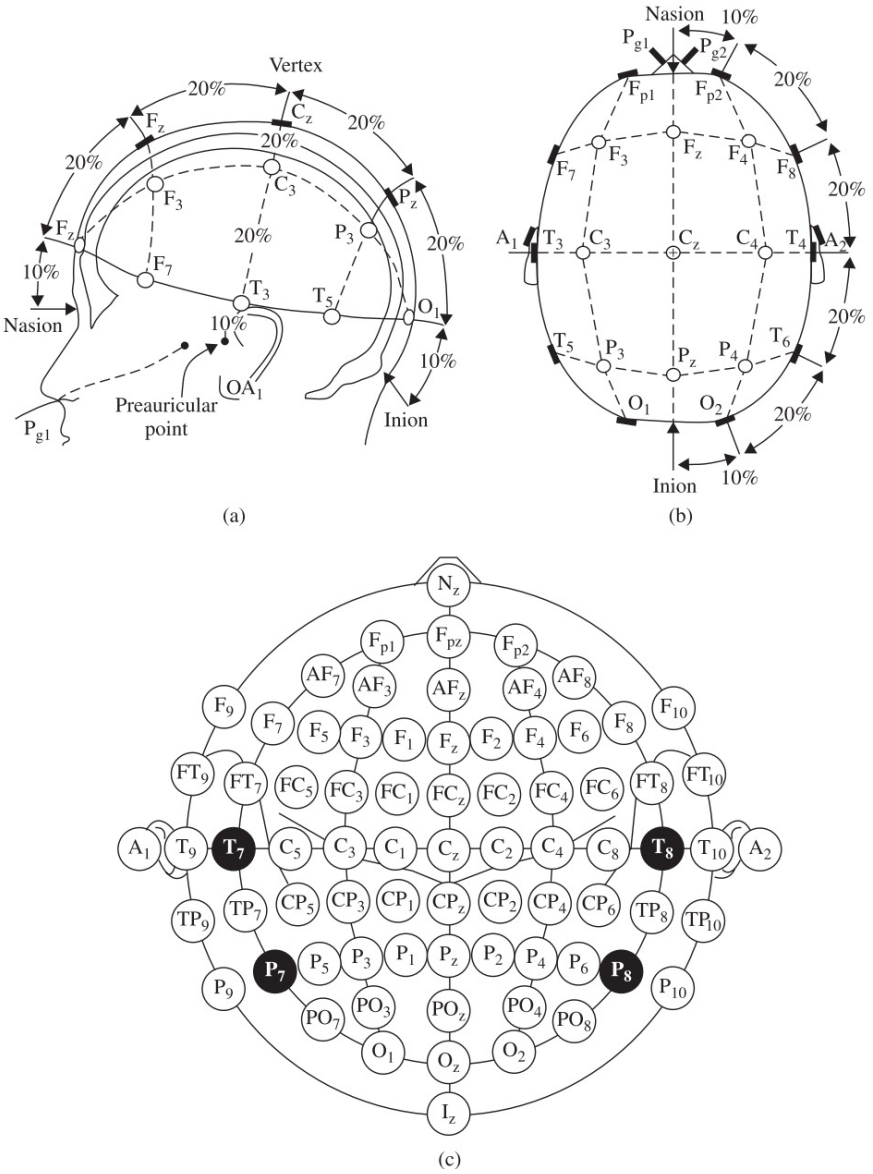


FIGURE 2.2: A diagrammatic representation of 10/20 electrode settings: (a) and (b) represent the 19 electrode placement and the three-dimensional measures, and (c) indicates a two-dimensional view of the 75 electrodes setup configuration including the reference electrodes (highlighted in black) based on the guidelines by the American EEG society (reproduced from (Sanei and Chambers, 2007)).

example, it associates time-locking neural activity to the onset of a specific stimulus (e.g., sound or picture) or a subject's response (e.g., finger movement, button press, vocalization, or attention shift). EEG also gives a readout of rapid background oscillatory patterns, such as gamma-band oscillations implicated in the detection of different brain disorders (Sanei and Chambers, 2007). Thus, the high temporal accuracy of EEG allows us to investigate various levels of neural activities in NDDs. To illustrate, EEG can reflect low-level activities (such as visual and auditory evoked potential) as well as higher-level cognitive processing, such as face perception, language, learning, and social cognition (Bowman and Varcin, 2018).

Another advantage of EEG over, for example, fMRI is that it reflects the neural activity directly. This means that the voltage fluctuations measured by EEG are direct reflections of biophysical phenomena at the neurons population-level (Cohen, 2014).

EEG can be collected from birth, where the neural activity is recorded by a set of electrodes attached to the infants' scalp (who are often held by a caregiver). This child-friendly nature of EEG is another essential advantage that can help in early detection, and therefore intervention and treatment (Bowman and Varcin, 2018). Last but not least, EEG recording sessions often have low operating costs, which are a significant strength to infants/children and sensitive subjects who may have difficulty dealing with the assessment, therefore requiring multiple recordings and sessions (Bowman and Varcin, 2018). These advantages of EEG make it suitable for NDDs detection where it may provide important clinical biomarkers for early risk assessment, diagnosing the disorder and monitoring its progression.

There are also different EEG shortcomings that make it not well suited to address all research questions. The most critical weakness of EEG is its low spatial resolution. Thus, the signal recorded from one electrode does not reveal only the neurons' activity directly under that electrode, but it reflects a mixture of neuronal activity from many brain regions close to that electrode. Therefore, EEG is not helpful in studies that need precise source localisation (though some methods help in measuring neurons localisation) (Cohen, 2014). Another set of studies that EEG may not be suited for includes measuring slow cognitive processes or the processes with variable time courses. In this case, the high temporal resolution of EEG becomes a limitation, and the relatively low temporal resolution of fMRI becomes more suited to study such slow processes (Cohen, 2014). Finally, although the EEG recording method is relatively cheap, using new EEG equipment with high spatial accuracy and improved signal quality may lead to the exact costs compared to MEG and MRI (around 200-600 USD per hour/subject) (Cohen, 2014).

2.2 Neurophysiological Activities

This section aims to identify the types of neurophysiological activities obtained from the EEG signals and help in detecting atypical brain functions associated with different NDDs. Two fundamentally different types exist for this purpose. The first type refers to the brain activity that does not change based on engagement in an event, such as motor imaginary and resting-state activities. The most common way to characterise this type of brain activity is by breaking down the oscillatory patterns into frequency bands that share physiological properties (Sanei and Chambers, 2007). Oscillatory brain activity will be described in Section 2.2.1.

The secondt type is typically generated in response to peripheral or external stimulations and observed after or during exposure to the stimuli. The changes in neurophysiologic signals resulting from perception and processing of stimuli are termed event-related potentials (ERPs). This type of brain activity will be discussed in Section 2.2.2, while the last section summarises the appropriate applications of each neurophysiological brain activity.

2.2.1 Oscillatory Brain Activity

Physiologically meaningful brain activity can be identified from changes in oscillatory brain activity. In the brain, the oscillation term refers to rhythmic fluctuations in the excitability of neurons (Cohen, 2014). These oscillations occur in different brain regions and change according to the state of subjects, for instance, between wake and sleep or between focusing on a specific mental task and idling. Brain oscillations may also be changed due to the presentation of visual or audio stimuli (Cohen, 2014). Different frequency bands are related to changes in the amplitude of oscillatory activity. The major five frequency bands include delta $0.5-4$ Hz, theta $4-8$ Hz, alpha and mu $8-13$ Hz, beta $13-30$ Hz and gamma >30 Hz (Sanei and Chambers, 2007). Nevertheless, these frequency ranges of the brain waves are inconsistent among studies, and this point will be further discussed in Section 2.3.4.2.

The delta-band is primarily found in the signals recorded during deep sleep, and also, it may be found in the waking state signals (Sanei and Chambers, 2007). Theta-band, on the other hand, appears in the signals recorded during consciousness slips towards drowsiness, and it plays a crucial role in infancy and childhood. It is abnormal to have larger contingents of theta-band activity in the waking adult and is usually caused by different pathological problems. The changes in the rhythm of the theta-band are examined for emotional and maturational research (Sanei and Chambers, 2007).

The alpha-band appears in the posterior half of the head and is often detected over the occipital brain region. The alpha-band usually appears as a sinusoidal-shaped signal, and in rare cases, it may manifest itself as sharp waves. Alpha waves are associated with a relaxed awareness without any attention or concentration. Subjects often produce some alpha waves while they close their eyes. These waves are usually reduced by opening the eyes, by anxiety, by hearing strange sounds, or in the case of mental concentration or attention. Mu-rhythm is a particular case of alpha with a frequency of about 10 Hz, and it can be observed over the sensorimotor cortex when a subject does not perform any movements (Sanei and Chambers, 2007).

The beta-band is often found in TD adults. This wave is associated with active attention, progressive thinking, and solving concrete problems. When a subject is in a panic state, a high-level beta-band may be acquired. Beta waves are usually encountered

over the frontal and central brain regions. Although the occurrence of gamma-band occurrence is rare, recognition of these rhythms can be used to confirm the presence of certain brain diseases, and they are usually located in the frontocentral brain area (Sanei and Chambers, 2007). Figure 2.3 depicts the normal brain rhythms with their usual amplitude levels.

Despite the difficulty of detecting the brain rhythms from the scalp EEGs, advanced signal processing methods can identify the desired waveforms within the EEGs and cater for their inherently linear/nonlinear statistics (Looney et al., 2015). Such methods will be described in Section 2.3.

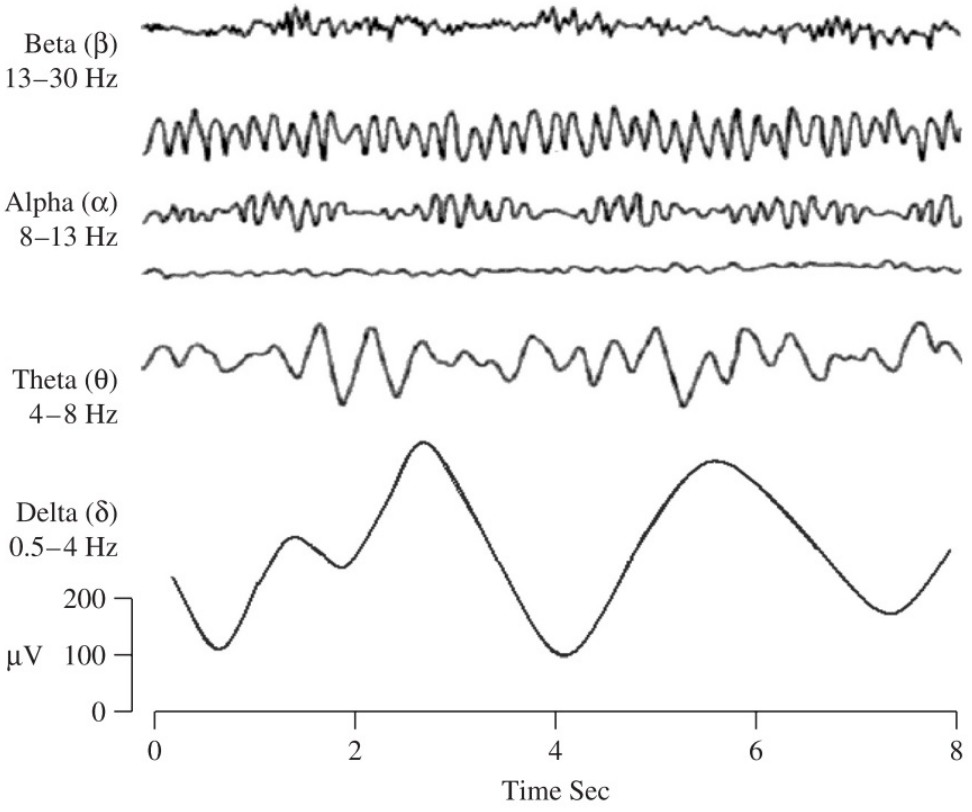


FIGURE 2.3: Four dominant normal brain rhythms (reproduced from (Sanei and Chambers, 2007)).

2.2.2 Event Related Potential

ERPs are specific patterns generated by the subject's brain after or during the presentation of preselected visual or audio stimuli. These patterns can typically be extracted from the recorded EEG signals by averaging many trials over temporal windows. ERPs are beneficial for answering questions about the timing of mental processes (Cohen, 2014).

Before using ERPs, it is essential to define their major components, as these components are the tools that can be used to determine abnormal brain activities associated with different brain disorders. An ERP component can be defined as a voltage deflection produced due to a specific neural process in a particular brain region. These components are usually characterised by their positive or negative polarity (amplitude) and its latency (Cohen, 2014). The most well-known ERP components are often used to make inferences about cognitive processes with clinical utility include P300, mismatch negativity (MMN), steady-state visual evoked potential (SSVEP), N170, and N400. Although each cognitive function can be reflected by a specific ERP component, a component that indicates one type of process might be beneficial for studying other processes. Studies on these components have led to their application in the investigation of clinical disorders, where they have demonstrated a significant promise (Duncan et al., 2009).

The simplest and most commonly used method to extract ERPs is the averaging method, i.e., by summing the voltage at each time point over trials and then dividing by the number of trials. However, to use ERPs as an investigation tool about the cognitive function, more sophisticated interpretations that reflect processes behind ERPs are needed, and the debate related to the neurophysiological mechanisms that produce ERPs (ERP generations) (Cohen, 2014) should be considered. There are different models in the literature about the emergence of ERPs from ongoing or oscillatory brain activity (Cohen, 2014). The two most well-known models are described as follows:

The first model is the additive model, which proposes that the ERP signal, elicited by external audio or visual stimuli, is added to the ongoing background oscillations. Thus, it causes an instantaneous change in the phase of background activity, which means that a transient peak in frequency (positive or negative) might be seen. However, based on this model, these frequencies are not related to the stimulus. Therefore, they might be attenuated entirely using the averaging method (Burgess, 2012; Cohen, 2014).

The second model is the phase alignment model, which suggests that the ERPs result from a sudden reset in the phases of background oscillations. Hence, a stimulus's presentation causes the background oscillations to shift their phases at a specific frequency band. Therefore, this model suggests that ERPs consist of changes in frequencies that should be addressed when using ERPs to investigate cognitive process characteristics. This change in frequencies can often be extracted using time-frequency analysis described in Section 2.3.3 (Burgess, 2012; Cohen, 2014).

2.2.3 General Observations

Different neurophysiological signals can be used to characterise abnormal EEG activities. The advantage of ERPs is their ability to detect atypical cognitive functions of

the NDDs population in response to a disorder-related task. For instance, a face perception task could reflect impaired social interaction and a restricted range of interests associated with ASD individuals. A disadvantage is that diagnosis depends on the presentation and perception of stimuli, and the subjects are thus required to pay attention to them. This scenario is impracticable when dealing with more severely impaired or younger individuals (such as infants) who may not perform tasks because of physical, cognitive, or developmental challenges.

On the other hand, oscillatory brain activity does not require subjects to respond to any stimuli. Thus, these activities are more appropriate for studying atypical maturational trajectories through early childhood. However, resting-state methods are not useful in the applications that aim to evaluate abnormal brain behavior associated with the clinical finding of impaired skills in different NDDs.

Choosing the proper neurophysiological signals depends on subjects' characteristics and the brain function under study. This thesis used the task-related signals (Chapter 4) and the resting-state EEG (Chapter 5-8) to study brain functions of individuals with ASD, at-birth HIE, and MDD, respectively.

2.3 Signal Processing

This section intends to review and discuss the most well-known EEG analysis methods used to reflect meaningful biomarkers inherent in the signals, which in turn help to identify brain abnormalities. The EEG signals are typically presented in the time domain. However, this domain does not reflect oscillatory brain activities, which can reflect different brain disorders, as discussed in Section 2.2.1. Thus, signal processing methods aim to transform EEG signals from the time domain to another representation that can reflect brain oscillation and, in turn, fulfills the objective in revealing neurophysiologic features that are relevant for NDDs prediction (Pachori and Patidar, 2014).

Nonstationarity and nonlinearity are two critical EEG characteristics that should be accounted for when choosing the appropriate signal processing method (Sanei and Chambers, 2007).

Nonstationarity can be observed by quantifying some properties of the signals at different time lags. The signals can be deemed stationary if their statistics, including mean value, variance, and frequency content, do not change over time (Cohen, 2014). Nevertheless, EEG properties generally change from segment to segment, and hence EEG is considered stationary only within short intervals, i.e. quasistationarity (Sanei and Chambers, 2007). The stationarity assumption of the signals could be held during a

normal brain condition, but it is not valid during physical and mental activities. Non-stationarity of the EEG signals can be observed during the change in wakefulness and alertness, eye blinking, and ERP signals (Sanei and Chambers, 2007).

This section reviews the signal analysis methods related to the nonstationary property, while the nonlinear analysis methods are discussed in the next section.

In general, EEG signals could be quantified in three different domains: (1) time domain, (2) frequency domain, and (3) time-frequency domain. A brief review of these domains is given in the following.

2.3.1 Time Domain Analysis

The EEG signals are often analysed in the time domain as the amplitude of neurophysiologic signals changes over time. Such change usually occurs time-locked to the presentation of an event. ERPs are a good example of the signals that can be characterised with the time-domain features.

Analysing an EEG signal in the time domain to reveal neurophysiologic changes is straightforward. Time series features, such as average, absolute minimum and maximum values, slopes, steepness, height, and width of predefined patterns can easily be computed. These time-domain features cannot be usually observed in a single trial and can rather be extracted by averaging many trials over temporal windows to enhance Signal to Noise Ratio (SNR) (Cohen, 2014).

Various studies have effectively extracted ERP components from the time domain to identify different NDDs (Sanei and Chambers, 2007). However, in most cases and even with ERPs, time-domain features are not enough, and it is beneficial to address frequency changes behind the ERPs as it provides a more in-depth analysis of the signal. Hence, transforming EEG signals to the frequency domain might be useful to reveal neurophysiological features of the data.

2.3.2 Frequency Domain Analysis

As discussed in Section 2.2.1, the EEG signal consists of a set of oscillations known as rhythms. Abnormal oscillatory activities within specific frequency bands are a promising candidate to reflect various brain disorders (Sanei and Chambers, 2007). Thus, a starting point for understanding functional brain abnormalities is investigating the frequency components inherent to the signals. In particular, power spectral density (PSD) has been frequently used as an informative frequency feature.

The Fourier methods are the most common frequency domain analysis approaches used to measure PSD. Fourier methods include fast Fourier transform (FFT) approaches

that have been used frequently in NDDs diagnosis research, such as (Martin-Brufau and Nombela Gomez, 2017). These methods involve decomposing a signal into infinite series of sinusoids to obtain the PSD. Such a decomposition assumes that the signal being processed is recorded from a linear and stationary process (Cohen, 2014), which is not the case for EEG, as explained earlier in this section. Applying Fourier spectral analysis to such signals induces additional harmonic components in the time domain to simulate the non-uniform signal and the deformed wave-profiles, which are the direct consequences of nonstationary and nonlinear effects, respectively (Huang et al., 1998). The existence of such spurious harmonic components spreads the energy over a broad spectrum in the frequency space causing misleading PSD distribution with no physical sense (Huang et al., 1998). Performing temporally localised frequency decomposition can address this issue, assuming that the signals are stationary within relatively short periods of time (Cohen, 2014).

In addition, time-varying changes in the frequency structure cannot be observed directly from the Fourier representation of the signals, which is another reason that requires performing temporally localised frequency decomposition method (Cohen, 2014). This type of decomposition is discussed next.

2.3.3 Time-Frequency Domain Analysis

EEG signals have specific properties in both the temporal and frequency domains, where the frequency spectrum is observed to vary over time, indicating the signals' nonstationarity. Therefore, it is highly suggested to use a hybrid of temporal and frequency domain methods.

Short-time Fourier transform (STFT) (Gabor, 1946) and wavelet transform (WT) (Mallat, 1989) are often used to extract the tempo-spectral features from nonstationary signals. Such transformation methods can capture sudden spectral variations of the signals. STFT has been used in various studies to identify different brain disorders such as epileptic seizures (Wang et al., 2019). STFT works as a contiguous and overlapping sliding window of arbitrary fixed length applied to the time series. Then, the Fourier transform is calculated at each time window to provide some localisation in time (Sweeney-Reed and Nasuto, 2007). However, the choice of window size introduces a situation known as the uncertainty principle, which is produced from the trade-off between bandwidth and time. Thus, choosing a small window provides good temporal resolution at the expense of frequency information, and conversely (Sweeney-Reed and Nasuto, 2007).

On the other hand, WT is a useful method that also provides time-frequency localisation of the signals. WT has been used frequently in the NDDs' diagnosis applications such as epilepsy and ASD (Ibrahim et al., 2018). WT reduces the window size problem

as it adapts the size of the window to the data (it uses short windows at high frequencies and longer windows at low frequencies) (Liang et al., 2005). However, the uncertainty problem still holds as it is not possible to guarantee that the chosen window size will co-occur with the time scales of interest (Sweeney-Reed and Nasuto, 2007).

Another time-frequency method for analysing nonstationary signals is the empirical mode decomposition (EMD). Unlike STFT and WT, EMD is self-adaptive according to the signal to be processed. This method can provide both good frequency and time resolution as it decomposes the signals into finite frequency-modulated components in the time domain (Liang et al., 2005).

EMD is the time-frequency approach adopted in this research. Thus, details about this method, the motivation for this choice, the EMD-based applications in the field of predicting different NDDs, challenges and knowledge gaps are discussed in the next section.

2.3.4 Empirical Mode Decomposition

EMD is a data-driven method proposed by Huang et al. (1998) to analyse a nonstationary and nonlinear signal in the time-frequency domain. There is an assumption that all signals contain a series of different intrinsic oscillation modes. Thus, EMD is performed to decompose a signal into its intrinsic mode functions (IMFs) sum of which is equal to the original signal (Huang et al., 1998). Each of these IMFs contains a frequency range, and these frequencies decrease from the first IMF to the last. As IMFs are locally narrow-band signals, they allow the calculation of instantaneous phases using the Hilbert transform (HT) (Bendat and Piersol, 1986). From the calculated phases, the instantaneous frequencies can be derived (Sweeney-Reed and Nasuto, 2007). Each IMF must satisfy two conditions:

1. The number of extrema and the number of zero crossings are either equal or differ at most by one.
2. The mean value of the envelope defined by the local maxima and local minima is zero at any given point.

The first condition implies a need of a narrow band requirement for a signal to be a stationary Gaussian process. The second condition is needed for abstaining instantaneous frequency from unwanted fluctuations produced by asymmetric waveforms (Huang et al., 1998). Figure 2.4 illustrates a visual example of IMFs, where an EEG signal has been decomposed into ten distinct IMFs.

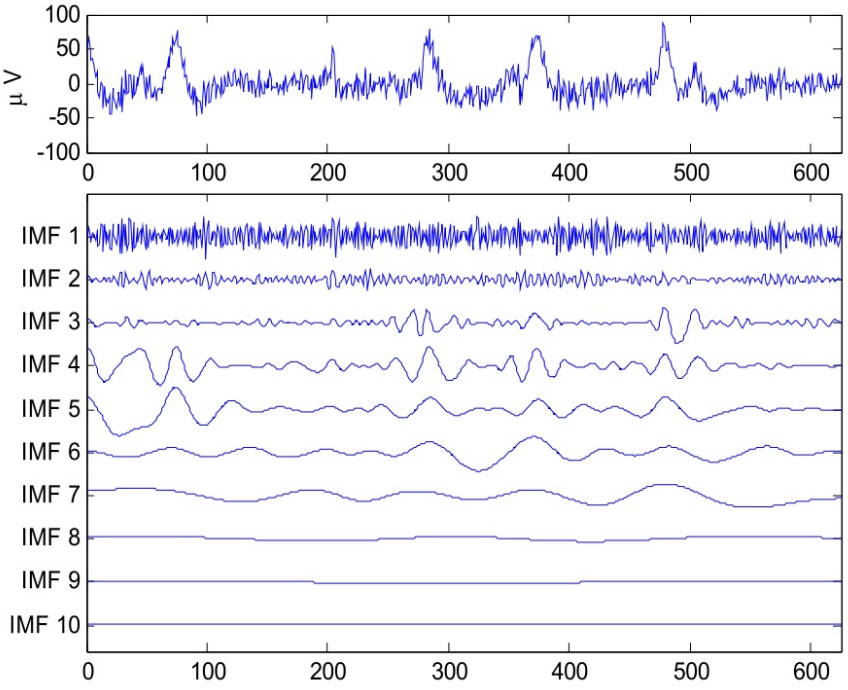


FIGURE 2.4: An example of an EEG signal and its corresponding IMFs (reproduced from (Huang et al., 2013)).

2.3.4.1 EMD Algorithm

EMD is an iterative algorithm to decompose a complex signal into finite components (IMFs). It removes the highest frequency oscillation from the time series in each repetition and produces a residue that contains lower frequency information. This process (known as the Sifting process) is repeated until only a trend remains, leading to adaptive decomposition based on the local time scale of the signal. Thus, intrinsic oscillations contained in the time series may be identified (Huang et al., 1998). The sifting process terminates when any of the following conditions is reached:

1. The residual signal energy becomes less than a predefined threshold.
2. The residual signal is a monotonic function that cannot be decomposed into more IMFs.

The resulting IMFs should sum to a composite nearly identical to the original signal given by:

$$x(t) = \sum_{i=1}^{n-1} IMF_i(t) + r_n(t) \tag{2.1}$$

where x is the original signal, i indexes the IMFs, n is the total number of IMFs, r_n is the residue at the end of the sifting process, and t is time (Huang et al., 1998).

Figure 2.5 illustrates the EMD algorithm where x is the original signal, d is the signal to be decomposed at each iteration, m represents components generated by the sifting process, which may be IMFs, and r is the residue or trend that remains at the end of the algorithm (Sweeney-Reed and Nasuto, 2007).

Figure 2.6 gives a schematic representation of the same algorithm. Panel a) shows the steps performed to produce the 1st IMF. Step 1 includes identifying the local maxima (red dots) and minima (green dots) of the signal. Step 2 includes linking the maxima (red line), interpolating the intervening points by a spline-fitted curve. The same process is done for the minima (green line). In Step 3, the mean of the maxima and minima splines is found (black line) and subtracted from the original signal. The residual is the 1st intrinsic mode function (black line). Panels b) and c) show the mean spline being subjected to Steps from 1 to 4 to obtain the 2nd and 3rd IMFs, respectively. The method is repeated until the mean spline is monotonic and no more IMFs can be extracted (Burgess, 2012).

Once all IMFs are identified, the instantaneous frequency of each IMF at each time point can be acquired by the HT (Huang et al., 1998). HT of a time series $x(t)$ is performed by shifting the phase angle of all components of the signal by ± 90 degree. The HT for an arbitrary signal $x(t)$ is given by Huang et al. (1998):

$$H[x(t)] = \frac{1}{\pi} P \int_{-\infty}^{\infty} \frac{x(u)}{t-u} du \tag{2.2}$$

in which P indicates the Cauchy principal value of the integral and $u = 1/t$. The HT and the original signal $x(t)$ are orthogonal. After utilising HT, the instantaneous phase can be derived from the analytic signal corresponding to the original signal $x(t)$ and defined as:

$$z(t) = x(t) + iH[x(t)] = a(t)exp[i\Theta(t)], \tag{2.3}$$

where the imaginary part H[x(t)] is the HT of x(t), a(t) and $\Theta(t)$ are the instantaneous amplitude and phase of the analytic signal z(t), respectively. The a(t) and $\Theta(t)$ are given by:

$$a(t) = [x^2(t) + H^2[x(t)]]^{1/2} \;\; ; \;\; \Theta(t) = arctan(\frac{H[x(t)]}{x(t)}) \tag{2.4}$$

The instantaneous frequency can then be obtained from the instantaneous phase where it is the time derivative of the phase given by:

$$w(t) = \frac{d\Theta(t)}{dt} \tag{2.5}$$

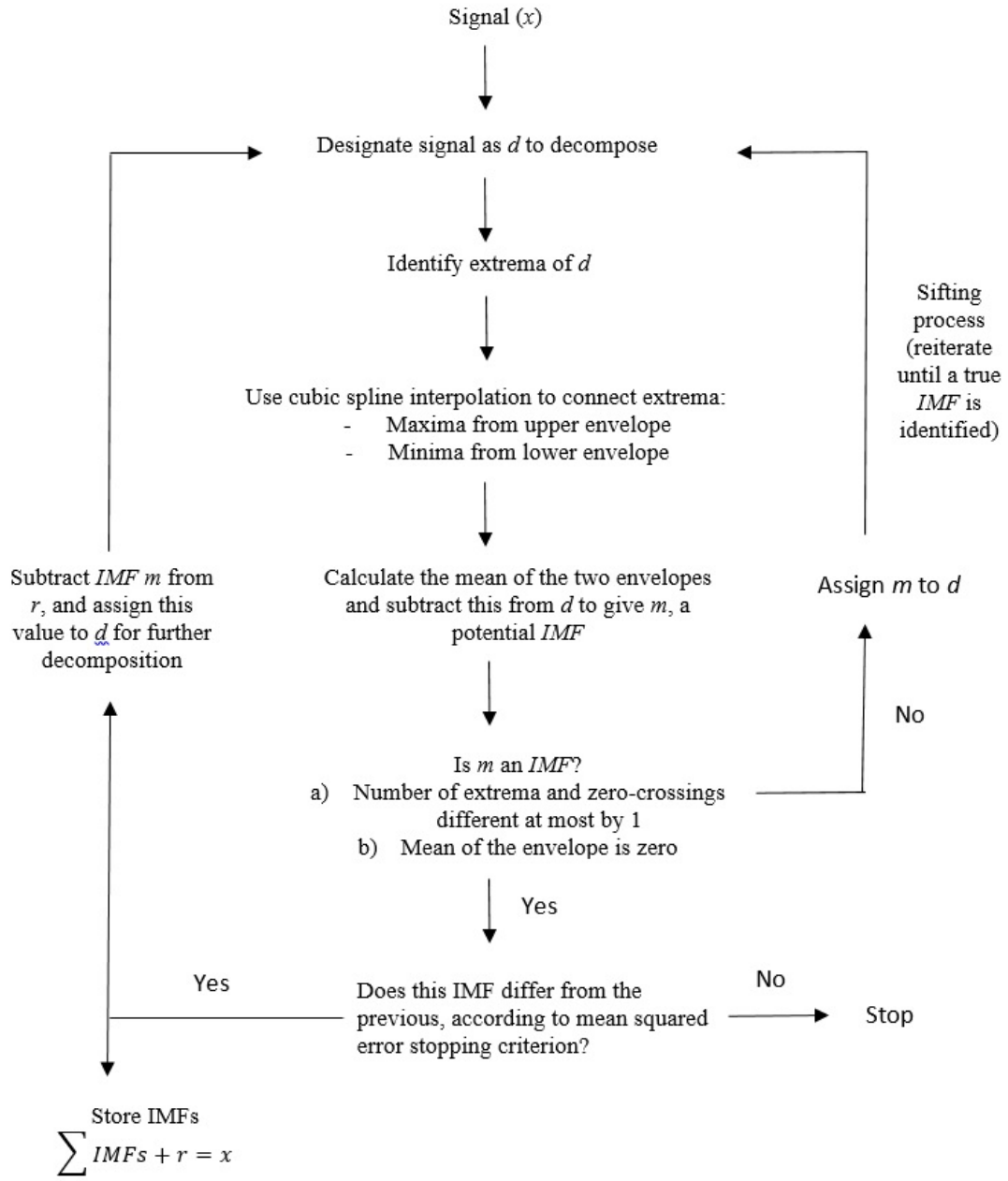


FIGURE 2.5: EMD algorithm (adapted from (Sweeney-Reed and Nasuto, 2007)).

2.3.4.2 Why EMD?

Instantaneous frequency, developed by Gabor (1946), is a proper solution to capture the tempo-spectral changes with good resolution. HT has often been performed in signal processing studies to compute the instantaneous frequency with an accurate estimation of its corresponding information. Nevertheless, instantaneous frequency only provides one value at each time, and real-time series data usually contain many intrinsic oscillations, which explains why instantaneous frequency is not defined for broadband signals, like the EEG (Burgess, 2012; Liang et al., 2005).

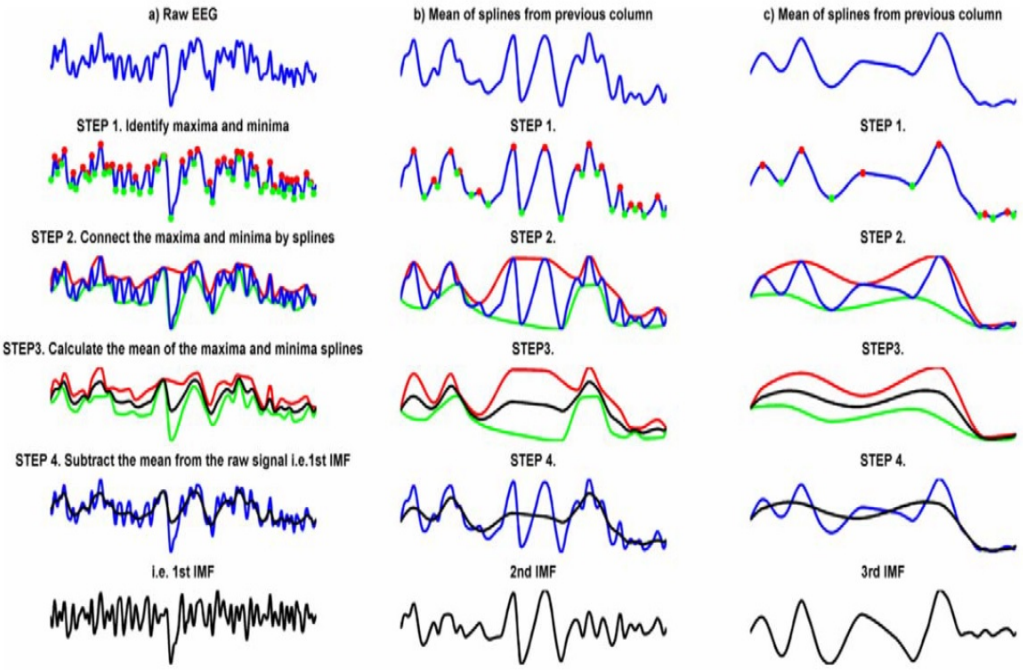


FIGURE 2.6: Schematic representation of the EMD algorithm (reproduced from (Burgess, 2012)).

Therefore, separating the intrinsic oscillations of a time series is required to estimate meaningful instantaneous information. Different approaches have been developed, but the most used methods include STFT, WT, and digital filters. However, all these methods are unsuitable for decomposing the signal optimally as they all rely on predefined frequency ranges (Burgess, 2012). For example, bandpass filters decompose EEG signals into frequency bands selected according to the traditional frequency bands, which are inconsistent among studies. For example, alpha-band was chosen to be from 8-12 Hz in (David et al., 2004), from 8-13 Hz in (Breakspear et al., 2004), from 8-14 in (Babiloni et al., 2006), or subdividing into 6-10 Hz and 10-14 Hz ranges in (Stam et al., 2003). These small changes in the alpha frequency cut-offs have been shown to influence findings (Sweeney-Reed and Nasuto, 2009).

The ranges of neural oscillations may also vary among subjects (specifically between infants and older individuals (Saby and Marshall, 2012), which make comparisons between subjects difficult and may lead to missing potentially meaningful subject's specific biomarkers (Sweeney-Reed and Nasuto, 2009). In addition, using filtering to obtain a narrow band signal could lead to spurious harmonics and negative frequencies, with no physical meaning, caused by the nonlinear and nonstationary nature of the EEGs (Huang et al., 1998).

The EMD method gives an interesting alternative to mitigate these problems as it adaptively decomposes a time series, alleviating the need for bandpass cut-offs. EMD decomposes a given signal by empirically identifying the physical time scales intrinsic to the data. Thus, each IMF has a single local frequency range, which is the desired property to compute the instantaneous frequency. Furthermore, the local symmetry requirement concerning the zero mean for an IMF leads to the avoidance of negative frequencies (Huang et al., 1998).

2.3.4.3 EMD Variants

The standard EMD (sEMD) is a univariate algorithm in nature. Thus, studies that employ the EMD method either decompose a single time series or do a sequential decomposition of multivariate signals. Sequential analysis of multivariate EEG leads to a non-identical number of IMFs across channels with different frequencies for a given IMF index. Such a scenario is problematic, as the IMFs used for comparison might relate to different frequency bands. For example, IMF4 for two channels might be the fourth of four IMFs for one channel but the fourth of eight IMFs for another. In such a case, the first would represent the low-frequency components in the signal, while the second would represent frequency components from the middle of the spectrum (Pachori, 2008)). This issue is known as mode-alignment. Another problem associated with sEMD is the phenomenon of mode-mixing, where similar frequencies appear across different IMFs (Wu and Huang, 2009).

Several extensions of EMD have been proposed to resolve these twin problems, and each of them has its own merits and demerits. These extensions include bivariate EMD (Rilling et al., 2007), trivariate EMD (Rehman and Mandic, 2010*a*), ensemble EMD (EEMD) (Wu and Huang, 2009), multivariate EMD (MEMD) (Rehman and Mandic, 2010*b*), and noise-assisted MEMD (NA-MEMD) (Rehman and Mandic, 2011).

The MEMD method improves the alignment of the corresponding IMFs from different channels and thus solving the mode-alignment problem, while the NA-MEMD algorithm extends the MEMD by adding a subspace containing a multivariate independent white noise to further address the mode-mixing problem. MEMD and NA-MEMD are the EMD variants that will be used in this work, and thus detailed descriptions of their algorithms will be given in Section 3.2.1 and Section 3.2.2, respectively.

2.3.4.4 EMD-based Applications to Predict NDDs

Different quantitative features have been extracted in the literature from the IMFs to characterise the neurological activity and reflect abnormal brain functions.

Linear time domain features are the straightforward measures used to characterise IMFs. Traditional statistical properties of the time series, such as mean, variance, and standard deviation values, can be used for this purpose. Djemili et al. (2016) calculated such features from the first four IMFs (resulting from the sEMD) to detect epileptic EEG signals from the normal ones. Hassan and Bhuiyan (2017) extracted statistical features from all IMFs (obtained from EEMD) to identify the different sleep stages.

Spectral features have also been used in some studies. Pachori (2008) introduced the calculation of the mean frequency (MF) from the IMFs (the first four IMFs decomposed using sEMD) to classify ictal and seizure-free intracranial EEG signals. Zahra et al. (2017) also computed instantaneous amplitude and instantaneous frequency for seizure detection, but they employed the MEMD to deal with the multichannel EEG recordings and extracted MF from the statistically significant IMFs.

Nonlinear time domain complexity measures have also been used to quantify the IMFs. Cui et al. (2015) calculated six entropies (the entropy concept will be discussed in Section 2.4.3) from selected IMFs (resulting from sEMD) to differentiate a mild cognitive impairment (MCI) group from the control group. Islam et al. (2018) computed Shannon and spectral entropy, described in Section 3.3, from selected IMFs (decomposed using sEMD) for sleep disorder diagnosis. Thanaraj et al. (2020) extracted five entropies from all IMFs resulting from the MEMD algorithm (excluding the first IMF and the residual mode) for the detection of schizophrenia.

Nonlinear synchronisation analysis between corresponding IMFs acquired from different brain regions has also been used in limited studies. Sweeney-Reed and Nasuto (2007) introduced this approach by using the phase-locking value (PLV) (described in Section 2.5.2) to calculate the phase synchronisation between all IMF combinations. Cho et al. (2017) examined the use of PLV for seizure detection by performing a comprehensive examination in the effects of different decomposition methods, including bandpass filtering, sEMD, MEMD, and NA-MEMD, on PLV estimation. They argued that PLV estimated using NA-MEMD could be used as a potential biomarker for seizure prediction.

2.3.4.5 Challenges and Knowledge Gaps

Some challenges, as well as gaps from the above EMD review, are discussed in the following:

1. The diagnosis of epilepsy is the most common application of EMD in EEG analysis to date in terms of brain disorders. Other studies include different aspects such as schizophrenia and MCI detection. Prediction of brain disorders such as ASD, CP, and cognitive function abnormalities using the EMD analysis method

did not gain much attention. Using EMD-based methods to analyse the brain signals of individuals with these disorders is interesting to address issues of the frequently used time-frequency analysis methods, including predefined cut-offs, nonstationarity and nonlinearity assumptions of the EEG. Thus, it is an open area to investigate whether EMD-based methods can identify new biomarkers to improve the detection of these disorders, which is the focus of this thesis.

2. There is a significant doubt on which EMD variant to use, especially with multichannel data. Various EMD extensions have been proposed to address the sEMD limitations. NA-MEMD seems to be the best alternative where it resolves the twin problem simultaneously. Nevertheless, NA-MEMD increases the signal dimensionality due to the added noise, which in turn increases the computational cost. Thus, the pros and cons of the EMD variate with respect to the underlying application should be considered.

3. Some EMD-based research argued that the low-order IMFs contain the most relevant features. Other studies used all IMFs to obtain the features or relied on selecting the statistically significant IMFs. Hence, it is still challenging to determine which IMF/IMFs contain(s) the relevant features that can discriminate well between the classes of interest.

   Identifying frequency ranges underlying the IMFs and finding their corresponding traditional brain waves (i.e., delta, theta, alpha, beta and gamma) is of great interest to understand brain abnormalities associated with various brain disorders and link the results with other findings. However, limited studies linked the IMFs with their corresponding brain waves, such as (Cho et al., 2017; Huang et al., 2013). Thus, this study aims to employ such a linking method to compare the findings with state-of-the-art results and identify biomarkers in a way that could be helpful to clinicians.

4. The IMF quantification features range from linear to nonlinear measures, and they rely on the type of component dependence they model (inter- or intra-). Choosing the right features depends strongly on the application and the EEG activity under study. Both linear and nonlinear measures will be used in this thesis, and justifications for their application will be provided when they are used in the following chapters.

2.4 Nonlinear Dynamical Analysis of EEG

As discussed earlier, brain activity exhibits nonlinear behavior, and EEG signals are considered the output of a nonlinear dynamical system. A historical review of nonlinear EEG analysis can be found in (Stam, 2005).

One approach to nonlinear time series analysis of EEG includes analysing the nonlinear behavior within the EEG signal (EEG complexity) by reconstructing, from time series of EEG, an attractor of the underlying dynamical system and characterising it in terms of its dimension (Sanei and Chambers, 2007). The other approach includes representing the nonlinear synchronisation between recordings from different brain regions. This section focuses on the nonlinear EEG complexity, while the nonlinear synchronisation between EEG channels is discussed in the next section.

Most research studying nonlinear EEG complexity focuses on introducing new analytical tools and developing nonlinear measures appropriate for nonstationary, noisy, and high-dimensional EEG data. Before discussing nonlinear complexity measures of EEG signals, a brief explanation of the conceptual framework of the nonlinear dynamics is given as follows.

2.4.1 Basic Concept of Dynamical Systems

A dynamical system is a model that changes its state over time and determines the system's evolution given only the initial state, which implies that these systems possess memory and the current state is a particular function of a previous state. Thus, a dynamical system is described by two things: states and dynamics. The state of a dynamical system is determined by the values of all the variables that describe the system at a particular moment in time, and the space made up of the state variables is called phase space or state space. The state of a system may be described by $\boldsymbol{m}$ variables, and thus it can be represented by a point in $\boldsymbol{m}$-dimensional phase space. System dynamics are the set of equations that determine how the system's state changes over time. This set of equations often consists of differential equations, one for each of the systems' variables. The line connecting consecutive states (points) in its phase space is defined as the system's trajectory, which is the path followed by the dynamical system as time progresses (Stam, 2005).

A dynamical system can be linear if all the equations describing its dynamics are linear; otherwise, it is nonlinear. Further, a dynamical system can be deterministic if the motion equations (which every system's future state must follow) do not contain any noise terms and stochastic otherwise. The neural networks of the brain are likely to be a chaotic system (Sanei and Chambers, 2007). A chaotic system is a deterministic dynamical system exhibiting irregular, seemingly random behavior (Ditto and Munakata, 1995). Deterministic chaos plays an important role as an effective tool for prediction and characterisation of the EEG signals (Sanei and Chambers, 2007).

A critical property of the chaotic systems is that, after long observation, the trajectory will converge to a subspace of the total phase space. This subspace is called the system's attractor since it "attracts" trajectories from all possible initial conditions (Stam, 2005).

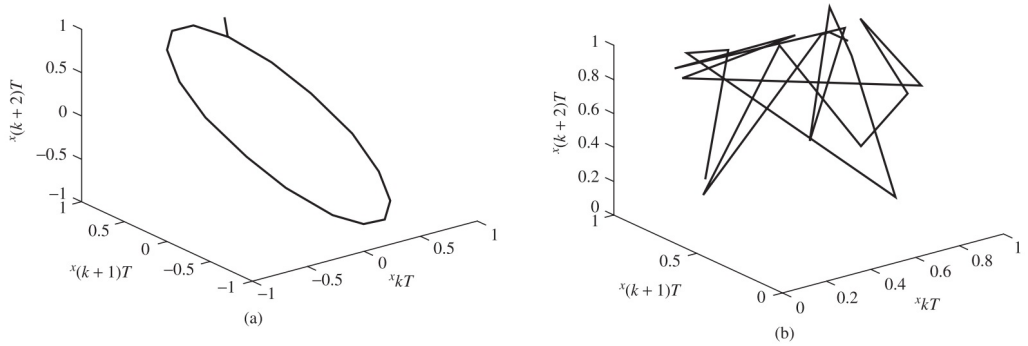


FIGURE 2.7: The attractors for (a) a sinusoid and (b) chaotic time sequence both started from the same initial point (reproduced from (Sanei and Chambers, 2007)).

The attractor, in chaotic systems, is a very complex object with fractal geometry. A chaotic system never repeats the same states, although its dynamics are confined to the attractor (Stam, 2005). Figure 2.7 shows an example of a chaotic or strange attractor compared with an attractor for a sinusoid.

2.4.2 Embedding: Reconstruction of Dynamics from Observations

The nonlinear EEG analysis approach aims to reveal, from observations, the underlying brain dynamics and their complexity. This means that the nonlinear analysis works in reverse, starting with the system's output (EEG signals) and working back to the state space, attractors and their properties. Thus, the first step toward nonlinear EEG analysis is to reconstruct an attractor in the phase space of the underlying system (the brain) from one or a few time series of observations. In other words, EEG signals need to be transformed from time-series representation into attractor representation which reflects the dynamics of the brain (Stam, 2005). However, due to the unknown mixing of the underlying system variables, the observations usually don't have a one-to-one correspondence with these variables. This problem can be solved using the embedding procedure, which allows reconstructing an equivalent attractor of the system under study. A frequently used method for embedding procedure is the delay embedding technique or just "embedding," which is based on the Takens' theorem (Takens, 1981).

Having an output signal, the embedding method constructs a set of delayed versions of the recorded signal, which are used as substitutes for the hidden internal signals. Plotting these delayed signals against each other gives the attractor in the phase space of the system (Semmlow, 2018). From a single signal x_t, a collection of signals is created, each being a delayed version of the original, as follows:

$$[x_t \;;\; x_{t+\tau} \;;\; x_{t+2\tau} \;;\ldots\ldots\ldots\ldots;\; x_{t+(m-1)\tau}] \tag{2.6}$$

where τ is the time delay and m is the embedding dimension.

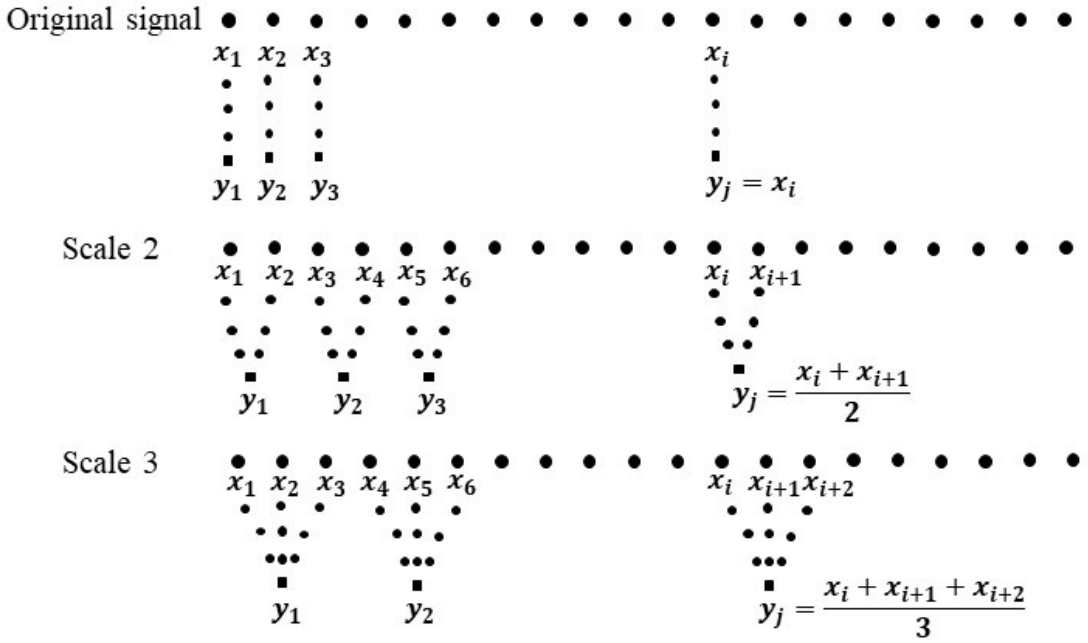
The proper choice of τ and m is a challenging task in nonlinear time series analysis. Too small delays result in similar constructed signals, and too large delays may ignore any deterministic structure of the sequence (i.e., they become independent of each other). On the other hand, Takens' theorem requires that the proper value of m should be at least $2d+1$ (where d is the true dimension of the attractor) to guarantee reconstruction of the phase space (Sanei and Chambers, 2007). However, the system (the brain in this research) order or dimension could not be inferred (Sanei and Chambers, 2007). Different approaches have been proposed to estimate τ and m. These methods usually fail when applied to real data. Moreover, the embedding dimension is strongly data-dependent as the maximum m is often limited by the length of the data, so it usually comes down to trial and error. Hence, the embedding parameters are optimal if they deliver the best possible estimates for the topological properties of the attractor or provide the most accurate prediction of the time series (Semmlow, 2018).

2.4.3 Characterisation of the Reconstructed Attractor

Once the attractor has been reconstructed, the next step toward nonlinear analysis is to characterise it quantitatively. Different measures have been introduced to accomplish this task, such as Lyapunov exponents, correlation dimension, and entropy (Stam (2005)). Short length, nonstationarity, and associated noise of the time series often raise issues in the computation of Lyapunov exponents and correlation dimension (Stam (2005)). Thus, the focus of this thesis is on entropy measures. Reviews of Lyapunov exponents and correlation dimension can be found in (Sanei and Chambers, 2007; Stam, 2005).

In physiological systems, entropy is a measure of signal complexity. Complexity can be defined as the amount of nonlinear information that a time series conveys over time (Omidvarnia et al., 2018). Different quantitive entropy measures have been proposed to search for or describe irregularity of the signals, such as sample entropy, permutation entropy and spectral entropy. These entropies provide a single number that can be used to compare the information content of different signals. Higher complexity is often expected in a more healthy condition as a representation of physiological reaction to the changing environment (Chu et al., 2017). The loss of such complexity may be of diagnostic value in certain disorders (Peng et al., 2009). Entropy measures have been widely applied to identify different brain abnormality conditions, such as schizophrenia detection (Thanaraj et al., 2020) and ASD prediction (Gani et al., 2020).

This section describes multiscale entropy, which is widely employed in investigating various brain disorders. Detail descriptions of sample entropy, permutation entropy, spectral entropy are given in Section 3.3 as they are the entropy measures used in this thesis.



$$y_j^{(\tau)} = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} x_i, 1 \leq j \leq {}^{N}/_{\tau}$$

FIGURE 2.8: MSE procedure (adapted from(Costa et al., 2005)).

As most biological signals are likely to carry information across multiple time scales, estimating the information content of the EEG signal over multiscale may reveal more information than the entropy of a broadband signal (Bosl et al., 2011). This fact gives rise to the proposal of the multiscale entropy (MSE) method (Costa et al., 2005).

The MSE quantifies the complexity in a time series by calculating the entropy measure, usually sample entropy, at multiple temporal scales using a coarse-graining procedure (Costa et al., 2005). Thus, the original EEG signal $\{x_1, \ldots, x_i, \ldots, x_N\}$ is coarse-grained into consecutive time series $\left\{y^{(\tau)}\right\}$. To this end, the original signal is first divided into non-overlapping windows of length τ. Then, the data points inside each window are averaged, so each coarse-grained time series is defined by $y_j^{(\tau)} = \left(\frac{1}{\tau}\right) \sum_{i=(j-1)\tau+1}^{j\tau} x_i, 1 \leq j \leq \frac{N}{\tau}$, where the length of each coarse-grained sequence is τ times shorter than the length N of the original signal. Entropy measues are then calculated for each series $\left\{y^{(\tau)}\right\}$. Figure 2.8 illustrates the coarse-graining process.

MSE has been adopted in a broad range of applications to characterise abnormal cortical dynamics of different physiological conditions such as ASD (Liu et al., 2017) and

ADHD (Chu et al., 2017). Nevertheless, the MSE method is not well adapted for studying nonlinear and nonstationary signals, such as EEG. This is due to its linear extraction of scales and the substantial effect of nonstationarities appearing on scales larger than those considered for MSE analysis (Peng et al., 2009). Therefore, it is essential to address these challenges by applying a data-driven analysis technique that is developed to deal with the nonlinear and nonstationary nature of EEG signals. Such a technique may improve quantification of the complexity of EEG signals (Peng et al., 2009). EMD-based methods are an adequate alternative for decomposing the data before calculating the entropy measures since it provides an adaptive way to analyse the signal without *a priori* assumptions of the oscillations it may possess (Peng et al., 2009).

2.4.4 Challenges and Knowledge Gaps

This section breifly reviewed the nonlinear analysis of EEG signals. Some challenges and knowledge gaps from this review in relation with the predection of different NDDs are listed in the following:

1. Different measures have been borrowed from chaos theory for the nonlinear analysis of EEG, such as entropies. However, only (Sajedi et al., 2013) employed such measures to quantify EEG signals for CP diagnosis, and they were not used for cognitive function prediction of at-risk infants to the best of the author's knowledge.

2. Combining these nonlinear measures may give a complete characterisation of the signal dynamics. Such a combination was employed in different cognitive neuroscience research kinds of (Thanaraj et al., 2020) and will also be investigated in this study.

3. EMD-based methods are considered potential candidates for achieving data-driven multi-scale operation by replacing the conventional coarse-graining scales with data-adaptive scales via IMFs. Thus, EMD-based approaches can overcome the nonstationarity and nonlinearity limitations of MSE analysis. Measuring entropies from the EMD-based domain has some applications in the EEG analysis of abnormal brain function as discussed in Section 2.3.4.4. However, this analysis approach was not employed to predict ASD, CP, and cognitive function. Investigating such an adaptive method to decompose the EEG signals could reflect new biomarkers for the prediction of these disorders, and this will be the focus of the thesis.

2.5 Nonlinear Synchronisation Analysis of EEG signals

While the previous section discussed the nonlinear analysis of a local time series, this section provides the basic concepts of nonlinear analysis in relations between two or more time series.

The brain can be visualised as a complex network of interacting and coupled subsystems. Higher brain functions, such as motor activities and cognition, depend strongly on information integration and efficient processing of this network. Various questions have been raised on how the functional interactions between different brain areas occur and how such interactions may be affected in different pathological conditions (Stam, 2005). These concepts are discussed briefly in the following.

2.5.1 The Brain as Network of Coupled Dynamical Systems

The synchronisation between oscillating systems is an important concept that describes how the functional interactions between different areas in the brain take place. Synchronisation can occur between regular, linear oscillators and irregular, chaotic systems. Considering the chaotic nature of the neural system, Boccaletti et al. (2002) defined synchronisation of chaos as the process of two or more dynamical systems adjusting a given property of their motion to a common behavior due to a coupling or to a forcing (periodical or noisy).

Two concepts related to the functional interactions between different areas in the brain have been identified: functional source and functional network (Stam, 2005). Functional source refers to the part or parts of the brain which contribute to the activity recorded at a single electrode. In contrast, the functional network refers to pair-wise correlations between all functional sources (Stam, 2005). Thus, nonlinear measures derived from single time series, which were discussed in Section 2.4.3, provide information about the dynamical complexity of the functional source, while synchronisation measures between signals provide information about the dynamics of the highest level functional network.

Typically, network dynamics in a pathological condition is accompanied by either abnormally higher or abnormally lower levels of synchronisation between functional sources compared to a healthy neurological system. Thus, quantifying the coupling among different functional sources is critical to observe deficits underlying a higher-level brain function (Peng et al., 2009).

2.5.2 Nonlinear Synchronisation Measures

Coherence is a well-known linear synchronisation measure that quantifies the squared cross-correlation between two signals (Pfurtscheller and Andrew, 1999; Hamedi et al., 2016). Coherence is based mainly on phase consistency; that is, the coherence between two signals is high when the phase difference between them tends to remain constant. Coherence is sensitive to both changes in power and phase relationships (Bendat and Piersol, 1980; Hamedi et al., 2016). This measure has been widely and successfully employed to quantify the synchronisation of brain processes in different brain disorders, such as (Kuhn-Popp et al., 2016; Ghaderi et al., 2017; Schwartz et al., 2017).

Nonlinear synchronisation methods are designed to detect the dynamics of EEG signals based on interacting chaotic oscillators. In contrast to coherence, phase synchronisation (PS) is a nonlinear synchronisation method that provides an amplitude-free measure of coupling between brain regions and is thus less susceptible to the effects of artifacts and inter-trial/inter-subject amplitude variability (Hamedi et al., 2016). This section reviews the phase-locking value (PLV) as the traditional PS measure.

PLV of EEG is a PS measure introduced by Lachaux et al. (1999) and has been employed in different cognitive neuroscience research kinds of (Cho et al., 2017; Sweeney-Reed et al., 2012). PLV quantifies the phase difference between two narrow-band time series. The phase of each signal is initially calculated using HT as described in equations (2.2 - 2.4). The instantaneous PLV between two signals of length N and phase $\Theta(t_n), n = 1, 2, ...N$ is then given by:

$$PLV = |exp[i\Delta\Theta(t_n)]| \tag{2.7}$$

where $\Delta\Theta(t)$ denotes the phase difference between two signals, $i = \sqrt{-1}$ and $exp[.]$ is the expected value. The PLV values range between zero and one in which zero reflecting the case of no phase synchrony between the two signals and one when they are strongly coupled.

Computing PLVs between pairs of signals recorded from nearby electrodes may lead to nonzero values arising from the same source contributing to both signals due to the volume conduction (VC) in channel space, i.e., through the tissues of the head, including the brain, cerebrospinal fluid, skull, and scalp (Sanei and Chambers, 2007). This type of artefact can easily be mistaken for phase-locking between different signals. Thus, another measure of phase-locking, which is zero in such linear mixing and nonzero when there is a consistent nonzero phase difference between the two signals, is recommended (Stam et al., 2007). Different synchronisation measures have been developed to address the VC effect, such as phase lag index (PLI) (Stam et al., 2007) and the imaginary part of coherency (Nolte et al., 2004). Further details about PLI are provided in

Section 3.4 as it is the chosen synchronisation measure employed in this work. PLI has been successfully used in various cognitive neuroscience research, such as the prediction of the severity of depression (Mohammadi and Moradi, 2021), Alzheimer's disease (AD) (Nobukawa et al., 2020) and ASD (Han et al., 2017).

2.5.3 Challenges and Knowledge Gaps

PS methods estimate the synchronisation between different brain regions over time and are suggested in the literature as interesting candidate measures considering the nonlinear and nonstationary properties of EEGs. PS measures quantify temporal relationships between the phases of different signals. Such quantification can further the understanding of impaired interactions between brain regions of individuals with different NDDs. Some challenges and knowledge gaps in this field are listed in the following:

1. To the best of the authors knowledge, there has been limited work in applying nonlinear synchronisation analysis to quantify EEG signals of at-risk infants for either CP or cognitive function prediction.

2. The EEG signals should be in a narrow-band representation to quantify the PS between them. Traditional decomposition of EEG using Fourier-based and band-pass filtering is often used for this purpose. As as discussed in Section 2.3.4.2, such decomposition methods require *a priori* selection of the filter cut-offs. This prior assumption may lead to missing potentially meaningful brain interactions where episodes of synchrony of interest may cross these artificial boundaries (Sweeney-Reed and Nasuto (2009)). EMD-based methods are ideally suited for such decomposition where they determine the relevant output oscillations in a data-driven manner so that synchrony events are unlikely to be missed (Looney et al. (2015)). Nevertheless, limited studies have explored the application of EMD-based synchronisation for NDDs prediction, as discussed in Section 2.3.4.4. Therefore, this will be the focus of this thesis.

3. Assessing the correlation between the complexity and the synchronisation measures is interesting as any change in signals' complexity, as a consequence of a particular brain disorder, could indicate a change in the synchronisation between them. Such an assessment has only been investigated in a few neuroscience applications, such as ASD and AD diagnosis (Edgar and Roberts, 2016; Nobukawa et al., 2020), and therefore, it will be explored in this study.

4. Combining the complexity and synchronisation features could provide a more detailed characterisation of the neural functions and thus assist in NDDs prediction. To the best of the author's knowledge, such a combination was only

employed in the recent research concerning the prediction of depression severity (Mohammadi and Moradi, 2021), and therefore, it will be investigated in this thesis.

2.6 Related Works

This section reviews advances in ASD classification as well as CP and cognitive function prediction after neonatal HIE, identifying some knowledge gaps to be addressed by this study.

2.6.1 ASD Diagnosis

ASD is a lifelong condition associated with societal costs related to services and lost productivity by individuals and their families (Zwaigenbaum and Penner, 2018). ASD is mainly related to social-communication deficits and restricted and repetitive behaviors.

Considering autism as a spectrum disorder reflects the heterogeneous nature of the core ASD symptoms and the fact that they fall along a continuum of severity. The term also indicates the presence of ASD subgroups, identified by distinct cognitive intellectual and/or language impairments, and the potential association of genetic and medical conditions or co-occurring other psychiatric disorders such as ADHD and anxiety (Ousley and Cermak, 2014). Early diagnosis (around or before 24 - 30 months of age) and subsequent early intervention is the most effective way of treating children with ASD in terms of maximising behavioral and cognitive outcomes (Bowman and Varcin, 2018).

Traditionally, ASD is diagnosed around 3.5 - 4 years of age, based on narrative behavioral interactions between the children and specialists besides parental questionnaires. Different assessment criteria were developed to standardise the methods for assessing parent- and caregiver-reported symptoms and clinical observation and determine the presence of an ASD and its severity. International classification of diseases, tenth edition (ICD-10) (World Health Organization, 1993) is the commonly-used diagnostic manual in the UK, followed by diagnostic and statistical manual of mental, fifth edition (DSM-5) (American Psychiatric Association, 2013; *Diagnostic criteria - a guide for all audiences.*, 2020). These assessment methods usually require an extended period to detect abnormalities, subjective and lacking biological evidence (Bowman and Varcin, 2018). Moreover, such a scenario cannot be applied for early diagnosis, as the children need to be at a certain age (typically > 3.5 years) to evaluate their behavior. Therefore, more emphasis has been recently given to diagnosing ASD based on neurological biomarkers, increasing evidence-based diagnostic accuracy (Bowman and Varcin, 2018).

Several studies have considered EEG metrics as potential biomarkers of ASD due to its high temporal resolution of tracking brain activity at resting-state via patterns of neural oscillations as discussed in Section 2.2.1 and in response to specific stimuli as mentioned in Section 2.2.2. The focus of this section is to review the literature that employed ERP-based measures to diagnose children with ASD, more precisely, in response to facial expression (FE) stimuli.

FE contribute significantly to nonverbal human communication relaying information about individuals' feelings, thoughts, and intentions. Emotion recognition abilities typically begin in early infancy, improving and developing through adulthood (Black et al., 2017). Six basic emotions (happy, sad, fear, anger, disgust, surprise) often contribute to the early emotional recognition with discrimination of these emotions reported in children aged five to seven months (Black et al., 2017). These emotions can be categorised into negative (i.e., sadness, anger, disgust, and fear) and positive (i.e., happiness and surprise) emotions (Black et al., 2017).

Different behavioral studies have been performed to investigate the face processing ability in individuals with ASD. The deficit in processing the six basic emotions was shown to be apparent across the developmental trajectory of the population with ASD (Uljarevic and Hamilton, 2013). Moreover, recognition of fear was shown to be worse than happiness (Uljarevic and Hamilton, 2013).

Black et al. (2017) and Monteiro et al. (2017) conducted systematic reviews of EEG-based ERP studies concerning FE processing in ASD. Most of the reviewed studies quantify the ERPs using measures of peak latency (the time from stimulus onset to the most positive or negative point within the defined time window of interest) and peak amplitude (the most negative or positive point within the time range of the ERP component of interest). N170 and P100 were the most commonly studied ERP components. Both the latency and amplitude of the N170 component were found to be atypical in ASD children. Furthermore, children with ASD showed different N170 latencies between fearful and happy expressions and between neutral and angry faces. On the other hand, some studies reported that ASD children had similar P100 compared to TD matched controls in response to FE. In contrast, differences in both latency and amplitude of the P100 component in response to the six basic emotions were concluded by others, as discussed in the reviews of Black et al. (2017) and Monteiro et al. (2017).

In addition to the heterogeneity obtained in the literature from quantifying the ERPs using latency and amplitude measures to study FE processing in ASD, these measures do not reflect the abnormal oscillations of the brain dynamics, which are relevant to characterise cognitive processes of various disorders. Limited studies used time-frequency analysis of EEGs to reflect tempo-spectral integration of cortical activity in processing FE in individuals with ASD, including phase synchronisation, coherence and PSD. All studies reported atypical cortical activities in children with ASD, with differences being

reported across the frequency spectrum. In the investigation of Yeung et al. (2014), children with ASD were found to have abnormal theta coherence compared to TD children. Specifically, they showed lower right frontal theta coherence in response to the positive emotions compared to neutral faces. Moreover, children with higher theta coherence appeared to have lower autistic symptomology.

Garcia Dominguez et al. (2013) employed the imaginary part of coherency and found that ASD children exhibited enhanced synchronisation during the post-stimulus time, particularly at lower frequency bands. However, they did not focus on the abnormalities in discriminating different emotional expressions. Jamal et al. (2014) used PS measurements to assess functional connectivity, and they found that these features can classify the children with ASD from their TD peers with 94.7% accuracy. Nevertheless, their exploration mainly combined the data from different stimuli. Thus, they tested the capability of the PS to differentiate between the ASD and TD children without giving attention to the effect of each FE (happy vs. fear vs. neutral). Though their features were based on time-frequency analysis, they did not discuss the oscillations of interest associated with their findings.

Khuntia et al. (2019) carried out multivariate pattern analysis using the classwise principle component analysis in both time and time-frequency domains. Classwise principle component analysis is a two-step classification algorithm developed to classify high dimensional noisy neural signals with fewer training samples (Das et al., 2007), including identifying the sparse non-informative subspace in the data and extracting the best features suited for the efficacy of the classifier performance in the remaining subspace. Classification performance reached around 81% in the time domain analysis and 84% in the time-frequency domain exploration regardless of the face and non-face stimuli (as they include a tree stimulus). In the time-frequency analysis, alpha and beta oscillations seemed to identify ASD children best.

Knowledge Gap/s in the Field

Controversy and knowledge gaps from the above EEG-based FE recognition review concerning ASD diagnosis are summarised in the following:

1. A significant degree of heterogeneity across FE recognition studies can be observed from the literature. While, for example, some research found that individuals with ASD have general deficits in recognising negative expressions, others found the abnormalities to be associated with the positive ones.

2. Most of the studies in the literature used latency and amplitude measures to characterise ERP activity for ASD diagnosis. Nevertheless, these measures do not reveal oscillatory brain activities, which are relevant to reflect different cognitive processes (as discussed in Section 2.2.1), and therefore could improve the prediction accuracy of ASD.

Limited studies used time-frequency measures to explore the oscillatory brain activities in response to FE stimuli. However, these studies analysed the signals using time-frequency methods that rely on a predefined basis, which may lead to missing potentially meaningful brain dynamics, as discussed in Section 2.3.4.2. Although EMD-based methods can solve this issue, no previous study in this field has investigated their relevance. Therefore, this thesis aims to use EMD-based time-frequency analysis to characterise ERP signals of individuals with ASD.

3. To the best of the author's knowledge, no single research exists exploring the nonlinear dynamics of ERP components to characterise the signals in response to FE in individuals with ASD. Investigating the nonlinear measures, such as entropy, of ERP signals is interesting and may reflect the abnormal EEG behavior of individuals with ASD and, therefore, aid in their diagnosis. Thus, using such nonlinear measures in the EMD-based domain is going to be the focus of Chapter 4.

2.6.2 CP Prediction in Newborns with HIE

HIE, as a consequence of perinatal asphyxia, is one of the common causes of neonatal brain injury. Perinatal asphyxia is a lack of oxygen and blood supply to the brain during the prenatal, intrapartum or postnatal period (Allen, 2012). HIE occurs approximately in 3 per 1000 term births and is considered one of the leading causes of neonatal death (Schie et al., 2015). Around 20–40% of HIE survivors have severe neurological disorders even when treated with hypothermia, which is the standard clinical treatment in most centers. CP is considered as the most frequent HIE outcome, and epilepsy co-occurring in 15–60% of children with CP (Byeon et al., 2015). Deficits other than CP or major disability include cognitive and behavioral impairments (Schie et al., 2015).

This section focuses on reviewing the aspects related to CP and available methods that assist in its early identification, while the methods related to the early prediction of cognitive function, as a consequence of HIE, are discussed in the next section.

According to Rosenbaum et al. (2007), CP is a group of developmental disorders in movement and posture, causing activity restrictions or disability that are attributed to disturbances occurring in the fetal or infant's brain. The "group" notation reflects the CP heterogeneity in terms of its types and severity of impairments. Children with CP may develop epilepsy, behavior disorders, poor saliva control, bladder problems and sleep disorders (Tonmukayakul et al., 2018).

Families and health care systems carry socio-economic costs to facilitate schools and community engagements for individuals with CP. Early identification of at-risk infants of developing CP later in life is important for appropriate early counseling and

planning of intervention strategies, which ultimately may lead to improved outcomes (Hadders-Algra, 2014).

Hadders-Algra (2014) provided a detailed review of the opportunities and challenges for early diagnosis and early intervention of CP. The study stated that the most frequently used assessment methods for early prediction of CP are (a) neurological and neuromotor assessments, (b) neuroimaging, and (c) neurophysiological tests. Even though neurological and neuromotor assessments have been widely used in diagnosing CP and their prediction value is generally good, they are subjective and often require a longitudinal series of tests to detect the abnormalities.

Alternatively, neuroimaging techniques have been used as promising tools for the early prediction of CP in high-risk infants (Hadders-Algra, 2014). Considerable literature also employed neurophysiological tests, particularly conventional grading electroencephalogram (cEEG) and amplitude-integrated electroencephalogram (aEEG), with infants at risk of adverse neurodevelopmental outcomes. Although these modalities were found to predict the outcomes well, their applications to predict CP (from at-risk infants) are lacking (Hadders-Algra, 2014). Moreover, interpretation of the prognostic value of neuroimaging and neurophysiological methods remains subjective.

On the other hand, quantitative EEG (qEEG) analysis could provide objective, reproducible and reliable biomarkers to characterise the brain activities related to CP. Apart from predicting CP from at-risk infants, limited qEEG studies have been conducted to investigate the neuropathology of individuals with CP. Spectral power, functional brain connectivity, particularly coherence, and complexity analysis of EEG signals are the most common measurements used in this field. Gao et al. (2016, 2017) assessed the temporal and spatial correlations of EEG signals in adolescent patients with CP and their neurotypical peers using microstate and omega complexity measures, respectively. These studies reported enhancement in the signals' temporal and spatial complexities of individuals with CP compared to the normal group. However, such linear-based measurements are not well adapted for EEG analysis due to the nonlinear and nonstationary characteristics of the signals.

Coherence measures and spectral power were employed in different studies to characterise EEG signals of children with CP. Koeda and Takeshita (1998) reported lower interhemispheric coherence (ICoh) for the alpha-band, higher ICoh for the theta-band, and higher intrahemispheric coherence (HCoh) for the delta-, theta-, and beta-band in the CPs compared to controls. Kulak and Sobaniec (2005) also showed lower ICoh for the alpha-band in the children with CP than controls. Moreover, Kulak et al. (2005) reported lower ICoh for the beta-band, higher ICoh for the theta- and delta-band and higher HCoh for the alpha-band in children with CP than the control group. However, coherence measures are affected by the VC and restricted by the assumption of signal's linearity and its low temporal resolution (Stam et al., 2007).

On the other hand, time-frequency analysis methods used in the previous CP studies to calculate PSD are limited by their stationarity assumption of the signals and their reliance on predefined traditional brain waves. Prior selection of the frequency ranges may lead to omitting potentially meaningful brain dynamics, specifically in the case of infants, due to the well-known variability between them and the older individuals in the neural oscillations of interest (Saby and Marshall, 2012).

Details on the identification of CP using nonlinear complexity analysis are limited to Sajedi et al. (2013), who adapted fractal dimension to measure the nonlinear complexity of EEGs. They reported a higher EEG complexity in children with CP and a classification accuracy of 94.8% in distinguishing between the CP and normal groups. Nevertheless, they estimated the fractal dimension from broadband EEGs and did not consider the series of different intrinsic oscillations inherent in the signals.

Knowledge Gap/s in the Field

The knowledge gaps from the above review concerning CP prediction are summarised in the following:

1. Although different qEEG research has been conducted to explore the brain activities related to CP, no single study exists using a qEEG measure that considers both the nonlinear and nonstationary characteristics of EEG. Using such measures is interesting as they may be more appropriate to reflect the abnormal EEG dynamics associated with CP brains and provide new biomarkers that aid in their prediction. Thus, the focus of this thesis (particularly Chapters 5 and 6) is to estimate the nonlinear complexity and synchronisation measures from the EMD-based domain to explore biomarkers for CP prediction.

2. Early identification of infants who are at high risk of developing CP later in life is recommended for appropriate early counseling and planning of intervention strategies Hadders-Algra (2014). However, Most previous qEEG explorations have been carried out in adolescents and children, and investigating qEEG to identify CP at infancy was rarely reported (George et al., 2020). Therefore, the aim of Chapters 5 and 6 is to quantify the EEG signals of at-risk infants who develop CP later in life.

2.6.3 Cognitive Function Prediction in Infants with HIE

Cognitive function impairment as a consequence of HIE is considered one of the most expected outcomes, which could also co-occur with other NDDs (Slaughter et al., 2016). Evidence suggests that children (with HIE at birth who did not develop CP) between

two-year-old and school-age are at high risk of developing learning disabilities, behavior and cognitive deficits (Schie et al., 2015; Schreglmann et al., 2016).

Cognitive impairment is generally associated with slow information processing, deficits in working memory, attention, and executive function. Consequently, a significant socio-economic impact can be generated on the affected individuals and their families, including education, social participation, employment, and quality of life. Thus, early prediction of cognitive deficits in at-risk infants could aid in providing them with early intervention and therefore improve their cognitive outcomes. So far, however, the diagnosis of cognitive impairments cannot be carried out before 3 to 5 years of age due to the slow rate of maturation of complex abilities such as executive functions and attention (Cainelli et al., 2021; He et al., 2018). Although abnormalities in circuitry formation begin early, they will manifest only when the system is no longer able to compensate for the constantly increasing demands of the surrounding environment (Cainelli et al., 2021).

Recently, there has been increased interest in investigating the brain function abnormality in early infancy to predict the cognitive outcomes. Neuroimaging techniques were used in several studies to identify infants at risk of cognitive impairment (He et al., 2018; Slaughter et al., 2016). Neurophysiological techniques, such as EEG, were also employed in the literature for the same purpose.

Kong et al. (2018) conducted a systematic review highlighting the two approaches adopted in predicting the cognitive outcomes. First is the analysis of EEG characteristics for the binary classification of individuals as either cognitively impaired or normal. Second is the analysis of EEG to estimate the relationship between its features and the individual's long-term cognitive score, which is based on clinical tests such as BSITD-III (Weiss et al., 2010) and Wechsler preschool and primary scale of intelligence III (WPPSI-III) (Volkmar, 2013). Compared with binary classification, predicting the cognitive scores reflects the level of cognitive impairment, rather than determining the group membership, which can be more challenging (Sui et al., 2020).

Limited studies have shown that early individual's EEG characteristics are related to the later cognitive clinical scores, and thus they can be used to predict the long-term cognitive outcome. Lloyd et al. (2021) used cEEG grading of preterm infants to predict their two-year neurodevelopmental outcome, which was based on the BSITD-III assessment. Moderate to high negative correlation was identified between EEG grades and the BSITD-III subscales (motor, cognitive and language). Nevertheless, the EEG grading system is considered subjective due to its dependency on the interpretation by clinicians, which may affect the accurate prediction of the cognitive outcome.

On the other hand, Suppiej et al. (2017) evaluated the relationship between spectral EEG features of preterm infants (within the first week of life) and developmental outcomes that were assessed at one year of age using the Griffiths scale of mental development (Griffiths, 1970). The Griffiths scale provides a global developmental quotient and sub quotients for the individual subscales (locomotor, personal-social, hearing and speech, hand and eye coordination, and performance). They reported a negative correlation between the delta-band and Griffiths developmental quotient. A positive correlation between alpha-, beta-band and developmental quotient was also reported.

Cainelli et al. (2021) also employed neonatal spectral EEG and reported a significant correlation between the frequency bands and the attention tasks that were evaluated at six years of age by WPPSI-III. West et al. (2005) quantified EEG of infants during the first four days after birth based on the measurements of the amplitude. Significant associations were reported between the amplitude-based measures and the mental developmental indices assessed by BSITD-II (Bayley, 1993) at 18 months. Thus, these studies employed objective features for predicting the cognitive outcomes and provided insights. However, spectral and amplitude-based measures are often limited by their linear and stationarity assumptions of the underlying signals. Spectral analysis is further limited by the *a priori* selection of the filter cut-offs. Therefore, it is interesting to explore other qEEG measures that consider these limitations and examine whether/how they could contribute to the course of early prediction of cognitive development.

Kuhn-Popp et al. (2016) investigated the relation between EEG coherence (measured at 14 months) and epistemic language skills that assessed at 48 months using a coding scheme for mental state terms (Bretherton and Beeghly, 1982). Correlational analyses results suggested significant associations between EEG coherence measures of the individuals' left hemisphere and their epistemic language skills at 48 months. Nevertheless, EEG coherence measures are associated with the known limitations of VC, the linearity assumption of the signals, and the low temporal resolution.

Knowledge Gap/s in the Field

Different studies were carried out to explore the relationship between early qEEG features and later cognitive function scores/severity. Although the significant insight of these studies, the employed features do not account for the nonlinearity and nonstationarity of EEGs. Using measures that address such critical properties of the signals is interesting as they are more appropriate for characterising brain dynamics and may reflect new biomarkers for early cognitive function prediction. Computing nonlinear measures from the EMD-based domain could be a good candidate for this purpose, and thus, this will be the focus of Chapter 7.

2.7 Summary

This chapter reviewed some existing EEG signal processing techniques and their NDDs applications. EMD-based methods have been suggested as the most suitable decomposition approach to study the time-frequency activities of the signals because they effectively consider the nonstationarity and nonlinearity of EEG. Nonlinear complexity and synchronisation analysis approaches have been reviewed and presented to quantify the nonlinear behavior of EEG signals and reflect brain function abnormalities associated with different NDDs. Some state-of-the-art applications of EEG signal processing for the prediction of ASD, CP, and cognitive function impairments have been reviewed, and knowledge gaps have been highlighted.

Chapter 3

Methodology

As outlined in Chapter 1, this thesis aims to utilise adaptive time-frequency nonlinear EEG analysis to explore quantitative biomarkers for abnormal brain dynamics. Particularly, the focus is to compute the nonlinear complexity and syhnchronisation measures from the EMD-based domain to explore quantitative features that could assist in predicting NDDs. The proposed analysis was carried out in a machine learning framework, and the aim of this chapter is to provide the overall process of this framework.

The layout of this chapter is ordered by the steps of the proposed methodology. Thus, Section 3.1 generally describes the preprocessing methods needed to be applied to the EEG signals to facilitate analyses. Section 3.2 presents the EMD-based algorithms used in this thesis to perform an adaptive tempo-spectral analysis, accounting for the non-stationarity and nonlinearity properties of the signals. Section 3.3 and Section 3.4 review the complexity and synchronisation features proposed to be extracted from the EMD-based domain to quantify the EEG signals, respectively. Section 3.5 provides details about statistical analysis methods utilised in this study to evaluate the proposed features and identify potential biomarkers. Finally, Section 3.6 describes the machine learning methods aimed to be used to explore how useful the observed biomarkers could be in the practical prediction.

3.1 Preprocessing methods

Preprocessing refers to any signals' reorganisation or transformation that occurs between collecting and analysing the data. Preprocessing methods are performed to attenuate noise and artefacts from the signals, such as eye-blinking, muscle activities, and other external or internal disturbing effects. Although most of these artefacts are removed by the hardware provided in advanced EEG machines, there is usually a remaining part that needs to be eliminated (Cohen, 2014). Different preprocessing have

been utilised for this purpose, but choosing the appropriate methods and steps depends on the experiment design, the equipment used to record the data and the planning analysis (Cohen, 2014; Sanei and Chambers, 2007).

Preprocessing methods were not applied on the dataset used in Chapters 4 and 8 as the preprocessing of these datasets was done previously in (Apicella et al., 2012; Jamal et al., 2014) and (Cavanagh et al., 2019), respectively. This section provides general descriptions of the preprocessing methods that were applied on the datasets of the infants with neonatal HIE used in Chapters 5-7, and specific details of these methods are provided in Chapter 5. EEGLAB (an open-source toolbox in MATLAB) was used for preprocessing, and the methods are presented below in the order of their application.

1. **Filtering:** Filtering can help remove low-frequency drifts and high-frequency artefacts from the signals. A finite impulse response (FIR) bandpass filter was used in this thesis for this purpose as it is preferred over infinite impulse response filters. FIR filters are more stable and less likely to introduce nonlinear phase distortions. However, FIR increases filter order, and thus its computational cost is a bit higher than IIR. Nevertheless, this is not a major limitation with modern computers (Cohen, 2014). To determine the FIR filter order, EEGLAB uses three times the lower frequency bound as the default order (Cohen, 2014).

2. **Bad-channels removal:** A bad channel is either completely flat or measures noise that is many orders of magnitude larger than the real brain signal (Cohen, 2014). The EEGLAB automatically identifies these channels. Removing a bad electrode entirely from the dataset may cause confusion when averaging across subjects because, for instance, one subject will have 19 electrodes while another will have 15 electrodes (Cohen, 2014). Thus, in this study, the bad channels were equally removed from the datasets of all subjects.

3. **Re-referencing:** Referencing is an issue for EEG as the voltage values recorded from each electrode should be related to a voltage value recorded from a chosen reference site. Averaged mastoids (the bone behind the ear) or earlobes are typical reference electrodes. Different referencing methods were discussed in the EEG research, and they generally concluded that no reference is perfect (Cohen, 2014).

   The signals can usually be re-referenced offline to any reference channel or channel combination even though they have been recorded online with a given reference (Cohen, 2014). This thesis used a common averaged reference (inside EEGLAB) for re-referencing. Thus, re-referencing was achieved by creating an average of all scalp electrodes and subtracting the resulting signal from each channel. Using this reference, amplitudes were reduced, and each channel contributed equally to the new reference.

4. **Epoching:** Epoching refers to the process of cutting the continuous EEG signal into segments to aid the analysis, particularly of the task-related datasets. Although epoching is not necessary for resting-state datasets, the signals can be segmented into non-overlapping segments of a few seconds to facilitate the analysis and reduce the effect of the nonstationary property of EEG (Cohen, 2014; Kulak et al., 2005; Sakkalis, 2011).

5. **Removing epochs because of artefacts:** Epochs that obviously contain artefacts rather than EEG, such as ocular artefact, can be automatically rejected through the EEGLAB toolbox, for example, by examining epochs with high variance compared to other epochs. Such artefacts can also be identified by visual inspection (Nunez et al., 2016). Not removing these artefacts hinders the ability of the independent components analysis algorithm, described next, to isolate typical artefacts such as eye blinks (Nunez et al., 2016).

6. **Independent Components Analysis (ICA):** Independent components analysis (ICA) is a source-separation method that decomposes the signals into a set of components that attempt to identify sources of variance in the data (Bell and Sejnowski, 1995). ICA provides a set of weights for all channels such that each component is a weighted sum of activity at all channels, and the weights are designed to isolate sources of brain electrical signals. Components can be judged as containing artefacts based on their time courses, topographies, and frequency spectra. Blink artefacts are the easiest to identify in which they have an anterior distribution and their time course is largely flat with occasional very high-amplitude spikes. Other components could identify line noise or muscle activities (Cohen, 2014).

   ICA is limited by the fact that the identification of the artefact component is inherently a subjective judgement (Nunez et al., 2016). In addition, the components are likely to contain both signal and noise, and thus, one should be cautious about removing a component that seems to contain a signal. In general, it is recommended to remove a component from the data only if it clearly contains artefact and no or very little signal (Cohen, 2014). EEGLAB toolbox provides active development of ICA, and thus it was used in this thesis to perform ICA.

3.2 EMD-based Analysis

EMD-based methods are introduced in the previous chapter (Section 2.3.4) as adaptive time-frequency analysis methods. These methods were chosen to be employed in this thesis due to their advantages in addressing the nonstationarity and nonlinearity properties of the EEG, as discussed in Section 2.3.4.2, and their challenges and gaps in the field of cognitive neuroscience, which are listed in Section 2.3.4.5.

Several extensions of sEMD have been proposed to address the mode-alignment and mode-mixing problems as discussed in Section 2.3.4.3. This section gives descriptions of the two EMD variants which are typically used for this purpose and that were employed in this work.

3.2.1 Multivariate EMD (MEMD)

Rehman and Mandic (2010*b*) introduced MEMD as a generic extension of the sEMD to deal with multidimensional data. While the sEMD finds the local mean using the average of upper and lower envelopes, this value cannot be computed directly for n-dimensional signals. Therefore, the multiple n-dimensional envelopes are generated by projecting the signal in n-variate space. Then, these projections are averaged to obtain the local mean.

The steps of MEMD are as follows:

1. Choose a suitable set of points for sampling on an $(n-1)$ space.
2. Calculate a projection, denoted by $\{P^{\theta_k}(t)\}_{t=1}^{T}$, of the input signal $\{v(t)\}_{t=1}^{T}$ along the direction vector X^{θ_k}, for all k (the whole set of direction vectors), giving $\{P^{\theta_k}(t)\}_{k=1}^{K}$ as the set of projections.
3. Find the time instants $t_j^{\theta_k}$ corresponding to the maxima of the set of projected signals $\{P^{\theta_k}(t)\}_{k=1}^{K}$.
4. Interpolate $[t_j^{\theta_k}, v(t_j^{\theta_k})]$ to get the multivariate envelope curves $\{e^{\theta_k}(t)\}_{k=1}^{K}$
5. For a set of K direction vectors, the mean $m(t)$ of the envelope curves is calculated as $m(t) = \frac{1}{K}\Sigma_{k=1}^{K} e^{\theta_k}(t)$
6. Extract the detail $c_i(t)$ using $c_i(t) = v(t) - m(t)$ (i is an order of IMF). If the detail $c_i(t)$ satisfies the stopping criterion for a multivariate IMF, apply the above procedure to $v(t) - c_i(t)$, otherwise apply it to $c_i(t)$

The sifting process for MEMD stops when all projected signals satisfy any stopping criterion used in sEMD. The Matlab code for MEMD that was used in this thesis is available from (Rehman and Mandic, 2009).

The MEMD algorithm improves the alignment of the corresponding IMFs from different channels and also works as a filter to isolate the inevitable noise in separate components. Thus, after decomposing the signal, the noisy components can be identified and removed (Rehman and Mandic, 2010*b*).

3.2.2 Noise-Assisted MEMD (NA-MEMD)

It has been found that the built-in adaptivity of EMD makes it behave almost as a dyadic (wavelet-like) filter bank when applied to the white Gaussian noise (Flandrin et al., 2004). Particularly, the EMD is capable of separating the white noise into IMF components having mean periods (estimated by counting the number of peaks (local maxima) in terms of the number of data points) nearly twice the value of the previous component, suggesting that the EMD is a dyadic filter (Wu and Huang, 2004). The dyadic filter bank structure was also confirmed when decomposing a multivariate channel of white Gaussian noise using MEMD method (Rehman and Mandic, 2011).

Rehman and Mandic (2011) proposed the NA-MEMD method utilising the dyadic filter bank property of MEMD to alleviate the mode-mixing problem. Thus, extra channels containing multivariate independent white noise are added to the input signal before being decomposed via MEMD. This helps to establish a uniformly distributed reference scale which, in turn, results in corresponding IMFs exhibiting a dyadic filter bank structure. It is important to note that the noise added in this way is never mixed with the original data, as it resides in a different subspace (Rehman and Mandic, 2011). Hence, NA-MEMD eliminates the mode-mixing problem and ensures that the IMFs associated with the original input signals are aligned and have the same information at the same level of decomposition. The details of the NA-MEMD method are as follows:

1. Create an uncorrelated white Gaussian noise time series (q-channel) of the same length as the input signals.
2. Add the noise channels (q-channel) created in step (a) to the input multivariate (n-channel) signal, obtaining an $(n+q)$-channel signal.
3. Process the resulting $(n+q)$-channel multivariate signal using the MEMD algorithm listed above to obtain multivariate IMFs.
4. From the resulting $(n+q)$-variate IMFs, discard the q channels corresponding to the noise, giving a set of n-channel IMFs corresponding to the original signal.

The Matlab code for NA-MEMD that extends the MEMD algorithm of Rehman and Mandic (2009) is available from (Yuexin et al., 2017).

3.3 Complexity Measures

Section 2.4.3 introduced the entropies as measures of signal complexity, quantifying the irregularity of the EEG and helping in the identification of abnormal brain dynamics. Section 2.4.4 discussed the limited applications of entropy measures in exploring the

brain activities of individuals with CP and cognitive impairments following at-birth HIE. Computing entropy features from the EMD-based domain to overcome the limitation of other time-frequency methods also did not gain much attention. Thus, the focus of this thesis is to compute entropy measures from the IMF components, which could be more suitable to reflect brain dynamics and reveal new biomarkers for ASD, CP and cognitive function prediction.

Particularly, sample entropy, permutation entropy and spectral entropy were used in this work. Thus, descriptions of these three measures are given next.

1. **Sample Entropy (SampEn)**

   SampEn was developed by Richman and Moorman (2000) to estimate the irregularity of a time series. It is a modification of approximate entropy (ApEn) (Pincus, 1991), improving its immunity to noise in the data (Richman and Moorman, 2000). Unlike other entropy measures (such as kolmogorov entropy) that require long data sequences to perform a precise estimation of signal complexity, ApEn is less sensitive to the signal length and can be used with short-length data (at least 1000 data points) (Napoli et al., 2020; Sanei and Chambers, 2007). This sensitivity is further improved in SampEn, which can be calculated from a minimum of 100 samples (Napoli et al., 2020).

   SampEn is defined as the probability that the signal of length N repeats itself within the tolerance of r for m points and also repeats itself for the next $m+1$ points (Sharma et al., 2015). For the time series $x(t)$ of length N, construct $N-m+1$ vectors $X_1, X_2, X_3, ..., X_N - m + 1$. Any vector X_m can be expressed as:

$$X_m(t) = \{x(t), x(t+1), \ldots, x(t-m+1)\} \ , \ 1 \leq t \leq N-m+1 \tag{3.1}$$

   where m is the embedding dimension. The SampEn is then given by:

$$SampEn(m, r, N) = -ln[A^m(r)/B^m(r)], \tag{3.2}$$

   where

$$A^m(r) = (N-m)^{-1}\Sigma_{t=1}^{N-m} C_t^{m+1}(r), \tag{3.3}$$

$$B^m(r) = (N-m)^{-1}\Sigma_{t=1}^{N-m} C_t^{m}(r), \tag{3.4}$$

$$C_t^m(r) = (N-m-1)^{-1} C_t \ , \ t = 1, 2, .., N-m. \tag{3.5}$$

   where $B^m(r)$ is the likelihood that $X_m(t_1)$ and $X_m(t_2)$ are matching for m points, while $A^m(r)$ is the likelihood that $X_m(t_1)$ and $X_m(t_2)$ will match for $m+1$ points. $C_t^m(r)$ is the probability of a vector $X_m(t_1)$ being similar to $X_m(t_2)$ within a tolerance r, C_t is the number of instances for which the distance between vectors $X(t_1)$ and $X(t_2)$ is smaller than r (Cui et al., 2015; Richman and Moorman, 2000).

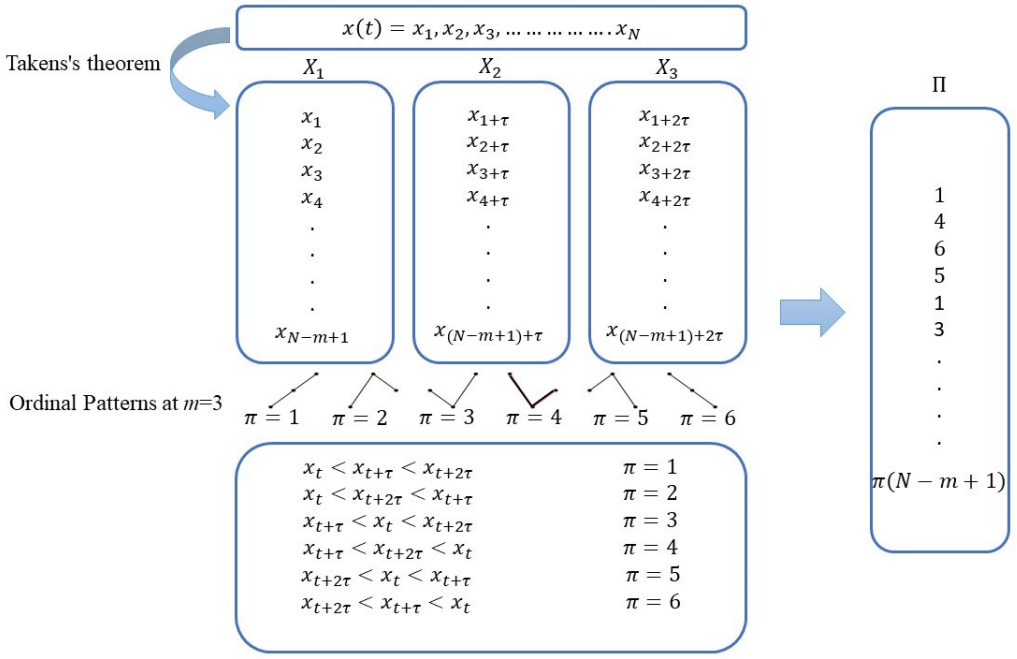


FIGURE 3.1: Ordinal pattern calculations at m=3.

In general, the less predictable or more irregular a time series, the higher its SampEn (Looney et al., 2015). For optimal estimation of SampEn, some studies have recommended the embedding dimension m= 2 or 3, and the tolerance r= 0.1 - 0.25 of the standard deviation of the signal (Bruhn et al., 2000; Pincus, 2001).

2. **Permutation Entropy (PEn)**

Bandt and Pompe (2002) proposed PEn to measure the irregularity of the signal by quantifying the occurrence of ordinal patterns within the time series. The PEn is simple, robust and, similar to SampEn, can be performed on short-length data (100 data points) (Napoli et al., 2020).

PEn measures the signal's complexity by comparing the neighboring values of multidimensional ordinal sequence vectors formed based on the embedding method. Figure 3.1 illustrates this process for time series $x(t)$ of length N and using the embedding dimension $m = 3$. From the set of embedding vectors (X_1, X_2, X_3), a new set Π of order patterns has been calculated by replacing each ordinal sequence with its corresponding permutation.

The probability of each ith permutation, $p(\pi_i)$, can be then obtained by dividing the number of occurrences of π_i in Π by the total number of elements in Π. The permutation entropy is the Shannon entropy of the probability distribution $p(\pi_i)$ given by:

$$PEn = -\sum_{i=1}^{K} p(\pi_i) \log(p(\pi_i)) \tag{3.6}$$

where K is the number of different permutations in Π, equal to $m!$.

The smaller the value of PEn, the less complex the time series is (Bandt and Pompe, 2002). PEn estimation depends on the selected values of embedding dimension m and time delay τ used to reconstruct the sequence vectors. Olofsen et al. (2008) suggested the values of $m = 3$ and $\tau = 1 - 2$ for PEn computation.

3. **Spectral Entropy (SpEn)**

   SpEn is a standard EEG complexity measure that computes the randomness of the signal spectrum. Thus, unlike SampEn and PEn, SpEn estimates the signal irregularity in the frequency domain. To this end, SpEn applies the Shannon entropy concept to the normalised PSD of the signal such that,

$$SpEn = -\sum_{i=1}^{N} p_i \log p_i \tag{3.7}$$

   where p_i is the probability of PSD at each frequency point i, and N is the total frequency points.

   SpEn can also be performed on short-length data, and it is an efficient way to reflect the degree of skewness in the frequency distribution (Cui et al., 2015). A high value of SpEn indicates a flat, uniform spectrum with a broad spectral content, and a low value of SpEn describes a spectrum with all the power condensed into a single frequency point (Cui et al., 2015).

3.4 Phase Syhnchronisation and Graph Theory

Phase synchronisation was discussed in Section 2.5.2 as nonlinear methods used to quantify the phase relations of electrophysiological brain signals. Such methods could reveal the underlying information exchange and relationship strength between each pair of the signal's sources.

Thus, PS has been chosen to be used in this thesis (Chapters 6 and 7) to characterise the functional interaction of the infants who developed CP and to predict their cognitive outcome. It was also employed in Chapter 8 to predict the depression severity of MDD individuals. Particularly, the Phase Lag Index, which is described next, was used to assess the PS between corresponding IMFs of each pair of EEG signals. Moreover, the values of PLI were estimated from the EMD-based domain, which could be more appropriate to explore brain dynamics, as discussed in Section 2.5.3, and therefore may reflect new biomarkers for disorders prediction.

Phase Lag Index (PLI)

PLI proposed by Stam et al. (2007) as a better alternative to the PLV (which was described in Section 2.5.2) to alleviate the VC effect since it provides a consistent and nonzero phase lag between two-time series that are not related to VC. The idea of PLI is based on disregarding phase difference that is centered around $0 \ (\mathrm{mod}\, \pi)$ by defining an asymmetry index for the phase difference distribution (Stam et al., 2007). Thus, PLI is obtained between two signals of length N and phase $\Theta(t_n), n = 1, 2, ...N$ by means of:

$$PLI = | < sign[\Delta\Theta(t_n)] > | \tag{3.8}$$

where "sign" is the signum function, and $-\pi < \Delta\Theta\,(t_n) \leq \pi$. The value of PLI ranges between zero and one, with zero referring to either no coupling between two signals or a coupling with phase difference centered around $0 \ (\mathrm{mod}\, \pi)$, whereas one indicating the two signals are perfectly phase-locked at a value of $\Delta\Theta$ different from $0 \ (\mathrm{mod}\, \pi)$ (Stam et al., 2007).

The PLI script used in this thesis is available from the *HERMES* toolbox (Niso et al., 2013), which provides MATLAB codes for all connectivity methods.

Graph Theory Analysis

Graph theory was used to represent the synchronisation between different brain regions. In the graph theory analysis, the brain is represented as a network where the nodes correspond to distinct EEG electrodes, and the edges represent the functional connections between them identified in this thesis by the PLI index.

The graph network can be characterised using a variety of measures, including nodal and global quantities. The nodal measures give insight into the node's properties in terms of its connectivity with the neighboring nodes, while the global metrics reveal the information flow of the whole network. Rubinov and Sporns (2010) provided a review of different graph theory matrices.

This study is mainly interested in quantifying the nodal properties of the signals to facilitate the comparison with the entropy measures as explored in Chapter 6. In particular, the node degree (ND) measure was utilised to quantify the connectivity strength of each node (channel) with other neighboring nodes. The ND of any channel a is calculated as the averaged PLI of that channel through other channels $b = 1, 2, , k(b \neq a)$, given by:

$$ND_a = \frac{1}{k-1} \sum_{b=1, b\neq a}^{\kappa} PLI_{ab} \tag{3.9}$$

where k represents the total number of channels.

3.5 Statistical Analysis

This thesis is aimed at finding novel biomarkers for autism, CP, and cognitive outcome prediction. Thus, according to Kuhn and Johnson (2013), the biological aspects of the analysis need a certain degree of scientific legitimacy, as well as prospective experimentation to further validate the results found in the data. Statistical analysis methods are designed to yield probabilistic statements about the biomarkers where they provide information on how these biomarkers differ (for example, lower or higher) from the typical groups (Kuhn and Johnson, 2013). Thus, statistical analysis was utilised to assess which features have individual associations with the outcome and provide probabilistic statements about the potential biomarkers. Then, to verify the statistical results and show how useful such biomarkers (if available) could be in practical prediction, they were considered within the classification/regression framework.

Different common statistical methods were used in this work, including one-way analysis of variance (ANOVA), Kruskal-Wallis and Pearson's correlation. Multiple comparison corrections were also used, including Bonferroni correction (Bonferroni, 2008) and Benjamini-Hochberg false discovery rate (BH-FDR) (Benjamini and Hochberg, 1995). Details of each approach are provided in the following chapters.

3.6 Machine Learning

Supervised machine learning approaches have been used to explore how useful the statistically significant features could be in practical prediction. Supervised machine learning takes a known set of labeled data (input) and known responses to the data (output) and trains a model to minimise a difference between the actual output and the target values and generate reasonable predictions on new data. Two supervised learning techniques were used in this thesis, namely classification (used in Chapters 4-6) and regression (employed in Chapters 7 and 8). Classification models predict discrete responses, i.e., classifying the input EEG signals into two classes (having NDDs or not). Regression techniques, in contrast, predict the continuous response (cognitive outcomes).

In general, the classification/regression module consists of training and testing classification/regression algorithm(s). Thus, the dataset is initially divided into two parts, one is saved for testing (the test set), and the rest (the training set) is used to build the models. Several classification and regression algorithms have been developed. Choosing the proper machine learning algorithm is partly based on trial and error. It also depends on the size and type of data under study, the insights one wants to get from the data, and how they will be used. A good machine learning model is one that generalises well from the training data to any data from the problem domain, allowing future

predictions on data the model has never seen (Theodoridis et al., 2010). Support vector machines (SVM), discriminant analysis (DA), k-nearest neighbors (kNN), and random undersampling boost (RUSBoost) are the classification algorithms used in this thesis (particularly in Chapters 4-6), while the ensemble boosted and bagged trees are the regression models used in Chapters 7 and 8. Further details about these machine learning models and the reasons behind choosing them are given in the following chapters.

Overfitting is a common problem in both regression and classification in which a model is too closely aligned to the training data and fails to fit testing data which may affect the accuracy of predicting future observations. Different strategies were used in this thesis to prevent overfitting, and brief descriptions of each of them are given in the following:

- **Cross-validation**

  Cross-validation is a powerful technique to prevent overfitting. Following this method, the data is partitioned into k subsets (folds) in regular k-fold cross-validation. The classification model is then iteratively trained on the k-1 folds and tested on the remaining k-fold (called the holdout fold). The overall error estimated is then averaged to provide an estimation of the model performance on unseen data (Theodoridis et al., 2010). The 10-fold cross-validation (i.e., k=10) is the most common technique used in applied machine learning to evaluate models and is reliable to be used with few samples (Molinaro et al., 2005). Leave-one-out cross-validation (LOOCV) is k-fold cross-validation taken to its logical extreme, with k equal to the number of data samples. LOOCV is also useful to be performed with a limited number of samples (Theodoridis et al., 2010). Both 10-fold and LOOCV were used in this thesis, and further details of their application are provided in the following chapters.

- **Handling imbalanced data problem**

  Data imbalance is common in different domains, including medical diagnosis, leading to the overfitting problem. It is defined as the existence of an over-representation of a given numeric value interval(s) or class(es) over another. The under-represented class or numeric value interval, in many cases, is the most relevant, and its wrong prediction may be costly. Thus, it is the combination of the skewed distribution and user preferences towards under-represented samples that form the basis for imbalanced learning tasks (Moniz and Branco, 2017).

  Standard machine learning algorithms bias the models toward the over-represented samples in imbalanced domains scenarios, leading to performance degradation. Thus, different techniques have been proposed to reduce this issue, including data sampling and ensemble learning (Gonzalez et al., 2020; Seiffert et al., 2010).

Ensemble learning is a technique that integrates two or more different models' predictions. Boosting and bagging are the most common strategies for ensemble learning. The boosting algorithm constructs a robust model from several (at least two) weak models. It builds a series of models from the training data, such that each model seeks to correct the errors of previous models. This process is repeated until the training set is predicted correctly. On the other hand, bagging operates oppositely by teaching several powerful models arranged in a parallel pattern to improve their predictions and then merging them (Gonzalez et al., 2020).

Data sampling balances the class distribution of the training data by randomly removing samples from the majority class (undersampling) or adding samples to the minority class (oversampling). Hybrid data sampling/boosting algorithms are also proposed to improve the performance of models trained on skewed data, including the random undersampling boost classifier (RUSBoost) and the synthetic minority oversampling boost (SMOTEBoost).

RUSBoost and SMOTEBoost outperform other procedures proposed to alleviate the class imbalance problem (Seiffert et al., 2010). Moreover, RUSBoost proved to perform better than SMOTEBoost, where SMOTE is an oversampling method, carrying the drawback of increasing model complexity and training time. SMOTEBoost further increases this drawback since boosting requires the training of several constructed models (Seiffert et al., 2010). RUSBoost was used in Chapters 5 and 6 to handle the imbalanced number of samples in the CP dataset, while ensemble boosted and bagged trees were used in Chapter 7 to control the skewed distribution of the cognitive scores of at-risk infants.

- **Feature selection**

  The number of features should be relatively small with respect to the number of training samples to ensure good generalisation performance of the model and thus, avoid overfitting. Therefore, selecting highly informative features from a larger pool of available ones is crucial to reduce data complexity and optimise the classification performance (Theodoridis et al., 2010).

  Two different feature selection techniques are known in the machine learning literature —Scalar feature selection and feature vector selection (Theodoridis et al., 2010). Scalar feature selection is employed independently of the classifier. The features are ranked in descending order using a score like Fisher's discriminant ratio. The top-ranked features are then selected to check a particular classifier's performance (Theodoridis et al., 2010). On the other hand, feature vector selection identifies the best combination of features based on several search techniques, such as sequential backward search and sequential forward search (Theodoridis et al., 2010).

The scalar feature selection method was used in this thesis due to its computational simplicity while achieving the ultimate goal of getting a reliable classification (Theodoridis et al., 2010). The steps of the scalar feature selection are as follows:

1. Normalising the features to zero mean and unit variance to remove the bias from features having high values.

$$\hat{x}_i = \frac{x_i - \bar{x}}{\sigma}, i = 1, 2, \ldots, N \tag{3.10}$$

where $\hat{x}_i$ is normalised value, N is the number of features, x_i is the feature i, $\bar{x}$ the mean and σ is the standard deviation.

2. Ranking the features in descending order according to the Fisher's discriminant ratio (f), which calculated as:

$$f = \frac{(\mu_1 - \mu_2)^2}{(\sigma_1^2 + \sigma_2^2)} \tag{3.11}$$

where μ_1 is the mean of the first class, μ_2 is the mean of the second class, σ_1^2 and σ_2^2 are the variance of the first and second class, respectively.

3. Computing the cross-correlations among the top-ranked feature (with the index i_1) and each remaining features. The index, i_2, of the second most important feature is computed as

$$i_2 = argmax\{a_1 C_j - a_2 |P_{i_1,j}|\}, j \neq i_1 \tag{3.12}$$

which incorporates the feature ranking value C for the jth feature, and the cross-correlation $(P_{i_1,j})$ between the best feature (i_1) and feature $j \neq i_1$. The parameters a_1, a_2 are user defined weighting factors chosen in this study as recomended by Theodoridis et al. (2010) to be 0.2 and 0.8, respectively.

4. Ranking the rest of the features according to

$$i_k = argmax\{a_1 C_j - \frac{a_2}{k-1} \sum_{r=1}^{k-1} |P_{i_r,j}|\}, j \neq i_r \tag{3.13}$$

for $r = 1, 2, \ldots, k-1$, and $k = 3, 4, \ldots, N$.

Chapter 4

Linear and Nonlinear Analysis of EEG Signals after Facial Stimuli Presentation in Children with Autism Spectrum Disorder

As per the first objective outlined in Chapter 1, the aim is to explore whether the nonlinear characteristics of task-related EEG signals, in particular during the presentation of FE stimuli, could improve the performance of ASD classification. Except for some minor modifications, the material in this chapter is identical to that presented in (Bakheet and Maharatna, 2021). The author's contribution to the work included all given materials, while Maharatna's contribution included supervision.

As pointed out earlier in Section 2.6.1, no previous research considered using the nonlinear dynamics of ERP components to characterise the signals in response to FE stimuli in individuals with ASD. In addition, the application of EMD-based methods in this field to overcome the traditional time-frequency methods limitation has not been investigated to the best of the author's knowledge.

This chapter employs the nonlinear features of ERP components to classify ASD and TD children. The linear features are also considered to reflect the temporal changes of the amplitude after stimuli presentation. Particularly, the SampEn has been employed to detect the EEG complexity in addition to three standard linear features, namely maximum (Max), minimum (Min), and standard deviation (Std).

The proposed features are extracted from the IMF components resulting from the MEMD method to reflect the overlapping time-frequency activity underlying ERP components. Statistical analysis is initially utilised to evaluate the proposed features. The entire set of features and the selected ones (using a feature selection technique) are then used to

train and test DA, SVM and kNN to show how useful they can be in practical ASD classification. Figure 4.1 presents the overall process of the proposed framework.

The quest here can be summarised as seeking to answer the following questions: (1) Would the proposed combination of linear and nonlinear features improve the classification of ASD compared to the previous works performed on a similar dataset? (2) Which IMF component(s) is/are the best to reveal ASD brain deficit, and what is/are its/their underlying frequency? (3) Which FE stimulus-evoked ERP response (i.e., neutral, happy or fearful expression) best discriminates between ASD and TD children?

The remainder of this chapter is structured as follows: Section 4.1 describes the EEG dataset used in this exploration. Section 4.2 presents the feature extraction scheme adopted to extract the linear and nonlinear measures from the IMF components. The statistical analysis method is provided in Section 4.3, while the feature selection scheme is given in Section 4.4. The classification procedure employed in this exploration is described in Section 4.5. The results are then presented in Section 4.6 and discussed in detail in Section 4.7. Section 4.8 concludes the chapter and suggests some future research directions.

4.1 Experimental Data Description

The dataset was taken from previous EEG-based ERP studies (Apicella et al., 2012; Jamal et al., 2014). It contains EEG signals from 24 subjects —12 ASD and 12 TD; age group 6 –13 years (mean age 10.2 and 9.7 years for ASD and TD, respectively). The ASD (particularly "Autistic Disorder") clinical diagnosis was based on the diagnostic and statistical manual of mental disorders, 4th edition, text revision (DSM-IV-TR) criteria, containing specific codes allowing comparisons between the DSM and the ICD manuals.

The EEG signals were acquired during the presentation of three FEs —neutral, happy and fearful. The experiment was done in 4 blocks, and in each block, 10 neutral, 10 happy and 10 fearful faces were presented twice at random order. Signals were recorded at 250 Hz using a 128-channel hydroCel geodesic sensor net, placed according to the arrangement presented in Figure 4.2. The acquired EEG data were segmented into 1000 ms epochs (150 ms baseline and 850 ms post-stimulus presentation) to focus on the time window surrounding the actual event. Epochs with signals over a threshold of 200 μV were considered artefacts and rejected. Data were band-pass filtered with cut-off frequencies from 0.5 Hz to 50 Hz to remove low-frequency drifts and high-frequency measurement noise using a fifth-order forward-backward Butterworth filter, and the baseline was corrected (Apicella et al., 2012; Jamal et al., 2014).

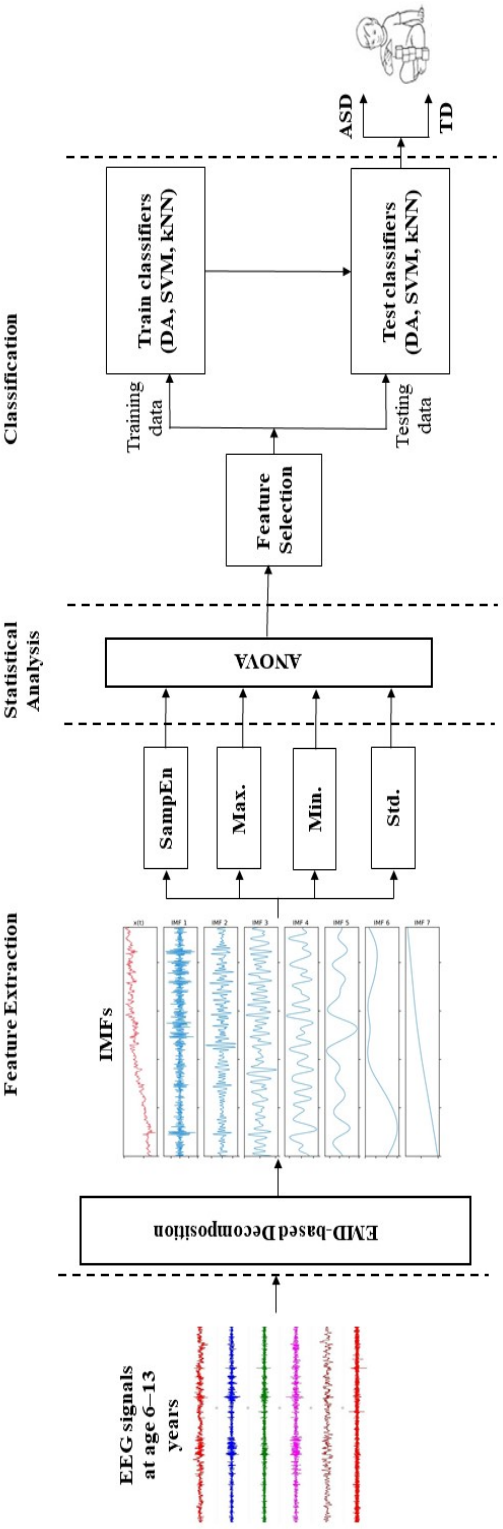


FIGURE 4.1: Block diagram of the proposed method.

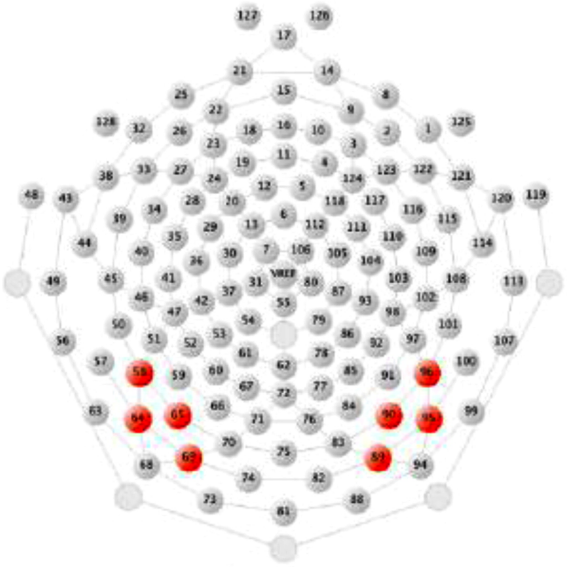


FIGURE 4.2: The 128-channel hydroCel geodesic sensor net arrangement. The right and left ROIs used in the present study are highlighted in red (adapted from (Apicella et al., 2012)).

Eight channels located in the fusiform gyrus area of the cortex were selected for the current study. The fusiform gyrus is believed to be the specific region for processing face features and emotions (Apicella et al., 2012). The right region of interest (ROI) includes electrodes no. 96 (P8/T6), 95 (P10), 90 (PO8) and 89, whereas left ROI includes electrodes no. 58 (T5), 64 (P9), 65 (PO7) and 69. These channels are highlighted in red in the electrode arrangement in Figure 4.2.

4.2 Feature Extraction Procedure

To characterise the multi-subject neural recordings collected from multiple channels over multiple trials, MEMD was used due to its ability to deal with multidimensional data while preserving the alignment of corresponding IMFs. Although NA-MEMD could further improve the time-frequency representation of the signals by alleviating the mode-mixing problem, it increases the computational cost significantly due to the added noise. Thus, this exploration used MEMD and investigated whether the IMF components contained mode-mixing to decide on the most appropriate method.

The proposed analysis method consisted of the following steps:

1. The epochs of each class of stimulus were averaged at the beginning of this analysis to improve the SNR and extract the ERP components.

2. For each class of stimulus, the data points of all subjects from each channel were stacked on top of each other. Hence, eight matrices were constructed (one for

each channel); each of them has the dimensionality of $N_s \times N_t$, where N_s denotes the number of subjects (which is 24) and N_t indicates the number of temporal samples (which is 250). Figure 4.3 illustrates this process. Such a procedure of combining signals from different sources and constructing a multivariate signal for MEMD analysis to acquire aligned IMFs has been adopted previously (Zahra et al., 2017).

3. The MEMD method was then applied to each matrix separately, as shown in Figure 4.3. Following this way, the datasets of all subjects for a specific channel were decomposed into the same number of IMFs.

4. As the number of IMFs may vary among channels, the lowest number of modes is often considered. However, in this exploration, the channels were decomposed into the same number of IMFs, which was eight for each class of stimulus. Figure 4.4 gives an example of the resulting IMFs from channel 1 of the first ASD subject collected during the happy stimulus presentation.

5. The frequencies of each IMF were then calculated using HT and plotted against amplitudes to inspect the signals, and it was found that IMF1 and IMF2 contained mode-mixing and high frequencies (>50 Hz) across all channels. Thus, IMF1 and IMF2 were excluded from further analysis. IMF8 was also ignored as it represents the residue mode, which might give unreal information about the signal (Huang et al., 1998).

   The frequencies of the remaining IMFs were localised approximately in the following ranges: IMF3 (30 –37 Hz), IMF4 (13 –20 Hz), IMF5 (8 –12 Hz), IMF6 (4 –6 Hz), and IMF7 (0.5 –2.5 Hz).According to the traditional ranges of the five physiological frequency bands (Sanei and Chambers, 2007), IMF3 to IMF7 frequencies belong to the gamma-, beta-, alpha-, theta- and delta-band, respectively.

6. SampEn, described in Section 3.3, was then computed over each IMF (IMF3 –IMF7) to depict their complexity in the time domain.

7. Standard linear features (Max, Min and Std) were also extracted from each IMF (IMF3–IMF7) to reflect the temporal changes of the amplitude after the stimuli presentation.

The whole analysis was carried out in the MATLAB software package R2018a.

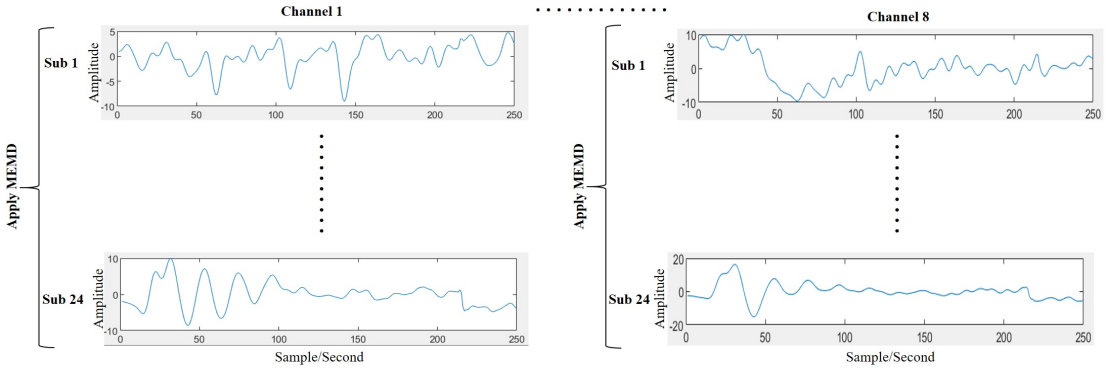


FIGURE 4.3: Simultaneous decomposition of the dataset recorded during the happy stimulus presentation.

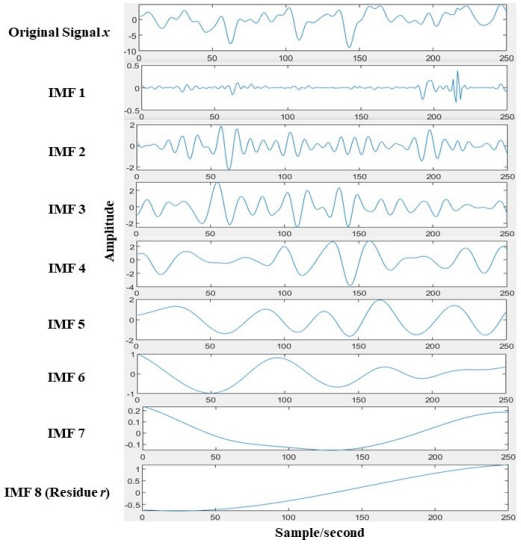


FIGURE 4.4: An example of the resulting IMFs from the MEMD method.

4.3 Statistical Analysis

Statistical analysis test was used to determine the capability of the proposed set of features to discriminate between the two classes (ASD and TD) and to choose the statistically significant ones for classification purposes. To this end, the one-way analysis of variance (ANOVA) was applied using MATLAB's statistics toolbox.

Generally, statistical tests include computing quantities (P-values) translated as statistically significant or non-significant in respect of a threshold known as the level of significance (α). The most common value of α is 0.05, and thus, the null hypothesis (which states that the two classes have the same distribution) could be rejected at $\alpha <= 0.05$.

This value of α needs to be adjusted (lowered) mainly when several independent tests are being performed simultaneously as the chosen value of α may be appropriate for each test, but not for the set of all tests (a problem known as multiple comparisons). Bonferroni correction (Bonferroni, 2008) is the straightforward approach for multiple

testing correction, in which the given α is divided by the total number of running tests to find a corresponding level of significance. In this study, the discriminative capability of each feature (SampEn, Max, Min, and Std) computed from each IMF was evaluated separately, and this process was performed simultaneously for each type of stimuli (60 P-values: 4 features $\times$ 5 scales $\times$ 3 stimuli). Hence, Bonferroni correction was employed, and the value of α was corrected from 0.05 to 0.0008. Thus, the feature has been considered significant if P-value$<= \alpha(= 0.0008)$.

4.4 Features Selection

The proposed method used four features extracted from the IMF signals of each EEG channel. Thus, the feature's number could be large and lead to the overfitting problem discussed in Section 3.6. Therefore, the scalar feature selection method based on Fisher's discriminant ratio, which was described in Section 3.6, was used in this exploration.

4.5 Machine Learning and Classification Process

To assess the proposed features for the practical ASD classification, *Classification Learner App* within the statistics and machine learning toolbox in MATLAB (Mathworks, 2021) was used to train DA, SVM, and kNN classifiers. For DA, two popular discriminant functions were investigated —linear (LDA) and quadratic (QDA). Three kernel functions were utilised for SVM —linear (L-SVM), quadratic (Q-SVM) and cubic (C-SVM). For kNN, a different number of nearest neighbor k was used, and the best accuracy was reported. More details about these classification algorithms are available in Appendix A.

In order to obtain the classifiers' performance, 75% of the total number of samples were randomly selected to be used for training, and the remaining 25% (which were unseen by the models during the training) were used for testing. The 10-fold cross-validation technique was used in the training process to prevent overfitting due to its reliability and lowest chance of introducing an undesired bias with few samples (Molinaro et al., 2005).

Thus, 90% of the training data samples were used for training, and 10% of the hold-out samples were used for validation. This procedure was repeated for each of the ten folds, and the classifier performance was then obtained by averaging the ten independent results. During the validation process, the hyperparameters were tuned using hyperparameter optimisation within the *Classification Learner App*. Different combinations of hyperparameters for a specific model type were investigated, and the model

TABLE 4.1: Details of the training and testing datasets.

Division	Training/validation	Testing
Percentage	75% (90% for training and 10% for validation)	25%
Samples	18 (16 for training and 2 for validation)	6

with the best classification accuracy was selected. Particularly, the *Covariance Structure* hyperparameter was tuned for DA, and *Number of Neighbors* and *Distance Metric* for kNN. More details about these hyperparameters can be found in (Mathworks, 2021). Table 4.1 shows the division of the dataset for the training, validation, and testing process.

The classification performance was evaluated using the conventional measures of accuracy, sensitivity and specificity. The abnormal class is typically called positive (P), and the normal one is known as negative (N). The correct classification of the abnormal condition is known as true positive (TP), and the accurate detection of the typical class is true negative (TN). Likewise, an incorrect classification can be of two types: classifying disordered as typical, which is known as false negative (FN), and classifying typical as abnormal, that is false positive (FP). Sensitivity and specificity are also called true positive rate (TPR) and true negative rate (TNR), respectively. These two measures and the accuracy (ACC) are given as follows:

$$TRP = \frac{TP}{TP + FN} \times 100\% \tag{4.1}$$

$$TNR = \frac{TN}{TN + FP} \times 100\% \tag{4.2}$$

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \tag{4.3}$$

4.6 Results

4.6.1 ANOVA Results

Table 4.2 presents the P-values of the proposed features for the happy, fear, and neutral datasets. The results indicate that SampEn features computed from IMF5 significantly differ between the ASD and TD groups in the happy and neutral datasets. SampEn extracted from IMF6 in the happy stimulus case also gave a low P-value indicating good discriminative capability between the two groups. The results also show that the Min and Std features extracted from IMF7 can significantly differentiate between

TABLE 4.2: P-values of the proposed features for the happy, fear, and neutral datasets. Significant features are indicated in boldface.

Component	Feature	Happy	Fear	Neutral
IMF3	**SampEn**	0.06	0.43	0.82
	Max	0.81	0.84	0.43
	Min	0.44	0.54	0.78
	Std	0.73	0.85	0.57
IMF4	**SampEn**	0.08	0.81	0.58
	Max	0.01	0.32	0.37
	Min	0.06	0.47	0.79
	Std	0.01	0.36	0.34
IMF5	**SampEn**	**0.000007**	0.04	**0.00001**
	Max	0.73	0.57	0.66
	Min	0.57	0.29	0.53
	Std	0.45	0.49	0.91
IMF6	**SampEn**	**0.00002**	0.16	0.84
	Max	0.25	0.70	0.77
	Min	0.18	0.14	0.31
	Std	0.62	0.54	0.76
IMF7	**SampEn**	0.05	0.80	0.48
	Max	0.01	0.05	0.004
	Min	**0.0007**	0.02	0.001
	Std	**0.0005**	0.02	0.002

ASD and TD signals in the happy stimulus case. All features computed from IMF3 and IMF4 gave high P-values indicating low discriminatory capability. Thus, IMF3 and IMF4 were excluded from the classification process. Moreover, the Max feature also achieved a high P-value in all comparisons. Therefore, it was eliminated from the training and testing matrices before feeding them into the classifiers.

4.6.2 SampEn Analysis Results

To show how the EEG complexity differs between ASD and TD children, the mean values of SampEn features extracted from IMF5 (for the happy and neutral stimuli), and IMF6 (for the happy stimulus) were computed for all channels. Table 4.3 presents the results. Compared to the TD group, the ASD group showed lower SampEn in almost all fusiform gyrus channels during FE processing tasks, indicating a complexity reduction of the fusiform gyrus brain activity, specifically in the alpha- and theta-band —the frequency bands corresponding to IMF5 and IMF6, respectively.

TABLE 4.3: Mean values of the channel-based SampEn features computed from IMF5 (for the happy and neutral stimuli) and IMF6 (for the happy stimulus) in each group (ASD and TD).

Channel no.	Group	IMF5		IMF6
		Happy	Neutral	Happy
96 (P8/T6)	**TD**	0.23	0.25	0.17
	ASD	0.18	0.19	0.16
95 (P10)	**TD**	0.23	0.25	0.16
	ASD	0.17	0.19	0.15
90 (PO8)	**TD**	0.25	0.24	0.19
	ASD	0.21	0.20	0.18
89	**TD**	0.23	0.24	0.17
	ASD	0.19	0.18	0.15
58 (T5)	**TD**	0.25	0.21	0.18
	ASD	0.20	0.22	0.15
64 (P9)	**TD**	0.21	0.21	0.17
	ASD	0.18	0.19	0.16
65 (PO7)	**TD**	0.25	0.23	0.22
	ASD	0.25	0.22	0.17
69	**TD**	0.23	0.24	0.19
	ASD	0.18	0.20	0.15

4.6.3 Machine Learning Results

DA, SVM, and kNN classifiers were trained on different dimensions of feature vectors to find the optimal pool of features that can best distinguish between the ASD and TD groups. The dimension of the vectors was as follows: Each feature from IMF5, IMF6 and IMF7 was first evaluated for classification separately. Thus, the feature vectors for each subject were $N_e \times N_f \times N_m$, where N_e is the number of channels, N_f is the number of features, and N_m is the number of IMFs, i.e., $(8 \times 1 \times 1 = 8)$. The classifiers were then trained with the combination of the three features (SampEn, Min and Std) computed from each IMF, and the dimensions of the feature vectors, in this case, were $(8 \times 3 \times 1 = 24)$. Additional investigations were obtained to explore whether combining IMF5, IMF6, and IMF7 could improve the classifiers' performance. In this case, the dimensions of the feature vectors were $(8 \times 1 \times 3 = 24)$ to explore each feature individually and $(8 \times 3 \times 3 = 72)$ for the combination of the three features (i.e., SampEn, Min and Std).

In general, the 10-fold cross-validation results of all classifiers showed enhancements in the classification performance when the classifiers were trained on the combination of the three features extracted from the last three IMFs (i.e., on the feature vector of size 72). Table 4.4 gives the detailed validation results of L-SVM to show an example of how the performance improved by combining all features, especially with the happy

TABLE 4.4: L-SVM performance using different feature vectors for all types of stimuli.

Component/s	Feature/s	Happy			Neutral			Fear		
		ACC	TPR	TNR	ACC	TPR	TNR	ACC	TPR	TNR
IMF5	**SampEn**	72.2%	78%	67%	72.2%	78%	67%	61.1%	78%	44%
	Min	55.6%	67%	44%	38.9%	56%	22%	55.6%	89%	22%
	Std	50%	67%	33%	44.4%	67%	22%	55.6%	78%	33%
	All	77.8%	89%	67%	72.2%	78%	67%	61.1%	78%	44%
IMF6	**SampEn**	77.8%	78%	78%	66.7%	67%	67%	50.0%	67%	33%
	Min	61.1%	44%	78%	72.2%	67%	78%	61.1%	44%	78%
	Std	55.6%	44%	67%	55.6%	44%	67%	61.1%	89%	33%
	All	88.9%	89%	89%	72.2%	67%	78%	38.9%	33%	44%
IMF7	**SampEn**	55.6%	78%	33%	44.4%	67%	22%	55.6%	44%	67%
	Min	77.8%	67%	89%	66.7%	67%	67%	61.1%	44%	78%
	Std	77.8%	56%	100%	72.2%	67%	78%	61.1%	44%	78%
	All	72.2%	67%	78%	66.7%	56%	78%	61.1%	44%	78%
IMF5–7	**SampEn**	88.9%	100%	78%	66.7%	78%	56%	44.4%	56%	33%
	Min	77.8%	67%	56%	61.1%	56%	67%	61.1%	56%	67%
	Std	83.3%	67%	100%	66.7%	56%	78%	55.6%	56%	56%
	All	**94.4%**	**89%**	**100%**	72.2%	67%	78%	61.1%	44%	78%

dataset. All other classifiers had similar improvements, and their results are available in Appendix B.

The classifiers were then trained on different subsets of the 72 features to prevent overfitting. The features on each subset were selected based on the scalar feature selection. In practice, one has to experiment with a different number of selected features and choose the one that results in the best classification performance (Theodoridis et al., 2010). Thus, the discrimination capability of the top 15 high-ranked features was initially explored, where the role of thumb is not to exceed the number of the training samples (which is 18) (Theodoridis et al., 2010). The top 10 and top 5 features were also investigated. Figure 4.5 gives the 10-fold cross-validation accuracies for all three stimuli when the classifiers were trained on the entire feature set and the top selected ones. Generally, it can be inferred that the classifiers' performances were enhanced by reducing the number of features in the happy, neutral, and fear datasets. The best classification accuracies were when the top 10 selected features were used to train almost all the classifiers. Visualisations of the training sets associated with the full and the reduced set of features for all stimuli are also shown in Figure 4.6. As seen, the two groups can often be better separated, especially in the happy dataset, using the reduced set of the highly selected features.

It can also be noticed from Figure 4.5 that the best discrimination between ASD and TD children, for all classifiers, was achieved using the happy dataset. The figure also shows that the classifiers' results for each stimulus type were close to each other, and the QDA achieved the worst performance in general. The best accuracy for the happy dataset was 100% when the top 10 or 15 features were used to train the LDA, kNN and all SVMs. The best accuracy was 88.9% for the neutral dataset when the top 10 or 15

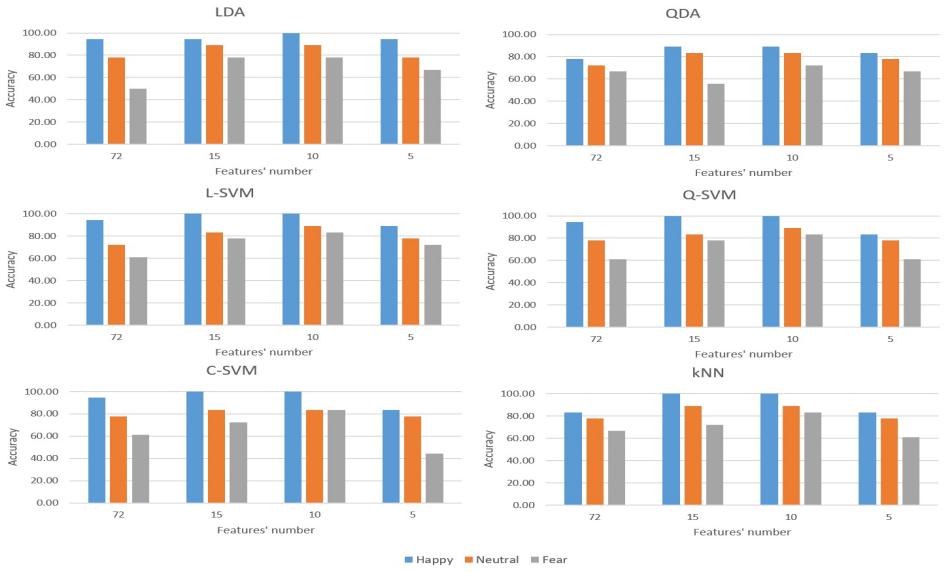


FIGURE 4.5: Accuracy of different classifiers with different groups of features for each stimulus.

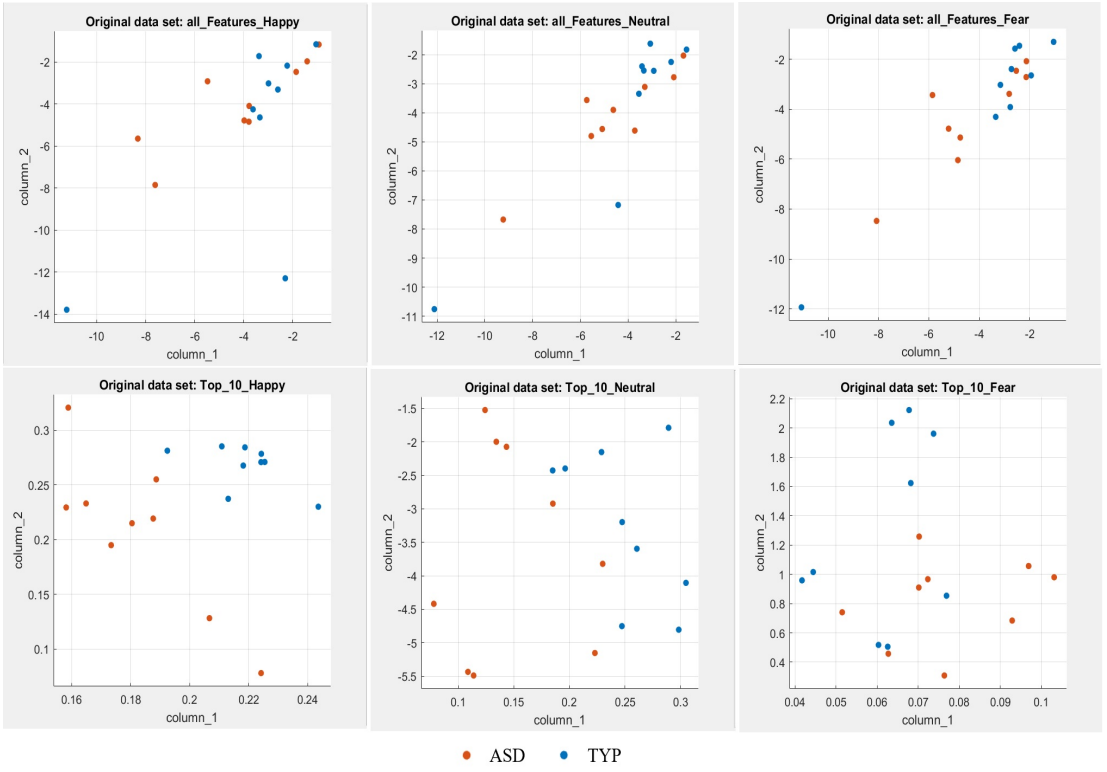


FIGURE 4.6: Scatter plots of the training sets associated with the full and the top high-ranked 10 features for all stimuli. Column_1 and column_2 represent arbitrary first and second features, respectively.

TABLE 4.5: Prediction performance of all classifiers on the happy, neutral, and fear testing data.

Classifier	Happy			Neutral			Fear		
	ACC	TPR	TNR	ACC	TPR	TNR	ACC	TPR	TNR
LDA	**100%**	**100%**	**100%**	33.3%	0%	66.7%	33.3%	33.3%	33.3%
QDA	66.7%	100%	33.3%	16.7%	0%	33.3%	50.0%	66.7%	33.3%
L-SVM	83.3%	66.7%	100%	16.7%	0%	33.3%	33.3%	33.3%	33.3%
Q-SVM	**100%**	**100%**	**100%**	16.7%	0%	33.3%	50.0%	100%	0%
C-SVM	83.3%	100%	66.7%	16.7%	0%	33.3%	33.3%	66.7%	0%
kNN	83.3%	100%	66.7%	33.3%	0%	66.7%	33.3%	33.3%	33.3%

FIGURE 4.7: Confusion matrix of the best classification results for each stimulus. Here class 1 depicts ASD, and class 2 represents TD.

features were used to train the LDA, kNN, L-SVM and Q-SVM. The best accuracy of the fear dataset was 83.3% when the top 10 features were used to train the kNN and all SVMs.

The trained models with the top 10 selected features were then used to predict the classes of the test datasets. This set of features was employed due to its best and consistent validation performance for all classifiers with the three stimuli datasets. Table 4.5 gives the testing results. Again the happy dataset reached the best classifiers' performance of 100% ACC, 100% TPR, and 100% TNR using the LDA and Q-SVM. Followed by the fear dataset with 50% ACC, 100% TPR, and 0% TNR using Q-SVM and 50% ACC, 66.7% TPR, and 33.3% TNR using QDA. The best performance for the neutral dataset was 33.3% ACC, 0% TPR, and 66.7% TNR using LDA and kNN. The confusion matrix of the best testing performance for each dataset is shown in Figure 4.7.

4.7 Discussion

This study addresses the abnormal brain function in FE processing in children with ASD over the fusiform gyrus brain region. It investigates the intrinsic EEG complexity (as indexed by SampEn measures) and the ERP alterations (through the standard linear features). Statistical results show significant differences between ASD and TD groups, especially in processing the happy stimulus through the SampEn features computed from IMF5 and IMF6. IMF5 and IMF6 have been found (through the HT analysis) to contain the alpha- and theta-band, respectively.

Mainly, the results show a decrease in the EEG complexity in the ASD group compared to the TD group. This result is in line with (Catarino et al., 2011) study, which demonstrated a reduction of EEG complexity in the ASD group in response to a visual matching task. The impairments in the alpha-band measurements of ASD signals in response to FE stimuli had also been reported previously. Khan et al. (2013) measured functional brain connectivity of MEG signals in the Fusiform Gyrus region while subjects were doing emotional recognition tasks. They found that children with ASD demonstrated specific alpha abnormalities compared to TD.

Statistical results also show significant differences between the two groups in processing the happy stimulus through the linear measurements computed from IMF7, which has been found (through the HT analysis) to include the delta-band frequencies. This finding indicates that the features extracted from the low-frequency component can significantly account for the linear ERP variations between the two groups. The role of the delta-band in recognising FE had been reported in different studies kinds of (Balconi and Lucchiari, 2006; Balconi and Pozzoli, 2009). In particular, Garcia Dominguez et al. (2013) found that those with ASD had increased coherence in the lowest frequency bands, delta and theta, over occipital channels, including the Fusiform Gyrus region, as viewing emotional faces compared to controls.

Different classifiers were trained and tested with various feature vectors constructed from IMF5-IMF7 to investigate their practical capability in ASD classification. Almost all validation results suggest that combining the nonlinear measurements (as indexed by SampEn) and the linear features (Min and Std) and extracting them from the combination of the last three IMFs (to include alpha-, theta- and delta-band) can improve the classifiers' performance. Moreover, reducing the number of features using the scalar feature selection further enhanced the results.

In general, the results of all employed classifiers are comparable in each case of stimuli during both the validation and testing procedures. The results show that the signals recorded during the happy stimulus presentation achieved the best classification performances. This finding suggests general impairment in the positive emotions recognition (as indexed by happiness) in children with ASD. This result is consistent with

(Yeung et al., 2014) finding, but it contradicts other studies that only found a deficit in recognition of negative emotions (such as fearful faces) (Corden et al., 2008; Yeung et al., 2019). This inconsistency might be due to the differences in the task demands that widely vary across studies (Yeung et al., 2014). Nevertheless, the testing results of the fear dataset indicate better performance than the neutral stimulus results. Particularly, the TPR was better in all classifiers. A good TPR is often important in a diagnostic test where children identified as having ASD should be highly likely to have the condition.

The results and methodology of this study have been compared to three previously published works that used the same dataset employed herein. Apicella et al. (2012) assessed the mean values of the ERP latencies and amplitudes using the ANOVA repeated measures of the three stimuli. They reported reduced amplitudes and delayed latencies of all the early ERPs in ASD children regardless of the FE stimuli. Thus, they suggested that ASD children similarly process all three faces compared to TD ones. However, they neither studied the frequency alteration nor nonlinear activity underlying ERP components. Jamal et al. (2014) reached 94.7% ACC in classifying the ASD and TD groups as the best performance with 85.7% TPR and 100% TNR using PS features and the SVM. However, this result was based on the combination of the datasets from the three stimuli. Thus, they did not explore the effect of each FE (happy vs. fear vs. neutral). They also did not discuss the oscillations of interest associated with their findings. Khuntia et al. (2019) employed classwise principle component analysis in both time and time-frequency domains (via STFT). Their exploration reached 81% in the time domain analysis and 84% in the time-frequency domain using alpha and beta oscillations. Nevertheless, they included the tree stimulus and hence their exploration was not mainly based on recognising the emotional faces. Moreover, they did not consider the nonlinear behavior of EEG signals.

This exploration has also been compared with two relevant EEG-based ERP studies that adopted another children dataset for ASD diagnosis. Yeung et al. (2014) used happy, sad, anger, disgust, fear, surprise, and neutral stimuli, and their samples included 18 children with ASD and 18 TD ranging in age from 9 –10 years old. They used ANOVA repeated measures to assess functional connectivity, as indexed by theta coherence, during FE recognition tasks. They found that the children with ASD exhibited abnormal patterns in the theta coherence associated with their FE recognition ability. Their results also suggested general impairment in recognition of positive emotions. Despite their significant findings, they only studied theta wave's connectivity and did not consider the nonlinear activity underlying ERP components.

Garcia Dominguez et al. (2013) used fearful and happy stimuli, and their dataset consists of 31 TD and 72 children with ASD (age ranges between 2 –4 years and 11 months). They trained and tested L-SVM with high dimensional feature vectors of the imaginary part of coherency of all frequency bands, and classification accuracy of 80% was reached. They also carried out ANOVA analysis and found that children with ASD

TABLE 4.6: Comparison of state-of-the-art methods employed on the FE-based ERP studies for ASD diagnosis.

Authors	Features	Dataset	Evaluation methods	Results	Frequency analysis	Emotional expression deficit
(Apicella et al., 2012)	Peak latency and amplitude of ERP components	Dataset used in this study	ANOVA	General delayed latencies and reduced amplitudes of all ERPs in ASD children regardless of the face expressions	Did not study	None
(Jamal et al., 2014)	Complex network parameters to measure brain connectivity	Dataset used in this study	SVM	94.7% ACC, 85.7% TPR, 100% TNR using the combination of the features of all the three stimuli	Did not mention	Did not study
(Khuntia et al., 2019)	Multivariate pattern of the time points	Dataset used in this study	CPCA	Around 81% in the time domain and 84% in the frequency domain analysis	Alpha and beta bands are best to identify ASD (spectrally processed using STFT)	None
(Yeung et al., 2014)	Functional connectivity (theta coherence)	TD (n = 18) and ASD (n = 18)	ANOVA	ASD children exhibited abnormal patterns when recognising different FEs	Theta coherence can identify ASD (they just explored theta wave using FFT)	Impairment in recognition of the happy FE in ASD
(Garcia Dominguez et al., 2013)	Functional connectivity (imaginary part of coherency)	TD (n = 31) and ASD (n = 72	ANOVA	ASD children exhibited enhanced synchronisation during FE recognition	Lower frequency bands are best to identify ASD (spectral analysis method did not mention)	Did not mention
			L-SVM	80% ACC was reached using feature vectors of the imaginary part of coherency of all frequency bands		
Current study	SampEn, Min, Std in the MEMD domain	TD (n = 12) and ASD (n = 12)	ANOVA	Significant differences were found for the happy stimulus using the features extracted from IMF5, IMF6 and IMF7	The combination of alpha, theta, and delta bands were found to best identify ASD (spectrally processed using MEMD and HT).	Impairment in recognition of the happy FE in ASD
			DA, SVM, and kNN	100% ACC, 100% TPR, 100% TNR was reached for the happy stimulus by combining all features and extracting them from IMF5-IMF7 combination		

exhibited enhanced synchronisation during post-stimulus, particularly at lower frequency bands. However, they did not report the differences in processing the happy and fearful expressions or consider the nonlinearity of the signals. Table 4.6 summarises all these studies.

From the above studies and the systematic review of (Monteiro et al., 2017), which reviewed other 14 articles, it can be argued that this exploration's novelty is mainly represented by including linear and nonlinear characteristics of ERPs. Almost none of the reviewed studies did consider the nonlinearity of the EEG signals when analysing the ERP components. Additionally, even though few articles studied the overlapping time-frequency activity in response to the recognition of emotional facial expression tasks (Garcia Dominguez et al., 2013; Khuntia et al., 2019; Yeung et al., 2014), they often spectrally analysed the signals using time-frequency methods that rely on predefined traditional brain waves. This limitation has been addressed in the proposed approach using the MEMD method, which decomposes the signals adaptively. Hence, all potentially meaningful brain dynamics are included in the analysis.

However, the advantages of the MEMD method have a price of being computationally complex, in which projection and interpolation operations, described in Section 3.2.1, require finding an interpolation surface on high dimensional data space during each iteration to identify envelopes. Such computational load could be unsuitable for more complex scenarios (Zhang et al., 2021).

Overall, the proposed framework classified ASD with superior accuracy to that of existing studies. Particularly, a high classification accuracy of 100% for the happy stimulus dataset was achieved. Nevertheless, this result cannot be compared likewise with Jamal et al. (2014)'s classification performance as they analysed the combination of the signals from different stimuli.

Therefore, the proposed framework could be instrumental in solving the late diagnosis problem of ASD. However, more work needs to be carried out with a larger number of samples to eliminate the possible effects of misclassification and establish the method's practical validity before putting it into clinical practice.

4.8 Conclusion

The proposed framework successfully classified children with ASD using their EEG signals collected over the Fusiform Gyrus brain region during FE recognition tasks. Linear and nonlinear measurements of IMFs were assessed using statistical analysis and a machine learning framework. High prediction accuracy of 100% for the happy stimulus dataset was achieved. Thus, the study reveals a general impairment in recognising the positive FE in children with autism. Moreover, EEG complexity, as indexed

by SampEn measure of the alpha and theta brain waves, and the linear features extracted from the delta-band components can be considered biomarkers of impaired FE processing in ASD children, which in turn can aid their prediction.

The next chapter investigates deploying the framework proposed here to reveal early abnormal brain dynamics in at-risk infants and predict their neurological outcomes.

Chapter 5

Prediction of Cerebral Palsy in Newborns with Hypoxic-Ischemic Encephalopathy Using Nonlinear Analysis of EEG signals

The previous chapter developed a framework that successfully classified children with ASD using the EEG signals collected during FE recognition tasks. This chapter aims to apply a similar method with the resting-state EEG signals recorded from infants with HIE to predict their CP outcomes as assessed at 24 months.

Except for some minor modifications, this chapter is based on the material given in (Bakheet et al., 2021). The author's contribution to the work included all materials presented in this chapter, while the materials and results related to the global graph-theoretic parameters are not included here as they belong to Alotaibi's contribution. Maharatna's contribution was supervision, Konn extracted the infants' EEG signals, and Vollmer provided the neurological evaluation information.

In addition to using the nonlinear characteristics of EEG, the previous chapter employed linear features (namely Max, Min, Std) to reflect the temporal changes of the amplitude after stimuli presentation, which was used to reveal the impaired FE processing in ASD children. Such temporal measurements are inappropriate to be used with the resting-state brain activity due to the well-known inter-trial/inter-subject amplitude variability. This variability could be modulated by task events (such as the presentation of FE in the previous chapter) in which activities become time-locked to the trial events. Hence, only nonlinear features of the EEG are explored in terms of their ability to discriminate between infants who developed CP and those with the normal

neuromotor outcome. For simplicity, "CP group" and "normal group" are used to refer to these two classes, respectively.

Section 2.6.2 discussed the use of EEG in CP prediction and highlighted the gaps in this field of study. It was shown that there is a lack of research that used a qEEG measure that considers the nonlinearity and nonstationary characteristics of EEG. Furthermore, all previous studies have been carried out in adolescents and children with CP, and the detailed exploration of using qEEG to identify CP at infancy was rarely investigated.

Therefore, this chapter investigates the effectiveness of nonlinear characteristics of EEG in a machine learning framework for the early prediction of CP. Specifically, SampEn, PEn and SpEn were used in this exploration to characterise infants' resting-state EEG. These features were extracted from the IMF components resulting from applying the EMD-based method to investigate the overlapping time-frequency brain dynamics.

Statistical analysis was initially utilised to evaluate the ability of the entropy measures to discriminate between the CP and normal groups. The significant features were used to train and test the RUSBoost classifier to show how useful they could be in practical CP prediction. The proposed methodology is depicted in Figure 5.1.

This chapter raises three research questions: (1) Can the intrinsic nonlinear features of at birth EEG discriminate between the CP and the normal groups as clinically assessed at 24 months? (2) Which IMF component(s) best predict the CP outcome, and what are their underlying frequencies? (3) Can combining the three entropy features (SampEn, PEn and SpEn) improve the classification performance?

The rest of the chapter is structured as follows: Section 5.1 describes the EEG dataset used in this exploration. Section 5.2 outlines the preprocessing methods used to generate artefact-free EEG. Section 5.3 describes the feature extraction scheme adopted to extract the nonlinear measures from the IMF components. The statistical analysis method is provided in Section 5.4, while the feature selection scheme is given in Section 5.5. The classification procedure and the reason behind choosing the RUSBoost classifier are described in Section 5.6. The results are then presented in Section 5.7 and discussed in detail in Section 5.8. Section 5.9 concludes the chapter and suggests some future research directions.

5.1 Experimental Data Description

EEG signals used in this study were collected from 30 neonates born at term equivalent age (38 –42 weeks of gestation) with HIE treated with hypothermia at UHS between February 2017 and August 2017. The infants were followed up (under the clinical follow-up program at the UHS) with a neurological examination at age 24 months by a

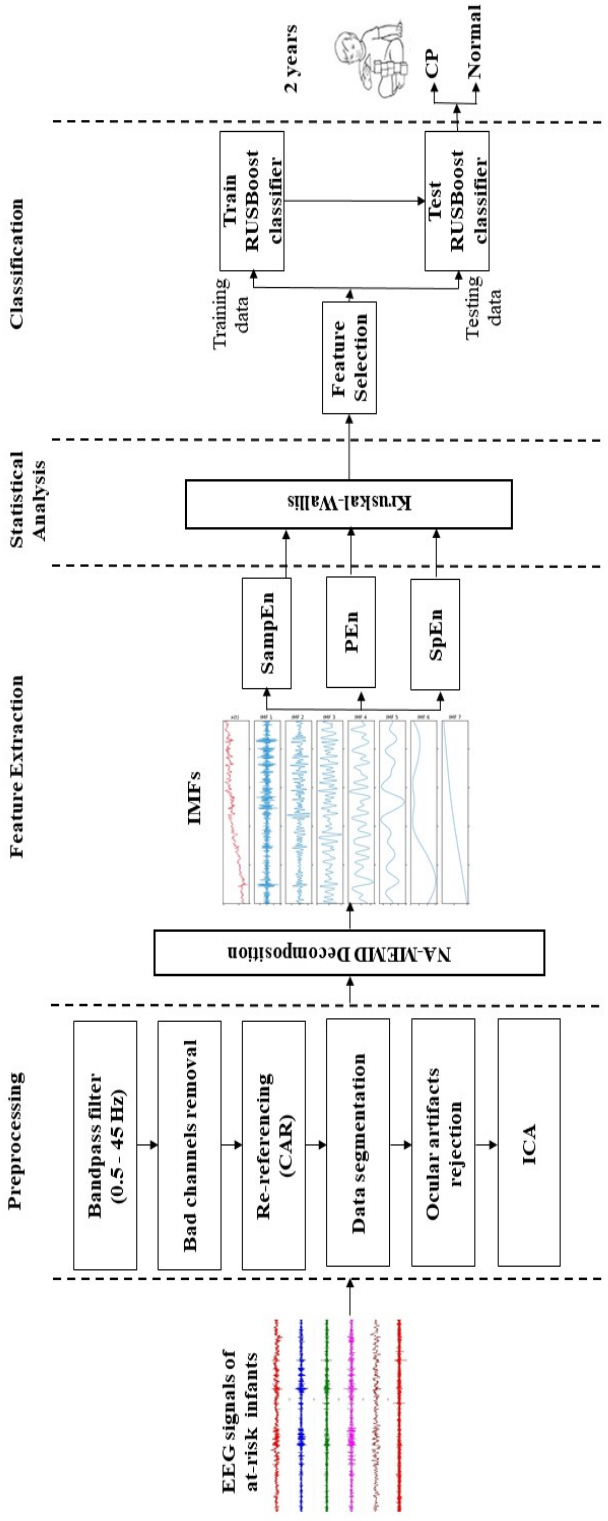


FIGURE 5.1: Block diagram of the proposed methodology for classifying the CP and the normal groups using the complexity features.

pediatric neurologist. Out of the 30 infants, 26 completed the follow-up assessment and were diagnosed with CP or normal neuromotor outcome. A diagnosis of CP was made according to the criteria of the surveillance of CP in Europe working group (Cans, 2000) in which 20 infants had normal neurology, while six developed a form of CP. Table 5.1 provides the neurological diagnosis of each infants.

The EEGs were recorded from the infants on the neonatal intensive care unit within the first seven days after birth, during a period of restfulness with eyes closed for at least 20 minutes. Nineteen surface electrodes (C3, C4, CZ, F3, F4, F7, F8, FZ, FP1, FP2, O1, O2, P3, P4, PZ, T3, T4, T5 and T6) were used according to the international 10-20 electrode placement, as presented in Figure 2.2 (a) and (b). Recordings were done by either a Nihon Kohden (sampling frequency 512 Hz, high-pass filter 0.08 Hz, low-pass filter 300 Hz) or XLTEK (sampling frequency 512 Hz, high-pass filter 0.1 Hz, low-pass filter 400 Hz) clinical video-EEG system. A consultant neurophysiologist examined all of the recorded EEGs visually and extracted the first continuous clip that was long enough without any clear significant artifact (the average length of the clips is approximately two minutes). Secondary analysis of anonymised, routinely collected clinical data was approved by the HRA and Health and Care Research Wales, HCRW (Reference ID 20/HRA/0260; IRAS project ID 278072 University Hospital Southampton R&D protocol number RHM CHI1047).

5.2 Data Preprocessing

The continuous resting-state EEG signals were preprocessed using EEGLAB, an open-source toolbox in MATLAB, to remove the remaining artifacts such as eye movement, muscle, heart activities and line noise. The block diagram of the preprocessing steps is depicted in Figure 5.1.

The signals were initially filtered using the FIR bandpass filter with 0.5 Hz and 45 Hz cut-off frequencies. Then, the bad channels were identified to make the data amenable for the analysis. EEGLAB automatically picks the bad channels based on two criteria: first, the flat channels, and second, the channels with a large amount of noise determined based on their standard deviation. Subsequently, seven bad channels were removed from each subject and not included in further analysis in any subject in the dataset. The remaining 12 channels were: C3, F3, F7, Fz, O1, O2, P3, P4, T3, T4, T5, and T6.

The signals were then re-referenced by a common averaged reference to reduce the confounding effects of the reference. After that, the continuous EEGs were segmented into epochs of 2s length according to the common approach followed for EEG resting-state

TABLE 5.1: the neurological outcomes evaluated at 24 months of the 30 neonates born with HIE.

Subject #	Neurology at age 24 months
subj_1	normal
subj_2	normal
subj_3	normal
subj_4	CP
subj_5	normal
subj_6	CP
subj_7	CP
subj_8	normal
subj_9	normal
subj_10	normal
subj_11	normal
subj_12	unspecificic signs
subj_13	normal
subj_14	normal
subj_15	normal
subj_16	normal
subj_17	normal
subj_18	normal
subj_19	incomplete follow-up
subj_20	normal
subj_21	normal
subj_22	unspecific signs
subj_23	CP
subj_24	normal
subj_25	normal
subj_26	CP
subj_27	CP
subj_28	unspecific signs
subj_29	normal
subj_30	normal

analysis (Kulak et al., 2005). The 2s epoch contains 1024 samples, and thus it is sufficient to compute the complexity measures used in this thesis (which can be calculated from a minimum of 100 samples (Napoli et al., 2020)) as discussed in Section 3.3.

Ocular artefact, particularly eye movement, was automatically detected through the EEGLAB toolbox by setting the threshold value equal to 55 μV because such artifacts are defined as having an amplitude greater than this value (Apicella et al., 2012). Thus, each epoch containing values above this threshold was marked as a bad epoch. After this step, the remaining epochs were visually inspected to determine whether they were contaminated by high frequency or line noise. The corrupted epochs were rejected and excluded from further analysis.

ICA was then applied, using the runICA algorithm implemented in EEGLAB, to remove the remaining artefacts from the signals, including eye blinks and muscle artefacts. Thus, the EEG signals from the 12 channels were separated into their 12 constituent independent components, as the general rule of ICA is to find the N independent components from the N linearly mixed-signal (input channel data). These independent components were then projected back to the EEGs using the estimated separating matrix after manually eliminating the artefact-related independent components

(Sanei and Chambers, 2007). Finally, a total of 12 channels, each with 30 artefact-free of 2s epochs per subject, were used in the next stage of the analysis.

5.3 Extracting Complexity Measures

The MEMD method was initially utilised to decompose the multidimensional EEG signals. The frequencies of each IMF were then identified using HT and plotted against amplitudes. Due to the mode-mixing found on different IMFs, the NA-MEMD was alternatively used to simultaneously deal with the mode-alignment and mode-mixing problems.

The proposed NA-MEMD-based analysis to characterise the multi-subject neural recordings collected over multi-channel and multi-epochs included the following steps:

1. For each channel, the data points from all infants were combined to obtain a multivariate signal. Hence, 12 different matrices were constructed (i.e., one matrix for each channel); each of them has the dimensionality of $N_s \times N_t \times N_e$, where N_s denotes the number of subjects (which is 26), N_t indicates the number of temporal samples (which is 1024), and N_e is the number of epochs of each subject (which is 30).

2. Each of the 12 matrices was reshaped into a two-dimensional time series of the dimension $[N_s \times N_e] \times N_t$ before decomposing it by the NA-MEMD algorithm. Following this step, each epoch was treated separately as an input variate to the NA-MEMD algorithm, as illustrated in Figure 5.2, and thus, the alignment of all IMFs was confirmed not only across infants but also across epochs. A similar process has been adopted previously by Hu and Liang (2011).

3. After the decomposition, different numbers of IMFs were produced from different electrodes. The EEG channels that yielded the lowest number of IMFs upon the decomposition gave 10 modes. Thus, to unify the number of IMFs among channels, the first 10 IMFs of each channel were considered. This way of alignment was performed because high-order IMFs (for example, IMF11) either contain very low energy or represent residual trends. In consideration of the uncertainty of longer trends, these final components were left out. Such alignment was performed in the literature by different studies, kinds of (Hassan and Bhuiyan, 2017). Figure 5.3 shows the extracted IMFs of a sample epoch from a channel that gave 10 IMFs.

4. The frequencies of each IMF were then acquired by HT and plotted against amplitudes to inspect the signals. Consequently, IMF1 and IMF2 were found to contain high frequencies (>50 Hz) across all channels. According to Huang et al.

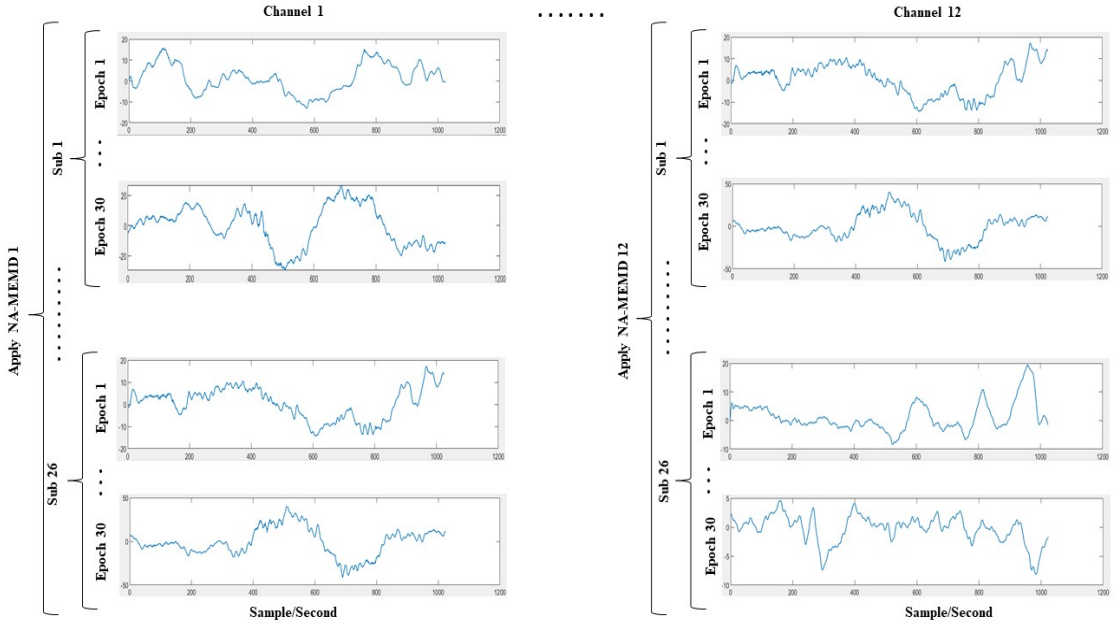


FIGURE 5.2: Simultaneous decomposition of the EEG signals.

(1998), such high-frequency components generally make the data more complicated and are probably not physical. Therefore, these modes were excluded from further analysis. IMF10 was also ignored as it represented the residue mode of some EEG channels, which might give unreal information about the signal. The scales of the remaining IMFs were localised approximately around the following ranges: IMF3 (30 –35 Hz), IMF4 (20 –25 Hz), IMF5 (10 –13 Hz), IMF6 (5 –8 Hz), IMF7 (3 –4 Hz), IMF8 (2 –3 Hz), and IMF9 (0.5 –2 Hz). Referring to the traditional brain waves (Sanei and Chambers, 2007), IMF3 to IMF6 frequencies belong to the gamma-, beta-, alpha- and theta-band, respectively, while IMF7, IMF8 and IMF9 all belong to the delta brain wave.

5. After IMF selection, the dimension of the dataset of each subject was: $N_c \times N_i \times N_e \times N_t$, where N_c is the number of channels which is 12, N_i is the selected number of IMFs, which is 7 (IMF3 –IMF9), N_e is the number of epochs which is 30, and N_t is the number of the samples which is 1024.

6. For each channel and each IMF, the proposed entropy measures, described in Section 3.3, were computed for each epoch. The calculated features were then averaged among the epochs to obtain one SampEn, one PEn, and one SpEn for each IMF signal. Thus, for each subject and each IMF signal, the final number of features was 36 (3 features $\times$ 12 channels). These features were then used to train and test the RUSBoost classifier.

The whole analysis was carried out in the MATLAB software package R2018a.

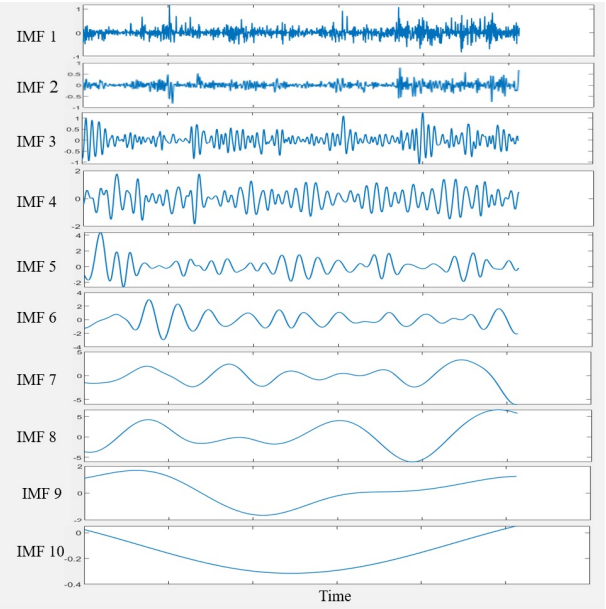


FIGURE 5.3: An example of a set of IMFs resulting from the NA-MEMD method.

5.4 Statistical Analysis

Kruskal-Wallis test (Theodorsson-Norheim, 1986) was adopted in this exploration to verify whether the discriminatory capability of the entropy features was statistically significant between the two classes (CP and normal). Kruskal-Wallis was chosen here due to the difference in sample sizes between the two groups, which suggests that ANOVA might not be the best choice to analyse these data. Therefore, It has been recommended that Kruskal-Wallis is better for such analysis (Hoffman, 2019).

Thus, each entropy feature computed from each IMF was evaluated using the Kruskal-Wallis test (21 P-values: 3 features $\times$ 7 scales). Bonferroni correction was initially employed to control these multiple comparisons, resulting in none of the null hypotheses being rejected. Nevertheless, Bonferroni is considered inappropriate when many hypotheses are simultaneously tested because the P-value will be reduced so severely that even discriminant features are unlikely to be identified as statistically significant (Chen et al., 2017; Cohen, 2014). Benjamini-Hochberg false discovery rate (BH-FDR) proposed by Benjamini and Hochberg (1995) as an alternative to the Bonferroni correction, improving the power of statistical inference with few type I errors (i.e., incorrectly rejecting the null hypothesis) instead of guarding against making any type I conclusion at all (Chen et al., 2017). Thus, the BH-FDR was used in this exploration.

BH-FDR works by controlling for the probability of type I errors within a distribution of P-values obtained via a statistical test (such as Kruskal-Wallis and ANOVA). Let $q = 0.05$ be the pre-specified upper bound of BH-FDR and m is the total number of comparisons. The first step is to compute index k :

$$k = \max\left\{i : p_{(i)} \leq \frac{i}{m}q\right\} \tag{5.1}$$

If k does not exist, reject no hypothesis, otherwise reject hypothesis of $H_i(i = 1, ..., k)$. BH-FDR method starts with comparing H_i from the largest to smallest P-value $(i = m, ..., 1)$ (Benjamini and Hochberg, 1995; Chen et al., 2017). Accordingly, the value of α was corrected from 0.05 to 0.0019. Thus, a feature was considered statistically significant if its P-value$<= \alpha(= 0.0019)$.

The BH-FDR script used is available from (Gerber, 2022) and Kruskal-Wallis test was carried out using the MATLAB statistics toolbox.

5.5 Features Selection

This exploration adopted three entropy features extracted from the multivariate IMF signal of each EEG channel. Hence, the feature's number might be greater than the number of samples (the 26 subjects) and thus might lead to the overfitting problem. Therefore, the scalar feature selection method based on Fisher's discriminant ratio was employed following the procedure described in Section 3.6.

5.6 Machine Learning and Classification Process

The choice of the classifier was guided by the class imbalance in the sample set, as described in Section 5.1 —20 neonates developed the normal neuromotor outcome and 6 developed CP. Hence, the RUSBoost classifier was adopted to discriminate between the two groups due to its efficiency in handling the class imbalance problem, as discussed in Section 3.6. The classifier training was performed using the *Classification Learner App* within the statistics and machine learning toolbox in MATLAB. More details about the RUSBoost algorithm are available in Appendix A.

While the RUSBoost classifier was employed to reduce the effect of imbalanced data, cross-validation was used to mitigate the limited number of samples problem. Particularly, the LOOCV method discussed in Section 3.6 was adopted with k equal to the number of subjects. LOOCV is useful when the same samples (subjects) need to be used for both training and testing due to their limited availability (Theodoridis et al., 2010).

Generally, LOOCV works as follows: given N instances (subjects here), it uses the feature vectors of $N - 1$ samples for training and the vector of the remaining sample for testing. This procedure is repeated N times, leaving out the data of a different instance

each time. Finally, the classifier performance is obtained by averaging the N independent results. Thus, the same dataset is utilised for training and testing, and, at the same time, the testing is carried out on samples that have not been seen in the training (Theodoridis et al., 2010).

The classifier's performance was evaluated using ACC, TPR, and TNR measures described in Section 4.5. In addition, the area under the curve (AUC) that quantifies the balance between sensitivity and specificity has been used. The AUC measures the area under the receiver operating characteristic (ROC) curve —a probability curve that plots TPR against the TNR at different classification thresholds. Generally, the value of AUC ranges between 0 to 1. AUC between 0.90 and 1 is considered excellent and good for values between 0.80 and 0.90. Acceptable when the value in the range 0.70-0.79, poor between 0.60 and 0.69, and has no discrimination for values between 0.50 and 0.59 (Huang et al., 2013).

5.7 Results

5.7.1 Kruskal-Wallis Results

Figure 5.4 presents the results of the Kruskal-Wallis test for SampEn, PEn, and SpEn, respectively, when computed from each IMF for all electrodes to discriminate between the CP and normal groups. Different embedding dimensions m and tolerances r were explored for SampEn estimation as suggested by (Pincus, 2001; Bruhn et al., 2000). The PEn was also calculated using time delay τ=1 and 2, and embedding dimension m=3 as recommended by Olofsen et al. (2008).

It can be noticed from the figure that the lowest P-value was obtained when all entropy features were calculated from IMF5, indicating good discriminatory capability between the CP and normal groups. The figure also shows the robustness of SampEn and PEn measures with the small changes of the embedding parameters. The P-values of SampEn, PEn, and SpEn estimated from IMF5 were significant and equal to 0.0019, 0.0013, and 0.0004, respectively.

To explore how the entropies calculated from IMF5 differ between CP and normal groups, the distributions of complexity measures of both groups were estimated as presented in Figure 5.5. Compared to the normal group, the CP group showed lower entropy values, indicating a complexity reduction, specifically in the alpha-band —the frequency band corresponding to IMF5.

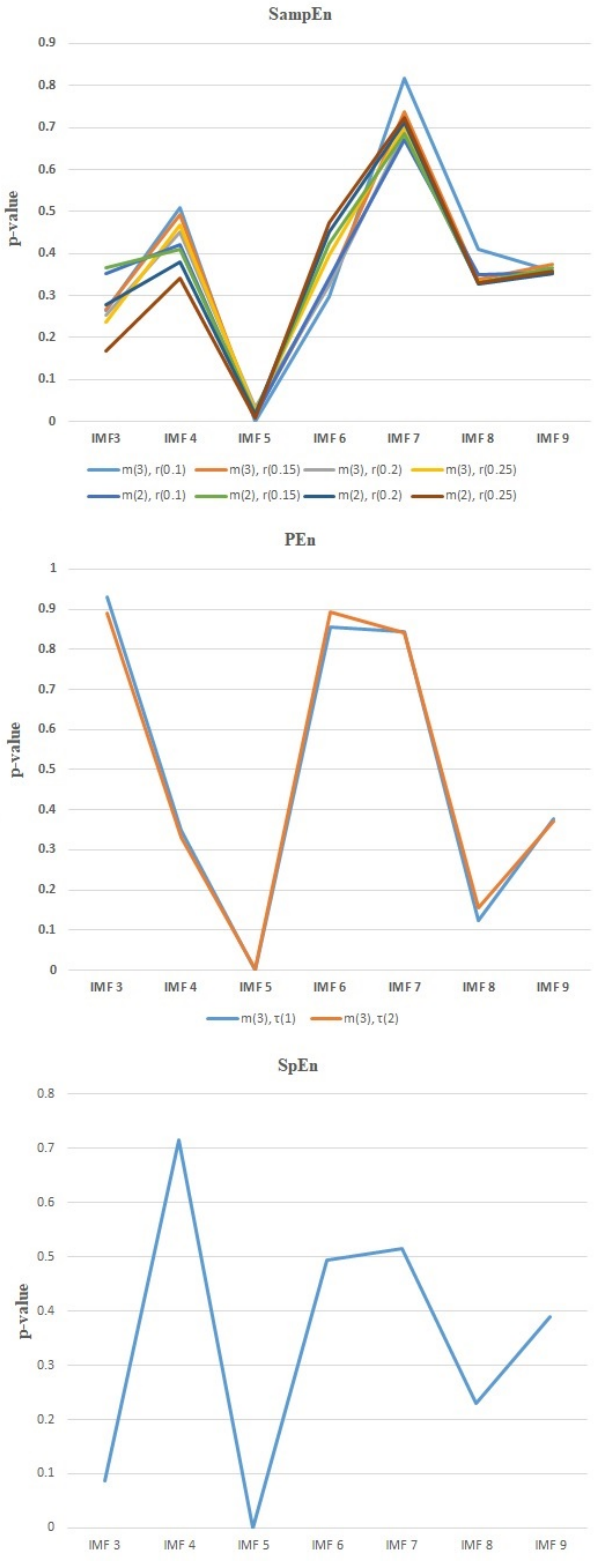


FIGURE 5.4: P-values of the entropy features comparing CP and normal groups for each IMF.

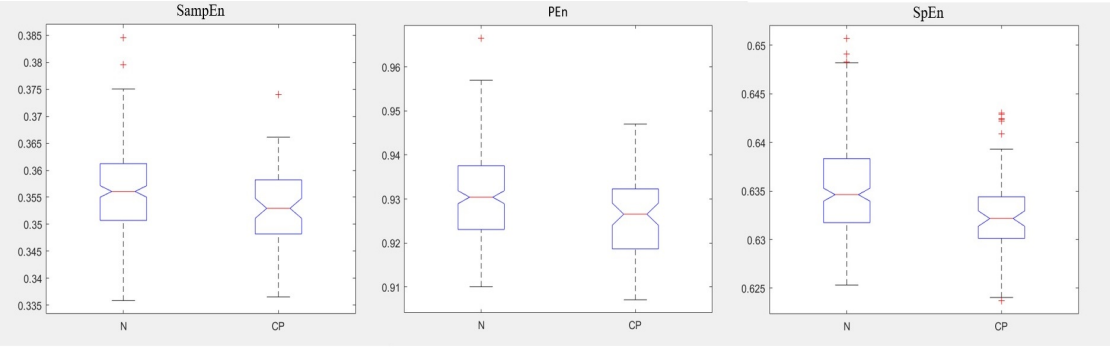


FIGURE 5.5: Box plots of the distribution of the complexity measures extracted from IMF5 (corresponding to the alpha-band) of CP and normal groups. The entropy values of CP show lower values compared to the normal group.

5.7.2 RUSBoost Results

Due to their significant discriminant capability between the CP and the normal groups, entropy features of all electrodes computed from IMF5 were selected to train and test the RUSBoost classifier.

Each entropy feature was initially used to train and test the classifier. The dimension of the feature vector was 12 (1 feature $\times$ 12 channels). Next, The classifier was evaluated on the combination of the three entropies, and the dimension of the feature vector, in this case, was 36 (3 entropies $\times$ 12 channels).

Table 5.2 gives the LOOCV results of the RUSBoost classifier. It can be inferred from the table that the classification performance was improved when the classifier was trained on the combination of the three entropies (i.e., on the feature vector of size 36). However, the number of features (the 36) is greater than the number of data samples (the 26 subjects). Thus, the classifier was trained on different subsets of the feature vector. For each subset, the features were selected using the scalar feature selection method. The performance of the top 20 high-ranked features was first investigated. The top 15, top 10, top 5, top 3, top 2, and top 1 features were also explored. It can be shown from Table 4.1 that the classifier's performance was generally enhanced by reducing the number of features. The best classification result was 84.6% ACC, 83% TPR, 85% TNR, and 0.87 AUC using the top 5 selected features.

5.8 Discussion

This study carried out a nonlinear analysis of the resting-state EEG signals in a machine learning framework to identify clinical biomarkers that could distinguish between the CP and the normal groups. Particularly SampEn, PEn, and SpEn in the NA-MEMD

TABLE 5.2: Performance of the RUSBoost classifier using different feature vectors extracted from IMF5.

Entropy Feature/s	ACC	TPR	TNR	AUC
SampEn (12 features)	50.0%	50%	50%	0.50
PEn (12 features)	61.5%	67%	60%	0.62
SpEn (12 features)	65.4%	67%	65%	0.67
SampEn, PEn, SpEn (36 features)	69.2%	67%	70%	0.71
SampEn, PEn, SpEn (top 20)	73.1%	67%	75%	0.65
SampEn, PEn, SpEn (top 15)	73.1%	67%	75%	0.75
SampEn, PEn, SpEn (top 10)	76.9%	83%	75%	0.79
SampEn, PEn, SpEn (top 5)	84.6%	83%	85%	0.87
SampEn, PEn, SpEn (top 3)	80.8%	80%	83%	0.82
SampEn, PEn, SpEn (top 2)	76.9%	80%	67%	0.73
SampEn, PEn, SpEn (top 1)	76.9%	80%	67%	0.77

domain were used to explore brain dynamics of children with neonatal HIE who developed CP.

The main challenge of this study was the limited and imbalanced number of samples in the dataset, as out of the 26 infants under prospective monitoring, only 6 developed CP at 24 months. LOOCV was employed to alleviate the limited number of samples problem, while the RUSBoost classifier was used to handle the imbalanced classes issue.

The statistical analysis results show that the best discriminatory capability of the SampEn, PEn, and SpEn was when they were extracted from IMF5 —the component corresponding to the alpha-wave (Sanei and Chambers, 2007). These results are robust to small changes in the entropy estimation parameters. Particularly, the CP group exhibits significantly lower entropies at the alpha-band compared to the normal one. This finding suggests that brain function is impaired in infants who developed CP with lower EEG complexity compared to their neurotypical peers.

Gao et al. (2016) also found differences in the alpha-band complexity of individuals with CP compared to controls. However, they reported a higher global complexity in EEG signals of CP. This inconsistency might be due to the differences in the complexity measures between the two studies. While the nonlinear complexity was employed here to measure the temporal irregularity of the signals, they used the linear omega complexity to assess the degree of synchronisation between spatially distributed brain areas. Furthermore, the age at investigation was different from our study. The exploration of Gao et al. (2016) was performed on male adolescents whilst this study investigated newborn brain complexity. Sajedi et al. (2013) also reported a higher EEG complexity in the children with CP using the fractal dimension as a nonlinear measure of time series irregularity. Nevertheless, they computed the complexity for the

broad-band signals (1-30Hz) and did not consider the complexity of each brain wave independently. Moreover, they reported enhanced complexity in the anterior brain area whilst this study characterised complexity for all regions.

Due to their significant Kruskal-Wallis results, entropies calculated from IMF5 were used to train and test the RUSBoost classifier. The results suggest that combining the three entropies (SampEn, PEn, and SpEn) can improve the classifier's performance. Moreover, selecting the highly informative features from this large pool using the scalar feature selection technique further enhanced the results and reached 84.6% ACC, 83% TPR, 85% TNR, and 0.87 AUC.

The methodology and findings of this investigation have been compared to state-of-the-art research which used qEEG analysis to characterise CP brain abnormalities. Table 5.3 summarises this comparison. The pros and cons of these existing techniques have also been discussed earlier in Section 2.6.2.

It can be seen from the table that a deficit in the alpha-band has been reported by almost all studies. Furthermore, all previous studies that we are aware of were carried out on adolescents and children with CP, and machine learning frameworks have been adopted in only two explorations (Koeda and Takeshita, 1998; Sajedi et al., 2013) and reached 91.7% ACC, 100% TPR, 83.3% TNR and 94.8% ACC, 92.5% TPR, 97.2% TNR, respectively. These results are slightly better than the classification performance of the proposed method 84.6% ACC, 83% TPR, 85% TNR.

In summary, this exploration's novelty resides mainly in using the nonlinear qEEG analysis to predict CP at infancy. The complexity measures employed in this investigation (entropies in the NA-MEMD domain) can handle most of the limitations raised with other qEEG measures used in the literature, including the assumption of linearity and stationarity of the EEG signals. Nevertheless, the NA-MEMD algorithm is limited by the high computational complexity associated with the MEMD method, as discussed in Section 4.7. The computational load is further growing with the NA-MEMD algorithm, including a subspace of multivariate independent white noise equal to the original multivariate signal.

5.9 Conclusion

The proposed framework successfully discriminated term-born infants with neonatal HIE who developed CP by two years from those with normal neuromotor outcome. Nonlinear characteristics of resting-state EEG (as indexed by SampEn, PEn, and SpEn) in the NA-MEMD were assessed using statistical analysis and a machine learning framework. Good classification performance of 84.6% ACC, 83% TPR, 85%TNR and

TABLE 5.3: Comparison of selected qEEG state-of-the-art methods used for CP classification.

Authors	Dataset	Features	Time-frequency	Evaluation method	Findings
(Gao et al., 2016)	Adolescent 14-22 years, CP (n=15) and normal (n=15)	**Microstate** to measure temporal correlation. **Omega complexity** to assess spatial correlation	FFT	Statistical analysis	**Microstate:** Higher temporal complexity in CPs compared to control. **Complexity:** Higher global omega complexity in CPs at the alpha-band compared to control
(Gao et al., 2017)	Adolescent 14-22 years, CP (n=15) and normal (n=15)	Detrended fluctuation analysis (DFA) to measure temporal correlation	FFT	Statistical analysis	The DFA exponents at alpha- and beta-band were significantly attenuated in the CPs compared to controls.
(Koeda and Takeshita, 1998)	Children 7-15 years, CP (n=12) and normal (n=15)	PSD and coherence measures	FFT	Statistical analysis	**PSD**: No significant difference in PSD. **Coherence**: Lower ICoh at the occipital region for alpha-band, higher ICoh at the frontal region for theta-band, and higher Hcoh at the left hemisphere for the delta-, theta-, and beta-band in the CPs compared to controls.
				DA classifier	Performance of 91.7% ACC, 100% TPR, 83.3% TNR has been reached using the statistically significant features.
(Kulak et al., 2005)	Children 6-14 years, CP (n=12) and normal (n=21)	PSD and coherence measures	FFT	Statistical analysis	**PSD**: Significant differences between the CP and control over the left and right hemispheres for the delta-, theta-, alpha- and beta-band. **Coherence**: Lower ICoh at the temporal, parietal and occipital regions for the alpha-band, lower ICoh at the frontal, central, parietal and occipital regions for the beta-band, higher ICoh at the frontal and temporal regions for the theta- and delta-band and higher HCoh at right hemisphere for the alpha-band in individuals with CP compared to control.
(Kulak and Sobaniec, 2005)	Children 6-15 years, CP (n=26) and normal (n=28)	PSD and coherence measures	FFT	Statistical analysis	**PSD**: Significant differences between the CP and control over the left and right hemispheres for the delta-, theta-, alpha- and beta-band. **Coherence**: Lower ICoh at the temporal, parietal and occipital regions for the alpha-band in the CPs compared to controls.
(Sajedi et al., 2013)	Children 4-14 years, CP(n=26) and normal(n=26)	PSD and fractal dimension to measure temporal complexity.	Welsh's method	Statistical analysis	**PSD**: A higher delta-band and lower theta and alpha powers were found in CPs compared to controls. **Complexity**: A higher EEG complexity at the interior region for the range (1-30Hz) in CPs compared to controls.
				Enhanced probabilistic neural network classifier	Performance of 94.8% ACC, 92.5% TPR, 97.2% TNR have been reached using the statistically significant features.
Current study	Term HIE infants, CP(n=6) and normal(n=20)	SampEn, PEn and SpEn	NA-MEMD	Statistical analysis	A lower EEG complexity at the alpha-band in CPs compared to controls.
				RUSBoost classifier	Performance of 84.6% ACC, 83% TPR, 85% TNR, 0.87 AUC was achieved using the combination of the three entropies.

0.87 AUC was reached through this approach. Moreover, the results indicate that the complexity features of the alpha-band are potential biomarkers for early CP prediction.

The next chapter employs the framework and the dataset used in this exploration to investigate the efficiency of the nonlinear synchronisation features to improve the performance of CP prediction.

Chapter 6

Complexity and Synchronisation Analysis of Electroencephalography Signals for Cerebral Palsy Prediction in Newborns with Hypoxic-Ischemic Encephalopathy

The previous chapter investigated the efficiency of at-birth EEG complexity analysis to predict the CP outcome at two years. The results established that the complexity of resting-state EEGs, particularly at the alpha-band component, can satisfactorily predict the outcomes.

This chapter extends the previous chapter investigation to examine whether at-birth EEG synchronisation can provide additional information to predict CP. This exploration is based on the fact that higher-level brain functions (motor activity here) depend strongly on the functional interactions between different brain areas, and such interactions may be affected due to the HIE condition.

Section 2.6.2 reviewed different studies that used functional connectivity analysis of EEG signals for CP prediction. It reported a lack of studies using synchronisation measures that can efficiently deal with the nonlinearity and nonstationarity of the signals.

Therefore, this chapter examines the effectiveness of employing PLI as a conduction-free nonlinear measure to characterise EEG synchronisation of infants (the dataset used in the previous chapter) and evaluate its ability to predict CP outcomes. It also compares the results of this investigation with the previous chapter's results, i.e., using the complexity measures.

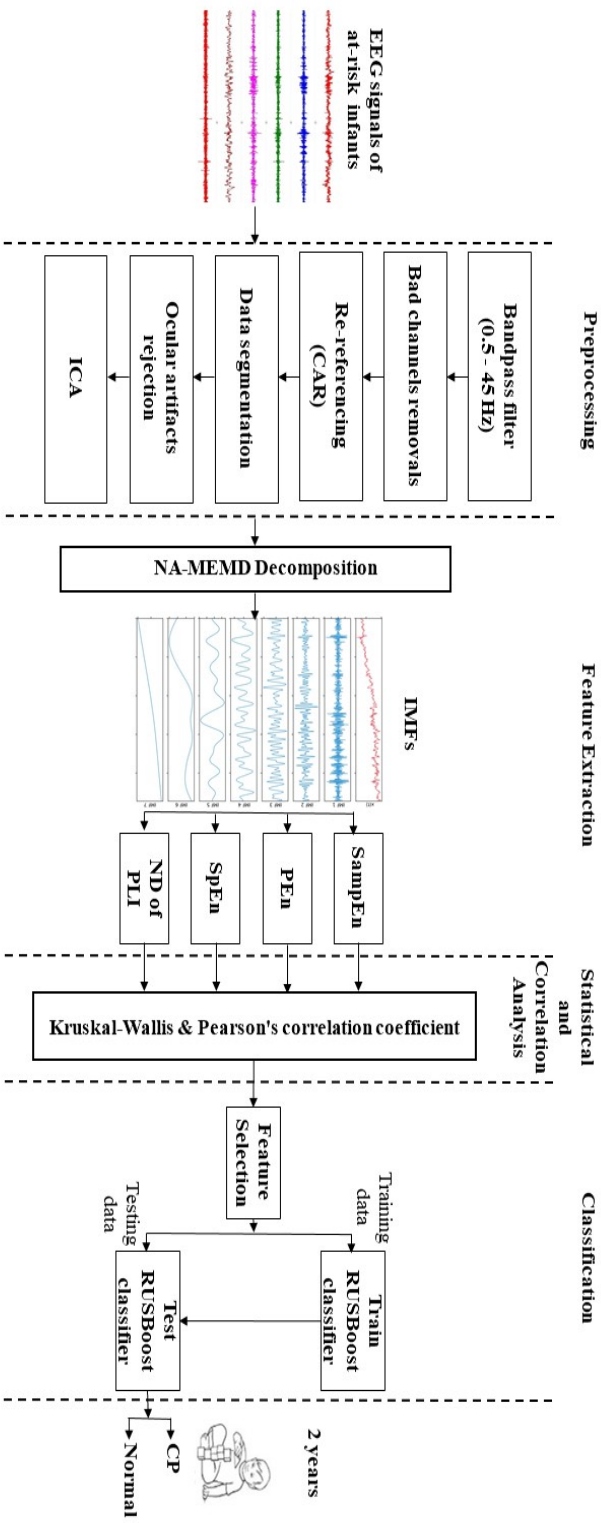


FIGURE 6.1: Block diagram of the proposed methodology for classifying CP and the normal groups using complexity and synchronisation features.

To this end, graph theory analysis, described in Section 3.4, was used to represent the connectivity between different brain regions. The edges of the brain network were identified in this exploration by the PLIs extracted from the IMF components of corresponding channels, resulting in a connectivity matrix for each IMF. ND measure, also illustrated in Section 3.4, was then utilised to quantify the connectivity strength of each channel with other neighboring nodes. ND was used as it measures the synchronisation at a channel level and thus facilitates the comparison with the entropy measures, which are also analysed (in this chapter) at each channel individually.

Statistical analysis is first performed to evaluate the effectiveness of the ND of PLI measures to differentiate between the CP and the normal groups. Correlation analysis between the ND of PLI and the entropy features is also assessed as it may reveal a more detailed picture of atypical EEG activities in the CP brains.

The classification performance of the RUSBoost classifier based on the ND of PLI features is then evaluated. Furthermore, the classifier is trained and tested on the combination of ND of PLI and entropy features to provide a more detailed characterisation of the neural functions leading to either CP or normal neuromotor outcomes. Figure 6.1 extends the methodology shown in Figure 5.1 by calculating the ND of PLIs in the NA-MEMD domain.

The novelty in this chapter can be summarised in the following: 1- Using the ND of PLIs to reflect abnormal functional connectivity associated with CP condition. 2- Computing the EEG complexity measures (SampEn, PEn and SpEn) and the EEG synchronisation (PLI) at channel level to compare their significance. 3- Applying a joint analysis of the EEG complexity and the EEG synchronisation measures to provide a more detailed characterisation of brain activity that could predict a CP outcome.

The remainder of this chapter is structured as follows: Section 6.1 extends the preprocessing procedure described in the previous chapter, identifying the brain region of each channel. Section 6.2 describes the extraction scheme used to extract the ND of PLI features. The statistical analysis method is provided in Section 6.3, while correlation analysis between the complexity and the synchronisation measures is given in Section 6.4. Section 6.5 and Section 6.6 discuss the feature selection method and the classification procedure employed, respectively. The results are then reported in Section 6.7, discussed in Section 6.8 and conclusions presented in Section 6.9.

6.1 Data Preprocessing

This exploration is an extension of the work presented in the previous chapter. Thus, the preprocessed signals of the previous chapter were used without any changes. As this investigation evaluates the entropies and the ND of PLI features at a channel level,

the twelve channels described in Section 5.2 were localised, as suggested by Sajedi et al. (2013), in the following brain lobes: the central (C3, T3, and T4), anterior (Fz, F3, and F7), and posterior (O1, O2, P3, P4, T5, and T6) lobes. Such localisation helps in comparing the findings with state-of-the-art results.

6.2 Extracting ND of PLI Measures

The feature extraction procedure of the current investigation is an extension to the process described in Section 5.3. Thus, following the NA-MEMD decomposition and IMF selection, the dimension of the dataset for each subject was: $N_c \times N_i \times N_e \times N_t$, where N_c is the number of channels which is 12, N_i is the selected number of IMFs, which is 7 (IMF3 –IMF9), N_e is the number of epochs which is 30, and N_t is the number of the samples which is 1024. The ND of PLI measures were then computed as follows:

1. The alignment of the selected IMFs (IMF3–IMF9) among channels was first investigated and confirmed.
2. For each IMF and each epoch, the PLI connectivity matrix was calculated between the signals from the twelve channels. The generated PLI matrices were then averaged over the epochs to get one connectivity matrix for each IMF signal.
3. The ND measures were then estimated from the connectivity matrices to quantify the connectivity strength of each channel using Equation 3.9, where k represents the total number of channels, which is 12 in this exploration.

   Thus, for each subject and each IMF signal, twelve ND features were used to train and test the RUSBoost classifier.

The whole analysis was carried out in the MATLAB software package R2018a.

6.3 Statistical Analysis

Kruskal-Wallis test was utilised to determine the ability of ND of PLI as well as SampEn, PEn and SpEn features computed from each IMF and each channel (336 P-values: 4 features $\times$ 7 scales $\times$ 12 channels) to discriminate between CP and normal groups. BH-FDR, described in Section 5.4, was employed to control for these multiple comparisons, and the value of α was corrected from 0.05 to 0.04. Thus, a feature was considered statistically significant if P-value$<= \alpha (= 0.04)$.

6.4 Correlation Analysis

Correlation analysis was utilised to evaluate the relationship between the complexity and the synchronisation measures. Thus, Pearson's correlation coefficient (R) was used in each group to determine the linear dependency between each entropy feature (SampEn, PEn, and SpEn) computed from each channel and the ND of PLI. This process was repeated for each IMF separately.

Theoretically, the value of R falls in the interval between $+1$ and -1, with 0 indicating no linear relationship, $+1$ refers to a perfect positive correlation as one variable increases, the other increase too, while -1 indicates a perfect negative correlation as one variable increases, the other decrease. The level of significance is also included in the correlation test where P-values$<=$ 0.05 correspond to a significant correlation in R and a low probability of observing the null hypothesis (that there is no relationship between the two variables).

BH-FDR correction was applied to these R-scores (504 P-values: 3 features $\times$ 7 scales $\times$ 12 channels $\times$ 2 groups), and no modification to the value of α was suggested. Thus, the correlation was considered significant if the if P-value$<= \alpha(= 0.05)$. The correlation analysis was carried out using the MATLAB statistics toolbox.

6.5 Features Selection

One of the aims of this investigation was to explore improving the classification performance of discriminating the CP and the normal groups by combining ND of PLI and the complexity features. Thus, 48 features for each subject and each IMF signal were initially used to train and test the classifier (4 features $\times$ 12 channels). This situation can lead to the overfitting problem as the number of features are greater than the number of samples (the 26 subjects). The scalar feature selection method described in Section 3.6 was employed for this purpose.

6.6 Machine Learning and Classification Process

This investigation used the same classifier as that employed in the previous chapter (RUSBoost) and followed a similar procedure as that described in Section 5.6 to evaluate its performance in discriminating between the CP and normal groups.

6.7 Results

6.7.1 Kruskal-Wallis Results

The discrimination capability of each entropy (SampEn, PEn, and SpEn) and ND of PLI features, when computed from each IMF and each channel to discriminate between the CP and normal groups, was evaluated. Embedding dimension m=3 and tolerance r=0.1 were used for SampEn computation, and m=3 and τ=1 for PEn calculation.

The results indicate significant group differences satisfied the criterion of P-value$<=$ $\alpha(= 0.04)$ in SampEn of F7 (P= 0.02), PEn of O2 (P=0.03) and P4 (P=0.02), and SpEn of P4 (P=0.017) and T6 (P=0.014) at the IMF5 component corresponding to the alpha-band. Figure 6.2 shows the distributions of these significant features for both groups. The CP group showed lower entropy values than the normal group, indicating a complexity reduction mainly observed in the alpha-band in all brain regions.

The results of ND of PLI also indicate significant group differences satisfied the criterion of P-value$<=$ $\alpha(= 0.04)$ in the ND of T4 (P=0.017) also for the IMF5 component. In line with the entropy results, it can be observed from the boxplot presented in Figure 6.2 that the ND of PLI was significantly lower in the CP group compared to the normal one.

Figure 6.2 presents the significant results. All Kruskal-Wallis results are available in Appendix C.

6.7.2 Correlation Analysis Results

To evaluate the relationship between the complexity and the synchronisation measures, correlation coefficients were estimated at each IMF and each channel between the entropy features (SampEn, PEn and SpEn) and the ND of PLI in the CP and the normal groups.

The results indicate a high correlation between entropies and ND of PLI passing through the P-value$<=$ $\alpha(= 0.05)$ criteria in different IMFs and all brain areas. Particularly, a high negative correlation was observed at the anterior, central and posterior regions in IMF4 and IMF5, corresponding to the beta- and alpha-band, respectively, in the normal group. This correlation was detected in the CP group in IMF4 at the three brain regions and in IMF5 only at the posterior area.

Moreover, high positive correlations passing through the P-value$<=$ $\alpha(= 0.05)$ were noticed at the three regions in IMF8, corresponding to the delta-band, in the normal group. This correlation was observed in the CP group at the central and posterior areas.

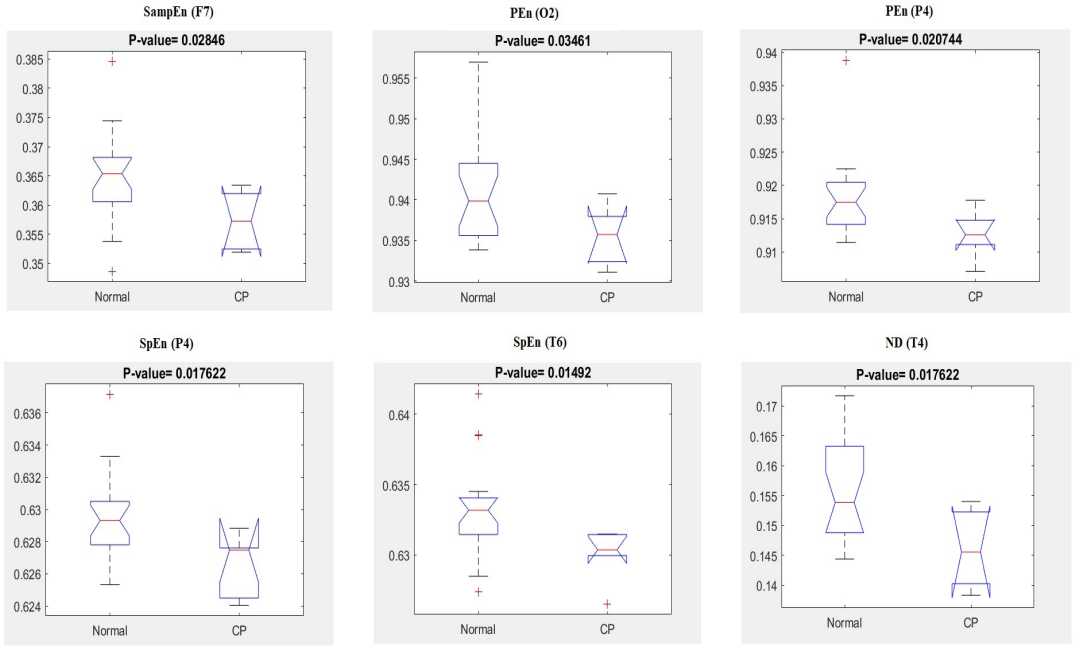


FIGURE 6.2: The box plots present the distribution of the significant entropy and ND of PLI features extracted from IMF5 to discriminate between the CP group and the normal one.

Figure 6.3 presents some of these significant results. All other results are available in Appendix C.

The scatter plots presented in Figure 6.3 show that the samples of the CP and the normal groups are well separated in almost all cases using the complexity and ND of PLI features, suggesting that their combination may improve the classification of these groups.

6.7.3 RUSBoost Results

The ND of PLI as well as the entropy features computed from IMF5, which gave significant discrimination between the two groups, were evaluated individually to train and test the RUSBoost classifier. Table 6.1 presents the classification performance results. This result indicates comparable performance of the ND of PLI feature compared to the entropy (SampEn, PEn and SpEn) features.

In addition, the classification performance of the combination of complexity and ND of PLI features was investigated. In this investigation, features extracted from IMF4, IMF5, and IMf8 were used due to their good separation between the two groups, which were observed from the scatter plots presented in Figure 6.3. Scalar feature selection was used, and the RUSBoost classifier was trained and tested on different subsets of the feature vector of 144 features (4 features (SampEn, PEn, SpEn and ND of PLI) $\times$ 12 channels $\times$ 3 IMFs). Table 6.2 presents the results of this exploration.

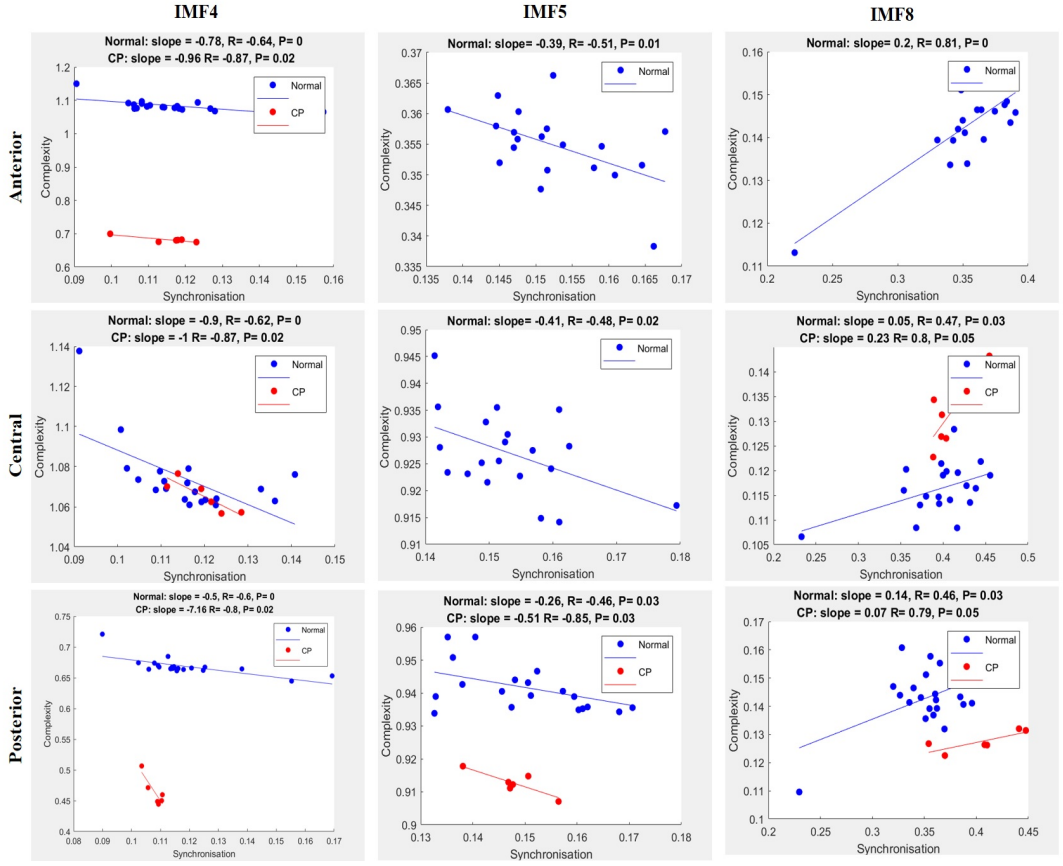


FIGURE 6.3: Scatter plots of the correlation between entropies and ND of PLI at IMF4, IMF5, and IMF8 at the three brain regions in the CP and the normal groups. The solid lines indicate the linear regression (blue: linear regression line for normal group, red: linear regression line for CP group). No significant correlation was detected in the CP group in some brain regions, and thus some panels have no representation of the CP correlation.

TABLE 6.1: Performance of the RUSBoost classifier using each of the significant SampEn, PEn, SpEn and ND of PLI features, all computed from IMF5.

Features	ACC	TPR	TNR	AUC
SampEn (F7)	50.0%	50%	50%	0.49
PEn (O2, P4)	69.2%	50%	75%	0.55
SpEn (P4, T6)	69.2%	50%	75%	0.58
ND of PLI (T4)	65.4%	33%	75%	0.59

The table shows improvements in the classification performance of the RUSBoost classifier resulting from the combination of the complexity (SampEn, PEn, SpEn) and ND of PLI features compared to using each feature individually. The improvement is noticed in almost all selected subsets of the combination. Particularly, the high-ranked top 10 and top 5 features achieved a best performance of 96.2% ACC, 100% TPR, 95% TNR, and 1.00 AUC.

As the high TPR is important in the clinical prognostic tests, the 100% TPR reached in this exploration is interesting in which infants identified as having the CP outcome

TABLE 6.2: Performance of the RUSBoost classifier using the combination of complexity and synchronisation features extracted from IMF4, IMF5, and IMf8.

Entropies and ND of PLI features	ACC	TPR	TNR	AUC
all 144	73.1%	67%	75%	0.77
top 25	88.5%	83%	90%	0.97
top 20	92.3%	100%	90%	0.98
top 15	96.2%	83%	100%	1.00
top 10	96.2%	100%	95%	1.00
top 5	96.2%	100%	95%	1.00
top 3	84.6%	80%	100%	0.90
top 2	80.8%	85%	67%	0.85
top 1	76.9%	80%	67%	0.84

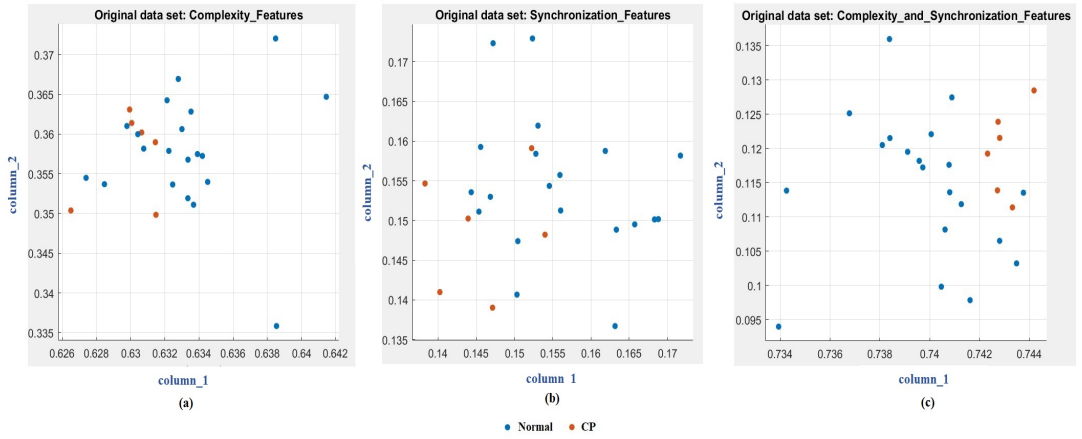


FIGURE 6.4: The scatter plots of the dataset of the CP and the normal groups with the (a) complexity features, (b) synchronisation features, and (c) their combination. The samples are better separated using the combination of two classes of features. Column_1 and column_2 represent arbitrary first and second features, respectively.

should be more likely to have the condition. Figure 6.4 shows the scatter plots of the dataset associated with the complexity features, ND of PLI measures, and their combination, respectively. As seen, the samples of the CP and the normal groups can often be better separated using the combination of two classes of features.

6.8 Discussion

This chapter explored the ability of the synchronisation and complexity measures in different brain regions to discriminate between the at risk-infants who developed CP by the age of two and those with the normal neuromotor outcome.

To this end, the ND of PLI and the entropy features in the NA-MEMD domain were assessed at each channel to discriminate between the CP and the normal groups. Furthermore, the relationship between the complexity and synchronisation features was evaluated to provide a joint analysis of the brain dynamics related to CP.

Following the previous chapter's strategies to minimise the machine learning challenges, the RUSBoost classifier was used to handle the imbalanced classes issue, and LOOCV was employed to reduce the effect of the limited number of samples.

The ND of PLI was used to quantify the nonlinear couplings between channels. Compared to the normal group, the averaged PLI value of the T4 channel at the alpha-band was smaller in the CP group. This finding is in line, to some extent, with previous EEG investigations that reported lower ICoh for the CP group in the alpha-band at the temporal derivations, implying hypoconnectivity in this brain areas (Kulak and Sobaniec, 2005; Kulak et al., 2005).

In line with the result of the previous chapter, significant reductions of the three entropies (SampEn, PEn, and SpEn) in the alpha-band (corresponding to IMF5) of different brain regions were confirmed in the CP compared to the normal group.

Regarding the relationship between the complexity and synchronisation features, different regional-and-band-specific dependencies were found in the CP and the normal groups. Mainly, significant positive correlations were observed at IMF8, corresponding to the delta-band, between the entropy and ND of PLI features in the three brain areas in the normal group and central and posterior regions in the CP group. This finding suggests that a high complexity of the signals at a brain region is linked to a significant interaction of that region with other brain areas and vice versa. In other words, the impairment in the interaction between different brain regions, which may be induced by the pathological condition of CP, leads to reducing the functional connectivity and complexity. Such a positive correlation was reported previously by Nobukawa et al. (2020) in the investigation of AD.

Furthermore, high negative correlations were identified in all brain regions in IMF4 and IMF5, corresponding to the beta- and alpha-band, respectively, in the CP and the normal groups. This correlation indicates that a decrease in the signals' complexity of a brain area is associated with an increase in that region's ability to network with other regions and conversely. Such a correlation was hypothesised in previous EEG abnormality investigations of AD (Nobukawa et al., 2020) and autism (Edgar and Roberts, 2016).

The negative correlation could be interpreted through the following mechanism. According to nonlinear dynamical theory, it is known that nonlinear coupled oscillations exhibit enhancement of complexity by the emergence of a chaotic state during the process of reducing the coupled strength for the state of complete synchronisation

(Pikovsky et al., 2001). This induced enhancement is attributed to perturbations to the stable orbit from each oscillation's behavior (Schweighofer et al., 2004; Nobukawa and Nishimura, 2016). Thus, during the process of decreasing coupled strength, the synchronisation reduces, and the complexity enhances.

The potential difference in the direction of the relationships, across different frequency bands, between the complexity and the synchronisation measures were reported previously in AD in the patient and control groups (Nobukawa et al., 2020). Correlation differences across frequency bands may reflect the different functional significance of activity in these brain frequencies. Detailed investigations of the regional and spectral influences on the relationships between complexity and synchronisation measures versus specific pathological conditions are motivated by such observations. Further exploration in this field is needed.

Due to their significant result, the ND of PLI and the entropy features calculated from IMF5 were used to train and test the RUSBoost classifier to show how useful it could be in practical CP prediction. The results indicate that the ND of PLI feature achieved comparable classification performance compared to the complexity features.

The relationship between the complexity and synchronisation features and the different slopes of these correlations between the CP and normal groups suggests that combining them can improve classification performance. Thus, additional exploration was performed by training and testing the RUSBoost classifier with the combination of SampEn, PEn, SpEn, and ND of PLI features. This investigation reached a high classification performance of 96.2% ACC, 100% TPR, 95% TNR, and 1.00 AUC.

Table 5.3 presents state-of-the-art research and shows the good predictive power of the coherence measure (the primary connectivity measure used in the literature) on the motor symptoms of CP. However, the PLI-based synchronisation approach used here can handle all limitations raised with coherence, including the assumption of linearity and stationarity of the EEG signals and the VC issue. Additionally, a joint analysis of the EEG complexity and EEG synchronisation features was explored to provide a more detailed characterisation of the neural functions that could lead to either CP or normal neuromotor outcomes. Better classification performance was achieved through the combination of the complexity and synchronisation features compared to (Koeda and Takeshita, 1998; Sajedi et al., 2013) (the only two existing studies that used machine learning to the best of the author's knowledge), which reached 91.7% ACC, 100% TPR, 83.3% TNR and 94.8% ACC, 92.5% TPR, 97.2% TNR, respectively.

However, the ND of PLI used in this exploration is limited by the fact that it cannot differentiate a node with a single high PLI from a node with lots of small PLIs. Thus, ND of PLI is a sensitive measure of node strength with nonhomogeneous degree distributions. Other measures such as "eigenvector centrality," which measures a node's

relative importance to the network, could be investigated in future work to address this issue (Hatlestad-Hall et al., 2021).

6.9 Conclusion

The proposed framework successfully discriminated resting-state EEGs of the term-born infants with neonatal HIE who developed CP by two years from those with normal neuromotor outcome. EEG complexity and EEG synchronisation measures in the NA-MEMD domain were assessed using statistical analysis and a machine learning framework. The relationship between the two classes of features in each group was reported, suggesting the complexity and synchronisation features of beta-, alpha-, and delta-band as biomarkers for the early prediction of CP. Combining these features improved the classification performance of the RUSBoost classifier, reaching 96.2% ACC, 100% TPR, 95% TNR, and 1.00 AUC.

Chapter 7

Cognitive Outcome Prediction in Infants With Neonatal Hypoxic-Ischemic Encephalopathy Based on Functional Connectivity and Complexity of the Electroencephalography Signals

As per the fourth objective outlined in Chapter 1, this chapter aims to explore whether the nonlinear analysis of resting-state EEG signals recorded from infants with HIE can provide biomarkers that aid in predicting their two-year cognitive outcomes assessed using BSITD-III (Weiss et al., 2010).

This chapter is based on the work presented in (Alotaibi et al., 2022) except for some minor modifications. The author's contribution to the work included all materials given in this chapter. The materials and results in the published paper related to the global graph-theoretic parameters derived from the weighted phase-lag index are not included in this chapter as they belong to Alotaibi's contribution. Maharatna's contribution included supervision, Konn's contribution consisted of extracting the EEG signals and Vollmer's contribution involved providing the neurological assessment information.

The BSITD-III consists of three scales; motor, language and cognitive, but the focus of this exploration will be on predicting the composite scores from the cognitive scale. Early prediction of such cognitive outcomes is of great interest as infants with HIE at

birth may not develop CP but are still at high risk of developing learning disabilities and cognitive deficits (Schie et al., 2015; Schreglmann et al., 2016).

Section 2.6.3 reviewed the application of qEEG analysis to predict the long-term cognitive function. As pointed out, there is a lack of studies that used qEEG measure(s) considering the nonlinearity and nonstationarity of EEGs. Therefore, this chapter explores the extent to which the nonlinear characteristics of the EEG can help identify the relationship between at-birth EEG recordings and later cognitive function scores/severity. Specifically, SampEn, PEn and SpEn were used to characterise EEG complexity, and ND of PLI was employed to measure EEG synchronisation of infants' resting-state signals. These features were extracted from IMFs resulting from NA-MEMD to take into account the nonstationarity of EEGs.

Correlation analysis was first utilised to evaluate the relationship between the nonlinear features and the cognitive scores. The significant features were used to train and test tree ensemble regression models (boosting and bagging) to evaluate their potential clinical application aiding in predicting cognitive outcomes. Figure 7.1 illustrates this exploration methodology.

This chapter raises the following research questions: (1) Can the complexity (SampEn, PEn and SpEn) and synchronisation (ND of PLI) features that are computed from the NA-MEMD domain of infants' EEG aid in predicting their two-year cognitive scores? (2) If yes, which IMF component(s) best predict the outcomes, and what are their underlying frequencies? (3) Can the significant feature(s) (if any) be used to train and test a regression model with good performance? (4) Can the combination of the significant features (if any) enhance the regression performance?

The rest of the chapter is organised as follows: Section 7.1 describes the EEG dataset used in this exploration. Section 7.2 presents the feature extraction procedure. The correlation analysis method is provided in Section 7.3, while the regression procedure is described in Section 7.4. The results are then presented in Section 7.5. Section 7.6 discusses the results and Section 7.7 concludes the chapter.

7.1 Experimental Data Description

The dataset of the 30 term-born infants with HIE used in Chapter 5 and Chapter 6 was used in this investigation. However, the outcomes of the neuropsychological assessment were considered rather than the assessment of the neurological examination employed in the previous two chapters.

The neuropsychological follow-up assessment was conducted at 24 months of age using the BSITD-III (Weiss et al., 2010). Among the 30 infants, 20 completed the two-year

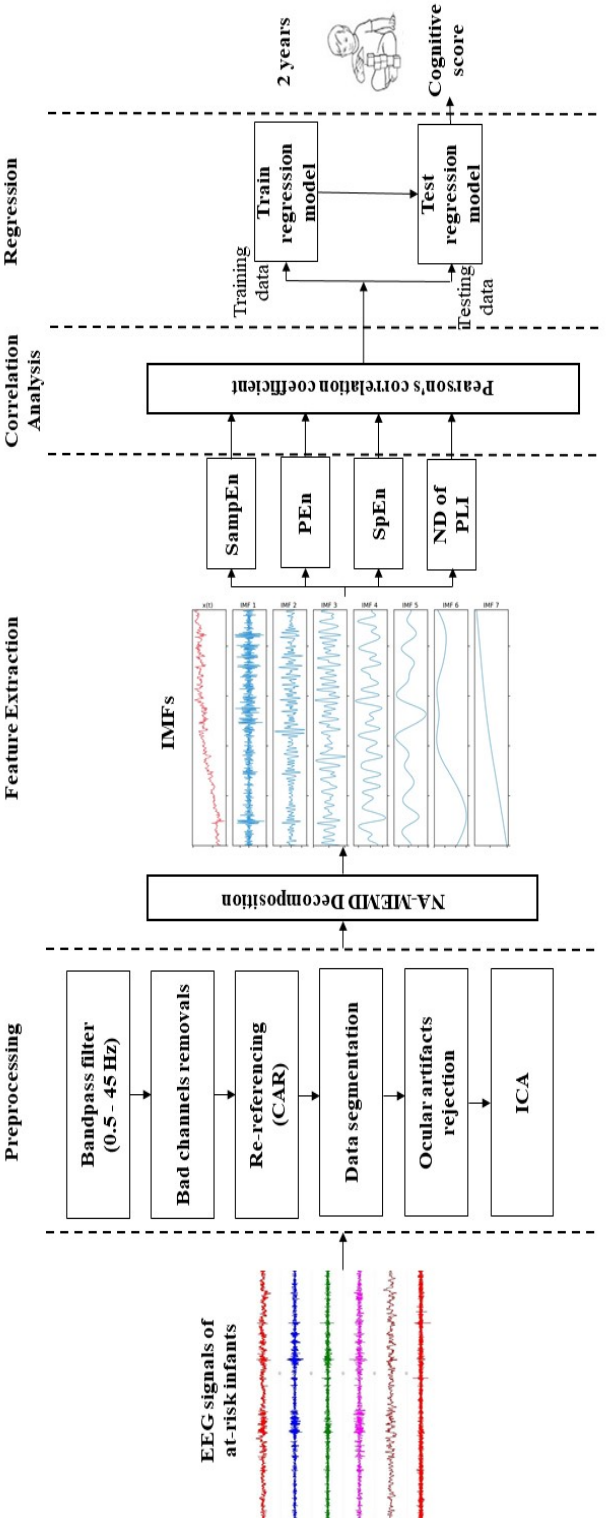


FIGURE 7.1: Block diagram of the proposed methodology for predicting the cognitive outcomes using the complexity and synchronisation features.

TABLE 7.1: Neuropsychological assessment at 24 months of the 30 infants born with HIE.

Subject #	Cognitive scores at age 24 months
subj_1	145
subj_2	Incomplete follow-up
subj_3	100
subj_4	Incomplete follow-up
subj_5	Incomplete follow-up
subj_6	Incomplete follow-up
subj_7	120
subj_8	95
subj_9	95
subj_10	140
subj_11	100
subj_12	120
subj_13	105
subj_14	Incomplete follow-up
subj_15	105
subj_16	130
subj_17	110
subj_18	140
subj_19	incomplete follow-up
subj_20	74
subj_21	105
subj_22	100
subj_23	Incomplete follow-up
subj_24	100
subj_25	125
subj_26	Incomplete follow-up
subj_27	Incomplete follow-up
subj_28	95
subj_29	Incomplete follow-up
subj_30	90

assessment. The BSITD-III cognitive scores for those infants ranged from 74 to 145. Infants with a cognitive score of < 80 are considered to have cognitive decline. Table 7.1 provides the cognitive score of each of the 20 infants who completed the assessment.

Before extracting the nonlinear features and starting this chapter's investigation, the EEG signals of those 20 neonates were preprocessed using a similar preprocessing procedure as that described in Section 5.2.

7.2 Feature Extraction

The feature extraction procedure of this exploration is similar to the process described in Section 5.3 (step 1 –3) except that the number of subjects in this exploration was 20 while it was 26 in Chapter 5's investigation. As the number of subjects and their identities differed between both explorations, the resulting IMFs and their frequencies also varied as the result of the NA-MEMD being a data-driven decomposition method.

Figure 7.2 illustrates the decomposition process of the current investigation. The number of IMFs, their instantaneous frequencies and the resulting feature vectors from the feature extraction procedure are described as follows:

1. The frequencies of each IMF were acquired by the HT, and it was found that IMF1 to IMF3 contained high frequencies (>50 Hz). Thus, these modes were excluded from further analysis. IMF10 was also ignored as it represented the residue mode of some EEG channels. The scales of the remaining IMFs were localised approximately around the following ranges: IMF4 (15 –26 Hz), IMF5 (10 –13 Hz), IMF6 (6 –8 Hz), IMF7 (3 –4 Hz), IMF8 (1.5 –3 Hz), IMF9 (0.5 –1.5 Hz). Therefore, IMF4 to IMF6 belong to the beta-, alpha- and theta-band, respectively, while IMF7 to IMF9 correspond to the delta-band.

2. After IMFs selection, the dataset dimension of each subject became: $N_c \times N_i \times N_e \times N_t$, where N_c is the number of channels which is 12, N_i is the selected number of IMFs which is 6 (IMF4 –IMF9), N_e is the number of epochs which is 30, and N_t is the number of the samples which is 1024.

3. Entropy features (SampEn, SpEn, and PEn) were then extracted from each IMF. Detailed descriptions of these measures are provided in Section 3.3.

4. ND of PLI features were also extracted from each IMF using Equation 3.9.

5. The extracted features were then used for the correlation analysis and to train and test the regression models.

The whole analysis was carried out in the MATLAB software package R2018a.

7.3 Correlation Analysis

Correlation analysis was utilised to evaluate the relationship between the nonlinear characteristics of at-birth EEGs and the cognitive scores assessed at 24 months. Pearson's correlation coefficient (R) was employed for this purpose to determine the linear dependency between the cognitive scores and SampEn, PEn, SpEn and ND of PLI features computed for each channel and each IMF.

The BH-FDR method, described in Section 5.4, was employed to control these multiple comparisons (288 P-values: 4 features $\times$ 6 scales $\times$ 12 channels), and the value of α was corrected from 0.05 to 0.03. Thus, the correlation was considered significant if the if P-value$<= \alpha(= 0.03)$. The significant features were then used to train and test the regression models described in the next section. The correlation analysis was carried out using the MATLAB statistics toolbox.

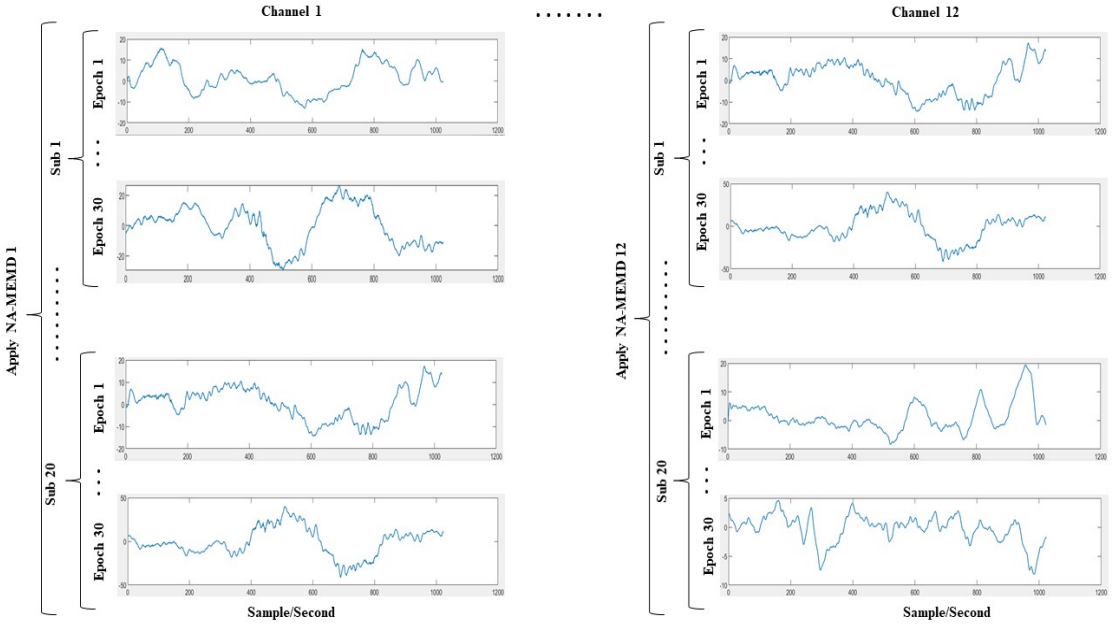


FIGURE 7.2: The proposed simultaneous decomposition method of the EEG signals.

7.4 Regression Model

Regression models were used in this exploration to predict the two-year cognitive scores from at-birth EEGs. The models fit the relationship between the two proposed sets of the nonlinear features (EEG complexity and EEG synchronisation) and the cognitive scores with the least possible error. Tree ensemble regression models were used to handle the data imbalance —the cognitive scores ranging between 74 and 145, and most of the scores being clustered above 95.

Mainly, boosting and bagging regression models, discussed in Section 3.6 were adopted in this investigation. Regression training and testing were performed using the *Regression Learner App* within the statistics and machine learning toolbox in MATLAB. Detailed descriptions of boosting and bagging models are available in Appendix A.

The ensembles regression models were trained and tested on different vectors constructed from the significant features (selected based on the correlation analysis), and LOOCV was used to prevent overfitting. In general, the prediction made by a regression model cannot be evaluated using the traditional classification performance metrics used in the previous chapters. In regression, the model outputs a continuous number (not classes), and hence, its performance is evaluated in terms of how close the output is to the actual value.

Therefore, the performance in this exploration was evaluated using error metrics especially designed for evaluating predictions made on regression problems. Root mean square error (RMSE), mean absolute error (MAE) and R-squared are commonly used

for this purpose (Pal, 2017). RMSE refers to the square root of the average squared difference between the predicted and actual scores, with lower RMSE indicating better performance. RMSE is given by:

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(\hat{y}_i - y_i)^2}{n}} \tag{7.1}$$

Where n is the total number of samples, y_i is the actual value, and $\hat{y}_i$ is the predicted value (Pal, 2017).

MAE is the absolute difference between the predicted value and the target one, and the smallest MAE value also refers to the best model's performance. MAE is given by:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{7.2}$$

Where n is the total number of samples, y_i is the actual value, and $\hat{y}_i$ is the predicted value (Pal, 2017).

R-squared, in contrast, determines how well the model predicts the specific score by comparing the learned model with the constant baseline model. The baseline model is built by taking the mean of training data and drawing the line on the mean. The value of R-squared is usually less than or equal to one where the higher value refers to a better fit between predicted and actual values. The formula of R-squared is given by:

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2} \tag{7.3}$$

Where y_i is the actual value, $\hat{y}_i$ is the predicted value, and $\bar{y}$ is the mean baseline model (Pal, 2017).

7.5 Results

7.5.1 Correlation Analysis Results

Correlation analysis was first carried out between the entropy features, computed from each IMF and each channel, and the cognitive scores of all subjects. Different embedding dimensions m, tolerances r, and time delays τ were explored for SampEn, and PEn estimations and the results indicate their robustness to small changes in the embedding parameters.

Table 7.2 presents the P-values of the correlation analysis using m=3 and r=0.1 for SampEn computation, and m=3 and τ=1 for PEn calculation. P-values of other embedding

TABLE 7.2: P-values of the correlation analysis of the entropy features. Significant features are shown in boldface.

	Channel	IMF4	IMF5	IMF6	IMF7	IMF8	IMF9
SampEn	C3	0.79	0.97	0.95	0.56	0.51	0.19
	F3	0.42	0.86	0.77	0.28	0.40	0.66
	F7	0.86	0.41	0.95	0.99	0.83	0.50
	Fz	0.38	0.61	0.07	0.13	0.71	0.27
	O1	0.61	0.45	0.41	0.93	0.32	0.43
	O2	0.27	0.52	0.97	0.50	0.07	0.81
	P3	0.33	0.75	0.33	0.68	0.42	0.71
	P4	0.33	0.92	0.68	0.46	0.86	0.18
	T3	0.59	0.42	0.82	0.18	0.98	0.84
	T4	0.73	0.77	0.66	0.51	0.79	0.48
	T5	0.28	0.65	0.60	0.74	0.68	0.27
	T6	0.50	0.71	0.48	0.32	0.95	0.07
PEn	**C3**	0.68	0.80	0.30	0.79	0.27	**0.01**
	F3	0.32	0.52	0.69	0.24	0.54	0.27
	F7	0.27	0.48	0.42	0.56	0.63	0.12
	Fz	0.39	0.47	0.37	0.32	0.81	0.89
	O1	0.49	0.39	0.09	0.98	0.85	0.95
	O2	0.18	0.98	0.66	0.80	0.62	0.27
	P3	0.31	0.84	0.86	0.26	0.76	0.73
	P4	0.12	0.42	0.93	0.47	0.65	0.37
	T3	0.49	0.33	0.51	0.15	0.40	0.28
	T4	0.88	0.33	0.52	0.24	0.99	0.28
	T5	0.13	0.92	0.44	0.56	0.61	0.24
	T6	0.59	0.89	0.48	0.73	0.43	0.40
SpEn	C3	0.49	0.49	0.99	0.42	0.44	0.74
	F3	0.55	0.55	0.90	0.08	0.12	0.93
	F7	0.83	0.83	0.10	0.99	0.21	0.55
	Fz	0.56	0.56	0.30	0.42	0.85	0.69
	O1	0.56	0.56	0.39	0.90	0.56	0.57
	O2	0.96	0.96	0.94	0.16	0.14	0.85
	P3	0.40	0.40	0.69	0.53	0.89	0.12
	P4	0.72	0.72	0.57	0.69	0.55	0.47
	T3	0.60	0.60	0.74	0.26	0.62	0.05
	T4	0.65	0.65	0.78	0.88	0.24	0.73
	T5	0.69	0.69	0.75	0.14	0.26	**0.03**
	T6	0.78	0.78	0.68	0.96	0.73	0.11

parameters are presented in Appendix D. Table 7.2 shows that the significant P-values generally were found in the IMF9 component of the left cerebral hemisphere. Particularly, the PEn calculated from channel C3 and SpEn computed from channel T5 exhibited significant correlations with the cognitive scores.

Figure 7.3 (a) and (b) presents the correlation plots between the entropies and cognitive

TABLE 7.3: P-values of the correlation analysis of the ND of PLI features. Significant features are shown in boldface.

	Channel	IMF4	IMF5	IMF6	IMF7	IMF8	IMF9
ND of PLI	C3	0.35	0.29	0.47	0.09	0.99	0.82
	F3	0.26	0.54	0.60	0.48	0.97	0.13
	F7	0.17	0.30	0.20	0.50	0.29	0.23
	Fz	0.09	0.66	0.36	0.48	0.44	0.18
	O1	0.16	0.27	0.34	0.79	0.46	0.79
	O2	0.06	0.94	0.81	0.15	0.54	0.11
	P3	0.09	0.72	0.57	0.36	0.98	0.23
	P4	0.13	0.21	0.60	0.41	0.96	0.49
	T3	0.54	0.36	0.30	0.24	0.65	0.28
	T4	0.36	0.85	0.05	0.55	0.11	0.05
	T5	0.06	**0.03**	0.94	0.51	0.13	0.10
	T6	0.13	0.37	0.94	0.17	0.82	0.54

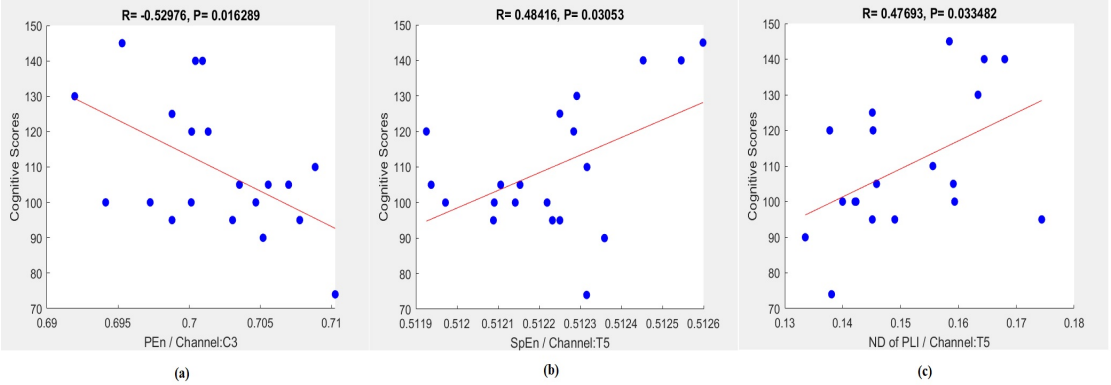


FIGURE 7.3: Scatter plots representing the correlation between cognitive scores and the significant features: (a) PEn, (b) SpEn and (c) ND of PLI.

scores. The figure shows that PEn of channel C3 exhibited a significant negative correlation of (R=-0.53, P=0.01). It also presents a high positive correlation demonstrated by the SpEn of channel T5 (R=0.48, P=0.03).

Correlation analysis between the ND of PLI (computed from each IMF and each channel) and the cognitive scores were also evaluated. The P-values of this analysis are presented in Table 7.3. The table shows that the ND of PLI extracted from IMF5 and channel T5 exhibited a significant P-value, indicating a high correlation with the cognitive scores. This correlation is also plotted in Figure 7.3 (c) showing positive dependency of (R=0.48, P=0.03).

7.5.2 Ensemble Regression Results

Table 7.4 presents the performance of LOOCV when the models were trained and tested on each of the significant features and their combination. It can be noticed from the

TABLE 7.4: Performance of the tree ensemble regression models using each significant nonlinear feature and their combination.

Features	RMSE	MAE	R-Squared	Regression algorithm
PEn (C3 and IMF9)	16.818	13.648	0.24	Bagged tree
	16.856	14.151	0.23	Boosted trees
SpEn (T5 and IMF9)	18.789	14.26	0.05	Bagged tree
	19.957	15.058	0.07	Boosted trees
ND of PLI (T5 and IMF5)	18.855	15.019	0.04	Bagged tree
	17.985	15.109	0.13	Boosted trees
PEn (C3 and IMF9), SpEn (T5 and IMF9) and ND of PLI (T5 and IMF5)	16.743	13.771	0.24	Bagged tree
	14.06	**11.876**	**0.47**	**Boosted trees**

table that the results were enhanced by combining the three significant features, i.e. when combining the EEG complexity and EEG synchronisation features.

To evaluate the error rate and gain a better understanding of whether it is acceptable, normalised RMSE can also be defined as the ratio of RMSE to the range or mean of the data where values closer to 0 represent good fitting models (Pal, 2017). Thus, based on the best performance and the mean of the cognitive scores, the normalised RMSE = 14.06/109.7=0.13, indicating an acceptable error rate.

Figure 7.4 gives the visualisation corresponding to the best result. The figure reveals small errors between the predicted and actual values for most individuals, indicating reasonable prediction values of the cognitive scores using the combination of the EEG complexity and EEG synchronisation features.

7.6 Discussion

This chapter explored the ability of the nonlinear analysis of EEG signals recorded from at-risk infants to predict their cognitive outcomes assessed at two years of age.

The significant challenge was the skewed distribution of the dataset, with most cognitive scores clustered above 95. Thus, tree ensembles regression models were adopted to deal with the imbalanced learning problem.

Correlation analysis was employed to assess the relationship between the nonlinear characteristics of at-birth EEGs and the cognitive scores. The results reveal a significant association in the delta-band component, represented by IMF9, between PEn and SpEn entropy features and the cognitive profile. Particularly, a significant positive correlation was observed between the SpEn extracted from the left-posterior brain area, corresponding to the T5 channel, and the cognitive scores. This correlation indicates that the lower SpEn is associated with a higher potential of cognitive impairments and vice versa. This result follows the general assumption of the signals' complexity reduction related to various pathological conditions (Catarino et al., 2011; Chu et al., 2017).

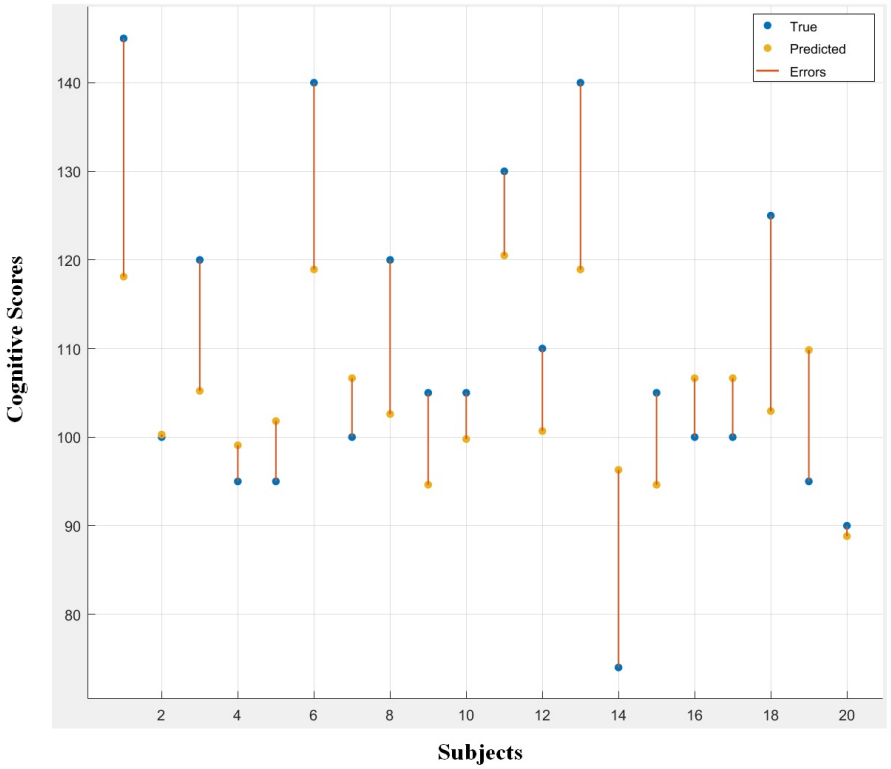


FIGURE 7.4: Response plot of the predicted cognitive scores versus the true ones.

On the other hand, PEn computed from the left-central lobe (corresponding to the C3 channel) was negatively correlated with the cognitive scores. This correlation suggests that the randomness behavior of the brain is negatively related to the level of cognitive function, i.e., an increase in the complexity of EEG signals leads to a reduction in cognitive outcomes. Even though this finding contradicts the assumption of decreased complexity in abnormal brain function, it is in line to some extent with Li et al. (2010) results where they reported an increase of ApEn in the abnormal EEG signals of neonates with HIE compared to normal term neonates. Although the reason for such inconsistency remains unclear, there is a general agreement that the cognition function of different pathological conditions involves alterations in the complex behavior of the brain.

The relationship between the delta-band characteristics and the cognitive function has been reported previously in several studies. Suppiej et al. (2017), for example, concluded that the high value of the delta power spectral in preterm infants correlated with poor cognitive outcomes assessed at one year of age. Increased delta activity in EEG of children suffering from learning disorders was also reported by Martnez-Briones et al. (2020) and further supported by Barttfeld et al. (2011) study, which suggested the difference in delta-band coherence between children having ASD and the control group.

The relationship between the ND of PLI features and the cognitive scores was also evaluated, and a positive dependency was observed through the ND of the left-posterior

region, corresponding to the T5 channel, computed at the alpha-band (represented by IMF5). Such a positive correlation suggests that the lower the degree of connectivity in this brain region, the lower the cognitive score and vice versa. This finding is in line with (Han et al., 2017) research which reported a lower left occipital-parietal PLI-based connectivity at the alpha-band in young ASD children compared to their typical peers. Moreover, (Han et al., 2017) proved that some of the connectivity parameters at the alpha-band were positively correlated with ASD symptom severity, indicating their potential application for early ASD prediction.

The correlation analyses also reveal the role of the left cerebral lobe, as all significant features were extracted from the signals recorded at the left hemisphere. Several studies, in addition to (Han et al., 2017), suggested the dependency of individuals' behavior on the neurophysiological processes of the left-brain area. For example, (Harmon-Jones et al., 2010) reported greater left than right asymmetric frontal cortical activity concerning social behavior. Paulus et al. (2013) showed that prosocial understanding is associated with higher left frontal cortical activation in two-year-old children. Moreover, (Kuhn-Popp et al., 2016) proved that coherence measures of the left hemisphere are the most important predictor of socio-cognitive skills. Hence, neural processes of the left hemisphere seem relevant to cognitive function development.

Tree ensemble regression models were used to evaluate the ability of the significant features to predict the cognitive outcomes. Hence, PEn and SpEn (at IMF9) and ND of PLI (at IMF 5), all computed from the left hemisphere, were used to train and test the regression models. The models were trained on each of these features as well as on their combination. The best performance was 14.06 RMSE (normalised RMSE was 0.2) and 0.47 R-squared using the boosted trees regression and the combination of all significant features. The methodology and findings of this investigation have been compared to the state-of-the-art research, which used qEEG analysis to predict cognitive outcomes. Table 7.5 summarises this comparison. The pros and cons of these existing studies have also been discussed in Section 2.6.3.

It can be observed from the table that Cainelli et al. (2021) and Suppiej et al. (2017) reported the correlation of early beta-, alpha- and delta-band with the later level of brain development, which (to some extent) is related to this exploration's findings. It can also be inferred from the table that the novelty of this investigation resides mainly in using nonlinear qEEG analysis to predict later cognitive outcomes using a machine learning framework besides traditional correlation analysis. All studies we were aware of adopted linear qEEG analysis, and an overall good predictive value has been achieved. However, linear-based features are not optimal to reveal the complex dynamics of the EEG signals. Moreover, state-of-the-art studies did not consider the adaptive time-frequency analysis of the signals. This limitation has been addressed using the NA-MEMD method.

TABLE 7.5: Comparison of the qEEG state-of-the-art methods employed for predicting cognitive outcomes.

Authors	Dataset	Features	Time-frequency	Evaluation method	Findings
(Lloyd et al., 2021)	57 preterm infants	EEG grading	None	Spearman's correlation coefficient	Moderate to large negative correlation between at-birth EEG grade and BSITD-III subscales.
(Suppiej et al., 2017)	21 preterm infants	Power spectral analysis	FFT	Spearman's correlation coefficient	Negative correlation between the delta spectral power and Griffiths scores developmental quotients (r=-0.68, P=0.015). Positive correlation between alpha and beta power spectral and Griffiths developmental quotients (R=0.61, P=0.032).
(Cainelli et al., 2021)	26 preterm infants	Power spectral analysis	FFT	Bayesian correlation	Significant associations between alpha-, beta-band and visual and auditory attention tests.
(West et al., 2005)	44 preterm infants	Amplitude	None	Linear regression	Positive correlation between mental developmental indices and continuity feature of EEG at different amplitude setting: 10 and 25 μV thresholds (R=0.19, P= 0.0032 and R =0.10, P=0.04 respectively).
(Kuhn-Popp et al., 2016)	32 infants	EEG coherence measures	FFT	Linear regression	Significant positive correlation between left hemisphere coherence and epistemic language at 48 months (R=0.59, P=0.003). Regression analyses showed, left-coherence scores are the most important predictor of epistemic state talk at 48 months
Current study	20 infants born with HIE	Entropies and ND of PLI	NA-MEMD	Pearson's linear correlation coefficient. and Tree ensembles regression models.	Significant correlation between PEn, SpEn, ND of PLI measured from left hemisphere and cognitive profile. Good regression performance of RMSE (14.06), MAE (11.87), and R-square (0.47) using the combination of the three significant features.

Nevertheless, the NA-MEMD method has a very high computational complexity, as discussed in the previous chapters (Sections 4.7 and 5.8). In addition, the analysis of the current chapter was conducted on a very small dataset of at-risk infants. Therefore, a further refinement of the proposed analysis with larger sample sizes is required to validate the findings.

7.7 Conclusion

This chapter investigated the potential of nonlinear analysis of at-birth EEG (recorded from infants with neonatal HIE) to predict cognitive outcomes at two years. Entropy features (SampEn, PEn, and SpEn) and ND of PLI were evaluated for this purpose using correlation analysis and a regression-based machine learning framework. Pearson's linear correlation showed significant correlations between PEn, SpEn, ND of PLI (calculated at the left hemisphere from alpha- and delta-band) and cognitive scores. The ensemble boosted trees regression model achieved reasonable prediction values of the cognitive scores using a combination of the nonlinear complexity and synchronisation features. Therefore, the findings suggest the potential of using the nonlinear EEG features derived from the left hemisphere and the alpha- and delta-band as biomarkers for early cognitive development prediction.

However, due to the limited number of samples used in this chapter, the proposed analysis needs to be validated on a larger dataset. Thus, a further investigation has been conducted in the next chapter to validate the suitability of the EEG synchronisation and complexity features computed from the NA-MEMD domain to serve as biomarkers for predicting cognitive scores using a bigger sample size of the MDD population.

Chapter 8

Depression Severity Prediction Based on Complexity and Synchronisation Analysis of the Electroencephalography Signals

This chapter aims to validate the method adopted in Chapter 7 on a much larger dataset to predict the cognitive outcomes from the EEG signals. In particular, it assesses the extent to which the EEG synchronisation (ND of PLI) and complexity (SampEn, PEn, and SpEn) features computed from the NA-MEMD domain could serve as potential biomarkers for predicting the cognitive outcomes from a bigger number of samples. To this end, a publicly available dataset of individuals with major depressive disorder (MDD) was used (Cavanagh and Allen, 2017). The dataset includes resting-state EEG recordings of depressed individuals associated with their Beck Depression Inventory (BDI) scores.

The clinical diagnosis of MDD patients includes all or some of the following symptoms: decreased pleasure or interest in daily activities, depressed or irritable mood, significant weight change, changes in activity patterns, changes to sleep patterns, fatigue or loss of energy, diminished concentration, feelings of guilt/worthlessness and thoughts of suicide (Palmer et al., 2015). BDI meets DSM-IV diagnostic criteria and has 21 items on the psychological, emotional, cognitive, and physical symptoms of a depressed person. Each item obtains a score from 0 to 3, and the total score of 21 items falls between 0 and 63 (Beck et al., 1996).

Brain imaging studies have attracted considerable attention recently, suggesting differences in induced task activation between individuals with MDD and controls when performing specified tasks (Cavanagh et al., 2019; Nelson et al., 2018; Webb et al., 2017).

The two groups also reported differences when measuring brain activities during a resting-state condition (Kaiser et al., 2016; Mohammadi and Moradi, 2021).

Brain activity characteristics have been used in different studies to classify the depression level as a binary concept between individuals with MDD and controls (Acharya et al., 2018; Bachmann et al., 2017, 2018; Cavanagh et al., 2019; Mohammadi et al., 2019). In addition, the literature has successfully reported the estimation of depression severity (using scores such as BDI) from fMRI (Yoshida et al., 2017), MEG (Jiang et al., 2016), and EEG signals (Mohammadi and Moradi, 2021). Some studies have also used non-neuronal features such as facial expression, speech and the combination of both video and audio recordings to predict depression severity (He and Cao, 2018; Jan et al., 2018).

This chapter investigates the possible association between SampEn, PEn, SpEn, and ND of PLI features, computed from the NA-MEMD domain, and the depression severity. These features are also used to train and test regression models to explore their practical ability to predict the BDI scores. Figure 8.1 illustrates this exploration methodology.

This chapter has been divided into seven sections: Section 8.1 describes the EEG dataset used in this investigation. Section 8.2 is then illustrated the feature extraction process. Correlation analysis is presented in Section 8.3, and the regression procedure is provided in Section 8.4. Section 8.5 presents the results, while Section 8.6 discusses them. Finally, Section 8.7 concludes the chapter.

8.1 Experimental Data Description

This exploration employed the dataset made available by Cavanagh and Allen (2017). Subjects included in this dataset were recruited from introductory psychology classes based on mass survey scores of BDI. Recruitment criteria involved (a) no history of head trauma or seizures, (b) no current psychoactive medication use, and (c) age 18 –25 years. A total of 46 (34 female) subjects were included in the high depressive symptomatology group, where they met the required criteria of having a stable high BDI ($\geq$13) between mass testing and experimental assessment. Out of those 46 participants, datasets of 44 subjects were used for the current investigation where two data files (ID numbers 571 and 572) were unavailable in the dataset (Cavanagh and Allen, 2017). Table 8.1 provides the BDI scores of the included subjects.

Scalp voltage was recorded from 64 electrodes using a Synamps^2 system (500 Hz sampling rate, band-pass filter 0.5 –100 Hz, impedances $< 10k\Omega$, and the online reference was a single channel placed between Cz and CPz). The dataset was preprocessed by identifying bad channels and subsequently interpolating and rejecting them, yielding 60 channels. Eye blinks were removed using ICA, and signals were referenced to averaged mastoids (Cavanagh and Allen, 2017; Cavanagh et al., 2019).

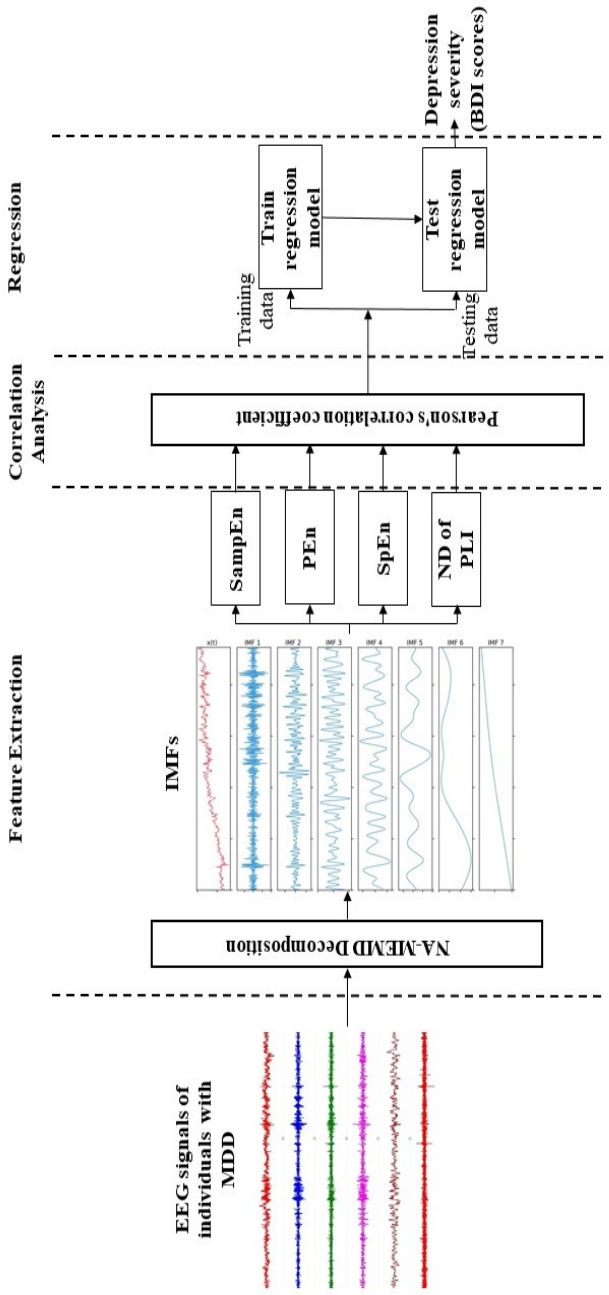


FIGURE 8.1: Block diagram of the proposed methodology for predicting the depression severity using the complexity and synchronisation features.

TABLE 8.1: BDI scores of the subjects

Subject #	Subject ID	BDI score
subj_1	558	29
subj_2	559	25
subj_3	561	27
subj_4	564	27
subj_5	565	24
subj_6	566	18
subj_7	567	24
subj_8	569	28
subj_9	586	22
subj_10	587	29
subj_11	590	22
subj_12	591	30
subj_13	592	27
subj_14	594	19
subj_15	595	20
subj_16	598	19
subj_17	602	27
subj_18	603	21
subj_19	604	22
subj_20	605	20
subj_21	606	20
subj_22	607	28
subj_23	608	23
subj_24	609	26
subj_25	610	15
subj_26	611	15
subj_27	612	25
subj_28	613	28
subj_29	614	19
subj_30	615	30
subj_31	616	16
subj_32	617	18
subj_33	618	24
subj_34	619	23
subj_35	620	27
subj_36	621	17
subj_37	622	14
subj_38	623	19
subj_39	624	23
subj_40	625	16
subj_41	626	14
subj_42	627	30
subj_43	628	19
subj_44	597	13

Thus, the dataset to be processed in this exploration included 44 subjects, each of them having resting-state clips (of approximately three minutes duration) recorded from 60 channels. However, analysing such high-dimensional signals is unsuitable for the NA-MEMD algorithm due to its high computational load discussed in Sections 4.7 and 5.8. Moreover, decomposing such large-scale signals may lead to unnecessary redundant decompositions, which in turn could bring interferences to the results (Chen et al., 2019; Zhang et al., 2021). Hence, To facilitate the analysis of the current investigation, a duration of one minute clip has been used, and a total of 19 electrodes (Fp1, Fp2, F3, F4, F7, F8, C3, C4, T7, T8, P3, P4, TP7, TP8, O1, O2, Fz, Cz, and Pz) have been selected. The locations of those electrodes have been chosen according to the 19 electrode placement of the 10/20 electrode settings.

8.2 Feature Extraction Procedure

1. The data points from all channels were combined for each subject to obtain a multivariate signal. Hence, a matrix of dimension $N_c \times N_t$ was constructed for each subject, where N_c denotes the number of channels (which is 19), and N_t indicates the number of temporal samples (which is 30000).

2. The NA-MEMD method was then applied to each matrix separately, as shown in Figure 8.2. In this way, the signals of each subject from all channels were decomposed into the same number of IMFs. Figure 8.3 gives an example of the resulting IMFs.

3. The frequencies of each IMF were then acquired by HT and plotted against amplitudes to inspect the signals. Consequently, IMF1 and IMF2 were found to contain very high frequencies (>60 Hz) across the channels of all subjects. Such highest frequency components make the data more complicated and are probably not physical (Huang et al., 1998). Therefore, these modes were excluded from further analysis. Components after IMF10 were also ignored as they contained very low frequencies (<0.5 Hz) or represented the residue mode. The scales of the remaining IMFs were localised approximately around the following: IMF3 (30 –55 Hz), IMF4 (14 –30 Hz), IMF5 (10 –13 Hz), IMF6 (7 –10 Hz), IMF7 (4 –6 Hz), IMF8 (2 –3 Hz), IMF9 (1 –2 Hz), and IMF10 (0.5 –1 Hz). Referring to the traditional brain waves (Sanei and Chambers, 2007), IMF3 to IMF5 frequencies belong to the gamma-, beta- and alpha-band, respectively. IMF6 and IMF7 frequencies belong to the theta-band, while IMF8, IMF9 and IMF10 belong to the delta brain wave.

4. After IMFs selection, the dataset dimension of each subject became: $N_c \times N_i \times N_t$, where N_c is the number of channels which is 19, N_i is the selected number of IMFs which is 8 (IMF3 –IMF10), and N_t is the number of the samples which is 30000.

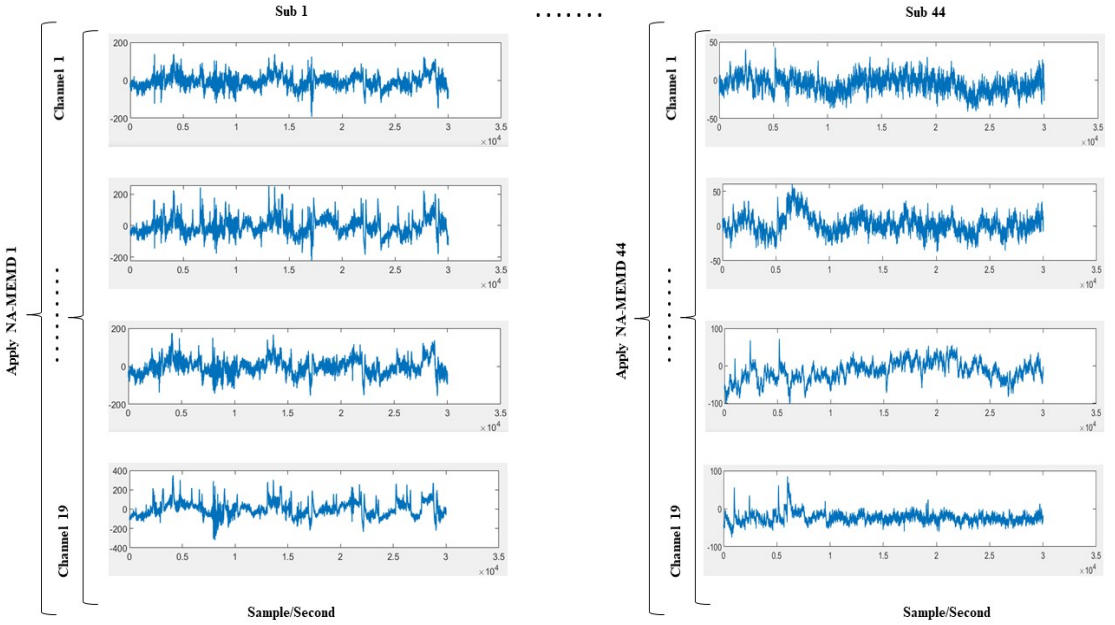


FIGURE 8.2: The proposed simultaneous decomposition method of the EEG signals.

5. Entropy features (SampEn, SpEn, and PEn) were then computed from each IMF of each channel using m=3 and r=0.2 for SampEn and m=3 and τ=1 for PEn estimations. Detailed descriptions of these measures are provided in Section 3.3.

6. ND of PLI features were also estimated from each IMF of each channel using Equation 3.9.

7. The extracted features were then used for the correlation analysis and to train and test the regression models.

The whole analysis was carried out in the MATLAB software package R2018a.

8.3 Correlation Analysis

Correlation analysis was employed to evaluate the relationship between the nonlinear EEG characteristics of individuals with MDD and their BDI scores. In particular, Pearson's correlation coefficient (R) was used to determine the linear dependency between the BDI scores and SampEn, PEn, SpEn and ND of PLI features computed for each channel and each IMF.

The BH-FDR method, described in Section 5.4, was employed to control the multiple comparisons, and the value of α was corrected from 0.05 to 0.04. Thus, the correlation was considered significant if the if P-value$<= \alpha(= 0.04)$. Significant features were

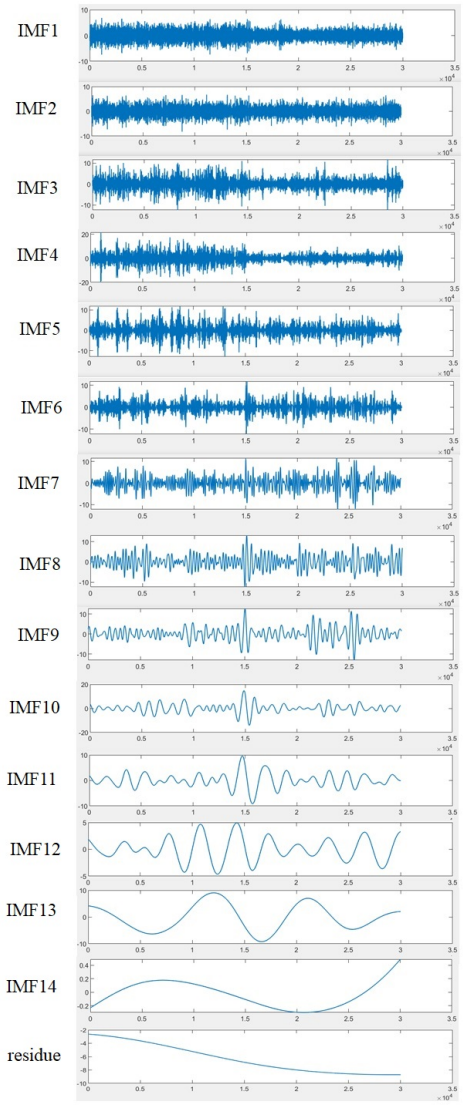


FIGURE 8.3: An example of the resulting IMFs from the NA-MEMD method.

then used to train and test the boosted and bagged regression models. The correlation analysis was carried out using the MATLAB statistics toolbox.

8.4 Regression Model

Ensemble boosted and bagged regression models, described in Section 3.6 and Appendix A, were used in the current investigation to predict the BDI scores from the EEG signals of MDD individuals. The models fit the relationship between the two sets of the nonlinear features (EEG complexity and EEG synchronisation) and the BDI scores with the least possible error. The training and testing of these models were performed using the *Regression Learner App* within the statistics and machine learning toolbox in MATLAB.

The regression models were trained and tested on different vectors constructed from the significant features (selected based on the correlation analysis), and LOOCV was used to prevent overfitting. RMSE and MAE, illustrated in Section 7.4, were used to evaluate the models' performance. The correlation coefficient (R) between actual BDI and predicted scores was also estimated to allow comparison with state-of-the-art findings.

8.5 Results

8.5.1 Correlation Analysis Results

In this exploration, the correlation analysis was estimated between the entropy features of the subjects, computed from each channel and each IMF, and their BDI scores. The SpEn metric reveals a positive relationship in the IMF4 component (which approximately corresponded to beta-band as illustrated in Section 8.2) of some channels (C3 (R=0.31, P=0.04), C4 (R=0.31, P=0.04), and Pz (R=0.30, P=0.04)). Table 8.2 presents all P-values of SpEn correlations, while the P-values of SampEn and PEn are provided in Appendix E. Figure 8.4 presents the plots of the significant correlations of the SpEn feature.

The relationship between the ND of PLI (computed from each channel and each IMF) and the BDI scores were also evaluated. Table 8.3 shows that the ND of PLI values estimated from IMF4 of different channels exhibited significant P-values, indicating high negative correlations with the BDI scores. Such correlations include C4 (R=-0.34, P=0.02), F4 (R=-0.38, P=0.01), F8 (R=-0.40, P=0.01), O2 (R=-0.30, P=0.04), P4 (R=-0.33, P=0.03), Pz (R=-0.39, P=0.01), and Tz (R=-0.40, P=0.01), TPz (R=-0.31, P=0.04). Noticeable positive correlations between ND of PLI features and BDI scores have also found in the IMF10 component (corresponded to delta-band) of different channels. This relationships include C3 (R=0.35, P=0.02), CZ (R=0.32, P=0.03), F7 (R=0.42, P=0.005), F8 (R=0.33, P=0.03), Fz (R=0.33, P=0.03), FP1 (R=0.37, P=0.01), FP2 (R=0.40, P=0.01), and Pz (R=0.36, P=0.02). Figure 8.5 shows correlation plots of some channels as examples to represent these relationships.

8.5.2 Ensemble Regression Results

Ensemble boosted and bagged trees regression models were trained and tested on the significant features estimated from IMF4 and IMF10. Thus, the selected features include the following: (from IMF4) SpEn: C3, C4, and Pz; ND of PLI: C4, F4, F8, O2, P4, Pz, T7, and TP7; (from IMF10) ND of PLI: C3, Cz, F7, F8, Fz, FP1, FP2, and Pz. Table 8.4 provides the LOOCV performance of the regression models. Boosted trees obtained

TABLE 8.2: P-values of the correlation analysis of the SpEn features.

Channel	IMF3	IMF4	IMF5	IMF6	IMF7	IMF8	IMF9	IMF10
C3	0.86	0.04	0.18	0.22	0.20	0.62	0.99	0.99
C4	0.83	0.04	0.47	0.16	0.33	0.63	0.97	0.97
Cz	0.76	0.29	0.80	0.13	0.96	0.90	0.78	0.63
F3	0.83	0.35	0.43	0.66	0.30	0.98	0.99	0.77
F4	0.86	0.14	0.51	0.80	0.21	0.87	0.75	0.98
F7	0.93	0.14	0.14	0.73	0.63	0.62	0.90	0.60
F8	0.88	0.27	0.10	0.63	0.38	0.73	0.94	0.58
FZ	0.94	0.21	0.49	0.66	0.23	0.92	0.73	0.94
FP1	0.60	0.32	0.72	0.05	0.73	0.57	0.95	0.80
FP2	0.58	0.19	0.49	0.06	0.51	0.69	0.83	0.95
O1	0.64	0.07	0.27	0.21	0.79	0.72	0.49	0.63
O2	0.73	0.06	0.73	0.48	0.86	0.89	0.27	0.81
P3	0.84	0.05	0.09	0.05	0.46	0.93	0.96	0.71
P4	0.94	0.05	0.22	0.08	0.54	0.74	0.80	0.59
Pz	0.84	0.04	0.11	0.05	0.71	0.91	0.80	0.61
T7	0.58	0.45	0.15	0.26	0.25	0.97	0.74	0.62
T8	0.32	0.10	0.27	0.41	0.55	0.99	0.84	0.83
TP7	0.97	0.15	0.12	0.05	0.40	0.69	0.51	0.57
TP8	0.80	0.06	0.32	0.10	0.64	0.72	0.66	0.76

TABLE 8.3: P-values of the correlation analysis of the ND of PLI features.

Channel	IMF3	IMF4	IMF5	IMF6	IMF7	IMF8	IMF9	IMF10
C3	0.39	0.19	0.80	0.78	0.43	0.28	0.39	0.02
C4	0.96	0.02	0.19	0.21	0.21	0.99	0.53	0.19
Cz	0.25	0.08	0.39	0.15	0.81	0.62	0.59	0.03
F3	0.92	0.19	0.50	0.95	0.23	0.77	0.33	0.16
F4	0.37	0.01	0.10	0.43	0.66	0.65	0.40	0.22
F7	0.64	0.30	0.35	0.08	0.77	0.75	0.06	0.00
F8	0.14	0.01	0.12	0.12	0.43	0.83	0.34	0.03
Fz	0.80	0.56	0.36	0.70	0.51	0.44	0.32	0.03
FP1	0.16	0.57	0.84	0.69	0.34	0.43	0.79	0.01
FP2	0.97	0.26	0.20	0.78	0.44	0.77	0.56	0.01
O1	0.57	0.05	0.47	0.92	0.17	0.73	0.51	0.31
O2	0.19	0.04	0.08	0.44	0.21	0.06	0.16	0.85
P3	0.74	0.12	0.55	0.99	0.53	0.90	0.03	0.33
P4	0.77	0.03	0.01	0.23	0.66	0.46	0.27	0.05
Pz	0.71	0.01	0.19	0.38	0.65	0.93	0.27	0.02
T7	0.54	0.01	0.13	0.76	0.08	0.11	0.16	0.21
T8	0.68	0.50	0.04	0.42	0.07	0.83	0.32	0.91
TP7	0.68	0.04	0.77	0.21	0.07	0.05	0.04	0.35
TP8	0.51	0.16	0.08	0.06	0.61	0.63	0.16	0.32

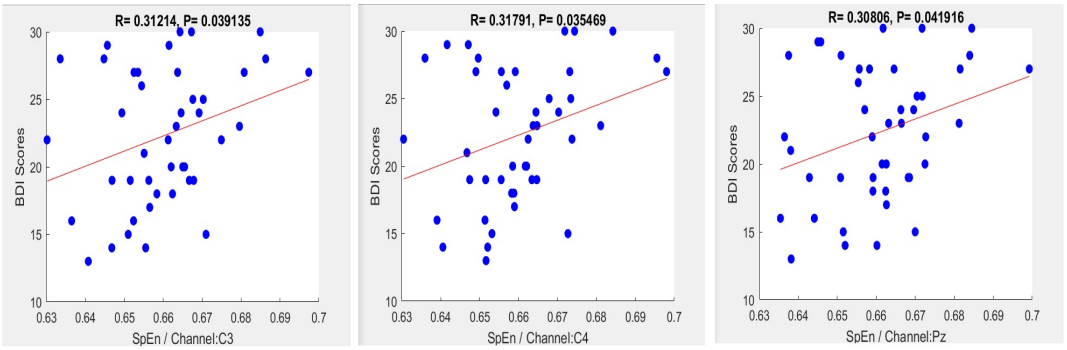


FIGURE 8.4: Scatter plots representing the correlations between the BDI scores and SpEn features estimated from IMF4.

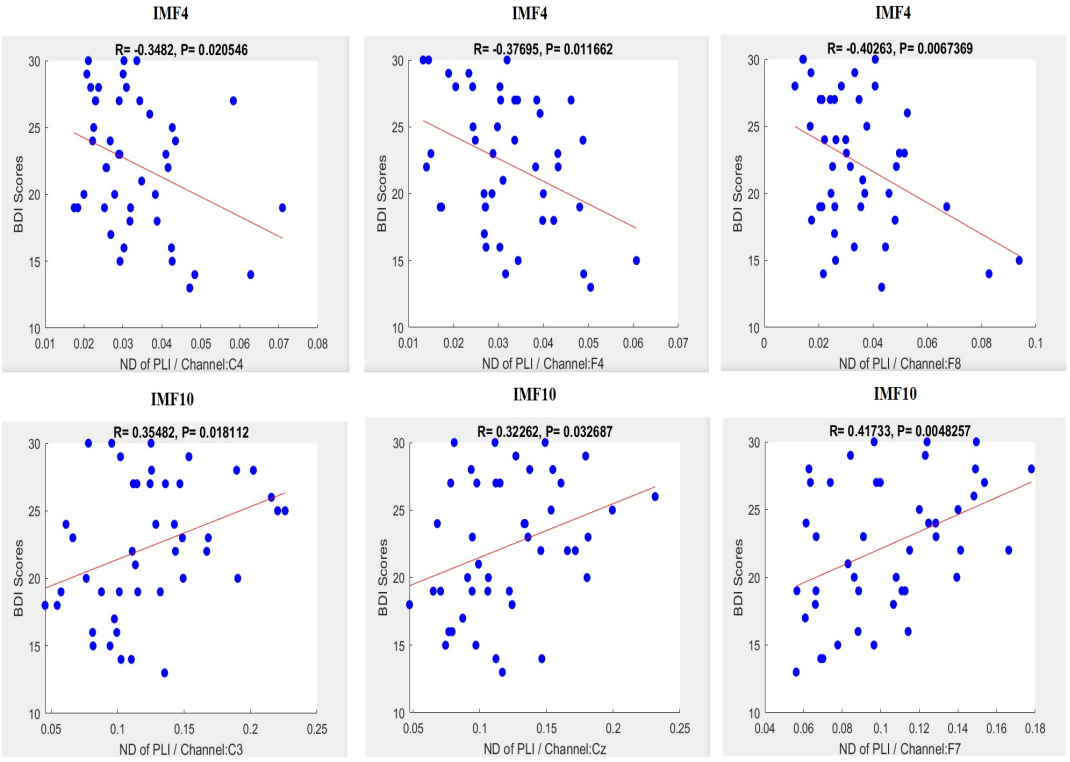


FIGURE 8.5: Scatter plots representing examples of the correlations between the BDI scores and ND of PLI features estimated from IMF4 and IMF10.

a better performance of RMSE (4.60) and MAE (3.72), while the correlation between actual BDI and predicted scores was $R = 0.42$; $P < 0.004$. Figure 8.6 shows the response plot corresponding to this result. The plot displays the predicted response (BDI scores) versus the subject number. The plot also illustrates the prediction errors, drawn as vertical lines between the predicted and true response.

TABLE 8.4: LOOCV performance of the regression models using significant nonlinear features.

Regression algorithm	RMSE	MAE	R
Bagged tree	4.91	4.07	0.32
Boosted tree	**4.60**	**3.72**	**0.42**

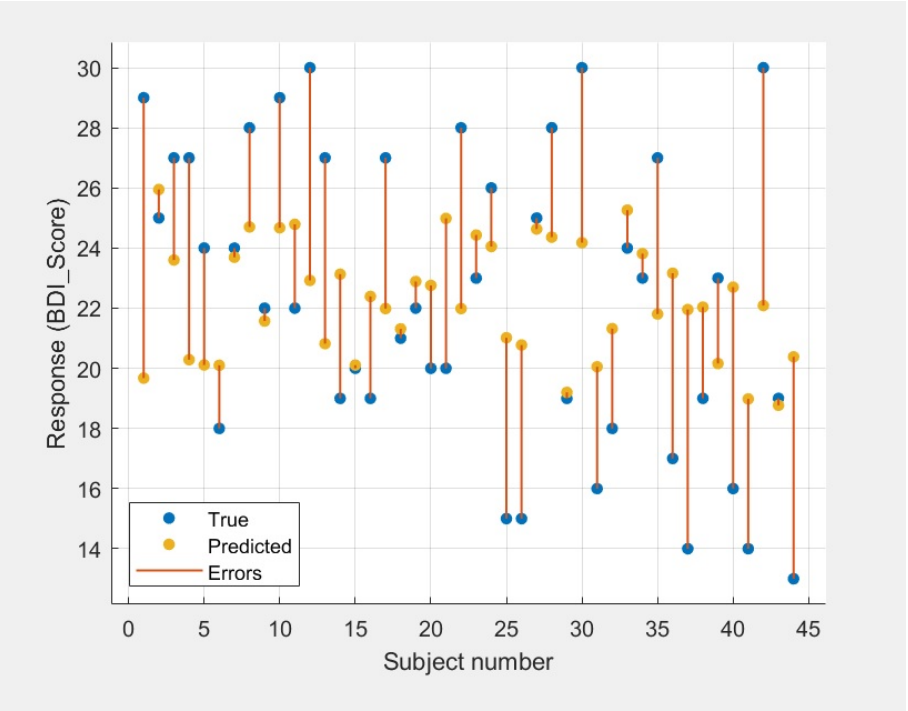


FIGURE 8.6: Response plot of the predicted BDI scores versus the subject number of the boosted tree model.

8.6 Discussion

This chapter investigated the efficacy of the EEG complexity and synchronisation features computed from the NA-MEMD domain to predict the BDI scores of individuals with MDD.

Correlation analysis was used to evaluate the relationship between the nonlinear characteristics of resting-state EEGs and the depression severity. The results reveal a significant positive correlation in the IMF4 component between SpEn entropy features and the BDI scores. This correlation indicates that as the level of depression increases, complexity and random behavior in the brain system increases too. Such a result is in line with (Mohammadi and Moradi, 2021) finding which also reported a significant positive relationship between complexity features and BDI scores of depressed individuals. Nevertheless, Mohammadi and Moradi (2021) found this correlation in the alpha-band while the current exploration identified it in IMF4, which approximately corresponded to the beta-band.

This discrepancy could be attributed to different decomposition methods used. While Mohammadi and Moradi (2021) used band-pass filters to decompose EEG signals into the traditional frequency bands, including delta (1-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), beta (13-30 Hz)), and gamma (30-45 Hz), this study used NA-MEMD which adaptively decompose the EEG including all oscillations involved in the signal. In addition, our findings could be, to some extent, associated with those of Mohammadi and Moradi (2021) where peaks of the beta wave in frequencies up to 20 Hz are known to have characteristics of an alpha-band state rather than ones for a beta-band (Sanei and Chambers, 2007).

The relationship between the ND of PLI features and the BDI scores was also estimated, and a negative dependency was observed through the ND of different channels, also computed at the IMF4 component. Such a negative correlation suggests that the lower the degree of connectivity, the higher the depression severity and vice versa. This finding is also in agreement with (Mohammadi and Moradi, 2021) research which reported a negative relationship between the ND of PLI features and depression severity. However, this correlation was also reported by Mohammadi and Moradi (2021) at the alpha-band while the current study found it in the beta-band.

In addition, a noticeable positive dependency between the ND of PLI features and the BDI scores was found in the IMF10 component, which corresponded to the delta brain wave. Leuchter et al. (2012) reported a relevant finding of the EEG analysis on depression, indicating that overall functional connectivity (measured by coherence) increases in patients with MDD compared with controls in the delta-band.

Ensemble boosted and bagged trees were used to evaluate the ability of the significant features to predict the depression severity. Thus, significant SpEn (at IMF4) and significant ND of PLI (at IMF4 and IMF10) were used to train and test the models. The boosted trees obtained the best performance reaching 4.60 RMSE, 3.72 MAE, and a correlation coefficient of $R = 0.42$ ($P < 0.004$) between actual BDI and predicted scores. This result has been compared with the previous studies carried out for predicting the depression scores of MDD individuals.

To the best of our knowledge, the recent exploration of Mohammadi and Moradi (2021) is the first research that estimates the depression severity using EEG signals. Their study is comparable to the current investigation, computing the global clustering coefficient (Rubinov and Sporns, 2010) and ND of PLI to quantify the functional connectivity of EEGs and Lempel-Ziv complexity (Aboy et al., 2006) and fuzzy entropy (Chen et al., 2007) to measure the complexity of the signals. However, they computed the synchronisation and complexity features from the brain waves estimated using traditional filter cut-offs, while similar features were calculated in this study from the IMF components resulting from the NA-MEMD method. Other studies, which we are aware of, for predicting the depression scores included adopting fMRI (Yoshida et al., 2017), MEG

TABLE 8.5: Comparison of state-of-the-art studies in the prediction of depression severity. "–" represents that the studies did not report the measures.

Study	Data	RMSE/MAE	normalised RMSE	R
(Yoshida et al., 2017)	fMRI	9.56/–	0.31	0.54
(Jiang et al., 2016)	MEG	–	–	0.68
(He and Cao, 2018)	Audio	9.99/8.19	–	–
Jan et al. (2018)	Visual	8.01/ 6.68	–	–
	Audio Visual	7.43/ 6.14	–	–
(Mohammadi and Moradi, 2021)	EEG	7.69/ 6.11	0.30	0.84
Current study	EEG	4.60/ 3.72	0.20	0.42

(Jiang et al., 2016), and non-neuronal features, including facial and speech expressions (He and Cao, 2018; Jan et al., 2018). Table 8.5 compares the results of these studies. The table shows that the current method provided acceptable results compared with the state-of-the-art research. Particularly, the boosted trees gained lower RMSE and MAE compared to earlier findings.

However, the boosted trees model achieved a lower correlation coefficient (R) between the predicted and actual scores than the previous studies. Nevertheless, one pitfall of correlation coefficient (R) is that it increases as predictors are added to the regression model, and this increase could be artificial when predictors are not actually improving the model's fit (Pal, 2017). Thus, RMSE is preferred for comparing the performance among different regression models. RMSE may also not be appropriate to compare different models that differ in the numbers and ranges of samples (Pal, 2017). To handle this issue, normalised RMSE, described in Section 7.5.2, was calculated for the boosted trees model of the current exploration and also for Mohammadi and Moradi (2021) and (Yoshida et al., 2017) based on the information they provided. The normalised RMSE results are provided in Table 8.5, indicating a lower error rate for the current method compared with Mohammadi and Moradi (2021) and (Yoshida et al., 2017) studies.

Overall, the results indicate that synchronisation and complexity features can characterise abnormal EEG activity, aiding in predicting depression severity. Estimating these features from the NA-MEMD domain further contributed to the analysis where the NA-MEMD method decomposes the signals adaptively and does not require a *priori* selection of the filter cut-offs. This property is significant because it can tackle the well-known frequency range variability between subjects. Furthermore, the adaptivity of the NA-MEMD method is important for both nonlinear and nonstationary properties of the EEG in which the decomposition is adapted to the local variations and hence can fully account for the underlying dynamics of the signals.

Nevertheless, the multivariate EMD algorithms are generally limited by the high computational complexity, which is prohibitive for analysing large-scale and complex signals (Zhang et al., 2021). The computational load is further increased with the NA-MEMD method due to the requirement of an additional subspace of multivariate independent white noise equalling the original multivariate signal. Although some algorithms have been proposed recently to reduce such computational loads, they still need improvements to make them reliable for a better decomposition performance (Zhang et al., 2021).

8.7 Conclusion

The exploration of this chapter was conducted to validate the method used in the previous chapter on a much larger dataset. Thus, an available dataset of depressed individuals was used to investigate the potential of nonlinear analysis of EEG to predict their BDI scores. Mainly, complexity features (SampEn, PEn, and SpEn) and ND of PLI estimated from the NA-MEMD domain were evaluated using correlation analysis and a regression-based machine learning framework. Pearson's linear correlation showed significant correlations between SpEn, ND of PLI (calculated from beta- and delta-band) and BDI scores. The ensemble boosted trees regression model achieved acceptable prediction performance compared to the state-of-the-art findings. Therefore, the results validated the potential of the proposed nonlinear analysis method to be used for depression prediction.

Chapter 9

Conclusions and Future Directions

This thesis was concerned with employing advanced nonlinear EEG analysis methods to identify biomarkers that could aid in NDDs prediction. It introduced analytical tools that combine sophisticated nonlinear signal processing methods to reveal abnormal brain activities in ASD children and at-risk infants (following HIE) some of whom developed CP by two years. The cognitive outcome of those infants was also considered. An additional investigation was conducted at the end of the thesis to validate the suitability of the proposed method to predict cognitive scores using a much larger dataset. An available dataset of depressed individuals and their depression severity scores was used for this purpose due to the limited availability of at-risk infants' datasets.

The work done through the thesis contributed to each of these strands, as discussed in the following section.

9.1 Summary of the Results

In Chapter 4 of this thesis, a machine learning framework was developed using a benchmark dataset of ASD children and their neurotypical peers recorded during FE recognition tasks. This dataset was initially used to validate the ability of the nonlinear analysis of EEG signals to reveal abnormal brain function. The features were calculated in the MEMD domain to adaptively consider the tempo-spectral activities underlying the EEG signals. Statistical analysis and different low complexity classifiers were used to assess the ability of the proposed features to discriminate ASD from TD children. High classification performance was reached using the combination of the linear and nonlinear features. Furthermore, EEG complexity of the alpha- and theta-band and the linear features of the delta-band were suggested as biomarkers of impaired FE processing in children with autism.

Due to the success of the framework proposed in Chapter 4, it was used as a foundation for the rest of the thesis. Thus, Chapter 5 employed a similar method on EEG signals of at-risk infants (born with HIE) to early identify their CP outcome. The nonlinear complexity features in the NA-MEMD domain were evaluated using statistical analysis and RUSBoost classifier. Good performance of 84.6% ACC, 83% TPR, 85%TNR and 0.87 AUC was reached for the classification of the infants who developed CP. Additionally, complexity features of the alpha-band were suggested as potential biomarkers for early CP prediction.

To study the abnormal brain functions related to the CP condition from another angle, Chapter 6 used nonlinear time series analysis to assess the synchronisation between EEG signals of different brain regions. The chapter also evaluated the relationship between the complexity and the synchronisation measures and confirmed their association. Additionally, a combination of these two classes of features was used to train and test the RUSBoost classifier, which reached a performance of 96.2% ACC, 100% TPR, 95% TNR, and 1.00 AUC, surpassing state-of-the-art research results. Thus, the combination of complexity and synchronisation features of beta-, alpha-, and delta-band were suggested as biomarkers for the early identification of CP.

Chapter 7 explored the potential usefulness of the proposed framework, including the complexity and synchronisation features, in predicting the cognitive outcome of at-risk infants. The features were initially evaluated for this purpose using correlation analysis. Significant correlations between the nonlinear features and the cognitive scores were identified. Following these correlations, two ensemble-based regression models were trained and tested on a combination of the significant features. Acceptable performance of RMSE (14.06)/normalised RMSE (0.13), MAE (11.87), and R-square (0.47) was reached through this exploration. The findings suggest that nonlinear features derived from the left hemisphere at the alpha- and delta-band can be considered biomarkers for early cognitive function prediction.

Due to the limited dataset used in Chapter 7, Chapter 8 conducted further analysis to validate the proposed method of using EEG complexity and synchronisation features estimated from the NA-MEMD domain to predict cognitive scores with larger sample size. To this end, an available dataset of MDD patients and their scores of depression severity was used. Correlation analysis indicated significant correlations between the nonlinear features and BDI scores at the beta- and delta-band. Following these correlations, boosted and bagged trees regression models were trained and tested on the significant features, and comparable performance with the state-of-the-art studies was reached, achieving 4.60 RMSE/0.2 normalised RMSE, 3.72 MAE, and a correlation coefficient of $R = 0.42$ between actual BDI and predicted scores. Thus, the findings validated the potential of the proposed method to analyse much larger samples of MDD individuals and predict their depression severity with acceptable performance.

Overall, the proposed machine learning framework contributed to the identification of brain function associated with different brain disorders. The novelty relied mainly on the fact that it allowed the nonlinear analysis from different perspectives combining the local complexity of the signals and the synchronisation behavior between them. Additionally, both nonlinear characteristics were observed in the EMD-based domain, allowing nonlinear and adaptive tempo-spectral analysis. Such a framework could be used alongside other diagnostic techniques to assist clinicians with an earlier prediction of CP and cognitive decline in at-risk populations, aiding in establishing tailored intervention programs at an early stage to improve the outcome. It could also be used in detecting ASD and estimating depression levels as the first step towards intervention/treatment.

9.2 Future Work

The work described in this thesis provides a novel framework and approach that can further be exploited and expanded to study functional brain dynamics in a vast array of brain disorders. Future prospects are outlined as follows:

- **From data perspective**

  The framework proposed and described in this thesis was tested on the dataset of at-risk infants with HIE to classify their CP or normal neuromotor outcome (a total of 26 infants) and predict their cognitive scores (20 infants) assessed at 24 months. It was also validated on two available datasets: a dataset of 24 subjects to classify children with ASD from their neurotypical peers; a dataset of 44 MDD individuals to predict their depression severity. Although the feasibility of the framework was positively assessed on these datasets, further assessment on a larger trial and a bigger number of samples of at-risk infants should be conducted to establish the method's practical validity in early predicting this population (which is the main aim of this thesis) before putting it into clinical practice.

  Furthermore, since the examined disorders are heterogeneous conditions including different subtypes and degrees of impairments, the datasets could be divided into subgroups based on their characteristics. Following this scenario, it would be interesting to assess the efficacy of the proposed approach for predicting different subgroups of disorders.

- **From methodology perspective**

  Although the proposed framework achieved good results using specific EEG metrics, including SampEn, PEn, SpEn and ND of PLI, investigating other advanced nonlinear measures is of great interest. For instance, eigenvector centrality and

clustering coefficient could be used to assess functional interaction between signals and Lempel-Ziv complexity and fuzzy entropy to quantify the signals' complexity. Selecting the most suitable metric is generally an empirical question, and thus, interrogating the data using several measures could be explored to study brain dynamics and understand how the estimates relate to one another.

Another possible area of future research would be investigating the effective (directed) connectivity between signals and localising them in the time-frequency domain via the EMD-based methods. Such an approach would aid in identifying the channels having stronger directed activities among other channels in the typical cases; and, hence, providing good localisation of the brain abnormalities following different NDDs.

- **From application perspective**

  The approach proposed in this thesis can be employed to explore the neurological properties of other NDDs such as ADHD and epilepsy. Such explorations would allow investigating the characteristics of these disorders and potentially identify the commonalities and differences between them.

Appendix A

Machine learning algorithms

A.1 Classification algorithms

A.1.1 DA

DA is a classification approach that assumes a Gaussian distribution of the data from each class and provides class separability by drawing a decision region between the different classes. In order to train the classifier, the fitting function computes the values of the distributions for the different classes. The trained classifier then predicts the response by maximising the ratio of the between-class variance and the within-class variance. Two commonly used DA models are LDA and QDA.

In LDA, the model assumes the same covariance matrix for each class but different means. Given M classes and the set of training samples represented by their feature vectors, for each class c_m, the model calculates the sample mean μ_m and covariance matrix Σ_m. The within-class scatter matrix S_w is then obtained as follows:

$$S_w = \sum_{m=1}^{M} \Sigma_m \tag{A.1}$$

To compute the between-class scatter matrix, the model first calculates the mean of the entire dataset μ and subtracts it from the centers of each class. Let B be a matrix keeping distance between class means and the mean of the entire dataset, $B = [\mu_1 - \mu, \mu_2 - \mu, \ldots, \mu_M - \mu]$. Then, the between-class matrix is then defined as:

$$S_b = \frac{1}{M} BB^T \tag{A.2}$$

The optimal projection matrix V is the one that maximise the ratio between the between-class scatter matrix and the within-class scatter matrix as follows:

$$\max_V J(V) = \frac{V^T S_b V}{V^T S_w V} \tag{A.3}$$

In QDA, however, the model assumes a different covariance matrix and means for each class. The model takes an empirical covariance matrix of each class by creating weighted classifiers for the different classes. Suppose N is a $K - by - M$ observations matrix, where M is the number of classes, and K is the number of observations. $N_{ki} = 1$ if observation k belongs to class c_i, and it is 0 otherwise. Thus, the mean is given by:

$$\widehat{\mu_m} = \frac{\sum_{k=1}^{K} N_{km} w_k x_k}{\sum_{k=1}^{K} N_{km} w_m} \tag{A.4}$$

where w_k are the weights. The estimate of the covariance matrix is given by:

$$\widehat{\Sigma} = \frac{\sum_{k=1}^{K} \sum_{k=1}^{K} N_{km} \left(x_k - \widehat{\mu_m}\right) \left(x_k - \widehat{\mu_m}\right)^T}{1 - \sum_{m=1}^{M} \frac{W_m^2}{W_m}} \tag{A.5}$$

where W_m is the sum of the weights for class m and W_m^2 is the sum of the squared weights for the same class (Mohanty et al., 2013; Halabi et al., 2020).

A.1.2 SVM

SVM is a well-known supervised classifier that defines the maximum class separation between two classes, known as the maximum margin hyperplane. The algorithm maps the observations into a feature space that can be separated using linear or nonlinear decision boundaries, depending on the kernel. It tries to maximise the distance between the separating hyperplane and the nearest training point(s) (the support vectors). The separating hyperplane in a two-dimension feature space which is given by:

$$f(x) = \left(x^T r + b\right)^d \tag{A.6}$$

where $d = 1$ for the linear kernel, and $d > 1$ for the nonlinear kernels. $r, x \in \mathbb{R}^2$ and $b \in \mathbb{R}^1$. The optimal solution r that maximising the distance between the hyperplane and the nearest training point(s) can be obtained by minimising the cost function

$$L(r, \xi) = \frac{1}{2}\|r\|^2 + C \cdot \sum_{m=1}^{Z} \xi_m \tag{A.7}$$

while satisfying the constraints:

$$\begin{gathered}(x_m r + b) \geq 1 - \xi_m \text{ for } y_m = +1 \\ (x_m r + b) \geq -1 + \xi_m \text{ for } y_m = -1 \\ \xi_m \geq 0 \ \forall m\end{gathered} \tag{A.8}$$

where $\|r\|^2 = r^T r$, C is a positive regularisation parameter selected by the user (a large value of C indicates a high penalty for classification errors), ξ_m is a measure of training error, Z is the number of misclassified samples, and y_m is the class label ($+1$ or 1 in the case of binary classification) for the $m - th$ sample (Mohanty et al., 2013; Hong et al., 2018).

A.1.3 kNN

kNN is a classification algorithm that assumes the similarity between new samples (from the testing set) and available samples (the training set) and classifies them based on this similarity. kNN estimates the similarity using different distance metrics, such as Euclidean, cosine and cubic metrics (Halabi et al., 2020). To calculate the distance for n-dimensional vectors, these metrics are given by:

$$\text{Euclidian: } d(x,y) = \sqrt{\sum_{i=1}^{n} (x_{i-} y_i)^2} \tag{A.9}$$

$$\text{Cubic: } d(x,y) = \sqrt[3]{\sum_{i=1}^{n} (x_{i-} y_i)^3} \tag{A.10}$$

$$\text{Cosine: } d(x,y) = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i^2} \tag{A.11}$$

kNN algorithm is summarised in the following:

1. Initialise k to a chosen number of neighbors.
2. For each sample in the data:
   (a) Calculate the distance between the test sample and the available sample.
   (b) Add the distance and the index of the sample to an ordered collection.
3. Sort the ordered collection of distances and indices from smallest to largest (in ascending order) by the distances.
4. Pick the first k entries from the sorted collection.

5. Get the labels of the selected k entries.

6. Return the mode of the k labels.

A.1.4 RUSBoost

RUSBoost is a hybrid sampling/boosting algorithm that is mainly well-suited to deal with the class imbalance problem.. It uses the random undersampling method that randomly removes samples from the majority class until the desired balance is achieved. The steps of RUSBoost are illustrated in Algorithm A.1.4.

Weak learner classification algorithm (*WeakLearn*) (Seiffert et al., 2010) is invoked with S_t, which returns a hypothesis h_t in each iteration. Then, the pseudo-error ϵ_t is calculated. The weight-update parameter α_t is then computed, and the distribution for the next iteration D_{t+1} is updated and normalised afterward. The final hypothesis $h_{final(x)}$ is returned from the algorithm representing a weighted vote of the T weak hypotheses. The decision tree classifiers are often used as *WeakLearn*, and hence the overall performance of the algorithm relies on the number of trees (Hassan and Bhuiyan, 2017; Seiffert et al., 2010).

Algorithm A.1.4 : RUSBoost.

Input: N instances $S = [(x_i, y_i)]$ of C classes with labels $y_i \in \omega, \omega = \{\omega_1, \omega_2, \ldots, \omega_C\}$ and $i = 1, 2, 3, \ldots, N$.
Weak learner algorithm, *WeakLearn*.
Integer T specifying the number of iterations.
Let, $B = \{(i, y) : i \in 1, 2, \ldots, m, y \neq y_i\}$
Initialise $D_1(i, y) = 1/|B|$ for $(i, y) \in B$
for $t = 1, 2, 3, \ldots, T$ **do**

1. Create temporary training data-set $S_t$with distribution D_t by random undersampling.
2. Call WeakLearn, providing it with S_t and distribution D_t.
3. Get back a hypothesis $h_t : X \times Y \to [0, 1]$
4. Calculate the pseudo-error ϵ_t of h_T :

$$\epsilon_t = \sum_{(i,y):y_i \neq y} D_t(i) \left(1 - h_t\left(x_i, y_i\right) + h_t\left(x_i, y\right)\right)$$

5. Compute the weight-update parameter, $\alpha_t = \epsilon_t / \left(1 - \epsilon_t\right)$.
6. Update distribution $D_t : D_{t+1}(i) = \frac{D_t(i)}{Z_t} \alpha_t^{\frac{1}{2}(1 + h_t(x_i, y_i) - h_t(x_i, y))}$ where $y_i \neq y$ and $Z_t = \sum D_t(i)$ is a normalisation constant chosen so that D_{t+1} becomes a distribution function.

end

Output: The final hypothesis:

$$h_{\text{final}}\left(x\right) = \arg\max_{y \in Y} \sum_{t=1}^{T} \left(\log \frac{1}{\alpha_t}\right) \times h_t(x, y)$$

A.2 Ensemble regression models

A.2.1 Bagging

Bagging is an ensemble learning strategy that combines the decision trees for a more robust outcome in either regression or classification techniques. It uses bootstrap statistical resampling to randomly generate subsets of the same size as the initial training set. Each subset is then treated as a new training set for the base learning algorithm. Homogeneous base-learners are generated in parallel as the subsets are sampled. Each learner within the Bagging predicts each case. These predictions are combined to make the final decision using the majority vote. The Bagging algorithm is given in Algorithm A.2.1 (Hu et al., 2021).

Algorithm A.2.1 : Bagging.

Input:

Data set $D = \{(x_1, y_1), (x_2, y_2), \cdots, (x_M, y_M)\}$.

Base learning algorithm L.

Number of learning round T.

Process:

For $t = 1, 2, \cdots, T$:

$D_t =$ booststrap samlple (D). #Generate a bootstrap sample from D

$h_t = L(D_t)$. #Train a base learner h_t from the a bootstrap sample

end

Output: $H(x) = \arg\max_y \sum_{t=1}^{T} 1(y = h_t(x))$

A.2.2 Boosting

Boosting improves predictive performance by training homogeneous base-learners (often decision trees) on a sequence of reweighted datasets. It first trains a learner from cases with initial weights. In the next learning rounds, the weights of cases are updated based on the performance of the previous learner. The correctly predicted cases gain lower weights, while cases with higher learning errors gain higher weights. Therefore, the cases with high learning errors receive more attention from the following learners.

Base learning algorithms are repeatedly trained according to the sampled training set with adjusted weights until the iteration reaches a pre-specified number. Each established base-learner is specialised to the sampled training set in the corresponding learning round. The final strong learner is produced using a weighted voting strategy. The Boosting algorithm is given in Algorithm A.2.2 (Hu et al., 2021).

Algorithm A.2.2 : Boosting.

Input:

Data set $D = \{(x_1, y_1), (x_2, y_2), \cdots, (x_M, y_M)\}$.

Base learning algorithm L.

Number of learning round T.

Process:

$D_1(i) = \frac{1}{M}$. #Initialise the weight distribution

For $t = 1, 2, \cdots, T$:

$h_t = L(D, D_t)$. #Train a base learner h_t from D using D_t

$\varepsilon_t = \sum_{i=1}^{M} D_t(i) [h_t(x_i) \neq y_i]$. #Measure the error of h_t

$\alpha_t = \frac{1}{2} \ln \frac{1 - \varepsilon_t}{\varepsilon_t}$. # Determine the weight of h_t

$$Z_t = \sum_{i=1}^{m} D_t(i) \times \begin{cases} e^{-\alpha_t} \text{ if } h_t(x_i) = y_i \\ e^{\alpha_t} \text{ if } h_t(x_i) \neq y_i \end{cases}$$

Z_t is a normalisation factor that enables D_{t+1} to be a distribution

$$D_{t+1}(i) = \frac{D_t(t)}{z_t} \times \begin{cases} e^{-\alpha_t} \text{ if } h_t(x_i) = y_i \\ e^{\alpha_t} \text{ if } h_t(x_i) \neq y_i \end{cases}$$

#Update the distribution

end

Output: $H(x) = \text{sign} \sum_{t=1}^{T} \alpha_t h_t(x)$

Appendix B

Classification results of different classifiers for ASD and TD discrimination

TABLE B.1: LDA performance using different feature vectors for all types of stimuli.

Component/s	Feature/s	Happy			Neutral			Fear		
		ACC	TPR	TNR	ACC	TPR	TNR	ACC	TPR	TNR
IMF 5	**SampEn**	72.2 %	78 %	67 %	72.2 %	67 %	78 %	50.0 %	44 %	56 %
	Min	61.1 %	56 %	67 %	61.1 %	67 %	56 %	55.6 %	67 %	44 %
	Std	55.6 %	67 %	44 %	55.6 %	56 %	56 %	38.9 %	56 %	22 %
	All	77.8 %	89 %	67 %	77.8 %	78 %	78 %	44.4 %	44 %	44 %
IMF 6	**SampEn**	88.9 %	89 %	89 %	72.2 %	67 %	78 %	44.4 %	56 %	33 %
	Min	61.1 %	56 %	67 %	66.7 %	56 %	78 %	55.6 %	44 %	67 %
	Std	61.1 %	56 %	67 %	61.1 %	56 %	67 %	55.6 %	44 %	67 %
	All	72.2 %	67 %	78 %	72.2 %	67 %	78 %	50.0 %	44 %	56 %
IMF 7	**SampEn**	61.1 %	78 %	44 %	44.4 %	67 %	22 %	50.0 %	44 %	56 %
	Min	77.8 %	78 %	78 %	72.2 %	67 %	78 %	50.0 %	44 %	56 %
	Std	72.2 %	56 %	89 %	72.2 %	67 %	78 %	55.6 %	44 %	67 %
	All	83.3 %	78 %	89 %	72.2 %	67 %	78 %	61.1 %	44 %	78 %
IMF 5–7	**SampEn**	88.9 %	100 %	78 %	66.7 %	78 %	56 %	50.0 %	56 %	44 %
	Min	72.2 %	67 %	78 %	50.0 %	44 %	56 %	55.6 %	44 %	67 %
	Std	83.3 %	78 %	89 %	66.7 %	67 %	67 %	55.6 %	56 %	56 %
	All	**94.4%**	**89 %**	**100 %**	77.8 %	78 %	78 %	50.0 %	56 %	44 %

TABLE B.2: QDA performance using different feature vectors for all types of stimuli.

Component/s	Feature/s	Happy			Neutral			Fear		
		ACC	TPR	TNR	ACC	TPR	TNR	ACC	TPR	TNR
IMF 5	SampEn	83.3 %	89 %	78 %	72.2 %	67 %	78 %	50.0 %	44 %	56 %
	Min	38.9 %	33 %	44 %	66.7 %	89 %	44 %	61.1 %	78 %	44 %
	Std	27.8 %	0 %	56 %	44.4 %	56 %	33 %	44.4 %	44 %	44 %
	All	61.1 %	67 %	56 %	66.7 %	67 %	67 %	61.1 %	44 %	78 %
IMF 6	SampEn	66.7 %	56 %	78 %	61.1 %	67 %	56 %	44.4 %	56 %	33 %
	Min	55.6 %	56 %	56 %	61.1 %	44 %	78 %	50.0 %	33 %	67 %
	Std	61.1 %	56 %	67 %	50.0 %	44 %	56 %	44.4 %	22 %	67 %
	All	66.7 %	56 %	78 %	61.1 %	56 %	67 %	44.4 %	33 %	56 %
IMF 7	SampEn	66.7 %	67 %	67 %	44.4 %	44 %	44 %	44.4 %	22 %	67 %
	Min	72.2 %	78 %	67 %	72.2 %	67 %	78 %	61.1 %	44 %	78 %
	Std	72.2 %	56 %	89 %	72.2 %	67 %	78 %	61.1 %	44 %	78 %
	All	66.7 %	67 %	67 %	66.7 %	67 %	67 %	61.1 %	56 %	67 %
IMF 5–7	SampEn	77.8 %	67 %	89 %	55.6 %	33 %	78 %	66.7 %	67 %	67 %
	Min	66.7 %	56 %	78 %	55.6 %	56 %	56 %	55.6 %	33 %	78 %
	Std	66.7 %	44 %	89 %	55.6 %	56 %	56 %	50.0 %	33 %	67 %
	All	**77.8 %**	**67%**	**89 %**	72.2 %	67 %	78 %	66.7 %	44 %	89 %

TABLE B.3: Q-SVM performance using different feature vectors for all types of stimuli.

Component/s	Feature/s	Happy			Neutral			Fear		
		ACC	TPR	TNR	ACC	TPR	TNR	ACC	TPR	TNR
IMF 5	SampEn	72.2 %	78 %	67 %	72.2 %	78 %	67 %	61.1 %	78 %	44 %
	Min	50.0 %	67 %	33 %	61.1 %	56 %	67 %	50.5 %	44 %	56 %
	Std	50.0 %	67 %	33 %	50.0 %	56 %	44 %	33.3 %	33 %	33 %
	All	72.2 %	78 %	67 %	72.2 %	78 %	67 %	61.1 %	78 %	44 %
IMF 6	SampEn	77.8 %	67 %	89 %	61.1 %	56 %	67 %	55.6 %	56 %	56 %
	Min	72.2 %	67 %	78 %	66.7 %	56 %	78 %	66.7 %	56 %	78 %
	Std	77.8 %	78 %	78 %	66.7 %	56 %	78 %	50.0 %	33 %	67 %
	All	83.3 %	78 %	89 %	77.8 %	67 %	89 %	44.4 %	44 %	44 %
IMF 7	SampEn	61.1 %	78 %	44 %	38.9 %	56 %	22 %	61.1 %	67 %	56 %
	Min	66.7 %	67 %	67 %	66.7 %	67 %	67 %	50.0 %	56 %	44 %
	Std	55.6 %	44 %	67 %	72.2 %	67 %	78 %	55.6 %	56 %	56 %
	All	66.7 %	78 %	56 %	66.7 %	67 %	67 %	61.1 %	78 %	44 %
IMF 5–7	SampEn	88.9 %	89 %	89 %	66.7 %	78 %	56 %	38.9 %	56 %	22 %
	Min	61.1 %	67 %	56 %	61.1 %	44 %	78 %	55.6 %	56 %	56 %
	Std	77.8 %	78 %	78 %	55.6 %	56 %	56 %	55.6 %	44 %	67 %
	All	**94.4 %**	**89 %**	**100 %**	77.8 %	67 %	89 %	61.1%	56%	67%

TABLE B.4: C-SVM performance using different feature vectors for all types of stimuli.

Component/s	Feature/s	Happy			Neutral			Fear		
		ACC	TPR	TNR	ACC	TPR	TNR	ACC	TPR	TNR
IMF 5	SampEn	55.6 %	56 %	56 %	72.2 %	78 %	67 %	61.1 %	67 %	56 %
	Min	50.0 %	56 %	44 %	61.1 %	56 %	67 %	55.6 %	44 %	67 %
	Std	33.3 %	44 %	22 %	50.0 %	56 %	44 %	50.0 %	44 %	65 %
	All	50.0 %	56 %	44 %	66.7 %	67 %	67 %	55.6 %	67 %	44 %
IMF 6	SampEn	72.2 %	67 %	78 %	61.1 %	44 %	78 %	61.1 %	67 %	56 %
	Min	61.1 %	67 %	56 %	72.2 %	78 %	67 %	55.6 %	67 %	44 %
	Std	66.7 %	78 %	56 %	50.0 %	44 %	56 %	61.1 %	89 %	33 %
	All	83.3 %	67 %	100 %	77.8 %	67 %	89 %	50.0 %	56 %	44 %
IMF 7	SampEn	61.1 %	78 %	44 %	50.0 %	67 %	33 %	61.1 %	67 %	56 %
	Min	55.6 %	56 %	56 %	55.6 %	44 %	67 %	50.0 %	56 %	44 %
	Std	44.4 %	33 %	56 %	55.6 %	44 %	67 %	61.1 %	56 %	67 %
	All	72.2 %	78 %	67 %	72.2 %	67 %	78 %	55.6 %	78 %	33 %
IMF 5–7	SampEn	83.3 %	89 %	78 %	72.2 %	78 %	67 %	44.4 %	56 %	33 %
	Min	66.7 %	67 %	67 %	72.2 %	78 %	67 %	61.1 %	56 %	67 %
	Std	77.8 %	67 %	89 %	61.1 %	56 %	67 %	55.6 %	44 %	67 %
	All	**94.4 %**	**89 %**	**100 %**	77.8 %	67 %	89 %	61.1 %	44 %	78 %

TABLE B.5: kNN performance using different feature vectors for all types of stimuli.

Component/s	Feature/s	Happy			Neutral			Fear		
		ACC	TPR	TNR	ACC	TPR	TNR	ACC	TPR	TNR
IMF 5	SampEn	77.8 %	89 %	67 %	72.2 %	89 %	56 %	66.7 %	78 %	56 %
	Min	66.7 %	67 %	67 %	61.1 %	78 %	44 %	61.1 %	78 %	44 %
	Std	44.4 %	44 %	44 %	55.6 %	56 %	56 %	50.0 %	67 %	33 %
	All	72.2 %	89 %	56 %	61.1 %	89 %	33 %	66.7 %	89 %	44 %
IMF 6	SampEn	77.8 %	89 %	67 %	72.2 %	78 %	67 %	44.4 %	56 %	33 %
	Min	66.7 %	67 %	67 %	61.1 %	44 %	78 %	66.7 %	67 %	67 %
	Std	72.2 %	67 %	78 %	61.1 %	56 %	67 %	66.7 %	44 %	89 %
	All	77.8 %	56 %	100 %	72.2 %	78 %	67 %	55.6 %	56 %	56 %
IMF 7	SampEn	61.1 %	67 %	56 %	55.6 %	67 %	44 %	55.6 %	44 %	67 %
	Min	77.8 %	78 %	78 %	72.2 %	67 %	78 %	61.1 %	44 %	78 %
	Std	72.2 %	56 %	89 %	72.2 %	67 %	78 %	55.6 %	56 %	56 %
	All	72.2 %	78 %	67 %	72.2 %	67 %	78 %	61.1 %	67 %	67 %
IMF 5–7	SampEn	77.8 %	100 %	56 %	72.2 %	78 %	67 %	66.7 %	67 %	67 %
	Min	72.2 %	78 %	67 %	66.7 %	67 %	67 %	55.6 %	67 %	44 %
	Std	72.2 %	89 %	56 %	66.7 %	67 %	67 %	66.7 %	67 %	67 %
	All	**83.3 %**	**100 %**	**67 %**	77.8 %	89 %	67 %	66.7%	78%	56%

Appendix C

Statistical analysis results of the complexity and synchronisation features

TABLE C.1: P-values of the Kruskal-Wallis test of the SampEn features.

Channel	IMF3	IMF4	IMF5	IMF6	IMF7	IMF8	IMF9
C3	0.54	0.18	0.43	0.76	0.72	0.72	0.20
F3	0.86	0.33	0.76	0.81	0.90	0.43	0.06
F7	0.72	0.06	0.02	0.11	0.81	0.16	0.13
Fz	0.58	0.30	0.07	0.05	0.63	0.30	0.58
O1	0.54	0.67	0.36	0.11	0.90	0.39	0.43
O2	0.90	0.54	0.33	0.76	0.47	0.95	1.00
P3	0.63	0.47	0.14	0.25	0.16	0.13	0.95
P4	0.33	0.39	0.43	0.33	0.30	0.72	0.27
T3	0.54	0.54	0.16	0.72	0.22	0.16	0.76
T4	0.30	0.76	0.90	0.95	0.90	0.95	0.54
T5	0.72	0.76	0.63	0.81	0.81	0.50	0.20
T6	0.36	0.86	0.18	0.07	0.54	0.76	0.58

TABLE C.2: P-values of the Kruskal-Wallis test of the PEn features.

Channel	IMF3	IMF4	IMF5	IMF6	IMF7	IMF8	IMF9
C3	0.58	0.72	0.13	0.72	0.39	0.67	0.95
F3	0.81	0.33	0.06	0.20	0.81	0.05	0.07
F7	0.63	0.13	0.07	0.50	0.10	0.54	0.72
Fz	0.76	0.81	0.30	0.76	0.14	0.14	0.58
O1	0.58	0.72	0.27	0.72	0.76	0.72	0.90
O2	0.95	0.47	0.03	0.86	0.81	0.10	0.20
P3	0.90	0.67	0.14	0.67	0.86	0.27	0.50
P4	0.63	0.22	0.02	0.47	0.11	0.10	0.20
T3	0.67	0.50	0.13	0.33	0.22	0.81	0.63
T4	0.90	0.86	0.20	0.13	0.10	0.67	0.67
T5	0.76	0.72	0.11	0.09	0.25	0.06	0.86
T6	0.47	0.72	0.06	0.39	0.86	0.90	0.67

TABLE C.3: P-values of the Kruskal-Wallis test of the SpEn features.

Channel	IMF3	IMF4	IMF5	IMF6	IMF7	IMF8	IMF9
C3	0.95	0.27	0.11	0.76	0.43	0.67	0.25
F3	0.72	0.30	0.30	0.58	0.72	0.14	0.10
F7	0.86	0.14	0.81	0.25	0.58	0.08	0.10
FZ	0.72	0.76	0.43	0.33	0.90	0.09	0.36
O1	0.43	0.18	0.10	0.25	0.58	0.33	0.10
O2	0.58	0.54	0.27	0.90	0.72	0.33	0.76
P3	0.95	0.81	0.05	0.81	1.00	0.05	0.39
P4	0.22	0.18	0.02	0.95	0.33	0.06	0.63
T3	0.27	0.25	0.07	0.25	0.58	0.22	0.07
T4	0.47	0.81	0.16	0.22	0.90	0.39	0.39
T5	0.43	0.76	0.10	1.00	0.08	0.13	0.33
T6	0.36	0.76	0.01	0.33	0.30	0.72	0.43

TABLE C.4: P-values of the Kruskal-Wallis test of the ND of PLI features.

Channel	IMF3	IMF4	IMF5	IMF6	IMF7	IMF8	IMF9
C3	1.00	0.50	0.72	0.63	0.27	0.72	0.18
F3	0.72	0.81	0.25	0.76	0.43	0.81	0.05
F7	0.36	0.58	0.36	0.18	0.72	0.18	0.13
Fz	0.20	0.63	0.72	0.30	0.72	0.72	0.50
O1	0.16	0.67	0.81	0.33	0.86	0.10	0.14
O2	0.27	0.58	0.27	0.27	0.54	0.76	0.39
P3	0.20	0.16	0.58	0.76	0.86	0.39	0.27
P4	0.36	0.95	0.36	0.18	0.67	0.39	0.50
T3	0.16	0.20	0.43	0.39	0.43	0.54	0.18
T4	0.14	0.95	0.02	0.20	0.81	0.81	0.81
T5	0.36	0.25	0.76	0.86	0.14	0.43	0.13
T6	0.72	0.67	0.86	0.81	0.81	0.06	0.90

TABLE C.5: P-values of the Pearson's correlation between the SampEn and ND of PLI features in the CP group.

Channel	IMF3	IMF 4	IMF 5	IMF 6	IMF 7	IMF 8	IMF 9
C3	0.99	0.44	0.36	0.06	0.22	0.82	0.76
F3	0.22	0.19	0.42	0.78	0.19	0.1	0.28
F7	0.5	0.79	0.41	0.53	0.77	0.63	0.06
Fz	0.28	0.55	0.58	0.7	0.52	0.17	0.62
O1	0.42	0.14	0.01	0.39	0.97	0.24	0.5
O2	0.41	0.1	0.82	0.08	0.4	0.22	0.89
P3	0.23	0.02	0.66	0.65	0.5	0.68	0.63
P4	0.31	0.61	0.96	0.73	0.22	0.83	0.69
T3	0.15	0.07	0.2	0.25	0.88	0.05	0.81
T4	0.4	0.26	0.4	0.18	0.2	0.91	0.91
T5	0.66	0.24	0.29	0.72	0.13	0.05	0.91
T6	0.34	0.95	0.07	0.49	0.24	0.27	0.29

TABLE C.6: P-values of the Pearson's correlation between the SampEn and ND of PLI features in the normal group.

Channel	IMF3	IMF4	IMF5	IMF6	IMF7	IMF8	IMF9
C3	0.26	0.03	0.19	0.39	0.32	0.03	0.9
F3	0.31	0.01	0.47	0.14	0.25	0.05	0.76
F7	0.12	0.03	0.03	0.05	0.12	0.00	0.22
Fz	0.32	0.01	0.01	0.44	0.86	0.58	0.14
O1	0.46	0.1	0.67	0.05	0.06	0.14	0.09
O2	0.16	0.53	0.28	0.79	0.1	0.03	0.2
P3	0.12	0.09	0.39	0.9	0.65	0.98	0.69
P4	0.05	0.04	0.37	0.82	0.23	0.03	0.06
T3	0.12	0.04	0.6	0.13	0.07	0.36	0.29
T4	0.54	0.43	0.91	0.74	0.86	0.32	0.07
T5	0.72	0.39	0.83	0.94	0.06	0.34	0.36
T6	0.45	0.17	0.61	0.08	0.89	0.22	0.58

TABLE C.7: P-values of the Pearson's correlation between the PEn and ND of PLI features in the CP group.

Channel	IMF3	IMF 4	IMF 5	IMF 6	IMF 7	IMF 8	IMF 9
C3	0.49	0.08	0.8	0.99	0.15	0.59	0.9
F3	0.15	0.06	0.98	0.31	0.34	0.69	0.35
F7	0.92	0.82	0.29	0.66	0.51	0.84	0.12
Fz	0.2	0.58	0.37	0.68	0.55	0.08	0.6
O1	0.66	0.18	0.31	0.07	0.76	0.19	0.27
O2	0.35	0.13	0.47	0.22	0.001	0.13	0.43
P3	0.26	0.2	0.11	0.36	0.12	0.93	0.78
P4	0.16	0.62	0.03	0.57	0.48	0.44	0.29
T3	0.13	0.02	0.09	0.27	0.9	0.42	0.92
T4	0.56	0.31	0.61	0.26	0.85	0.05	0.56
T5	0.64	0.47	0.63	0.49	0.06	0.67	0.51
T6	0.38	0.53	0.82	0.8	0.45	0.3	0.97

TABLE C.8: P-values of the Pearson's correlation between the PEn and ND of PLI features in the normal group.

Channel	IMF 3	IMF4	IMF5	IMF6	IMF7	IMF8	IMF9
C3	0.41	0.003	0.028	0.73	0.69	0.18	0.09
F3	0.34	0.002	0.19	0.15	0.11	0.04	0.08
F7	0.26	0.001	0.1	0.07	0.45	0.14	0.2
Fz	0.57	0.002	0.004	0.07	0.98	0.96	0.95
O1	0.71	0.06	0.95	0.85	0.51	0.028	0.08
O2	0.16	0.2	0.03	0.08	0.76	0.06	0.07
P3	0.18	0.05	0.58	0.18	0.33	0.3	0.07
P4	0.13	0.12	0.58	0.06	0.5	0.6	0.14
T3	0.17	0.07	0.13	0.13	0.08	0.76	0.43
T4	0.79	0.59	0.74	0.73	0.5	0.71	0.28
T5	0.95	0.25	0.62	0.37	0.33	0.28	0.1
T6	0.33	0.04	0.05	0.92	0.93	0.4	0.76

TABLE C.9: P-values of the Pearson's correlation between the SpEn and ND of PLI features in the CP group.

Channel	IMF3	IMF4	IMF5	IMF6	IMF7	IMF8	IMF9
C3	0.4	0.13	0.66	0.93	0.12	0.73	0.89
F3	0.37	0.02	0.49	0.33	0.53	0.42	0.41
F7	0.98	0.5	0.06	0.98	0.56	0.23	0.008
Fz	0.4	0.47	0.27	0.41	0.45	0.46	0.67
O1	0.99	0.06	0.008	0.69	0.7	0.37	0.08
O2	0.75	0.16	0.26	0.24	0.06	0.86	0.89
P3	0.31	0.05	0.26	0.49	0.11	0.93	0.5
P4	0.74	0.35	0.16	0.71	0.35	0.14	0.29
T3	0.3	0.32	0.58	0.3	0.85	0.09	0.36
T4	0.66	0.25	0.38	0.24	0.76	0.27	0.74
T5	0.75	0.82	0.44	0.5	0.07	0.12	0.39
T6	0.98	0.39	0.15	0.82	0.25	0.18	0.52

TABLE C.10: P-values of the Pearson's correlation between the SpEn and ND of PLI features in the normal group.

Channel	IMF3	IMF4	IMF5	IMF6	IMF7	IMF8	IMF9
C3	0.97	0.008	0.01	0.06	0.96	0.02	0.35
F3	0.47	0.001	0.19	0.527	0.1	0.02	0.61
F7	0.9	0.003	0.48	0.07	0.18	0.00002	0.07
Fz	0.42	0.005	0.001	0.06	0.9	0.2	0.49
O1	0.51	0.08	0.39	0.14	0.38	0.02	0.06
O2	0.68	0.28	0.1	0.07	0.17	0.03	0.07
P3	0.78	0.05	0.2	0.21	0.34	0.6	0.41
P4	0.22	0.0008	0.08	0.08	0.23	0.006	0.06
T3	0.43	0.09	0.13	0.66	0.08	0.11	0.08
T4	0.97	0.73	0.72	0.2	0.74	0.25	0.09
T5	0.61	0.14	0.64	0.08	0.21	0.03	0.12
T6	0.79	0.02	0.01	0.38	0.89	0.29	0.07

Appendix D

Correlation analysis results between the entropy features and the cognitive scores using different embedding parameters

TABLE D.1: P-values of the correlation analysis of the SampEn features using embedding dimension m=3. Significant features are shown in boldface.

Tolerance (r)	Channel	IMF4	IMF5	IMF6	IMF7	IMF8	IMF9
0.15	C3	0.97	0.96	0.62	0.59	0.54	0.18
	F3	0.56	0.94	0.47	0.41	0.37	0.71
	F7	0.96	0.59	0.72	0.94	0.74	0.57
	Fz	0.46	0.36	0.15	0.33	0.77	0.25
	O1	0.55	0.53	0.47	0.94	0.42	0.42
	O2	0.37	0.62	0.56	0.36	0.08	0.81
	P3	0.51	0.85	0.38	0.76	0.42	0.77
	P4	0.40	0.75	0.79	0.70	0.88	0.18
	T3	0.56	0.67	0.54	0.19	0.95	0.61
	T4	0.63	0.98	0.99	0.56	0.59	0.46
	T5	0.33	0.67	0.64	0.66	0.61	0.31
	T6	0.61	0.86	0.26	0.33	0.80	0.07
0.2	C3	0.87	0.94	0.62	0.59	0.61	0.17
	F3	0.82	0.90	0.46	0.68	0.37	0.76
	F7	0.87	0.63	0.44	0.95	0.72	0.60
	Fz	0.49	0.41	0.19	0.52	0.82	0.22
	O1	0.68	0.85	0.56	1.00	0.47	0.40
	O2	0.50	0.48	0.48	0.36	0.09	0.82
	P3	0.71	0.90	0.32	0.63	0.45	0.87
	P4	0.49	0.81	0.87	0.90	0.87	0.25
	T3	0.43	0.66	0.48	0.21	0.99	0.42
	T4	0.59	0.83	0.94	0.67	0.52	0.53
	T5	0.40	0.49	0.58	0.55	0.57	0.30
	T6	0.73	0.78	0.27	0.43	0.73	0.08
0.25	C3	0.79	0.92	0.62	0.58	0.63	0.17
	F3	0.93	0.84	0.48	0.79	0.38	0.76
	F7	0.70	0.87	0.33	0.99	0.70	0.65
	Fz	0.56	0.36	0.21	0.60	0.86	0.21
	O1	0.81	0.93	0.55	0.95	0.45	0.36
	O2	0.62	0.47	0.51	0.37	0.09	0.88
	P3	0.76	0.81	0.26	0.59	0.46	0.96
	P4	0.68	0.72	0.90	0.96	0.94	0.26
	T3	0.35	0.59	0.47	0.20	0.99	0.39
	T4	0.56	0.82	0.98	0.79	0.46	0.56
	T5	0.52	0.43	0.50	0.53	0.58	0.30
	T6	0.85	0.61	0.29	0.46	0.71	0.09

TABLE D.2: P-values of the correlation analysis of the SampEn features using embedding dimension m=2. Significant features are shown in boldface.

Tolerance (r)	Channel	IMF4	IMF5	IMF6	IMF7	IMF8	IMF9
0.1	C3	0.72	0.77	0.71	0.53	0.46	0.21
	F3	0.92	0.94	0.51	0.52	0.36	0.66
	F7	0.78	0.56	0.55	0.89	0.81	0.52
	Fz	0.61	0.41	0.18	0.41	0.69	0.27
	O1	0.51	0.77	0.52	0.99	0.37	0.43
	O2	0.51	0.58	0.51	0.34	0.08	0.80
	P3	0.99	0.92	0.34	0.74	0.42	0.71
	P4	0.57	0.71	0.85	0.74	0.94	0.19
	T3	0.41	0.62	0.54	0.19	0.98	0.83
	T4	0.57	0.90	0.96	0.54	0.65	0.49
	T5	0.66	0.54	0.71	0.63	0.65	0.28
	T6	0.61	0.80	0.27	0.40	0.77	0.07
0.15	C3	0.57	0.89	0.68	0.51	0.55	0.19
	F3	0.62	0.88	0.51	0.72	0.38	0.71
	F7	0.63	0.89	0.34	0.94	0.75	0.58
	Fz	0.70	0.37	0.20	0.62	0.78	0.25
	O1	0.74	0.99	0.50	0.97	0.44	0.41
	O2	0.63	0.49	0.52	0.32	0.08	0.81
	P3	0.91	0.78	0.27	0.64	0.43	0.77
	P4	0.71	0.69	0.92	0.93	0.89	0.19
	T3	0.37	0.53	0.50	0.22	0.96	0.61
	T4	0.56	0.79	0.96	0.66	0.61	0.47
	T5	0.73	0.51	0.54	0.49	0.62	0.32
	T6	0.87	0.63	0.30	0.46	0.73	0.08
0.2	C3	0.60	0.70	0.65	0.54	0.62	0.18
	F3	0.55	0.88	0.53	0.78	0.38	0.76
	F7	0.60	0.90	0.27	0.98	0.73	0.61
	Fz	0.63	0.41	0.19	0.64	0.83	0.21
	O1	0.85	0.95	0.49	0.98	0.48	0.39
	O2	0.71	0.48	0.56	0.34	0.09	0.82
	P3	0.86	0.67	0.25	0.61	0.46	0.87
	P4	0.73	0.58	0.92	0.97	0.88	0.25
	T3	0.38	0.48	0.46	0.22	0.99	0.41
	T4	0.61	0.87	0.91	0.77	0.54	0.54
	T5	0.72	0.50	0.48	0.48	0.58	0.31
	T6	0.97	0.61	0.34	0.46	0.72	0.08
0.25	C3	0.62	0.61	0.64	0.59	0.64	0.17
	F3	0.52	0.89	0.56	0.80	0.38	0.75
	F7	0.62	0.82	0.26	0.99	0.70	0.66
	Fz	0.67	0.50	0.19	0.63	0.86	0.21
	O1	0.89	0.92	0.47	0.97	0.45	0.35
	O2	0.69	0.50	0.59	0.37	0.09	0.88
	P3	0.87	0.61	0.24	0.59	0.46	0.96
	P4	0.75	0.50	0.89	0.98	0.95	0.27
	T3	0.38	0.44	0.44	0.21	0.99	0.39
	T4	0.62	0.89	0.88	0.82	0.47	0.57
	T5	0.78	0.58	0.51	0.50	0.58	0.31
	T6	0.94	0.59	0.36	0.46	0.71	0.09

TABLE D.3: P-values of the correlation analysis of the PEn features using embedding dimension m=3. Significant features are shown in boldface.

Time delay (τ)	Channel	IMF4	IMF5	IMF6	IMF7	IMF8	IMF9
2	**C3**	0.67	0.83	0.28	0.57	0.34	**0.02**
	F3	0.34	0.53	0.76	0.34	0.39	0.38
	F7	0.32	0.50	0.46	0.45	0.68	0.14
	Fz	0.40	0.48	0.31	0.51	0.93	0.96
	O1	0.50	0.43	0.10	0.71	0.84	0.99
	O2	0.18	0.98	0.59	0.75	0.86	0.25
	P3	0.33	0.97	0.88	0.29	0.65	0.79
	P4	0.14	0.50	0.79	0.60	0.31	0.54
	T3	0.49	0.32	0.51	0.16	0.75	0.32
	T4	0.89	0.38	0.74	0.17	0.98	0.28
	T5	0.16	0.93	0.41	0.58	0.72	0.17
	T6	0.54	0.97	0.56	0.84	0.29	0.33

Appendix E

Correlation analysis results between the SampEn and PEn features and the BDI scores

TABLE E.1: P-values of the correlation analysis of the SampEn features.

Channel	IMF3	IMF4	IMF5	IMF6	IMF7	IMF8	IMF9	IMF10
C3	0.22	0.32	0.47	0.92	0.33	0.76	0.60	0.92
C4	0.46	0.77	0.54	1.00	0.38	0.73	0.80	0.69
Cz	0.29	0.20	0.22	0.55	0.88	0.56	0.76	0.78
F3	0.15	0.10	0.55	0.19	0.42	0.82	0.58	0.30
F4	0.72	0.84	0.75	0.21	0.27	0.64	0.37	0.74
F7	0.29	0.35	0.56	0.17	0.85	0.61	0.63	0.05
F8	0.50	0.55	0.92	0.28	0.76	0.53	0.73	0.11
Fz	0.39	0.19	0.79	0.15	0.29	0.83	0.35	0.89
FP1	0.21	0.13	0.81	0.05	0.46	0.75	0.55	0.26
FP2	0.61	0.54	0.52	0.06	0.46	0.82	0.32	0.30
O1	0.71	0.36	0.58	0.23	0.53	0.18	0.36	0.37
O2	0.60	0.52	0.36	0.34	0.94	0.70	0.21	0.17
P3	0.38	0.36	0.85	0.13	0.75	0.66	0.96	0.77
P4	0.34	0.78	0.66	0.27	0.61	0.94	0.73	0.37
Pz	0.48	0.43	0.84	0.19	0.89	0.80	0.83	0.73
T7	0.05	0.04	0.56	0.81	0.52	0.84	0.55	0.06
T8	0.46	0.39	0.53	0.53	0.50	0.69	0.31	0.37
TP7	0.29	0.21	0.45	0.23	0.28	0.62	0.90	0.17
TP8	0.27	0.31	0.48	0.47	0.52	1.00	0.23	0.18

TABLE E.2: P-values of the correlation analysis of the PEn features.

Channel	IMF3	IMF4	IMF5	IMF6	IMF7	IMF8	IMF9	IMF10
C3	0.16	0.19	0.13	0.66	0.70	0.58	0.52	0.37
C4	0.15	0.10	0.14	0.59	0.74	0.47	0.68	0.51
Cz	0.20	0.15	0.20	0.50	0.79	0.50	0.81	0.55
F3	0.17	0.19	0.13	0.42	0.80	0.65	0.84	0.65
F4	0.17	0.16	0.14	0.46	1.00	0.52	0.99	0.77
F7	0.24	0.14	0.06	0.49	0.67	0.88	0.97	0.43
F8	0.11	0.12	0.04	0.31	0.59	0.53	0.99	0.76
Fz	0.18	0.29	0.17	0.51	0.97	0.47	0.95	0.80
FP1	0.12	0.25	0.27	0.35	0.65	0.41	0.98	0.83
FP2	0.20	0.18	0.24	0.29	0.58	0.25	0.75	0.46
O1	0.24	0.22	0.12	0.79	0.85	0.56	0.90	0.94
O2	0.24	0.22	0.16	0.61	0.95	0.46	0.66	0.80
P3	0.21	0.19	0.16	0.76	0.61	0.97	0.62	0.51
P4	0.19	0.13	0.08	0.67	0.74	0.59	0.72	0.76
Pz	0.21	0.19	0.14	0.61	0.88	0.40	0.82	0.43
T7	0.20	0.22	0.16	0.42	0.72	0.78	0.87	0.82
T8	0.16	0.09	0.06	0.54	0.71	0.95	0.86	0.55
TP7	0.19	0.23	0.12	0.63	0.76	0.89	0.68	0.29
TP8	0.19	0.16	0.22	0.73	0.88	0.81	0.85	0.67

References

Aboy, M., Hornero, R., Abasolo, D. and Alvarez, D. (2006), 'Interpretation of the lempel-ziv complexity measure in the context of biomedical signal analysis', *IEEE Transactions on Biomedical Engineering* **53**(11), 2282–2288.

Acharya, U. R., Oh, S. L., Hagiwara, Y., Tan, J. H., Adeli, H. and Subha, D. P. (2018), 'Automated eeg-based screening of depression using deep convolutional neural network', *Computer Methods and Programs in Biomedicine* **161**, 103–113.

Allen, K. A. (2012), 'Hypoxic Ischemic Encephalopathy: Pathophysiology Experimental Treatments', **11**(3), p125–133.

Alotaibi, N., Bakheet, D., Konn, D., Vollmer, B. and Maharatna, K. (2022), 'Cognitive outcome prediction in infants with neonatal hypoxic-ischemic encephalopathy based on functional connectivity and complexity of the electroencephalography signal', *Frontiers in Human Neuroscience* **15**.

American Psychiatric Association (2013), *Diagnostic and Statistical Manual of Mental Disorders*.

Apicella, F., Sicca, F., Federico, R. R., Campatelli, G. and Muratori, F. (2012), 'Fusiform Gyrus responses to neutral and emotional faces in children with Autism Spectrum Disorders: a High Density ERP study', *Behavioural Brain Research* pp. p1–8.

Babiloni, C., Brancucci, A., Vecchio, F., Arendtnielsen, L., Chen, A. and Rossini, P. (2006), 'Anticipation of somatosensory and motor events increases centroparietal functional coupling: An EEG coherence study', *Clinical Neurophysiology* **117**(5), p1000–1008.

Bachmann, M., Lass, J. and Hinrikus, H. (2017), 'Single channel eeg analysis for detection of depression', *Biomedical Signal Processing and Control* **31**, 391–397.

Bachmann, M., Peske, L., Kalev, K., Aarma, K., Lehtmets, A., pik, P., Lass, J. and Hinrikus, H. (2018), 'Methods for classifying depression in single channel eeg using linear and nonlinear signal analysis', *Computer Methods and Programs in Biomedicine* **155**, 11–17.

Bakheet, D., Alotaibi, N., Konn, D., Vollmer, B. and Maharatna, K. (2021), 'Prediction of cerebral palsy in newborns with hypoxic-ischemic encephalopathy using multivariate eeg analysis and machine learning', *IEEE Access* pp. 1–1.

Bakheet, D. and Maharatna, K. (2021), 'Linear and nonlinear analysis of intrinsic mode function after facial stimuli presentation in children with autism spectrum disorder', *Computers in Biology and Medicine* **133**(2), p104376.

Balconi, M. and Lucchiari, C. (2006), 'EEG correlates (event-related desynchronization) of emotional face elaboration : A temporal analysis', *Neuroscience Letters* **392**, p118–123.

Balconi, M. and Pozzoli, U. (2009), 'Arousal effect on emotional face comprehension Frequency band changes in different time intervals', *Physiology & Behavior* **97**(3-4), p455–462.

Bandt, C. and Pompe, B. (2002), 'Permutation Entropy : A Natural Complexity Measure for Time Series', *Physical review letters* **88**(17), 174102.

Barttfeld, P., Wicker, B., Cukier, S., Navarta, S., Lew, S. and Sigman, M. (2011), 'A big-world network in ASD: Dynamical connectivity analysis reflects a deficit in long-range connections and an excess of short-range connections', *Neuropsychologia* **49**(2), p254–263.

Bayley, N. (1993), *Bayley Scales of Infant Development (2nd ed.)*, Psychological Corporation.

Beck, A. T., Steer, R. A., Brown, G. K. et al. (1996), 'Manual for the beck depression inventory-ii'.

Bell, A. J. and Sejnowski, T. J. (1995), 'An information-maximization approach to blind separation and blind deconvolution', *Neural Computation* **7**, 1129–1159.

Bendat, J. S. and Piersol, A. G. (1980), *Engineering applications of correlation and spectral analysis*, Wiley-Interscience, New York.

Bendat, J. S. and Piersol, A. G. (1986), Random data - analysis and measurement procedures - second edition (revised and expanded).

Benjamini, Y. and Hochberg, Y. (1995), 'Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing', *Journal of the Royal Statistical Society* **57**(1), 289–300.

Berardi, N., Pizzorusso, T. and Maffei, L. (2000), 'Critical periods during sensory development', *Current Opinion in Neurobiology* **10**(1), p138–145.

Black, M. H., Chen, N. T., Iyer, K. K., Lipp, O. V., Bölte, S., Falkmer, M., Tan, T. and Girdler, S. (2017), 'Mechanisms of facial emotion recognition in autism spectrum disorders: Insights from eye tracking and electroencephalography', *Neuroscience and Biobehavioral Reviews* **80**, p488–515.

Boccaletti, S., Kurths, J., Osipov, G., Valladares, D. L. and Zhou, C. S. (2002), 'The synchronization of chaotic systems', *Physics Reports* **366**, p1–101.

Bonferroni, C. E. (2008), *The Concise Encyclopedia of Statistics*, Springer, New York, NY.

Bosl, W. J., Tager-Flusberg, H. and Nelson, C. A. (2018), 'EEG Analytics for Early Detection of Autism Spectrum Disorder: A data-driven approach', *Scientific Reports* **8**, 6828.

Bosl, W., Tierney, A., Tager-Flusberg, H. and Nelson, C. (2011), 'EEG complexity as a biomarker for autism spectrum disorder risk', *BMC Medicine* **9**(1), 18.

Bowman, L. C. and Varcin, K. J. (2018), 'The Promise of Electroencephalography for Advancing Diagnosis and Treatment in Neurodevelopmental Disorders', *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* **3**(1), p7–9.

Breakspear, M., Williams, L. M. and Stam, C. J. (2004), 'A Novel Method for the Topographic Analysis of Neural Activity Reveals Formation and Dissolution of Dynamic Cell Assemblies'', *Journal of Computational Neuroscience* **16**(1), p49–68.

Bretherton, I. and Beeghly, M. (1982), 'Talking about internal states: The acquisition of an explicit theory of mind', *Developmental Psychology* **18**(6), p906–921.

Bruhn, J., Ropcke, H. and Hoeft, A. (2000), 'Approximate entropy as an electroencephalographic measure of anesthetic drug effect during desflurane anesthesia', *Anesthesiology* **92**(3), p715–726.

Burgess, A. P. (2012), 'Towards a Unified Understanding of Event-Related Changes in the EEG: The Firefly Model of Synchronization through Cross-Frequency Phase Modulation', *PLoS ONE* **7**(9), p1–21.

Byeon, J. H., Kim, G. H., Kim, J. Y., Sun, W., Kim, H. and Eun, B. L. (2015), 'Cognitive dysfunction and hippocampal damage induced by hypoxic-ischemic brain injury and prolonged febrile convulsions in immature rats', *Journal of Korean Neurosurgical Society* **58**(1), p22–29.

Cainelli, E., Vedovelli, L., Lucia, I., Mariani, C., Bisiacchi, P. S. and Suppiej, A. (2021), 'Neonatal spectral EEG is prognostic of cognitive abilities at school age in premature infants without overt brain damage', *European Journal of Pediatrics* **180**(3), p909–918.

Cans, C. (2000), 'Surveillance of cerebral palsy in Europe: a collaboration of cerebral palsy surveys and registers. Surveillance of Cerebral Palsy in Europe (SCPE).', *Developmental Medicine and Child Neurology* **42**(12), p816–824.

Catarino, A., Churches, O., Baron-Cohen, S., Andrade, A. and Ring, H. (2011), 'Atypical EEG complexity in autism spectrum conditions: A multiscale entropy analysis', *Clinical Neurophysiology* **122**(12), p2375–2383.

Cavanagh, J. F. and Allen, J. J. B. (2017), 'Depression Rest'.
URL: *http://predict.cs.unm.edu/downloads.php*

Cavanagh, J. F., Bismark, A. W., Frank, M. J. and Allen, J. J. B. (2019), 'Multiple dissociations between comorbid depression and anxiety on reward and punishment processing: Evidence from computationally informed eeg', *Computational psychiatry* **3**, 1–17.

Chen, J., Sun, H. and Xu, B. (2019), 'Improvement of empirical mode decomposition based on correlation analysis', *SN Applied Sciences* **1**.

Chen, S.-Y., Feng, Z. and Yi, X. (2017), 'A general introduction to adjustment for multiple comparisons', *Journal of Thoracic Disease* **9**(06), p1725–1729.

Chen, W., Wang, Z., Xie, H. and Yu, W. (2007), 'Characterization of surface emg signal based on fuzzy entropy', *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **15**(2), 266–272.

Cho, D., Min, B., Kim, J. and Lee, B. (2017), 'EEG-Based Prediction of Epileptic Seizures Using Phase Synchronization Elicited from Noise-Assisted Multivariate Empirical Mode Decomposition', *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **25**(8), p1309–1318.

Chu, Y. J., Chang, C. F., Shieh, J. S. and Lee, W. T. (2017), 'The potential application of multiscale entropy analysis of electroencephalography in children with neurological and neuropsychiatric disorders', *Entropy* **19**(8), p1–13.

Cioni, G., Inguaggiato, E. and Sgandurra, G. (2016), 'Early intervention in neurodevelopmental disorders: Underlying neural mechanisms', *Developmental Medicine and Child Neurology* **58**, p61–66.

Cohen, M. X. (2014), *Analyzing neural time series data : theory and practice*.

Corden, B., Chilvers, R. and Skuse, D. (2008), 'Avoidance of emotionally arousing stimuli predicts socialperceptual impairment in Asperger's syndrome', *Neuropsychologia* **46**(1), p137–147.

Costa, M., Israel, B., Medical, D. and Goldberger, A. (2005), 'Multiscale entropy of biological signals', *Phys. Rev. E* **71**, 021906.

Cui, D., Wang, J., Bian, Z., Li, Q., Wang, L. and Li, X. (2015), 'Analysis of entropies based on empirical mode decomposition in amnesic mild cognitive impairment of diabetes mellitus', *Journal of Innovative Optical Health Sciences* **08**(05), p1550010.

Das, K., Osechinskiy, S. and Nenadic, Z. N. (2007), 'A classwise pca-based recognition of neural data for brain-computer interfaces', *Annu Int Conf IEEE Eng Med Biol Soc.* pp. 6520–3.

David, O., Cosmelli, D. and Friston, K. J. (2004), 'Evaluation of different measures of functional connectivity using a neural mass model', *NeuroImage* **21**(2), p659–673.

De Felice, A., Ricceri, L., Venerosi, A., Chiarotti, F. and Calamandrei, G. (2015), 'Multifactorial Origin of Neurodevelopmental Disorders: Approaches to Understanding Complex Etiologies', *Toxics* **3**(4), p89–129.

Diagnostic criteria - a guide for all audiences. (2020).
URL: *https://www.autism.org.uk/advice-and-guidance/topics/diagnosis/diagnostic-criteria/all-audiences*

Ditto, W. and Munakata, T. (1995), 'Principles and Applications of Chaotic Systems', *Communications of the ACM* **38**(11), p96–102.

Djemili, R., Bourouba, H. and Amara Korba, M. C. (2016), 'Application of empirical mode decomposition and artificial neural network for the classification of normal and epileptic EEG signals', *Biocybernetics and Biomedical Engineering* **36**(1), p285–291.

Duncan, C. C., Barry, R. J., Connolly, J. F., Fischer, C., Michie, P. T., Naatanen, R., Polich, J., Reinvang, I. and Van Petten, C. (2009), 'Event-related potentials in clinical research: Guidelines for eliciting, recording, and quantifying mismatch negativity, P300, and N400', *Clinical Neurophysiology* **120**(11), p1883–1908.

Edgar, J. C. and Roberts, T. P. L. (2016), 'Joint Analysis of Band-Specific Functional Connectivity and Signal Complexity in Autism', *Journal of autism and developmental disorders* **45**(2), p444–460.

Ernst, M., Torrisi, S., Balderston, N., Grillon, C. and Hale, E. A. (2015), 'fMRI Functional Connectivity Applied to Adolescent', *Annual Review of Clinical Psychology* **11**, p361–377.

Falkmer, T., Anderson, K., Falkmer, M. and Horlin, C. (2013), 'Diagnostic procedures in autism spectrum disorders: a systematic literature review', *European Child & Adolescent Psychiatry* **22**(6), p329–340.

Flandrin, P., Rilling, G. and Gonalves, P. (2004), 'Empirical mode decomposition as a filter bank', *IEEE Signal Processing Letters* **11**, 112–114.

Gabor, D. (1946), 'Theory of communication. Part 1: The analysis of information', *Journal of the Institution of Electrical Engineers - Part III: Radio and Communication Engineering* **93**(26), p429–441.

Gani, E., Handayani, N., Harke Pratama, S., Faadhilah Afif, N., Aziezah, F., Christy Keintjem, A., Haryanto, F. and Suprijadi (2020), 'Brainwaves Analysis Using Spectral Entropy in Children with Autism Spectrum Disorders (ASD)', *Journal of Physics: Conference Series* **1505**(1), p8–13.

Gao, F., Jia, H., Wu, X., Yu, D. and Feng, Y. (2016), 'Altered resting-state eeg microstate parameters and enhanced spatial complexity in male adolescent patients with mild spastic diplegia', *Brain Topography* **30**(2), p233–244.

Gao, F., Wu, X., Feng, Y. and Jia, H. (2017), 'Attenuation of temporal correlations of neuronal oscillations in patients with mild spastic diplegia', *Scientific reports* **7**(1), 14966.

Garcia Dominguez, L., Stieben, J., Perez Velazquez, J. L. and Shanker, S. (2013), 'The Imaginary Part of Coherency in Autism: Differences in Cortical Functional Connectivity in Preschool Children', *PLoS ONE* **8**(10).

George, J. M., Pagnozzi, A. M., Bora, S., Boyd, R. N., Colditz, P. B., Rose, S. E., Ware, R. S., Pannek, K., Bursle, J. E., Fripp, J., Barlow, K., Iyer, K., Leishman, S. J. and Jendra, R. L. (2020), 'Prediction of childhood brain outcomes in infants born preterm using neonatal MRI and concurrent clinical biomarkers (PREBO-6): Study protocol for a prospective cohort study', *BMJ Open* **10**(5).

Gerber, E. M. (2022), 'FDR (False Discovery Rate)'.
URL: *https://www.mathworks.com/matlabcentral/fileexchange/71734-fdr-false-discovery-rate*

Ghaderi, A. H., Nazari, M. A., Shahrokhi, H. and Darooneh, A. H. (2017), '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* **8**(4), 267–278.

Gonzalez, S., Garcia, S., Del Ser, J., Rokach, L. and Herrera, F. (2020), 'A practical tutorial on bagging and boosting based ensembles for machine learning: Algorithms, software tools, performance study, practical perspectives and opportunities', *Information Fusion* **64**, p205–237.

GOV.UK (2021), 'Family Resources Survey: financial year 2019 to 2020'.
URL: *https://www.gov.uk/government/statistics/family-resources-survey-financial-year-2019-to-2020/family-resources-survey-financial-year-2019-to-2020*

Griffiths, R. (1970), *The Abilities of Young Children: A Comprehensive System of Mental Measurement for the First Eight Years.*, Child Development Research Center, London.

Hadders-Algra, M. (2014), 'Early diagnosis and early intervention in cerebral palsy', *Frontiers in Neurology* **5**(09), p1–13.

Halabi, N. E., Abi, R., Daou, Z., Achkar, R., Hayek, A., Boercsoek, J. and Engineering, C. (2020), Comparative Study for Classification Methods to Predict and Detect Epilepsy Seizure, *in* 'Conference on Biomedical Engineering', pp. p1–6.

Hamedi, M., Salleh, S.-H. and Noor, M. N. (2016), 'Electroencephalographic motor imagery brain connectivity analysis for bci: A review', *Neural Comput.* **28**(6), 999–1041.

Han, J., Zeng, K., Kang, J., Tong, Z., Cai, E., Chen, H., Ding, M., Gu, Y., Ouyang, G. and Li, X. (2017), 'Development of Brain Network in Children with Autism from Early Childhood to Late Childhood', *Neuroscience* **367**, p134–146.

Harmon-Jones, E., Gable, P. A. and Peterson, C. K. (2010), 'The role of asymmetric frontal cortical activity in emotion-related phenomena: A review and update', *Biological Psychology* **84**(3), p451–462.

Hassan, A. R. and Bhuiyan, M. I. H. (2017), 'Automated identification of sleep states from EEG signals by means of ensemble empirical mode decomposition and random under sampling boosting', *Computer Methods and Programs in BiomedicineC* **140**, p201–210.

Hatlestad-Hall, C., Brua, R., Syvertsen, M. R., Erichsen, A., Andersson, V., Vecchio, F., Miraglia, F., Rossini, P. M., Renvall, H., Taubll, E., Maest, F. and Haraldsen, I. H. (2021), 'Source-level eeg and graph theory reveal widespread functional network alterations in focal epilepsy', *Clinical Neurophysiology* **132**(7), 1663–1676.

He, L. and Cao, C. (2018), 'Automated depression analysis using convolutional neural networks from speech', *Journal of Biomedical Informatics* **83**, 103–111.

He, L., Li, H., Holland, S. K., Yuan, W., Altaye, M. and Parikh, N. A. (2018), 'Early prediction of cognitive deficits in very preterm infants using functional connectome data in an artificial neural network framework', *NeuroImage: Clinical* **18**, p290–297.

Hoffman, J. (2019), *Chapter 25. Analysis of Variance I. One-Way*, Elsevier, pp. 391–417.

Hong, K. S., Khan, M. J. and Hong, M. J. (2018), 'Feature Extraction and Classification Methods for Hybrid fNIRS-EEG Brain-Computer Interfaces', *Frontiers in Human Neuroscience* **12**(6), p1–25.

Hu, M. and Liang, H. (2011), 'Intrinsic mode entropy based on multivariate empirical mode decomposition and its application to neural data analysis', *Cognitive Neurodynamics* **5**(3), p277–284.

Hu, X., Mei, H., Zhang, H., Li, Y. and Li, M. (2021), 'Performance evaluation of ensemble learning techniques for landslide susceptibility mapping at the Jinping county, Southwest China', *Natural Hazards* **105**(2), p1663–1689.

Huang, J. R., Fan, S. Z., Abbod, M. F., Jen, K. K., Wu, J. F. and Shieh, J. S. (2013), 'Application of multivariate empirical mode decomposition and sample entropy in EEG signals via artificial neural networks for interpreting depth of anesthesia', *Entropy* **15**(9), p3325–3339.

Huang, N., Shen, Z., Long, S., Wu, M., SHIH, H., ZHENG, Q., Yen, N., Tung, C. and Liu, H. (1998), 'The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis', *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **454**, p903–995.

Ibrahim, S., Djemal, R. and Alsuwailem, A. (2018), 'Electroencephalography (EEG) signal processing for epilepsy and autism spectrum disorder diagnosis', *Biocybernetics and Biomedical Engineering* **38**(1), p16–26.

Iemmi, V., Knapp, M. and Ragan, I. (2017), 'The autism dividend - reaping the rewards of better investment', *National Autism Project* .
URL: *https://nationalautistictaskforce.org.uk/national-autism-project/autism-dividend/*

Islam, M. R., Rahim, M. A., Akter, H., Kabir, R. and Shin, J. (2018), 'Optimal IMF selection of EMD for sleep disorder diagnosis using EEG signals', *ACM International Conference Proceeding Series* pp. 96–101.

Jamal, W., Das, S., Oprescu, I.-A., Maharatna, K., Apicella, F. and Sicca, F. (2014), 'Classification of autism spectrum disorder using supervised learning of brain connectivity measures extracted from synchrostates', *Journal of Neural Engineering* **11**(4), 046019.

Jan, A., Meng, H., Gaus, Y. F. B. A. and Zhang, F. (2018), 'Artificial intelligent system for automatic depression level analysis through visual and vocal expressions', *IEEE Transactions on Cognitive and Developmental Systems* **10**(3), 668–680.

Jiang, H., Popov, T., Jylänki, P., Bi, K., Yao, Z., Lu, Q., van Gerven, M. and Jensen, O. (2016), 'Predictability of depression severity based on posterior alpha oscillations', *Clinical Neurophysiology* **127**(4), 2108–2114.

Kaiser, R. H., Whitfield-Gabrieli, S., Dillon, D. G., Goer, F., Beltzer, M., Minkel, J., Smoski, M., Dichter, G. and Pizzagalli, D. A. (2016), 'Dynamic resting-state functional connectivity in major depression', *Neuropsychopharmacology* **41**(7), 1822–30.

Khan, S., Gramfort, A., Shetty, N. R., Kitzbichler, M. G., Ganesan, S., Moran, J. M., Lee, S. M., Gabrieli, J. D., Tager-Flusberg, H. B., Joseph, R. M., Herbert, M. R., Hamalainen, M. S. and Kenet, T. (2013), 'Local and long-range functional connectivity is reduced in concert in autism spectrum disorders', *Proceedings of the National Academy of Sciences of the United States of America* **110**(8), p3107–3112.

Khuntia, A. T., Divakar, R., Apicella, F., Muratori, F. and Das, K. (2019), 'Visual processing and attention rather than face and emotion processing play a distinct role in ASD: an EEG study', *bioRxiv* p. 517664.

Koeda, T. and Takeshita, K. (1998), 'Electroencephalographic Coherence Abnormalities in Preterm Diplegia', **8994**(97), p51–56.

Kong, A. H. T., Lai, M. M., Finnigan, S., Ware, R. S., Boyd, R. N. and Colditz, P. B. (2018), 'Background EEG features and prediction of cognitive outcomes in very preterm infants : A systematic review', *Early Human Development* **127**, p74–84.

Kuhn, M. and Johnson, K. (2013), *Applied predictive modeling*.

Kuhn-Popp, N., Kristen, S., Paulus, M., Meinhardt, J. and Sodian, B. (2016), 'Left hemisphere EEG coherence in infancy predicts infant declarative pointing and preschool epistemic language', *Social Neuroscience* **11**(1), p49–59.

Kulak, W. and Sobaniec, W. (2005), 'Quantitative EEG analysis in children with hemiparetic cerebral palsy.', *NeuroRehabilitation* **20**(2), p75–84.

Kulak, W., Sobaniec, W. and Bockowski, L. (2005), 'EEG spectral analysis and coherence in children with hemiparetic cerebral palsy', **11**(9), 449–455.

Lachaux, J.-p., Rodriguez, E., Martinerie, J. and Varela, F. J. (1999), 'Measuring Phase Synchrony in Brain Signals', *Hum Brain Mapping* **8**(4), p194–208.

Leuchter, A. F., Cook, I. A., Hunter, A. M., Cai, C. and Horvath, S. (2012), 'Resting-state quantitative electroencephalography reveals increased neurophysiologic connectivity in depression', *PLoS One* **7**(2), e32508.

Li, L., Chen, W., Shao, X. and Wang, Z. (2010), 'Analysis of amplitude-integrated EEG in the newborn based on approximate entropy', *IEEE Transactions on Biomedical Engineering* **57**(10), p2459–2466.

Liang, H., Bressler, S. L., Buffalo, E. A., Desimone, R. and Fries, P. (2005), 'Empirical mode decomposition of field potentials from macaque V4 in visual spatial attention', *Biological Cybernetics* **92**(6), p380–392.

Liu, T., Chen, Y., Chen, D., Li, C., Qiu, Y. and Wang, J. (2017), 'Altered electroencephalogram complexity in autistic children shown by the multiscale entropy approach', *NeuroReport* **28**(3), 169–173.

Lloyd, R. O., Toole, J. M. O., Livingstone, V., Filan, P. M. and Boylan, G. B. (2021), 'Can EEG accurately predict 2- year neurodevelopmental outcome for preterm infants ?', *Archives of Disease in Childhood - Fetal and Neonatal Edition* **106**(5), p535–541.

Looney, D., Hemakom, A. and Mandic, D. P. (2015), 'Intrinsic multi-scale analysis: A multi-variate empirical mode decomposition framework', *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **471**(2173), p20140709.

Mallat, S. G. (1989), 'A theory for multiresolution signal decomposition: the wavelet-trepresentation', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **11**(7), p674–693.

Martin-Brufau, R. and Nombela Gomez, M. (2017), 'Bioelectrical markers of ADHD: enhancement of direct EEG analysis', *Electronic Journal of Research in Educational Psychology* **15**(1), p185–200.

Martnez-Briones, B. J., Fernndez-Harmony, T., Garfalo Gmez, N., Biscay-Lirio, R. J. and Bosch-Bayard, J. (2020), 'Working memory in children with learning disorders: An eeg power spectrum analysis', *Brain Sciences* **10**(11), p1–19.

Mathworks (2021), 'Statistics and Machine Learning Toolbox User's Guide R2018b'. **URL:** *https://uk.mathworks.com/help/stats/*

Mohammadi, Y., Hajian, M. and Moradi, M. H. (2019), Discrimination of depression levels using machine learning methods on eeg signals, *in* '2019 27th Iranian Conference on Electrical Engineering (ICEE)', pp. 1765–1769.

Mohammadi, Y. and Moradi, M. H. (2021), 'Prediction of Depression Severity Scores Based on Functional Connectivity and Complexity of the EEG Signal', *Clinical EEG and Neuroscience* **52**(1), 52–60.

Mohanty, N., John, A. L.-s., Manmatha, R. and Rath, T. M. (2013), *Shape-Based Image Classification and Retrieval*, Vol. 31, Elsevier B.V.

Molinaro, A. M., Simon, R. and Pfeiffer, R. M. (2005), 'Prediction error estimation : a comparison of resampling methods', *Bioinformatics* **21**(15), p3301–3307.

Moniz, N. and Branco, P. (2017), 'Evaluation of Ensemble Methods in Imbalanced Regression Tasks', *Proceedings of Machine Learning Research* **74**, p129–140.

Monteiro, R., Simoes, M., Andrade, J. and Branco, M. C. (2017), 'Processing of facial expressions in autism: a systematic review of eeg/erp evidence', *Review Journal of Autism and Developmental Disorders* **4**, p255–276.

Napoli, N. J., Demas, M., Stephens, C. L., Kennedy, K. D., Harrivel, A. R., Barnes, L. E. and Pope, A. T. (2020), 'Activation Complexity: A Cognitive Impairment Tool for Characterizing Neuro-isolation', *Scientific Reports* **10**(1), 1–20.

Nelson, B. D., Infantolino, Z. P., Klein, D. N., Perlman, G., Kotov, R. and Hajcak, G. (2018), 'Time-frequency reward-related delta prospectively predicts the development of adolescent-onset depression', *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* **3**(1), 41–49.

Niso, G., Brua, R., Pereda, E., Gutierrez, R., Bajo, R., Maest, F. and Del Pozo Guerrero, F. (2013), 'Hermes: Towards an integrated toolbox to characterize functional and effective brain connectivity', *Neuroinformatics* **11**.

Nobukawa, S. and Nishimura, H. (2016), 'Chaotic resonance in coupled inferior olive neurons with the Llinás approach neuron model', *Neural Computation* **28**(11), p2505–2532.

Nobukawa, S., Yamanishi, T., Kasakawa, S. and Nishimura, H. (2020), 'Classification Methods Based on Complexity and Synchronization of Electroencephalography Signals in Alzheimer's Disease', *Frontiers in Psychiatry* **11**, p1–12.

Nolte, G., Bai, O., Wheaton, L., Mari, Z., Vorbach, S. and Hallett, M. (2004), 'Identifying true brain interaction from eeg data using the imaginary part of coherency', *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology* **115**, 2292–307.

Nunez, M. D., Nunez, P. L. and Srinivasan, R. (2016), Electroencephalography (eeg): Neurophysics, experimental methods, and signal processing.

Olofsen, E., Sleigh, J. W., Dahan, A. and Zealand, N. (2008), 'Permutation entropy of the electroencephalogram : a measure of anaesthetic drug effect', *British Journal of Anaesthesia* **101**(6), p810–821.

Omidvarnia, A., Mesbah, M., Pedersen, M. and Jackson, G. (2018), 'Range entropy: A bridge between signal complexity and self-similarity', *Entropy* **20**(12).

Ousley, O. and Cermak, T. (2014), 'Autism Spectrum Disorder: Defining Dimensions and Subgroups', *Current developmental disorders reports* **1**(1), p20–28.

Pachori, R. B. (2008), 'Discrimination between Ictal and Seizure-Free EEG Signals Using Empirical Mode Decomposition', *Research Letters in Signal Processing* pp. p1–5.

Pachori, R. B. and Patidar, S. (2014), 'Epileptic seizure classification in EEG signals using second-order difference plot of intrinsic mode functions', *Computer Methods and Programs in Biomedicine* **113**(2), p494–502.

Pal, R. (2017), Chapter 4 - validation methodologies, *in* R. Pal, ed., 'Predictive Modeling of Drug Sensitivity', Academic Press, pp. 83–107.

Palmer, S. M., Crewther, S. G., Carey, L. M. and , T. S. P. T. (2015), 'A meta-analysis of changes in brain activity in clinical depression', *Frontiers in Human Neuroscience* **8**.

Paulus, M., Kühn-Popp, N., Licata, M., Sodian, B. and Meinhardt, J. (2013), 'Neural correlates of prosocial behavior in infancy: Different neurophysiological mechanisms support the emergence of helping and comforting', *NeuroImage* **66**, p522–530.

Peng, C. K., Costa, M. and Goldberger, A. L. (2009), 'Adaptive data analysis of complex fluctuations in physiologic time series.', *Adv Adapt Data Anal.* **1**(1), p61–70.

Pfurtscheller, G. and Andrew, C. (1999), 'Event-related changes of band power and coherence: methodology and interpretation', *Journal of clinical neurophysiology : official publication of the American Electroencephalographic Society* **16**(6), 512519.

Pikovsky, A., Rosenblum, M. and Kurths, J. (2001), *Synchronization: A Universal Concept in Nonlinear Sciences*, Cambridge University Press, Cambridge.

Pincus, S. M. (1991), 'Approximate entropy as a measure of system complexity', *Proceedings of the National Academy of Sciences of the United States of America* **88**(6), p2297–2301.

Pincus, S. M. (2001), 'Assessing serial irregularity and its implications for health', *Annals of the New York Academy of Sciences* **954**, p245–267.

Rehman, N. and Mandic, D. P. (2009), 'Function MEMD applies the "Multivariate Empirical Mode Decomposition" algorithm (Rehman and Mandic, Proc. Roy. Soc A, 2010) to multivariate inputs.'.
URL: *https://www.commsp.ee.ic.ac.uk/ mandic/research/emd.htm*

Rehman, N. and Mandic, D. P. (2010*a*), 'Empirical Mode Decomposition for Trivariate Signals', *IEEE Transactions on Signal Processing* **58**(3), p1059–1068.

Rehman, N. and Mandic, D. P. (2010*b*), 'Multivariate empirical mode decomposition', *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **466**(4), p1291–1302.

Rehman, N. and Mandic, D. P. (2011), 'Filter bank property of multivariate empirical mode decomposition', *IEEE Transactions on Signal Processing* **59**(5), p2421–2426.

Richman, J. S. and Moorman, J. R. (2000), 'Physiological time-series analysis using approximate entropy and sample entropy', *Am J Physiol Heart Circ Physiol.* **278**(6), p2039–2049.

Rilling, G., Flandrin, P., Gonalves, P. and Lilly, J. (2007), 'Bivariate Empirical Mode Decomposition', *IEEE Signal Processing Letters* **14**(12), p936–939.

Rosenbaum, P., Paneth, N., Leviton, A., Goldstein, M., Bax, M., Damiano, D., Dan, B. and Jacobsson, B. (2007), 'A report: The definition and classification of cerebral palsy april 2006', *Developmental medicine and child neurology. Supplement* **109**, 8–14.

Rubinov, M. and Sporns, O. (2010), 'Complex network measures of brain connectivity: Uses and interpretations', *NeuroImage* **52**(3), 1059–1069. Computational Models of the Brain.

Saby, J. N. and Marshall, P. J. (2012), 'The utility of EEG band power analysis in the study of infancy and early childhood', *Dev Neuropsychol* **37**(3), p253–273.

Sajedi, F., Ahmadlou, M., Vameghi, R., Gharib, M. and Hemmati, S. (2013), 'Linear and nonlinear analysis of brain dynamics in children with cerebral palsy', *Research in Developmental Disabilities* **34**, p1388–1396.

Sakkalis, V. (2011), 'Review of advanced techniques for the estimation of brain connectivity measured with EEG/MEG', *Computers in Biology and Medicine* **41**(12), p1110–1117.

Sanei, S. and Chambers, J. (2007), *EEG signal processing*, John Wiley & Sons.

Schie, P. E. M. V., Schijns, J. and Becher, J. G. (2015), 'Long-term motor and behavioral outcome after perinatal hypoxic-ischemic encephalopathy Original article Long-term motor and behavioral outcome after perinatal hypoxic-ischemic encephalopathy', *European Journal of Paediatric Neurology* **19**(3), p354–359.

Schreglmann, M., Helps, S., Hart, D. and Vollmer, B. (2016), 'Are 2-year old children who underwent therapeutic hypothermia for neonatal hypoxic-ischemic encephalopathy at risk for behavioral problems?', *Neuropediatrics* **47**, FV04–09.

Schwartz, S., Kessler, R., Gaughan, T. and Buckley, A. W. (2017), 'Electroencephalogram Coherence Patterns in Autism: An Updated Review', *Pediatr Neurol.* (67), 7–22.

Schweighofer, N., Doya, K., Fukai, H., Chiron, J. V., Furukawa, T. and Kawato, M. (2004), Chaos may enhance information transmission in the inferior olive, *in* 'Proc Natl Acad Sci U S A', Vol. 101, pp. p4655–4660.

Seiffert, C., Khoshgoftaar, T. M., Van Hulse, J. and Napolitano, A. (2010), 'RUSBoost: A hybrid approach to alleviating class imbalance', *IEEE Transactions on Systems, Man, and Cybernetics Part A:Systems and Humans* **40**(1), p185–197.

Semmlow, J. (2018), Stochastic, Nonstationary, and Nonlinear Systems and Signals, *in* J. Semmlow, ed., 'Circuits, Signals and Systems for Bioengineers', third edit edn, Academic Press, chapter 10, pp. p449–489.

Sharma, R., Pachori, R. B. and Acharya, U. R. (2015), 'Application of entropy measures on intrinsic mode functions for the automated identification of focal electroencephalogram signals', *Entropy* **17**(2), p669–691.

Siuly, S. and Zhang, Y. (2016), 'Medical Big Data: Neurological Diseases Diagnosis Through Medical Data Analysis', *Data Science and Engineering* **1**(2), p54–64.

Slaughter, L. A., Susan, B.-m., Hintz, R. and Dvorchik, I. (2016), 'Early Conventional MRI for Prediction of Neurodevelopmental Impairment in Extremely-Low-Birth-Weight Infants', **110**, p47–54.

Spittle, A., Orton, J., Anderson, P., Boyd, R. and Doyle, L. W. (2012), Early developmental intervention programmes post-hospital discharge to prevent motor and cognitive

impairments in preterm infants, *in* A. Spittle, ed., 'Cochrane Database of Systematic Reviews', Vol. 12, John Wiley & Sons, Ltd, Chichester, UK, p. CD005495.

Stam, C. J. (2005), 'Nonlinear dynamical analysis of EEG and MEG: Review of an emerging field', *Clinical Neurophysiology* **116**(10), p2266–2301.

Stam, C. J., Nolte, G. and Daffertshofer, A. (2007), 'Phase Lag Index : Assessment of Functional Connectivity From Multi Channel EEG and MEG With Diminished Bias From Common Sources', **28**, p1178–1193.

Stam, C. J., van der Made, Y., Pijnenburg, Y. A. L. and Scheltens, P. (2003), 'EEG synchronization in mild cognitive impairment and Alzheimer's disease', *Acta Neurologica Scandinavica* **108**(2), p90–96.

Sui, J., Jiang, R., Bustillo, J. and Calhoun, V. (2020), 'Neuroimaging-based Individualized Prediction of Cognition and Behavior for Mental Disorders and Health : Methods and Promises', *Biological Psychiatry* **88**(11), p818–828.

Suppiej, A., Cainelli, E., Cappellari, A., Trevisanuto, D., Balao, L., Grazia, M., Bono, D. and Bisiacchi, P. S. (2017), 'Spectral analysis highlight developmental EEG changes in preterm infants without overt brain damage', *Neuroscience Letters* **649**(10), p112–115.

Sweeney-Reed, C. M. and Nasuto, S. J. (2007), 'A novel approach to the detection of synchronisation in EEG based on empirical mode decomposition', *Journal of Computational Neuroscience* **23**(1), p79–111.

Sweeney-Reed, C. M. and Nasuto, S. J. (2009), 'Detection of neural correlates of self-paced motor activity using empirical mode decomposition phase locking analysis', *Journal of Neuroscience Methods* **184**(1), p54–70.

Sweeney-Reed, C. M., Riddell, P. M., Ellis, J. A., Freeman, J. E. and Nasuto, S. J. (2012), 'Neural Correlates of True and False Memory in Mild Cognitive Impairment', *PLoS ONE* **7**(10), e48357.

Takens, F. (1981), Detecting strange attractors in turbulence, *in* D. Rand and L. Young, eds, 'Lecture Notes in Mathematics', warwick 1980 edn, Vol. 898, Springer, Berlin, Heidelberg, pp. p366–381.

Thanaraj, P., Noel, A., Raj, J., Balasubramanian, P. and Chen, Y. (2020), 'Schizophrenia detection using MultivariateEmpirical Mode Decomposition and entropy measures from multichannel EEG signal', *Biocybernetics and Biomedical Engineering* **40**(3), p1124–1139.

Theodoridis, S., Pikrakis, A., Koutroumbas, K. and Cavouras, D. (2010), *An Introduction to Pattern Recognition: A MATLAB Approach*, 1st edn, Elsevier.

Theodorsson-Norheim, E. (1986), 'Kruskal-Wallis test: BASIC computer program to perform nonparametric one-way analysis of variance and multiple comparisons on ranks of several independent samples', *Comput. Methods Programs Biomed* **23**(1), p57–62.

Tonmukayakul, U., Shih, S. T. F., Bourke-taylor, H., Imms, C., Reddihough, D., Cox, L. and Carter, R. (2018), 'Systematic review of the economic impact of cerebral palsy', *Research in Developmental Disabilities* **80**, p93–101.

Uljarevic, M. and Hamilton, A. (2013), 'Recognition of Emotions in Autism : A Formal Meta-Analysis', *Journal of Autism and Developmental Disorders* **43**, p1517–1526.

Volkmar, F. R., ed. (2013), *WPPSI-III – Wechsler Preschool and Primary Scale of Intelligence, Third Edition*, Springer New York, New York, NY.

Wang, X., Gong, G., Li, N. and Qiu, S. (2019), 'Detection Analysis of Epileptic EEG Using a Novel Random Forest Model Combined With Grid Search Optimization', *Frontiers in Human Neuroscience* **13**, p1–52.

Webb, C. A., Auerbach, R. P., Bondy, E., Stanton, C. H., Foti, D. and Pizzagalli, D. A. (2017), 'Abnormal neural responses to feedback in depressed adolescents', *J Abnorm Psychol* **126**(1), 19–31.

Weiss, L., Oakland, T. and Aylward, G. (2010), *Bayley-III Clinical Use and Interpretation*.

West, C. R., Battin, M. R., Williams, C. E., Dezoete, J. A. and Harding, J. E. (2005), '413 Early Quantitative Electroencephalographic Measures of Continuity are Associated with Neurodevelopmental Outcome at 18 Months in Preterm Infants', *ESPR EUROPEAN SOCIETY FOR PEDIATRIC RESEARCH* **58**(425).

World Health Organization (1993), *The ICD-10 classification of mental and behavioural disorders : diagnostic criteria for research*.

Wu, Z. and Huang, N. E. (2004), 'A study of the characteristics of white noise using the empirical mode decomposition method', *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* **460**, 1597 – 1611.

Wu, Z. and Huang, N. E. (2009), 'Ensemble Empirical Mode Decomposition: A Noise-Assisted Data Analysis Method', *Advances in Adaptive Data Analysis* **01**(01), p1–41.

Yeung, M. K., Han, Y. M., Sze, S. L. and Chan, A. S. (2014), 'Altered right frontal cortical connectivity during facial emotion recognition in children with autism spectrum disorders', *Research in Autism Spectrum Disorders* **8**(11), p1567–1577.

Yeung, M. K., Lee, T. L. and Chan, A. S. (2019), 'Impaired Recognition of Negative Facial Expressions is Partly Related to Facial Perception Deficits in Adolescents with High-Functioning Autism Spectrum Disorder', *Journal of Autism and Developmental Disorders* **50**(5), p1596–1606.

Yoshida, K., Shimizu, Y., Yoshimoto, J., Takamura, M., Okada, G., Okamoto, Y., Yamawaki, S. and Doya, K. (2017), 'Prediction of clinical depression scores and detection of changes in whole-brain using resting-state functional mri data with partial least squares regression', *PLoS One* **12**(7).

Yuexin, W., Yi, Z., Steven, S., Peng, X. and Dezhong, Y. (2017), 'Function NAMEMD applies the "Noise Assisted Multivariate Empirical Mode Decomposition" algorithm to multivariate inputs.'.
URL: *https://www.pudn.com/detail/4041428*

Zahra, A., Kanwal, N., Rehman, N., Ehsan, S. and Mcdonald-maier, K. D. (2017), 'Seizure detection from EEG signals using Multivariate Empirical Mode Decomposition', *Computers in Biology and Medicine* **88**, p132–141.

Zhang, J., Feng, F., Marti-Puig, P., Caiafa, C. F., Sun, Z., Duan, F. and Sol-Casals, J. (2021), 'Serial-emd: Fast empirical mode decomposition method for multi-dimensional signals based on serialization', *Information Sciences* **581**, 215–232.

Zwaigenbaum, L. and Penner, M. (2018), 'Autism spectrum disorder: Advances in diagnosis and evaluation', *BMJ* **361**, p1–16.