DOCTOR OF PHILOSOPHY

A Nonlinear Manifold Learning & Dynamical System Approach in Characterising Alzheimer's Disease

Gunawardena, Rajintha

Award date:
2023

Awarding institution:
Coventry University

Coventry University
Centre for Computational Science and Mathematical Modelling

A Nonlinear Manifold Learning & Dynamical System Approach in Characterising Alzheimer's Disease

Shenal Rajintha Alexander Samarathunge Gunawardena

Director of Studies: Dr. Fei He

Secondary Supervisor: Dr. Matthew England

A thesis submitted in partial fulfilment of the University's requirements for the degree of Doctor of Philosophy

October 16, 2023

Acknowledgements

I would like to thank my parents for their encouragement and support throughout my university life. Without them, their prayers and the sacrifices they have made for me, none of this would have been possible.

To my director of studies Dr. Fei He, my gratitude and my thanks for having given me this opportunity, for continuous guidance, for having faith in me and immense advice given during my PhD.

Thank you to all my friends at CSM who were there to support and advise me during my research and the writing of my thesis.

Last but not least I thank you, my GOD, the Father, Son and the Holy Spirit, for all of this was through thy grace. Thank you for all the counsel and wisdom blessed upon me.

Abstract

Alzheimer's disease, the leading cause of cognitive decline in older individuals, imposes a significant burden on healthcare and the economy. Currently, around 47 million people worldwide are affected by neurocognitive disorders, with a projected triple by 2050. Therefore, a need for cost-effective methods for early diagnosis is vital. AD is a degenerative neurological condition marked by brain disruptions and cognitive decline. Furthermore, the electrophysiology of brain cortical activity is shown to change before physical symptoms occur, suggesting the potential for early detection and intervention.

The electroencephalogram (EEG) is a cost-effective and non-invasive technique used to examine the electrophysiology of cortical activity and is extensively studied in relation to neurodegenerative diseases like Alzheimer's disease. It provides insights into the functioning and integrity of neural circuits indirectly. Additionally, the EEG can detect abnormalities in physiological processes that disrupt brain networks at an early stage, preceding clinical symptoms and visible structural changes in neuroimaging scans like Magnetic Resonance Imaging (MRI).

Extensive research has revealed the nonlinear nature of brain activity's electrophysiology, with complex temporal interactions among brain regions even during rest. Therefore, it is essential to consider spatial and temporal nonlinearities in EEG analysis to assess brain connectivity and dynamic interactions. This thesis examines changes in the brain cortex of individuals with mild to moderate Alzheimer's disease. Specifically, it aims to contribute to Alzheimer's disease characterisation through novel applications of nonlinear methods for connectivity and dynamic analysis of cortical interactions using resting-state EEG.

Using kernel-based nonlinear manifold learning techniques, Isomap and GPLVM (Isomap-GPLVM), a novel measure of linear and nonlinear connectivity is derived. This measure helps identify significant changes in nonlinear connectivity between spe-

cific brain regions in mild to moderate Alzheimer's disease. Isomap-GPLVM analysis uncovers significant connectivity differences between occipital bipolar channels and other regions (parietal, centro-parietal, and fronto-central) in Alzheimer's disease and a group of healthy controls. Furthermore, connectivity changes between fronto-parietal EEG channels and the rest of the channels are found to be crucial for Alzheimer's disease diagnosis. These results align with previous studies using functional MRI (fMRI), resting-state fMRI, and EEG, supporting links to resting-state functional networks in the brain.

Using Isomap-GPLVM analysis, significant changes in statistical dependencies between EEG channel pairs are further investigated for directed dynamic nonlinear dependencies using transfer entropy. This novel application of transfer entropy in characterising Alzheimer's disease with resting-state EEG uncovers increased intra-hemispheric information flow between parietal-occipital and centro-parietal-occipital regions, predominantly in the left hemisphere. These findings suggest a potential compensatory mechanism. These findings are consistent with previous studies utilising resting-state EEG and resting-state fMRI.

Cross-frequency interactions between different frequency ranges enable the integration of information from various brain regions. Previous studies have identified changes in cross-frequency interactions within the EEG associated with neurodegenerative diseases like Alzheimer's disease. The findings from using transfer entropy to examine directed nonlinear dependencies (direction of information flow) between important EEG channels are used to learn dynamic nonlinear input-output time-series models. These models are then analysed in the frequency-domain to examine cross-frequency interactions at higher-order nonlinearities. Data-driven modelling and analysis methods from control systems engineering, system identification and frequency response analysis, are used for this purpose.

Contents

List of Figures

List of Tables

List of Abbreviations

AD Alzheimer's Disease

AIC Akaike's Information Criteria

ARX Auto-Regressive with eXogenous input

BIC Bayesian Information Criteria

CF Cross-Frequency

CFC Cross-Frequency Coupling

CT Computerised Tomography

EC Eyes-close

EEG Electroencephalogram

EFC Effective Connectivity

EO Eyes-open

FC Functional Connectivity

fMRI functional MRI

FPE Final Prediction Error

FRF Frequency Response Function

FRO Forward Regression OLS

GC Granger Causality

GFRF Generalised Frequency Response Function

HC Healthy Control

iFRO iterative FRO

KDE Kernel Density Estimator

KNN k-nearest-neighbour estimators

LTI Linear time-invariant

MCI Mild Cognitive Impairment

MEG Magnetoencephalography

MRI Magnetic Resonance Imaging

NARX Nonlinear Auto-Regressive with eXogenous input

NARX Nonlinear Auto-Regressive with eXogenous input

NOFRF Nonlinear Output Frequency Response Function

OLS Orthogonal Least Squares

PD Parkinson's Disease

PDF Probability Density Function

PET Positron Emission Tomography

rsfMRI resting-state fMRI

SC Structural Connectivity

SVM-MCV Linear SVM classification with Monte-Carlo cross-validation

Chapter 1

Introduction

1.1 Background

Alzheimer's Disease (AD), the primary cause of cognitive decline in older individuals, imposes a considerable burden on both the healthcare system and the economy. Currently, around 47 million individuals worldwide are affected by neurocognitive disorders, and it is projected that this number will triple by the year 2050 [2]. AD is a degenerative neurological condition that progresses over time and is marked by the accumulation of amyloid plaques, neurofibrillary tangles, and widespread disruptions in brain function [250, 252]. AD is a neurodegenerative disease where patients experience a gradual decline in memory, executive function, and other cognitive skills, eventually reaching a point where they are unable to perform daily tasks independently [114]. Unfortunately, the underlying causes of AD remain largely unclear, with around 70% of the associated risks believed to be genetic, many of which are directly linked to specific genes [25]. Nevertheless, clear visual indications can be observed in individuals diagnosed with AD, as their brain size and shape undergo significant changes compared to healthy brains [2]. Furthermore, the electrophysiology of brain cortical activity is shown to change before physical symptoms occur [242, 62, 22, 222]. This suggests the possibility of early detection and initiation of diagnostic procedures to prevent or at least delay the progression of the brain towards the typical AD stage.

The diagnosis of important neurological disorders such as AD in early stages and the accurate disease progress characterisation can be vital for the treatment and im-

provement of the patient's life [75, 128, 30]. However, the current methods of diagnosis depend mainly on neuro-pathological examinations through mental status and monitoring of neuronal activity using neuroimaging scans [128]. These methods are expensive [113] and a need for a complementary cost-effective precise diagnosis method is critical [147]. The Electroencephalogram (EEG) is one such method. The EEG is non-invasive, and economical and has been previously widely researched in the context of neurodegenerative diseases [242, 62, 20]. Through the EEG, the behaviour and integrity of the underlying neural circuits can be indirectly studied [195]. Furthermore, EEG has the ability to detect abnormalities in disordered physiological processes that impact brain networks at an earlier stage, prior to the emergence of clinical symptoms and before structural changes become visible in neuroimaging scans such as the Magnetic Resonance Imaging (MRI) [173, 222]. This is because, the high temporal resolution of the EEG allows for accurate recording of the dynamically changing oscillations that form the foundation of brain activity [20, 197]. These oscillations appear to undergo changes in the early stages of AD [20, 110]. To advance the early diagnosis of AD, it is crucial to first characterise the disease in terms of brain connectivity (how different brain regions are connected functionally) and the dynamic changes that take place in how different brain regions interact [242, 245, 62, 64, 20, 7].

Earliest work related to the dynamical analysis of EEG in AD patients [246, 244], investigated the nonlinear dynamics of the EEG measures, and showed the nonlinear indices computed for the global brain electrical activity, exhibiting specific patterns of dysfunction in dementia. In the case of AD, when analysing the nonlinear dynamics of the EEG measures, researchers have demonstrated a reduction in the nonlinear complexity [122, 5]. Several comprehensive reviews on connectivity measures and dynamical analysis of EEG shed light on the importance of nonlinear techniques for the characterisation and diagnosis of neurodegenerative diseases [242, 62, 7]. Hence, there is an increasing interest in nonlinear methods in the analysis of EEG for characterising AD [23]. Consequently, the research presented in this thesis will focus on examining changes in the brain cortex of individuals with mild to moderate AD. More specifically, this work aims to contribute towards the disease characterisation of AD in terms of novel applications of nonlinear methods for connectivity analysis and dynamic analysis of the complex interactions between different cortical regions of the brain using resting-state EEG.

1.2 Motivation

The nonlinear nature of the electrophysiology of brain activity has been well studied. Through surrogate testing methods, the initial confirmation of the presence of nonlinear dynamics in the EEG, which was previously debatable, was confirmed [242, 42]. The temporal interactions between brain regions even at rest show a significant level of complexity [65, 87]. Therefore, it is crucial for methods that assess brain connectivity and dynamical interactions of brain regions to account for spatial and temporal nonlinearities across the EEG.

Assessing brain connectivity can be achieved by measuring the statistical dependence between corresponding EEG channels [20] and is characterised by different measures of (dis)similarity [185, 73, 66, 256, 6, 80, 14, 257, 64, 43]. Some of these measures are capable of analysing nonlinear structures present both locally and globally within the EEG data [280, 63, 185, 20, 43, 62] and are frequently used in various approaches for selecting EEG channels [15]. However, these measures do not have the same interpretation and can even mean the opposite [63, 185, 43]. Therefore, developing a generic measure of (dis)similarity, that accounts for spatio-temporal nonlinear interactions between EEG channels, is important for brain connectivity analysis and channel selection. This is one of the aims of the research presented in this thesis. This can be achieved by learning (dis)similarity information within the EEG by using kernel-based nonlinear manifold learning [249, 150]. The resulting kernel matrix from these methods is used as a novel measure of linear and nonlinear connectivity (statistical dependencies or functional connectivity) between EEG channels. Using this methodology, nonlinear connectivity changes between specific brain regions (EEG channel-pairs) that can distinguish mild to moderate AD compared to Healthy Control (HC) will be selected and used for analysing the nonlinear dynamic interactions between these brain regions. Selecting such important EEG channel-pairs is crucial for reducing computational complexity in implementing nonlinear dynamic analysis methods.

Conflicting impacts of ageing on brain functions have been documented, where older individuals exhibit decreased activity in certain brain regions but increased activity in others [213]. These findings challenge the conventional belief that ageing is solely associated with a straightforward pattern of cognitive and neural decline. As pointed out in [243], this contrasts with previous research indicating a general deterioration of both structural and functional brain integrity in AD. The work presented in this

thesis aims to explore these compensatory mechanisms in AD by examining the directed dynamic nonlinear dependencies or information flow between the selected EEG channel-pairs mentioned above. This analysis will determine the disruption or gain of dynamic nonlinear dependencies in relation to the direction of information flow between EEG channel-pairs. To achieve this, transfer entropy [228, 199] from the field of information theory, will be used. In the context of characterising AD, especially compensatory mechanisms, using the resting-state EEG, this application of transfer entropy is novel.

The coordination of the underlying spatio-temporal neural activity across different scales has become a key question in neuroscience [48, 115]. The dynamical interactions between larger groups of neurons modulating the activity of local neuronal dynamics have been shown to give rise to Cross-Frequency (CF) interactions [48, 51, 126]. Therefore, CF interactions have been hypothesised as the main carrier mechanism through which global and local processes interact enabling the integration of information from different brain regions [119]. In line with this, CF interactions have been reported in studies using electrophysiological data such as the EEG [126, 57, 77, 222, 115, 83]. Furthermore, certain changes in CF interactions within the EEG have been previously reported concerning neurodegenerative diseases such as AD [271, 80, 117].

In the field of control systems, dynamic analysis of linear and nonlinear systems is carried out in the frequency domain via frequency response analysis methods [32, Chapter 1]. As such, from the system input-output data, a time-series black box model is identified. This is known as system identification [32, Chapter 1]. The identified (learnt) model is then analysed in the frequency domain, using frequency response analysis methods. Furthermore, it should be emphasised that, given the identification process is done appropriately, unlike the black-box time-series models identified, a model's frequency response is a unique solution [32, Chapter 6]. Thus, the frequency response characteristics remain unchanged and unique for all local solutions. Therefore, using the frequency response analysis methods for dynamical systems, the CF interactions within the system can be analysed [32, Chapter 6]. System identification and frequency response analysis have been used to understand the nonlinear dynamics in complex engineering systems for control and fault diagnosis [32, Chapter 1]. These techniques have also been successfully applied in neuroscience to study the interactions in different brain regions using electrophysiological data [102, 104, 37, 99]. One of the main aims of the research presented in this thesis will focus on the novel application of the above mentioned control systems approach to characterise AD. As

previously mentioned, from applying transfer entropy to the selected EEG channel-pairs, the loss/gain of dynamic nonlinear dependencies with respect to the direction of information flow is examined on selected EEG channel-pairs. This knowledge will be used to build time-series models using system identification. The frequency response analysis of these models would enable the understanding of complex CF interactions between the selected channel-pairs in an input-output sense. Thus, the dynamic nature of the cortical electrical activity in mild to moderate AD can be characterised in the frequency domain. Furthermore, this would enable the understanding of the CF interactions involved with the compensatory mechanisms in AD.

1.3 Aims and Objectives

The main ambition of the research presented here is to apply nonlinear analysis techniques to characterise mild to moderate AD in comparison to HC. This is in relation to brain connectivity, directed nonlinear dependencies and nonlinear frequency-domain analysis. This involves the novel application of techniques from manifold learning, information theory and control systems engineering. Thus, this work would contribute towards formulating a complete data-driven framework for dynamic analysis of cortical neural activity. This would also contribute towards the early diagnosis and disease progress characterisation of neurodegenerative diseases such as AD. The wider key objectives and the novel contributions from this research are broken down as follows;

1. Develop a novel generic measure of (dis)similarity that accounts for spatial and temporal nonlinearities in EEG data, to identify connectivity changes between EEG channels specific to AD.
   (a) Develop a novel measure of linear and nonlinear connectivity between EEG channels using kernel-based nonlinear manifold learning.
   (b) Characterise mild to moderate AD in terms of brain connectivity.
   (c) Identify important EEG channel-pairs that can distinguish well between AD and HC.
   (d) Identify any biophysical links between the results.
2. Investigate compensatory mechanisms in AD by examining dynamic nonlinear dependencies respective to the direction of information flow between selected

(from 1. above) EEG channel-pairs.

(a) Identify an appropriate methodology for the application of transfer entropy, from information theory to EEG data avoiding spurious nonlinear dependencies.

(b) Utilise this methodology to assess disruptions or gains in dynamic nonlinear dependencies with respect to the direction of information flow.

(c) Characterisation of compensation mechanisms involved in AD using directed nonlinear dependencies. In this aspect, the application of transfer entropy to the resting-state EEG is novel.

(d) Identify any biophysical links between the results.

3. Novel application of a control systems approach to characterise CF interactions involved in the loss/gain of dynamic nonlinear dependencies with respect to the direction of information flow between the selected channel-pairs (from 1.).

(a) Apply system identification to the selected channel-pairs to build input-output time-series models based on the knowledge gained from 2. The direction of information flow between the channels is used to determine which channel is to be the input and the output.

(b) Analyse the frequency response characteristics of the identified models to understand CF interactions between the selected channel-pairs in the frequency-domain.

(c) Novel characterisation of AD and the compensatory mechanisms involved in the frequency-domain.

Overall, the research aims to assess brain connectivity and dynamic nonlinear dependencies, understand compensatory mechanisms and investigate CF interactions involved, to characterise mild to moderate AD using the resting-state EEG data. Thus, the work aims to contribute towards the development of a complete data-driven framework for the analysis and characterisation of neurodegenerative diseases.

1.4 Thesis overview

The thesis is organised into 6 chapters in the following manner;

- Chapter 2 overviews different brain connectivity methods while pointing out the issues and how these issues are mitigated in this thesis to characterise complex abnormal dynamics in AD. The chapter points out the dynamic nonlinear nature of the brain electrical activity and highlights the importance of considering methods that can capture nonlinear interactions between brain regions. Consequently pointing out what nonlinear methods will be used, and why these methods are suitable for characterising brain electrical activity in AD.

- Chapter 3 focuses on a novel brain connectivity analysis method that generalises (dis)similarity while accounting for spatio-temporal nonlinearities. The chapter demonstrates how kernel-based manifold learning can be used as a measure of spatio-temporal functional connectivity between EEG channels to determine the important inter-relationships in characterising patients with mild to moderate AD. The methodology presented can determine changes in cortical (EEG channel) inter-relationships that are crucial in distinguishing AD patients from HCs. Furthermore, the results reported in this chapter are consistent with other previous studies while linking connectivity changes to functional networks.

- Chapter 4 presents the novel application of transfer entropy to understand the compensation mechanisms involved in AD using resting-state EEG. This chapter explores these compensatory mechanisms by examining the directed dynamic nonlinear dependencies or information flow between the selected EEG channel-pairs from Chapter 3. A specific surrogate testing method to mitigate spurious nonlinear dependencies is used. This analysis will determine the disruption or gain of dynamic nonlinear dependencies in relation to the direction of information flow between EEG channel-pairs. The findings from this chapter are comparable to other studies based on resting-state EEG and rsfMRI.

- Chapter 5 is a novel application of system identification and frequency response analysis to characterise CF interactions in mild to moderate AD. Building upon the findings of Chapter 4, data-driven models are constructed to capture the nonlinear dynamics between pairs of channels. These models are then examined in the frequency domain to identify noteworthy alterations in CF interactions between the AD and HC groups. This chapter analyses directed broadband CF interactions between cortical regions at an EEG sensor level.

- Chapter 6 summarises the findings from this work, and potential future research

directions based on these findings are emphasised in the concluding remarks.

1.5 Summary of contributions and research outputs

Novel contributions from the work presented in this thesis for characterising AD using machine-learning, information theory and methods from controls systems engineering are listed below. The research undertaken aims to contribute towards the advancement of a comprehensive data-driven framework for analysing and characterising neurodegenerative diseases.

1. **Chapter 3** : This chapter introduces a novel generic measure of (dis)similarity for brain connectivity analysis and channel selection. This is achieved by learning the (dis)similarity information within the EEG using kernel-based nonlinear manifold learning. The focus is on functional connectivity changes and, thereby, EEG channel selection.

2. **Chapter 4**: An explanatory analysis using transfer entropy to identify changes in directed dynamic nonlinear dependencies or changes in information flow in AD is presented in this chapter. Transfer entropy is applied in a sliding window fashion to account for the time-varying behaviour of the brain. This type of analysis in relation to understanding the compensatory mechanisms involved in AD, using the resting-state EEG is novel.

3. **Chapter 5**: In this chapter a novel application of system identification and frequency response analysis methods from control systems engineering is presented. The methods mentioned above are used to characterise the cortical activity in AD in relation to changes in CF interactions involved in the compensatory mechanisms.

Based on the work presented in Chapter 3 on using kernel-based manifold learning for brain connectivity analysis. A conference paper, 'Kernel-based Nonlinear Manifold Learning for EEG Functional Connectivity Analysis with Application to Alzheimer's Disease. 2022 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)' [93] has been published. Furthermore, a journal paper based on the same Chapter has been published in the journal Neuroscience [94]; 'Kernel-based Nonlinear Manifold Learning for EEG-based Functional Connectivity Analysis and Channel Selection with

Application to Alzheimer's Disease'.

Chapter 2

Nonlinear Dynamical Analysis of the EEG and Neurodegenerative Diseases

2.1 Introduction

Computerised Tomography (CT) scans [116], MRI [137], and Positron Emission Tomography (PET) [194] are promising neuro-imaging techniques to aid with the early diagnosis of AD. These imaging techniques detect structural and functional changes in the brain [231]. Magnetoencephalography (MEG) and EEG can be used to identify functional changes in the cerebral cortex [242]. With similar diagnostic sensitivity and specificity with neuro-imaging techniques [8, 242], MEG and EEG have emerged as promising tools for the diagnosis of AD.

Certain outer areas of the brain are functionally coupled with the central regions of the brain (Fig. 2.1) [112, 215, 81, 59]. Thus, the complex processors within the centre of the brain (sub-cortical layers) interact with the surface layers (cortical layers, Fig. 2.1A). The EEG, recorded at the scalp level, reflects grossly summed currents of the electrical fields generated by neural activity in cortical neural circuits [195, 218]. The EEG is widely used to study the functional state of the outer as well as the central layers of the brain [242, 20, 27]. Therefore, analysing hidden structures within EEG data, in general, and in the context of AD, is important and has gained considerable

attention [242, 64, 218]. This is achieved using connectivity analysis.

The electrical activity of the brain can be divided into distinct frequency bands that correspond to various states of neural function. These bands are named delta, theta, alpha, beta and gamma [21]. These frequency bands, also known as brain waves, offer significant insight into the operational dynamics of the brain [48, 233]. Several recent studies have demonstrated the usefulness of EEG biomarkers in diagnosing and monitoring the progression of AD [85, 158, 22, 125]. Individuals with AD typically exhibit decreased power in the frequency bands alpha and beta and increased power in the frequency bands theta and delta in various brain regions [110]. Abnormal changes in brain connectivity measures have also been observed [64, 85, 176, 158, 22, 135]. [125] found specific neural biomarkers associated with cognitive function in AD patients, including changes in the power spectrum of low-frequency oscillations in the occipital area and altered signal complexity in the parietal and occipital regions. They also determined that spectral density features and entropy were key EEG biomarkers in differentiating between HC and patients with AD and mild cognitive impairment.

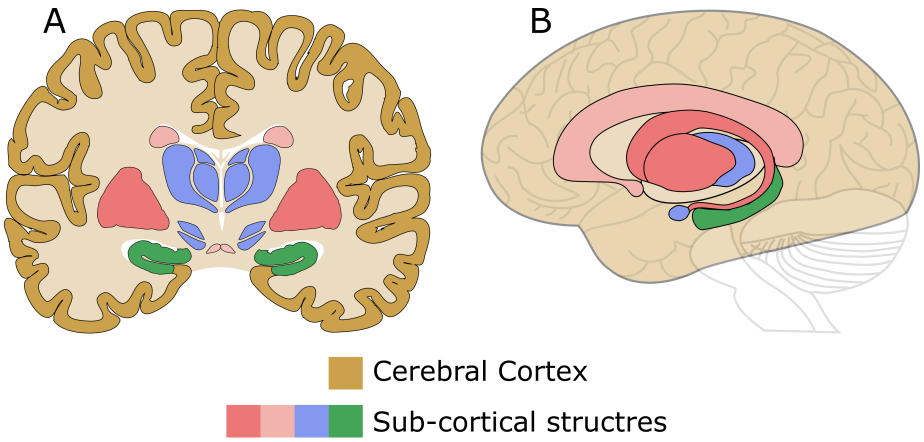


Figure 2.1: ***Basic anatomy of the brain highlighting the cerebral cortex and the basal ganglia.*** ***A)*** *front cross-section of the brain, showing the cerebral cortex (cortical layers) in a dark yellow and structures related to the sub-cortical layers highlighted in various other colours.* ***B)*** *illustrates a side outer view of the brain. The cerebral cortex is shown in dark yellow (opaqued) and the internal structures of the brain relating to the sub-cortical layers are shown in various colours.*

Brain connectivity analysis has become a crucial aspect of computational neuroscience research to better understand the brain in terms of its inter-relationships between brain regions or neuronal groups [26, 202, 7]. Connectivity, in neuroscience,

refers to the physical, statistical or causal links between different parts of the brain or neuronal groups [78, 264]. There are three types of connectivity: structural, functional, and effective. Structural connectivity (SC) pertains to the physical or anatomical connections between brain regions or neuronal populations [60]. Therefore, SC measures the tangible physical links between the components within the network. Functional connectivity (FC), essentially, captures the statistical dependence between physiological time-series from spatially separated neuronal groups or cortical regions. This is done either in the time-domain, frequency-domain or the time-frequency domain [63, 279, 26, 185, 7]. On the other hand, effective connectivity (EFC) looks into the influence of one brain network entity (neuronal population or cortical region) over another [84]. Unlike FC measurements, which are non-causal, EFC is directional and sometimes depends on a dynamic model [84].

The above mentioned analysis methods can be conducted at different scales, from between individual neurons to groups of neurons and between cortical regions [222]. This study focuses on EEG based FC and EFC, to understand the changes in cortical activity (via FC) and then differences in both linear and nonlinear cortical dynamics (via EFC) between AD and HC cohorts. EEG based FC/EFC refers to a collection of measurements that quantify connectivity between different low-level networks at a higher level. These low-level networks consist of interconnected neurons that span an area of the cortex that is larger than $1cm^2$ and can vary in size based on factors such as the local density of cells, depth of the region, and the direction of current flow within the region [187, 195].

2.2 Functional connectivity

FC assesses the statistical dependence between brain regions (EEG/MEG channels) [20] and is characterised by different measures of (dis)similarity [185, 73, 66, 256, 6, 80, 14, 257, 64, 43], such as distance measures, entropy and mutual information [43, 185, 63]. Some of these measures can be used to analyse nonlinear structures present locally and globally within the EEG data [280, 63, 185, 20, 43, 62] and are often used in many EEG channel selection approaches [15]. However, regardless of structural connectivity, brain regions functionally connected under one measure do not necessarily imply the same with another measure, as they could even be disconnected [63, 185, 43].

Similarity or dissimilarity between two variables (EEG channels), in general, express

the degree to which the two objects are respectively alike/related or different/distinct [148, 237]. Local similarities refer to the relatedness or correlation between nearby data points. This entails that data points closer together in space, time, or any other relevant dimension tend to have similar values or characteristics. Conversely, global dissimilarities refer to a lack of correlation or differences between data points that are far apart from each other. It is important that FC measures account for both local similarities and global dissimilarities within the EEG [63, 185, 43].

Dauwels *et al.* [63] showed that various (dis)similarity measures could be correlated to each other, such as in the application of early diagnosis of AD. These correlated measures can often be grouped, and a measure from each group is sufficient to analyse the structures within the data [63]. Therefore, the development of a generic measure of (dis)similarity is important for analysing brain FC [43]. In this study, a novel EEG FC analysis method is introduced (Chapter 3), attempting to generalise FC by learning the spatial and temporal structures within the data and quantifying this information as a generic measure of (dis)similarity. This can be used to assess linear and nonlinear FC, between EEG channels. This is achieved using kernel-based nonlinear manifold learning to determine a subset of important channel inter-relationships that can discriminate well between AD and HC groups. This sub-set of channels will then be used for dynamical analysis (EFC) to identify dynamical changes due to neurodegeneration, e.g. AD, on global and local brain dynamics.

2.3 Effective Connectivity

EFC aims to understand the direction of information flow and how one neuronal population or cortical region influences another. For this type of analysis, mathematical models (hypothesis-driven) [84, 42] and statistical techniques for data-driven models [41, 99] are used to examine the strength and direction of interactions and to infer underlying neural dynamics. In contrast, the analysis of FC is mostly descriptive [84]. Both hypothesis-driven mathematical models and data-driven models that infer EFC can be defined within the frameworks of dynamic systems theory–dynamic systems approach [247, 41, 84, 42].

A systems perspective, also referred to as a systems viewpoint or systems approach, involves the observation of all the individual procedures, and sub-processes as a unified entity. This viewpoint considers only the primary inputs and outputs, with a focus

on the overall system as a whole [134]. A system whose state evolves with time as a function of a previous state is a dynamical system [216, 217]. Therefore, these systems possess memory. Thus, a dynamical system is defined by its state in time and the dynamics which describe the changes of the state over time. The state of a dynamical system is given by the quantities of all the variables that describe the system at a particular point in time. The dynamics of the state are described by a set of differential equations (continuous-time representation of a dynamical system) or difference equations (also called a mapping function for discrete-time representation of a dynamical system). The structures of these representations are either determined using physical laws (hypothesis-driven) that govern the system [216, 217] or can be identified using data-driven methods [164, 44]. Dynamic systems theory is the study of long-term behaviour in dynamic systems.

In neuroscience, concerning model-based effective connectivity, hypothesis-driven modelling techniques are informed using neuro-anatomical connectivity (connectome) studies to define the structure of the model [247, 84]. Two popular hypothesis-driven modelling methodologies are dynamic causal modelling (DCM) and structural equation modelling (SEM) [49]. DCM uses a Bayesian approach to evaluate model efficacy, while structural equation modelling (SEM) is a generalized linear modelling framework that combines path modelling with factor analysis [247]. Both DCM and SEM consider the brain as a deterministic nonlinear or linear system.

In contrast to DCM and SEM, which depend on prior knowledge about connectivity, several data-driven modelling techniques based on Granger causality (Granger Causality (GC)) [90] have been developed for measuring effective connectivity [232]. A time-series can be predicted using its own past values, GC quantifies the improvement in the quality of prediction of one time-series by incorporating the past values of another [274, 90]. This helps to determine if the predicted time-series was impacted by the past of the other time-series used in the prediction, which can uncover the direction and the strength of information flow between the two series being examined [90, 285, 103]. Furthermore, GC has recently been extended to handle nonlinearities [285, 103]. Directed Transfer Function (DTF) [132] and Partial Directed Coherence (PDC) [224] are methods that are based on GC for data-driven Effective Connectivity (EFC). However, these two methods are purely linear and cannot capture nonlinear interactions.

Completely model-free measures for EFC do exist [24] such as transfer entropy [228, 266] which has gained much interest in the field of neuroscience to infer EFC, however, most EFC measures are based on hypothesis-driven or data-driven dynamic models [84, 24, 99]. Nevertheless, the primary benefit of using transfer entropy to infer EFC is that it doesn't rely on a specific model for the interaction between the two systems being studied. This makes it advantageous for exploratory analyses compared to hypothesis-based or data-driven methods. This is especially useful when trying to identify unknown non-linear interactions. In essence, the sensitivity of transfer entropy to higher-order correlations (nonlinear interactions) [193] provides an advantage in these scenarios.

When deciding between model-based and data-driven techniques for a specific problem, it is important to consider their distinct underlying assumptions. Model-based approaches, such as DCM, rely on well-defined biophysical models of neuronal dynamics. As such, care needs to be taken when selecting the most suitable model or a combination of models and exploring various parameters to test a predetermined hypothesis [223]. However, a major drawback of model-based techniques is the uncertainty involved in predefining these parameters and the potentially vast number of parameter combinations that need to be considered. Established methodologies can be helpful in determining the most appropriate model [186]. Nevertheless, it is possible that no single model is sufficient, and multiple models could be equally suitable for a given dataset [223]. Conversely, data-driven methods do not assume any specific spatial or temporal relationships as this information is obtained from the data itself. Data-driven methods can be employed to assess connectivity even when prior knowledge about the underlying structure is unavailable, and this is true in the case of abnormal brain dynamics [67] such as in AD. Therefore, in characterising EFC and abnormal brain dynamics in AD, model-free and data-driven model-based methods are more suitable.

Data-driven model-based EFC is strongly linked to system identification–data-driven dynamic modelling techniques used in control systems theory [227, 99]. Once an appropriate model is identified, frequency response analysis can be used to understand the dynamics of the learnt model in the frequency domain. Through frequency response analysis the linear frequency profile and the nonlinear frequency interactions (couplings) within the modelled system can be analysed in depth [32, Chapter 6][100, 101]. In general, it is challenging to establish predetermined governing equations for brain

activity [92], especially in the case of abnormal brain dynamics [67]. Therefore, the control systems perspective to data-driven dynamic modelling and analysis of EEG signals has recently attracted significant interest [103, 104, 157, 91, 99]. The application of system identification and frequency response analysis to understand the changes in the cortical dynamics in the context of AD is novel and, thus will be one focus of the research undertaken. As such, novel applications of techniques for Functional Connectivity (FC) analysis, model-free EFC and data-driven EFC will be presented in this thesis for characterising AD.

In the work presented in this thesis, more specifically, important channels-pairs that have significant changes in statistical dependencies between HC and AD groups will be first selected using FC (Chapter 3). Once these channel-pairs are identified, an exploratory analysis will be conducted using transfer entropy to examine the nonlinear dependencies in directed information flow between each pair of channels (Chapter 4). The aim is to gain insight into which selected channel-pairs display significant differences in directed information flow in AD. This prior knowledge regarding the information flow and its direction will be utilised to construct data-driven dynamic input-output time-series models through system identification (Chapter 5). Subsequently, the dynamics of these models will be analysed using frequency response analysis to understand the cortical nonlinear dynamic changes occurring in AD.

The next section will briefly highlight the developments in the analysis of cortical dynamics using the EEG in the context of neurodegenerative diseases such as AD, highlighting the significance of using methods that can capture nonlinearities. This will be followed by a section reviewing the nature of linear and nonlinear dynamic systems in the time and frequency domains and how a control systems approach is suitable for brain dynamic analysis.

2.4 Brain Dynamic Analysis Using the EEG in Relation to Neurodegenerative Diseases

The EEG is understood as an outcome of nonlinear deterministic dynamics, with the possibility of being a chaotic process [244, 41]. The EEG is commonly observed as a continuous spectrum, and the irregularity or the aperiodic nature of EEG signals [246] can be assumed, using dynamic systems theory [247, 41, 84], to be the result of

a low-dimensional nonlinear deterministic system. As such, it may be inferred that parsimonious models are to be used for the explanation of EEG complexity [41, 84]. Understanding these dynamics leads to a comprehensive interpretation of the functional relationship between cortical and sub-cortical neural circuits. A clear change in these dynamics is observed when analysing patients with neurodegenerative diseases such as AD [205, 242, 141]. Thus, the significance of comprehensive methodologies to further analyse the nonlinear dynamics in the brain is clearly apparent [242, 42]. Data-driven nonlinear dynamic analysis of EEG measures has evolved from nonlinear indices such as correlation embedding, Lyapunov exponents and entropy to model-based nonlinear time-series analysis [242]. As such, the development of nonlinear time-series analysis of EEG measures is a clear direction in the development of EEG analysis [242].

In light of the above discussion and section 2.3, the control systems perspective on dynamic analysis of EEG signals becomes more apparent. In nonlinear system identification, usually, the dynamic models that are learnt or identified are nonlinear auto-regressive time-series models [52]. The link between the nature of these models and Lyapunov exponents has been well established [181]. Since the Lyapunov exponent is a significant factor when differentiating the dynamic profiles between healthy individuals, AD and Parkinson's Disease (PD) patients [242, 205, 122, 64], the application of system identification and frequency response analysis takes an interesting viewpoint in the analysis of cortical dynamics and CF interactions between cortical regions. The application of this approach to the dynamic analysis of EEG, the early diagnosis of neurodegenerative diseases such as AD and PD, is novel.

Nonlinear systems, in both the time and frequency domains, exhibit significantly more complexity compared to linear systems. Consequently, in order to understand this intricate behaviour and effectively analyse dynamic nonlinear systems in the frequency domain, the subsequent section will provide an overview of the characteristics of such systems in both the time and frequency domains.

2.5 The Dynamic Nature of Linear and Nonlinear Systems

If a system maintains consistent dynamic properties over time, it is referred to as time-invariant. Systems that adhere to the superposition principle and are time-

invariant are known as linear time-invariant (Linear time-invariant (LTI)) systems [163]. Furthermore, if an LTI system is causal, it means that the output at a specific time is solely determined by the input up to that time. In this case, at time t, the system can be effectively described as the convolution between the impulse response $h(t)$ of the system and the input signal $u(t)$.

$$y(t) = \int_{\tau=0}^{\infty} h(\tau)u(t-\tau)d\tau \tag{2.1}$$

where $y(t)$ represents the output of the system, and τ represents a time delay index. The impulse response, $h(t)$, characterises the system's behaviour in the time domain, representing its response when stimulated by a unit impulse. Traditional systems theory, based on LTI systems, is a well-established field of study. However, it's important to note that the LTI property of a system is typically an approximation, which is often justifiable.

Nonlinear systems, on the other hand, encompass systems that do not adhere to the superposition principle, and their behaviour is significantly more intricate. In the case of LTI systems, the frequency components of the output response mirror those of the input signal. Conversely, in nonlinear systems, the output response exhibits a broader range of frequency components compared to the corresponding input signal. This occurs because nonlinear systems possess the capability to transfer energy between frequency components, including those that are absent in the input signal [144]. Describing nonlinear systems in the time-domain involves extending the concept of convolutional integrals used in LTI systems, as represented by eq.(2.1), to a series of multidimensional convolution integrals referred to as the Volterra series. The Volterra series explains a category of stable nonlinear systems at the zero equilibrium, within the vicinity of the equilibrium point, as depicted in eq.(2.2) below.

$$\begin{cases} y(t) = \sum_{n=1}^{N} y_n(t) \\ y_n(t) = \int_{-\infty}^{+\infty} \cdots \int_{+\infty}^{-\infty} h_n(\tau_1, \cdots, \tau_n) \prod_{n}^{i=1} u(t-\tau)d\tau_i \end{cases} \tag{2.2}$$

where $h_n(\tau_1, \cdots, \tau_n)$ refers to the n^{th} order Volterra kernel, which represents the time domain characteristics of the nonlinearity specific to the n^{th} order. $y(t)$ denotes the

output of the system, $u(t)$ represents the input to the system, and $y_n(t)$ is referred to as the n^{th} order nonlinear output or the output produced by the n^{th} order nonlinearity of the system. $n = 1, \cdots, N$, where N represents the highest order of nonlinearity considered [221].

Equations (2.1) and (2.2) represent linear and nonlinear systems in the time domain. These representations are useful for examining system dynamics, including transient response analysis, for both linear and nonlinear systems. While in nonlinear systems, this includes the study of bifurcations, limit cycles and chaotic regimes [189]. By applying the Fourier transform to these time-domain representations, the frequency-domain representations are obtained. In the frequency-domain, the relationships between input frequencies and the output amplitude and phase-shift can be studied for both linear and nonlinear systems. For nonlinear systems, this enables a comprehensive study of nonlinear phenomena, such as energy transfer from input frequencies to output frequencies that are not present in the input signal–generation of harmonics and inter-modulations that result in CF interactions at various orders of nonlinearity. The following subsections will outline the frequency-domain representations of linear and nonlinear systems.

2.5.1 Linear systems in the frequency-domain

Frequency domain analysis of linear systems is conducted using the well-established technique of linear frequency response functions Frequency Response Function (FRF)s. The linear FRF provides a comprehensive depiction of the steady-state dynamics of a linear system and is unique regardless of the time-domain model employed to represent the system [163]. It is the quantitative measure of the output spectra resulting from an input stimulus. It enables observation of the system dynamics across a range of frequencies and is represented as a bode plot, showing the relationship between output magnitude, phase, and input frequency [68]. Essentially, it represents the ratio of the output spectrum to the input spectrum. In the case of LTI systems, the output frequency response $Y(j\omega)$, for all frequencies ω, can be explicitly characterised for any input signal, given the knowledge of the FRF as

$$Y(j\omega) = G(j\omega)U(j\omega) \tag{2.3}$$

where $Y(j\omega)$ and $U(j\omega)$ are the input and output spectra, respectively, and $G(j\omega)$ is the FRF which is a nonparametric representation of a linear system in the frequency-domain.

2.5.2 Frequency-domain representations of nonlinear systems

In linear systems, it is widely known that the output frequency components correspond precisely to the frequency components present in the input signal [163, 68]. However, this relationship does not hold true in the case of nonlinear systems, where the output frequency response exhibits significantly more intricate behaviour. For example, when considering a nonlinear system with a single-frequency input component, ω_1. The resulting output may include not only the original input frequency component ω_1, but also its super-harmonics, such as $2\omega_1$ and $3\omega_1$, as well as sub-harmonics, such as $\frac{\omega_1}{2}$ and $\frac{\omega_1}{3}$, and so forth. However, when the input comprises multiple frequency components, such as ω_1, ω_2, and ω_3, the output exhibits a broader range of possible frequency components. These components encompass the original input frequencies ω_1, ω_2, and ω_3, the super-harmonics, sub-harmonics, and frequencies resulting from the inter-modulations between the input frequencies. Examples of such inter-modulations could be $\omega_1 - \omega_2$, $\omega_1 - \omega_2 + \omega_3$, $\omega_1 + \omega_3$, and numerous others. Consequently, nonlinear systems possess distinct characteristics in comparison to linear counterparts. Specifically, the output frequency response of nonlinear systems is considerably more diverse, featuring an abundance of frequency components beyond those present in the input spectra. This is known as the nonlinear phenomenon, where energy is transferred from the input frequency modes to other frequency modes [144, 204, 206, 146, 145]. This subsection provides an overview of the frequency-domain representations of nonlinear systems, namely the Generalized Frequency Response Functions (Generalised Frequency Response Function (GFRF)s) and the Nonlinear Output Frequency Response Functions (Nonlinear Output Frequency Response Function (NOFRF)s). These approaches extend the concept of linear FRFs, to the nonlinear instance.

2.5.3 Output frequency response of non-linear systems

In subsection 2.5.1, the output frequency response properties of linear systems are defined by eq.(2.3). However, as mentioned earlier, due to the intricate nature of the output frequency response in nonlinear systems, eq.(2.3) does not hold in the nonlinear

context. The time-domain representation of nonlinear systems can be expressed using the Volterra series, eq.(2.2). Building upon this representation, Lang and Billings [145] derived an expression to characterise the output frequency response $Y(j\omega)$ of a nonlinear system as

$$\begin{cases} Y(j\omega) = \sum_{n=1}^{N} Y_n(j\omega) \quad \forall\omega \\ Y_n(j\omega) = \dfrac{1/\sqrt{n}}{(2\pi)^{n-1}} \displaystyle\int_{\omega=\omega_1+\cdots+\omega_n} H_n\left(j\omega_1, \cdots, j\omega_n\right) \prod_{n}^{i=1} U(j\omega_i) \quad d\sigma_{n\omega} \end{cases} \tag{2.4}$$

where N represents the maximum order of nonlinearity under consideration. $Y_n(j\omega)$ denotes the frequency characteristics of the n^{th} order nonlinear output (i.e., the output frequency response of the n^{th} order nonlinearity), while $U(j\omega)$ represents the frequency spectrum of the input. The n^{th} order GFRF, denoted as $H_n\left(j\omega_1, \cdots, j\omega_n\right)$, provides a description of the dynamic characteristics of the n^{th} order nonlinearity in the frequency domain. $d\sigma_{n\omega}$ represents an infinitesimally small region within the hyperplane $\omega = \omega_1 + \cdots + \omega_n$. Consequently, the GFRFs serve as direct extensions of the linear Frequency Response Functions (FRFs) to the nonlinear case. Furthermore, the GFRF remains unique irrespective of the specific time-domain model employed to represent the corresponding nonlinear system. The expression in eq.(2.4) above represents the natural extension of eq.(2.3), which characterises the output frequency response of linear systems, to the nonlinear case.

The generation of the output frequencies of a nonlinear system, as shown by eq.(2.4) is the sum over the output frequencies contributed by each order of nonlinearity, $Y_n(j\omega)$. Thus by considering the 'output frequency range' f_{Y_n} (the frequency space) of each nonlinear order $Y_n(j\omega)$, the 'frequency range' f_Y of the output response, $Y(j\omega)$, of a nonlinear system is much greater [146, 145] such that

$$f_Y = \bigcup_{n=1}^{N} f_{Y_n}. \tag{2.5}$$

2.5.4 Generalised Frequency Response Functions (GFRFs)

The intricate nature of nonlinear systems, as previously discussed, surpasses the complexity observed in linear systems. This distinction becomes apparent when exam-

ining the time-domain representation of nonlinear systems using the Volterra series, eq.(2.2), in contrast to the linear counterpart in eq.(2.1). In the case of nonlinear systems, the time-domain dynamics of each order of nonlinearity are described by separate multidimensional Volterra kernels. The Fourier transform of the time-domain impulse response of a linear system, eq.(2.1), corresponds to the linear FRF [162]. Similarly, George introduced the concept of GFRFs in [86]. These GFRFs are defined as the Fourier transform of the Volterra kernels, represented by the terms $h_n(\tau_1, \cdots, \tau_n)$ in eq.(2.2). Consequently, the GFRF of an n^{th} order nonlinearity can be expressed as

$$H_n\left(j\omega_1, \cdots, j\omega_n\right) = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} h_n(\tau_1, \cdots, \tau_n) \times e^{-(\omega_1\tau_1 + \cdots + \omega_n\tau_n)j} \; d\tau_1, \cdots, d\tau_n \tag{2.6}$$

Hence, the notion of GFRFs serves as a direct extension of the linear FRF to the nonlinear case. As evident from eq.(2.6), GFRFs are multidimensional frequency functions that capture the intricate dynamics associated with each order of nonlinearity.

In linear systems, the FRF can be explicitly used to characterise the output frequency response, eq.(2.3). However, this approach does not hold for nonlinear systems due to the involvement of high-dimensional frequency functions associated with each order of system nonlinearities [145, 129]. To elucidate the complex relationship between GFRFs, and the system's output frequency response, Lang and Billings [145] derived the expression presented in eq.(2.4). As previously discussed, this expression describes the output frequency response of nonlinear systems in terms of the GFRFs. It illustrates how the dynamics of the n^{th} order nonlinearity (represented by the n^{th} order GFRF) interact with the input spectrum to generate the output frequency response of the corresponding nonlinearity, denoted as $Y_n(j\omega)$. Consequently, the summation of all the output spectra from all nonlinearities yields the actual output frequency response, $Y(j\omega)$, of the system.

As mentioned earlier, nonlinear systems exhibit a distinctive characteristic whereby energy is transferred from the input frequencies to other frequencies that are absent in the input signal. This phenomenon arises from the interplay between the n^{th} order nonlinear dynamics and the input spectrum, as discussed above. Thus, the comprehensive explanation of this phenomenon extends beyond the capabilities of the GFRFs [144]. To address this, Lang and Billings introduced the concept of NOFRFs in [144]. These NOFRFs comprehensively capture the energy transfer from the input to differ-

ent orders of system nonlinearities and, consequently, the generation of new frequency components. The NOFRFs can be viewed as another extension of the linear FRF to the nonlinear domain and serve as a complement to the GFRFs.

2.5.5 Non-linear Output Frequency Response Functions (NOFRFs)

Lang and Billings [144] introduced the concept of NOFRFs to elucidate the energy transfer phenomenon in nonlinear systems. Additionally, in [144], the authors proposed another concept that serves as a natural extension of the input spectrum $U(j\omega)$ to the n^{th} order nonlinear case, and it is expressed as

$$U_n(j\omega) = \frac{1/\sqrt{n}}{(2\pi)^{n-1}} \int_{\omega=\omega_1+\cdots+\omega_n} \prod_n^{i=1} U(j\omega_i) \;\; d\sigma_{n\omega} \tag{2.7}$$

where $U_n(j\omega)$ is the nonlinear composition of the input $U(j\omega)$ to the n^{th} order nonlinearity and is related to the Fourier transform, $FT\{\ \}$, of $u^n(t)$ such that

$$U_n(j\omega) = \frac{1/\sqrt{n}}{(2\pi)^{n-1}} \times FT\left\{u^n(t)\right\} \tag{2.8}$$

The significance of $U_n(j\omega)$ lies in its role in elucidating the energy transfer phenomena observed in nonlinear systems and, thus, the definition of the NOFRFs, as shown in [144], which will be examined in the following discussions.

The second expression in eq.(2.4), the definition for $Y_n(j\omega)$, can be written in terms of eq.(2.7).

$$Y_n(j\omega) = \frac{\int_{\omega=\omega_1+\cdots+\omega_n} H_n\left(j\omega_1, \cdots, j\omega_n\right) \prod_n^{i=1} U(j\omega_i) \;\; d\sigma_{n\omega}}{\int_{\omega=\omega_1+\cdots+\omega_n} \prod_n^{i=1} U(j\omega_i) \;\; d\sigma_{n\omega}} \times \left(\frac{1/\sqrt{n}}{(2\pi)^{n-1}} \int_{\omega=\omega_1+\cdots+\omega_n} \prod_n^{i=1} U(j\omega_i) \;\; d\sigma_{n\omega} \right) \tag{2.9}$$

where the expression in the parenthesis is $U_n(j\omega)$. Therefore, $Y_n(j\omega)$ can be re-written as

$$Y_n(j\omega) = G_n(j\omega) U_n(j\omega) \tag{2.10}$$

where

$$G_n(j\omega) = \frac{\int_{\omega=\omega_1+\cdots+\omega_n} H_n\left(j\omega_1, \cdots, j\omega_n\right) \prod_n^{i=1} U(j\omega_i) \;\; d\sigma_{n\omega}}{\int_{\omega=\omega_1+\cdots+\omega_n} \prod_n^{i=1} U(j\omega_i) \;\; d\sigma_{n\omega}}. \tag{2.11}$$

In eq.(2.11) above $G_n(j\omega)$ is defined as the n^{th} order NOFRF and is only valid in the frequency space ω where

$$\int_{\omega=\omega_1+\cdots+\omega_n} \prod_n^{i=1} U(j\omega_i) \;\; d\sigma_{n\omega} \neq 0. \tag{2.12}$$

Hence, the output frequency response $Y(j\omega)$ of a nonlinear system, as depicted in eq.(2.4), can be explicitly represented by using the NOFRFs $G_n(j\omega)$, where $n = 1, \cdots, N$, by introducing eq.(2.10) into eq.(2.4), yielding:

$$Y(j\omega) = \sum_{n=1}^{N} Y_n(j\omega) = \sum_{n=1}^{N} G_n(j\omega) U_n(j\omega) \tag{2.13}$$

where N represents the maximum order of nonlinearity to be taken into account. Therefore, eq.(2.13) can be defined as the characterisation of the output frequency response, $Y(j\omega)$, of a nonlinear system based on the NOFRFs [144]. It is evident that this representation of $Y(j\omega)$ in nonlinear systems shares similarities with the description of linear systems in eq.(2.3). The authors of [144] presented three significant properties regarding the NOFRFs representation of $Y(j\omega)$, which can be summarised as follows:

1. The NOFRFs can effectively describe $Y_n(j\omega)$ in a manner comparable to how the linear FRF characterises the output frequency response of linear systems, as shown in eq.(2.3) in subsection 2.5.1. Thus, the characterisation provided in eq.(2.13) exhibits a similar nature.

2. The valid frequency range ω of the n^{th} order NOFRF $G_n(j\omega)$, as given by eq.(2.12), corresponds to f_{Y_n} defined in eq.(2.5), representing the possible range of output frequencies contributed by the n^{th} order nonlinearity.

Lang and Billings [144] evidently described the energy transfer phenomena and the generation of new frequency components using the concept of NOFRFs. In brief, this energy transfer mechanism is summarised as follows;

- The nonlinear composition of $U_n(j\omega)$ from $U(j\omega)$ produces the possible frequency

components f_{Y_n} of $Y_n(j\omega)$.

- Within the frequency range f_{Y_n}, the NOFRF $G_n(j\omega)$ acts upon $U_n(j\omega)$ to produce the n^{th} order output frequency response $Y_n(j\omega)$ of the system.

- By aggregating the effects of the output frequency responses of all nonlinearities, $Y_n(j\omega)$ for $n = 1, \cdots, N$, the resulting output frequency response $Y(j\omega)$ is obtained, eq.(2.13).

- Therefore, $Y(j\omega)$ encompasses a greater number of frequency components compared to the corresponding input excitation $U(j\omega)$.

The dependence of the NOFRFs $G_n(j\omega)$ (for $n \geq 2$) on the frequency domain characteristics of the n^{th} order nonlinear dynamics can be observed from eq.(2.6). Specifically, it relies on the GFRF $H_n(j\omega_1, \cdots, j\omega_n)$ and the input spectrum $U(j\omega)$. This reliance of $G_n(j\omega)$ on the input spectrum reflects the fact that the behaviour of nonlinear systems in the frequency domain is generally influenced by both the system properties and the corresponding input [146, 145].

Considering the analysis of nonlinear dynamics between EEG channels, the NOFRFs take an interesting outlook, specifically eq.(2.13). This is because NOFRFs are a set of one-dimensional frequency functions which are weighted normalised average of the corresponding hyper-dimensional GFRFs [144]. Furthermore, the NOFRFs can be accurately determined easily [95] and directly show the energy transference from the frequencies in one EEG channel to another at different orders of nonlinearities. This allows an in-depth analysis of the complex CF interactions occurring in the cortical electrical activity of the brain (at an EEG sensor level). In the context of neurodegenerative diseases, this type of EEG analysis is novel. More specifically, considering a pair of EEG channels, $u(t)$ and $y(t)$ and the respective spectra, $U(j\omega)$ and $Y(j\omega)$. Through system identification, an input-output model can be learnt. Observing this nonlinear model of the EEG channel-pairs in the frequency domain, using NOFRFs, the energy transference from the input channel frequencies, $U(j\omega)$, to the different orders of nonlinearities $Y_n(j\omega)$ can be observed [144]. Thus, CF interactions between the frequencies of the channels $U(j\omega)$ and $Y(j\omega)$ can be examined at different orders of nonlinearity using $Y_n(j\omega)$. System identification and frequency response analysis techniques can be used for non-stationary systems which are time-varying [118], which is the case in EEG signals [101, 100, 103].

2.6 Chapter Summary

This chapter provides an overview of various brain connectivity techniques, FC and EFC, and addresses the challenges associated, along with the solutions proposed in this thesis, to effectively analyse complex abnormal dynamics in AD. It emphasises the dynamic and nonlinear nature of brain electrical activity and underlines the significance of employing methods capable of capturing nonlinear interactions. Furthermore, it specifies the specific nonlinear methods that will be utilized and explains why these methods are well-suited for characterising brain electrical activity in AD.

In relation to FC, the main issue is that the various (dis)similarity measures that quantify FC have different interpretations. Thus, there is a need for a generic measure of (dis)similarity in order to identify which EEG channel-pairs are important to distinguish AD from HC. Furthermore, to characterise AD in regard to the dynamic interactions or directed information flow, EFC, between cortical brain regions (EEG channels), the chapter highlights the need to use data-driven methods such as transfer entropy (model-free method) and data-driven model-based methods. Finally, the chapter overviews the dynamic nature of linear and nonlinear systems in the time and frequency domains to point out how well-suited a control systems approach (system identification and frequency response analysis) is to characterise the complex CF interactions in AD.

Chapter 3

Kernel-based Nonlinear Manifold Learning for EEG Functional Connectivity Analysis with Application to Alzheimer's Disease

3.1 Introduction

To diagnose and characterise neurological disorders, dynamical, causal and cross-frequency coupling analysis using the EEG has gained considerable attention. The selection of important EEG channels can be crucial to reduce the computational complexity in implementing these methods and improving classification accuracy. In neuroscience, measures of (dis)similarity between EEG channels are often used as FC features, and important channels are selected via feature selection. Developing a generic measure of (dis)similarity is important for FC analysis and channel selection. In this study, learning of (dis)similarity information within the EEG is achieved by using kernel-based nonlinear manifold learning. The focus is on FC changes and, thereby, EEG channel selection. Isomap and Gaussian Process Latent Variable Model (Isomap-GPLVM) are employed for this purpose. The resulting kernel (dis)similarity matrix is used as a novel measure of linear and nonlinear FC between EEG channels. The analysis of EEG from HC and patients with mild to moderate AD is presented as a case

study. Classification results were compared with other commonly used FC measures. Our analysis shows significant differences in FC between bipolar channels of the occipital region and other regions (i.e. parietal, centro-parietal, and fronto-central) between AD and HC groups. Furthermore, our results indicate that FC changes between EEG channels along the fronto-parietal region and the rest of the channels are important in diagnosing AD. Our results and its relation to functional networks are consistent with those obtained from previous studies using functional MRI (fMRI), resting-state fMRI (rsfMRI) and EEG.

3.2 Functional Connectivity Analysis and EEG Channel Selection

In-depth dynamical analysis, such as the analysis of linear and nonlinear dynamic relationships between EEG channels, causality, and cross-frequency coupling analysis, has received much interest [218, 242, 120, 102, 99, 104]. However, some of these methods can often incur high computational complexity. Consequently, in practice, to reduce the computational complexity, improve classification accuracy and gain prior knowledge on which underlying cortical regions might be important in AD, the selection of important EEG channels from high dimensional EEG data is vital [15]. Furthermore, to select channels to perform nonlinear dynamical analysis, the channel selection method should be able to account for nonlinear dependencies between channels [218, 242].

For most EEG channel selection techniques, features from the channels are first extracted, and important channels are selected via feature selection. These feature selection methods can be categorised into the following three groups [15]. a) *Filtering methods*: Independent evaluation criteria, including FC measures, are used for channel selection. Depending on the criteria, these are often only based on single or pairwise EEG channel(s). Filtering methods are good at eliminating irrelevant and redundant features. b) *Wrapper methods*: Subsets of features are generated based on a method of choice. Each subset is evaluated using a classification algorithm to select a subset of channels. These are based on greedy search algorithms aiming to find the best possible combination of features. c) *Embedded Methods*: These techniques simultaneously perform feature selection and classification. For example–LASSO-based feature selection, logistic regression and decision tree are techniques that come under embedded methods. Ranking of features can be easily done using embedded methods.

In this study, learning spatio-temporal linear and nonlinear (dis)similarities within the data is achieved using kernel-based nonlinear manifold learning (dimensionality reduction). The focus is on the differences in EEG FC between healthy and patient groups including the selection of important EEG channels. The motivation behind using a kernel-based method is, the learnt (dis)similarity information is reflected in the kernel, as a generic measure of distance [226] (pairwise comparisons between EEG channels). In this work, the kernel matrix is evaluated using Gaussian Process Latent Variable Model (GPLVM) [151]. Robust kernel Isomap [56] is used as an initialisation method, for GPLVM (Isomap-GPLVM). This enables the learning of both local similarities and global dissimilarities within the EEG data and embedding this information in the reduced-dimension manifold (latent space) [152]. Furthermore, since dimensionality reduction is used to reduce the temporal dimension, temporal structures within the data are taken into account in the latent space. Considering the above, the kernel matrix evaluated using Isomap-GPLVM can be regarded as a more objective (dis)similarity measure containing information on both linear and nonlinear spatio-temporal EEG inter-relationships. It is a generalisation of different functional connectivity measures [226] and can be a better alternative to using various (dis)similarity measures [63, 185, 43]. Based on this novel FC measure, this chapter introduces an EEG channel selection method to determine which channel inter-relationships are more important for in-depth neural dynamical analysis, such as understanding the effect of neurodegeneration on global and local brain dynamics. This work presents the analysis of EEG data from a cohort of age-matched healthy controls (HC) and patients with mild to moderate AD as a case study.

Participant-specific kernel matrices are obtained using Isomap-GPLVM. Linear SVM classification with Monte-Carlo cross-validation (SVM-MCV) is used to assess, how well the proposed FC measure can differentiate between HC and AD groups. FC analysis is presented for both eyes-open (EO) and eyes-closed (EC) conditions. Linear SVM-MCV is also used to rank the selected pairwise features. Therefore, the proposed channel selection method is a hybrid form [15] of the aforementioned categories of feature selection methods. Specifically, the proposed approach is an integration of filtering and embedded methods. The channel pairs chosen using this approach can be linked to other EEG studies in the literature considering connectivity analysis in AD. This chapter aims to introduce and demonstrate the efficacy of the method proposed by comparing it with other commonly used FC measures.

This chapter is organised as follows. Specifics about the EEG data, participants and the pre-processing steps are provided in Section 3.3. Section 3.4 discusses the manifold learning methodology via Isomap-GPLVM and the use of related kernel-based (dis)similarity matrix for the classification of EEG data, which are measured from a group of AD patients and an age-matched healthy control cohort. This section also presents the linear SVM and Monte-Carlo cross-validation procedures. Section 3.5 presents the results obtained followed by a discussion of the results in Section 3.6. Limitations of the study and possible improvements to the methodology are discussed along with the concluding remarks in Section 3.7.

3.3 Data

The research undertaken includes a total of 20 AD cases and 20 age and gender-matched healthy controls (HC) (less than 70 years of age), which are selected based on clinical and radiological diagnostic criteria as described in [37]. Task-free EEG recordings that require minimal cooperation of AD patients are used; typically, this group of patients can have difficulty engaging with or following cognitive tasks. The details of experimental design, diagnosis confirmation, data acquisition and EEG electrode configuration are provided in [37]. All AD participants were in the mild to moderate stage of the disease at the time of EEG recordings.

The Sheffield Teaching Hospital memory clinic team, focusing mainly on young-onset memory disorder, recruited all AD participants. AD participants were diagnosed between 1 month and 2 years before data collection. The diagnosis was made using a series of psychological tests, medical history, neuro-radiological examinations and neurological examinations. High-resolution structural magnetic resonance imaging (MRI) was used to eliminate other causes of dementia in all participants. The age and gender-matched HC participants were recruited, whose structural MRI scans and neuropsychological tests were normal. This study was approved by the Yorkshire and The Humber (Leeds West) Research Ethics Committee (reference number 14/YH/1070). All participants gave their informed written consent.

3.3.1 EEG Data

A modified 10–10 overlapping 10–20 international system of electrode placement method was used to acquire EEG recordings. All EEG data were recorded using an

XLTEK 128-channel headbox with Ag/AgCL electrodes placed on the scalp at a sampling frequency of 2 kHz. A common referential montage with linked earlobe reference was used. During the 30 minutes of EEG recording, participants were encouraged not to think about anything specific. All participants had their eyes-open (EO) for 2 minutes and then eyes-closed (EC) for 2 minutes, in repeat, during the 30-minute recording. The EEG data were reviewed by an experienced neurophysiologist on the XLTEK review station with time-locked video recordings (Optima Medical LTD). Subsequently, from the resting-state EEG recordings, three 12-second artefact-free epochs under EO and EC conditions were isolated.

From the referential montage, the following 23 bipolar channels are produced for the analysis: F8–F4, F7–F3, F4–C4, F3–C3, F4–FZ, FZ–CZ, F3–FZ, T4–C4, T3–C3, C4–CZ, C3–CZ, CZ–PZ, C4–P4, C3–P3, T4–T6, T3–T5, P4–PZ, P3–PZ, T6–O2, T5–O1, P4–O2, P3–O1 and O2–O1. The bipolar channels are obtained by simply subtracting the two common referenced signals involved. In summary, three 12-second epochs of EO EEGs are collected from 20 HC and 20 AD participants and used in this study.

It should be noted, that a bipolar montage is preferred in several studies [73, 253] due to evidence of inter-hemispheric disconnection in patients with AD [122]. Furthermore, Nunez *et al.* in [195] explains in detail that the EEG bipolar montages, given the bipolar electrode pairs are sufficiently close (1-3cm), can be effective in improving the spatial resolution of the EEG due to better estimation of localised electric fields along the scalp surface. Bipolar channels estimate the instantaneous electric field along the scalp surface midway between the pair of electrodes [195]. To avoid confusion, from now on any bipolar channel(s) will be referred to as *EEG channel(s)*, or just *channel(s)* unless otherwise specified.

3.3.2 Pre-processing Tasks

In this study, since the high-dimensional temporal structures of the multi-channel EEG are examined, the use of filters would pose a major issue due to the phase-related distortions induced [167]. Therefore, firstly, convert the time-series EEG data to the frequency domain using the Fast Fourier Transform (FFT). Thus, unwanted frequency components can be easily removed with minimal phase distortions. Thereafter, inverse-FFT is used to reconstruct the time-domain signals without the unwanted frequency

components. The analysis in this work is performed using EEG frequencies between 2 and 100 Hz. Frequencies below 2 Hz, are not used to avoid low-frequency artefacts due to eye-blinking and slow movements. Furthermore, frequency components around 50Hz (49.5-50.5Hz) are also removed to avoid any contamination by AC power line noise. After removing the unwanted frequency components, the reconstructed time-domain signals are then down-sampled to 200Hz.

3.4 Methods

This chapter introduces a novel measure of FC, a methodology that employs kernel-based manifold learning to identify important channel inter-relationships (channel pairs) within the EEG data for the case of AD. Manifold learning is a nonlinear dimensionality reduction technique that learns a lower-dimensional representation from high-dimensional data [152]. EEG data comprises multi-channel time-series data, which is high-dimensional spatio-temporal data. Kernel-based manifold learning can reduce the temporal dimension and learn both linear and nonlinear spatial and temporal structures within the EEG data. The kernel matrix obtained from such manifold learning methods will contain this information as a measure of (dis)similarity and will be named a kernel (dis)similarity matrix. This matrix can be used as a general measure of spatio-temporal functional connectivity.

Manifold learning techniques that maintain local similarities in the lower-dimensional space (also called latent space) entail a mapping from the data space to the latent space [150, 152]. This ensures that data points relatively close in the data space are positioned close together in the latent space. Kernel principal component analysis (KPCA), locally linear embedding (LLE), t-SNE, and Isomap are examples of such techniques. In contrast, techniques that involve a mapping from the latent space to the data space preserve global dissimilarities; that is, two points that are relatively distant in the data space are guaranteed to be distant in the latent space [150, 152]. Generative topographic mapping, density networks, and GPLVM are examples of these techniques. Among these, GPLVM is a kernel-based method that preserves global dissimilarity. The kernel in kernel-based manifold learning techniques, such as Isomap and GPLVM, captures the nonlinear structures within the data in a non-parametric fashion.

The methodology proposed in this study combines the strengths of both local and

global (dis)similarity preservation by utilising GPLVM and Isomap. Specifically, Isomap is used as an initialisation method for GPLVM; it is named–Isomap-GPLVM. Furthermore, since manifold learning is performed to reduce the temporal dimension, the method takes into account the temporal structures present within the EEG data. Consequently, the spatio-temporal local similarities and global dissimilarities within the EEG data are preserved in the latent space. The resulting kernel matrix from GPLVM provides a generic measure of (dis)similarity between EEG channels, capturing the preserved information.

3.4.1 Gaussian Process Latent Variable Model (GPLVM)

As a probabilistic nonlinear manifold learning technique, a GPLVM [151] learns the mapping of a high-dimensional observed dataset $\mathbf{Y} \in \mathbb{R}^{N\times D}$ from the corresponding low-dimensional latent positions $\mathbf{X} \in \mathbb{R}^{N\times Q}$, $Q < D$, i.e. a mapping from $\mathbf{X} \rightarrow \mathbf{Y}$, using a Gaussian process (GP) [229]. Here $\mathbf{Y} = [\mathbf{y}_1, \cdots, \mathbf{y}_N]^T$, $\mathbf{y}_i \in \mathbb{R}^D$ and $\mathbf{X} = [\mathbf{x}_1, \cdots, \mathbf{x}_N]^T$, $\mathbf{x}_i \in \mathbb{R}^Q$.

In principal component analysis (PCA), the mapping $\mathbf{X} \rightarrow \mathbf{Y}$ is governed by the dominant eigenvectors of the covariance matrix [251]. GPLVM is a probabilistic manifold learning method, which is a nonlinear generalisation of PCA [151], where the probabilistic mapping $\mathbf{X} \rightarrow \mathbf{Y}$ is governed by a kernel matrix $\mathbf{K} \in \mathbb{R}^{N\times N}$ [151]. The marginal log-likelihood of the data $\mathbf{Y}$ given the latent positions $\mathbf{X}$ [151, 150] is

$$L = -\frac{DN}{2}\ln(2\pi) - \frac{D}{2}\ln(|\mathbf{K}|) - \frac{1}{2}\operatorname{tr}\left(\mathbf{K}^{-1}\mathbf{Y}\mathbf{Y}^T\right), \tag{3.1}$$

where $\mathbf{K}(\mathbf{X}, \mathbf{X})$ is a positive semi-definite matrix. The i^{th} row and j^{th} column of $\mathbf{K}(\mathbf{X}, \mathbf{X})$ is given by $k(\mathbf{x}_i, \mathbf{x}_j)$ where $k(\cdot, \cdot)$ is the kernel/covariance function with a set of hyper-parameters θ. The use of a kernel function allows the nonlinear functional mapping from $\mathbf{X}$ to $\mathbf{Y}$ and provides a probabilistic nonlinear latent variable model [150]. In GPLVM, maximising L is done with respect to both $\mathbf{X}$ and θ, therefore, the optimal estimates for $\mathbf{X}$ and θ are obtained jointly. This is a highly complex optimisation with the possibility of multiple local minima [151]. As such, an appropriate initialisation of the latent positions, $\mathbf{X}$, is critical to guide the optimisation of GPLVM [36]. Which initialisation method to use depends on the specific application [150].

In this chapter, the objective is to learn the spatial and temporal structures within the

EEG data taking into account both local similarities and global dissimilarities. Since GPLVM only preserves global dissimilarities [152], initialising $\mathbf{X}$ with respect to local similarities within the data is appropriate [152, 150, 36]. Isomap has previously been successfully applied in spatio-temporal motion capture data to build latent spaces for controlling a robotic hand [255]. Furthermore, the use of Isomap to initialise GPLVM has been reported to have superior performance in motion capture data [36]. Therefore, Isomap is deemed an appropriate method to determine the initial latent positions. Section 3.4.2 provides further information about the specific Isomap variant used and its superiority over other methods that learn local similarities.

Covariance Function in GPLVM

In a GP, the covariance function $k(\cdot, \cdot)$ determines what type of functions can be learned [3]. Furthermore, it is the covariance function that defines the regions of similarity and dissimilarity between the input variables [211]. Therefore, in GPLVM, $k(\cdot, \cdot)$ defines the regions of similarity and dissimilarity between the latent positions $\mathbf{x}_i \in \mathbf{X}$.

In this study, the Radial Basis Function (RBF) (also called the squared exponential kernel) [211] is used as the covariance function for GPLVM. This is due to its inherent properties and the ability to clearly interpret its hyper-parameters [3]. The RBF kernel has the universal approximating property [182] and can be integrated against most functions to obtain a smooth mapping from $\mathbf{X} \rightarrow \mathbf{Y}$ [3, 211]. The RBF covariance function is given by

$$k\left(\mathbf{x}_i, \mathbf{x}_j\right) = \sigma^2 \exp\left(-\frac{\left\|\mathbf{x}_i - \mathbf{x}_j\right\|^2}{2l^2}\right), \tag{3.2}$$

where l and σ are the length-scale and the output-variance hyper-parameters, $\theta = [l, \sigma]$. Here, the length-scale l determines how quickly the similarity between $\mathbf{x}_i$ and $\mathbf{x}_j$ drops off as the distance between the latent positions increases [3, 211].

3.4.2 Isomap as an initialisation for GPLVM

Isomap [249] aims to preserve the geometry within nonlinear data by using the geodesic distances (along the surface of the high dimensional data) between the data points. It approximates the geodesic distances using weighted neighbourhood graphs to project high-dimensional data to a lower-dimensional representation, preserving shape

information [249]. This is the reason for choosing Isomap over methods such as KPCA and t-SNE as the initialisation for GPLVM. The robust kernel Isomap variant [56] is used in this study.

Robust kernel Isomap approximates the geodesic distance to project the data into the latent space (i.e. lower-dimensional representation) while preserving topological stability and providing a method for eliminating critical outliers [56]. The data points are projected, according to how close the points are, in the data space (i.e. preserving local similarities). In the analysis of EEG data, robustness to noise is vital as this could affect the local similarities and the geodesic distance calculations. This is the main reason for utilising robust kernel Isomap, instead of the competing LLE method and its variants.

3.4.3 Kernel-Based nonlinear Manifold Learning of High-dimensional EEG Data Using Isomap-GPLVM

Isomap-GPLVM is applied individually to the EEG data of each AD and HC participant by reducing the temporal dimension. Following the definition of the data space $\mathbf{Y} \in \mathbb{R}^{N \times D}$ (in Section 3.4.1), here $N = 23$ (23 EEG channels, Section 3.3.1 and D is the temporal dimension to be reduced. The associated latent space of each AD and HC participant will be $\mathbf{X} \in \mathbb{R}^{N \times Q}$. The resulting kernel matrix, $\mathbf{K}(\mathbf{X}, \mathbf{X}) \in \mathbb{R}^{23 \times 23}$ of each participant, quantifies the spatio-temporal (dis)similarity information between the 23 respective EEG channels as a generic measure of similarity. Fig. 3.1A illustrates this.

Robust Kernel Isomap is a technique that approximates the geodesic distance between data points to project them onto a lower-dimensional representation (latent space), while preserving local similarities (how close data points are) [56, 152]. The size of the resulting lower-dimensional representation, denoted as Q, is determined by the user. In the proposed approach, first, apply Robust Kernel Isomap to the high-dimensional data to obtain an initial estimate of the lower-dimensional representation. Then use GPLVM to refine the lower-dimensional representation based on the global dissimilarities between data points (how far apart data points are). This will result in the final latent-space $\mathbf{X}$. The kernel matrix, $\mathbf{K}(\mathbf{X}, \mathbf{X})$, from GPLVM, governs the mapping between the latent space $\mathbf{X}$ and the high-dimensional data space $\mathbf{Y}$ (see Section 3.4.1). This kernel matrix $\mathbf{K}$ will reflect both local similarities and global dissimilarities

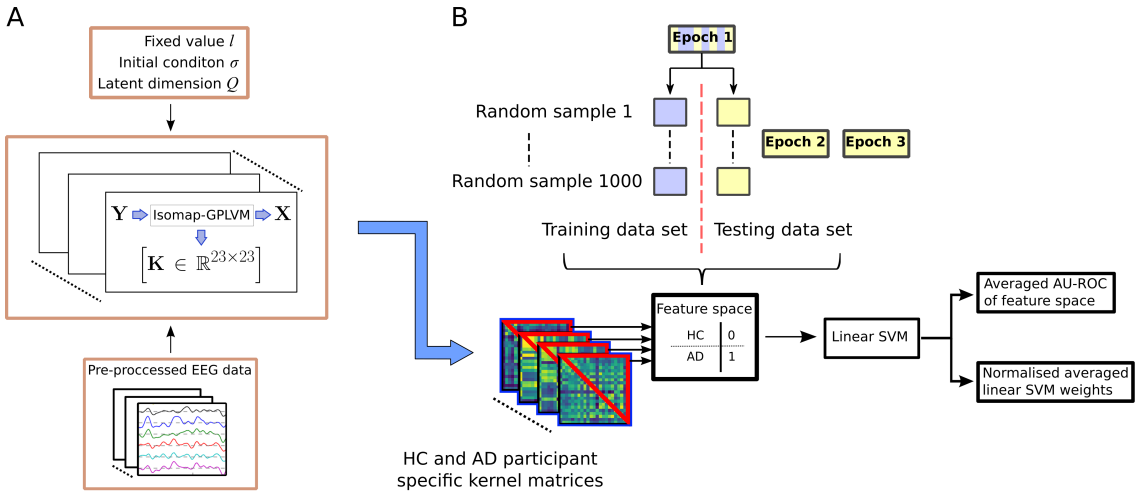


Figure 3.1: ***The Isomap-GPLVM method for evaluating the kernel (dis)similarity matrices.*** ***A) Isomap-GPLVM****: EEG data of each participant is first pre-processed via FFT filtering, to remove unwanted frequency components and normalise the data (zero mean and unit variance). Then participant specific kernel (dis)similarity matrices are evaluated using Isomap-GPLVM. From the EEG data,* $\mathbf{Y}$*, Isomap-GPLVM learns the spatio-temporal local similarities and global dissimilarities within the data (see Section 3.4.3). This information is embedded in the latent space* $\mathbf{X}$ *and is reflected in the kernel matrix* $\mathbf{K}$ *(see Section 3.4.1). The best set of values for* l*,* σ *and* Q*, from a grid search, are chosen based on how well the kernel (dis)similarity matrices are distinguishable from HC to AD.* ***B) Linear SVM-MCV*** *is used to assess this. The set of values for* l*,* σ *and* Q *that produce the best average AU-ROC from the testing set is chosen. All channel pairwise similarities from the kernel (dis)similarity matrices are used as features. This study uses EEG data from 20 HC and 20 AD participants. Three epochs of EEG data from each participant are available (see Section 3.3.1). From Epoch 1, 10 HC and 10 AD participants are chosen randomly for the training set, and the remaining 10 HC and 10 AD are used for the testing set. Epochs 2 and 3 are used for the testing set. The feature space has two classes, AD and HC. The classification is binary – AD is denoted as 1 and HC as 0. A linear SVM classifier is used on the feature space to determine which channel pairs (inter-relationships) are better at distinguishing between groups.*

that are learnt from the original data and embedded in $\mathbf{X}$, and refer to it as the kernel (dis)similarity matrix. The use of the RBF covariance function allows to quantify this information in $\mathbf{K}$ as a generic measure of similarity. By reducing the temporal dimension of the data, temporal information is naturally incorporated into the kernel (dis)similarity matrix.

Given the choice of the latent dimension, Q, the best set of values, from a grid search, for $\theta = [l, \sigma]$ is chosen based on how well the kernel (dis)similarity matrices are distinguishable from HC to AD. Linear SVM with Monte-Carlo cross validation (SVM-MCV)

is used to assess this (see Section 3.4.4 and Fig. 3.1B). The initial conditions that result in the highest average area under the receiver operator curve (AU-ROC) are chosen. The grid search is done using the search ranges $l = [2, 100]$ and $\sigma = [2, 30]$ for several latent dimensions, $Q = [5, 8]$.

It was found in this study that fixing the length-scale l of the RBF covariance function $k(\cdot, \cdot)$ in eq. (3.2), in the maximising of L in eq. (3.1), produces consistently better average AU-ROC results from applying SVM-MCV (Fig. 3.1B and Section 3.4.4) across latent dimensions $Q = [5, 8]$. Similar behaviour has been reported in [212] when using GPLVM with a back-constrained [152] likelihood to preserve local similarities. The appropriate fixed value of l and the initial condition of σ in $k(\cdot, \cdot)$, in maximising L, is found using a grid search method. Therefore, in other words, the fixed value for the length-scale l, as mentioned above, leads the manifold learning method to produce a kernel matrix where its similarity measure is optimised for the differentiation of HC and mild to moderate AD EEG data. The complete Isomap-GPLVM methodology is illustrated in Fig. 3.1.

It should be noted that the pre-processed 23-channel EEG data of each participant contains 2400 time samples (see Section 3.3.1), $\mathbf{Y} \in \mathbb{R}^{23\times 2400}$. This can be nearly perfectly represented (recovered from $\mathbf{X} \rightarrow \mathbf{Y}$ with a 95% confidence) in a latent space $\mathbf{X} \in \mathbb{R}^{23\times Q}$, with $Q \geq 5$ using Isomap-GPLVM. To achieve the same recovery accuracy, the linear principal component analysis, requires a latent dimension of $Q = 20$.

3.4.4 Linear SVM and Monte-Carlo cross-validation (SVM-MCV) procedure

This study comprises of 20 HC and 20 AD participants (see Section 3.3.1). From each participant, three 12-second epochs of EO EEGs are used. The kernel (dis)similarity matrices of the EEG data are produced for each AD and HC participant using Isomap-GPLVM, for all three epochs. The pairwise (dis)similarity measures are used as features to assess how well it can distinguish between the HC and AD groups.

Due to the relatively small number of participants, linear SVM is preferred, as it has been shown to be effective with small datasets [184, 166, 281]. Furthermore, it provides a globally optimum solution and the number of features does not affect the classification complexity [184, 166, 127]. The use of a Monte-Carlo cross-validation strategy is preferred because of its better performance with smaller data samples [234]

and the asymptotically consistent property for linear (classification) models [236]. Additionally, some AD participants could easily be detected, while others might not. This depends on the severity of neurodegeneration of the participants with mild to moderate AD used in this study. Since such information is not explicitly available, a randomised cross-validation strategy is used to obtain a fair balance in the linear SVM weights [236, 276, 277]. Furthermore, this also implies that including features from more than one epoch of each participant in the training dataset could increase the risk of participant-specific biases in the classifier. In this chapter, the aim is to find the group differences in FC between HC and AD and thereby perform channel selection. Therefore, the use of only one epoch in the training dataset and the rest in the testing set, ensures the generalising capability of the linear SVM classifier and the weightings skewed towards the most significant predictors.

Monte Carlo cross-validation is used where, from the *first epoch*, 10 HC and 10 AD participants are randomly picked for the training set. The remaining 10 HC and 10 AD participants from the *first epoch* are used for testing. 1000 such random samples are taken to generate 1000 different training and testing sets. As discussed previously, the 2^{nd} and 3^{rd} epochs of all participants will also be included in the testing dataset (Fig. 3.1B). AU-ROC from the 1000 *testing sets* is used as a metric to determine the performance of the linear SVM classification. This procedure of linear SVM-MCV is illustrated in Fig. 3.1B. The AU-ROC is preferred when considering the cost of misclassification, especially in medical diagnosis, as it helps to minimise the likelihood of misdiagnosis [288, 111, 50].

3.4.5 Kernel (dis)similarity matrix analysis

After the initial condition that produces the highest average AU-ROC is determined, SVM-MCV is used to analyse the associated kernel (dis)similarity matrices and rank the pairwise channel FC changes between HC and AD. The ranking is done using the absolute values of the normalised average of the linear SVM weights (normalised average linear SVM weights) resulting from the 1000 training sets (Section 3.4.4). The averaged weights are normalised so that the highest absolute weight is 1 (Fig. 3.1B). Due to the linearity in the classification method used, according to the superposition principle, the averaging of the linear SVM weights can be easily interpreted. Two approaches are used when implementing SVM-MCV:

(A) **Global EEG FC analysis**. All pairwise (dis)similarities are used, as in Fig. 3.1B. This identifies the best channel pairwise comparisons that can distinguish between HC and AD considering the global EEG interactions.

(B) **Channel-specific EEG FC analysis**. Each row of all kernel matrices forms a channel-specific feature space, as shown in Fig. 3.2. SVM-MCV is applied to each feature space individually. This ranks channel pairs considering a specific channel and its connectivity with the rest of the EEG to identify any significant region-specific FC changes between the HC and AD groups [192, 18, 19].

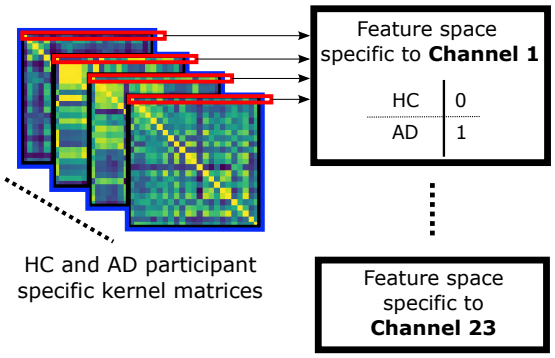


Figure 3.2: ***SVM-MCV: Channel-specific EEG FC analysis***. *All corresponding rows from HC and AD kernel matrices are grouped into channel-specific feature spaces. Each feature space has two classes, i.e. AD (1) and HC (0). Individual linear SVM classifiers are used on each feature space to determine which EEG channels, considering only its connectivity with the rest of the EEG, are better at distinguishing between the groups. The same SVM-MCV approach (as described in Section 3.4.4 and Fig. 3.1B), is now applied to each individual channel-specific feature space.*

3.4.6 Software packages used

The Isomap-GPLVM methodology is implemented in Python. GPLVM is applied using the package 'GPflow' [175]. Robust kernel Isomap and the various distance measures (Euclidean, Bray-Curtis, Correlation) are implemented using, the 'Scikit-learn' package [203]. The 'Dyconnmap' package [170], is used for the FC measures PLV, iPLV, PLI and iCoherence. Multiple hypothesis testing (in Results Section 3.5) is done using the Mann–Whitney U test, using the 'SciPy' package [268], and the Benjamini-Hochberg [29] false discovery rate controlling method, using the 'MultiPy' package [209].

3.5 Results

The Isomap-GPLVM method introduced in Section 3.4 is applied to the EO and EC EEG data (Section 3.3.1) separately. The best fixed value l and the initial condition σ for $k(\cdot,\cdot)$ to maximise L, is determined via a grid search using the three 12-second epochs according to the procedure explained in Section 3.4.3, using SVM-MCV (see Section 3.4.4 and Fig. 3.1B).

Table 3.1 illustrates, the selected Q and l values for EO and EC conditions. Given the choices for l, σ and Q, from the participant-specific kernel (dis)similarity matrices evaluated (Fig. 3.1 A), the channel inter-relationships (FC) that are able to differentiate well, between AD and HC groups are presented in this section, for both EO and EC conditions. The FC analysis is done in two approaches: global EEG FC changes and channel-specific EEG FC changes (see Section 3.4.5).

Table 3.1: *Selected latent dimension and fixed length-scale values for EO and EC conditions*

condition	Latent dimension Q	Fixed length-scale l	Average AU-ROC
EO	8	66.5	0.73
EC	8	83.5	0.77

Fig. 3.3 illustrates the bipolar montage EEG channels used in this work (see Section 3.3) on a 10–20 international standard electrode placement map. The mid-points between the 10–20 EEG overlap with certain 10–10 EEG electrode positions [130]. Therefore, the EEG channels used in this work (Fig. 3.3) measure the scalp electrical activity at those overlapping positions (see Section 3.3). The corresponding underlying cortical regions of these positions [219] are used as location markers. Table 3.2 shows the 23 bipolar montage EEG channels used and the respective underlying cortical regions.

It should be noted that the EEG has a low spatial resolution. EEG bipolar channels measure the propagated electrical activity on the overlying scalp regions (Table 3.2). Therefore, in this study when results are presented with respect to the cortical region it does not refer to the explicit activity in the actual brain cerebral cortex.

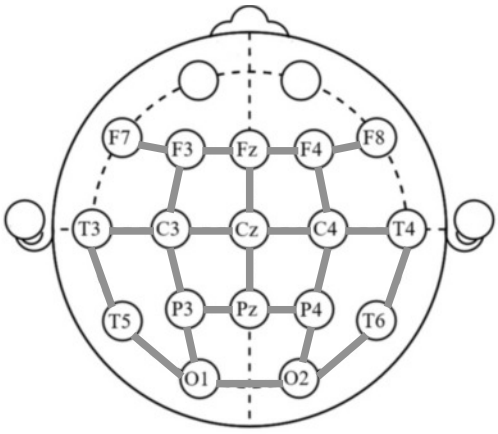


Figure 3.3: ***All the 23 channels, bipolar montage****. EEG bipolar montage channels mapped into a 10–20 international standard arrangement. The bold grey lines connecting any two EEG electrodes indicate that these two electrodes result in a bipolar channel. Bipolar channels give an estimate of the instantaneous electric field along the scalp surface midway between the pair of electrodes.*

Table 3.2: *List of all 23 channels of the scalp EEG bipolar montage and the corresponding underlying cortical regions*

Channel index and name		Corresponding cortical region
0	O1-O2	Occipital (O)
1	P4-O2	
2	P3-O1	
3	T5-O1	
4	T6-O2	
5	P3-PZ	Parietal (P)
6	P4-PZ	
7	T3-T5	Temporal (T)
8	T4-T6	
9	C3-P3	Centro-Parietal (CP)
10	C4-P4	
11	CZ-PZ	
12	C3-CZ	
13	C4-CZ	
14	T3-C3	Centro-Temporal (CT)
15	T4-C4	
16	F3-C3	Fronto-Central (FC)
17	FZ-CZ	
18	F3-FZ	
19	F4-FZ	
20	F4-C4	
21	F7-F3	Frontal (F)
22	F8-F4	

3.5.1 Comparison of kernel (dis)similarity against commonly used functional connectivity measures

To demonstrate the efficacy of the proposed FC measure, comparisons with commonly used FC measures are presented here. Table 3.3 shows the comparison of the AU-ROC

values for the SVM-MCV global EEG FC analysis for both EO and EC conditions. It is evident that, when considering all the pairwise kernel (dis)similarity measures (global EEG FC analysis), the proposed Isomap-GPLVM based FC measure has a considerably higher AU-ROC than other measures. This is especially true for the EC condition.

Table 3.3: *Comparison of SVM-MCV global analysis with AU-ROC values from the proposed FC measure against commonly used FC measures, under EO and EC conditions.*

FC measure	AU-ROC EC condition	AU-ROC EO condition
Isomap-GPLVM	**0.77 ± 0.07**	**0.73 ± 0.03**
Euclidean	0.61 ± 0.04	0.62 ± 0.04
Bray-Curtis	0.60 ± 0.02	0.61 ± 0.03
Correlation	0.68 ± 0.05	0.67 ± 0.04
PLV	**0.74 ± 0.04**	**0.70 ± 0.04**
iPLV	0.59 ± 0.05	0.64 ± 0.05
PLI	0.58 ± 0.06	0.63 ± 0.05
iCoherence	0.62 ± 0.04	0.58 ± 0.04

Fig. 3.4 illustrates the performance of the Isomap-GPLVM based FC measure with respect to the channel-specific EEG FC analysis (see Section 3.4.5). The distribution of the averaged AU-ROCs from all channel-specific feature spaces (Fig. 3.2) of the respective FC measures, is shown as a box-plot in Fig. 3.4. The channel-specific EEG FC analysis is used to identify the important FC changes, between HC and AD, with respect to a specific cortical region. In both the EO and the EC conditions, the average AU-ROC of each channel-specific feature space is directly compared, between the proposed method and other FC measures. In both conditions, it was observed that half of the channel-specific feature spaces (Fig. 3.2) from Isomap-GPLVM attained higher AU-ROC values, than the corresponding feature spaces in other FC measures. From the remaining half, most feature spaces matched the performance of corresponding feature spaces in other FC measures, while the rest underperformed. The data for all the average AU-ROCs of all channel-specific feature spaces for all FC measures used, in this comparison is not provided here. However, what is mentioned above is reflected in the box-plots in Fig. 3.4. Therefore, in general, with respect to the channel-specific EEG FC analysis, the proposed method improves the overall result under both conditions. In the EO condition, Isomap-GPLVM performs significantly better compared to other FC measures.

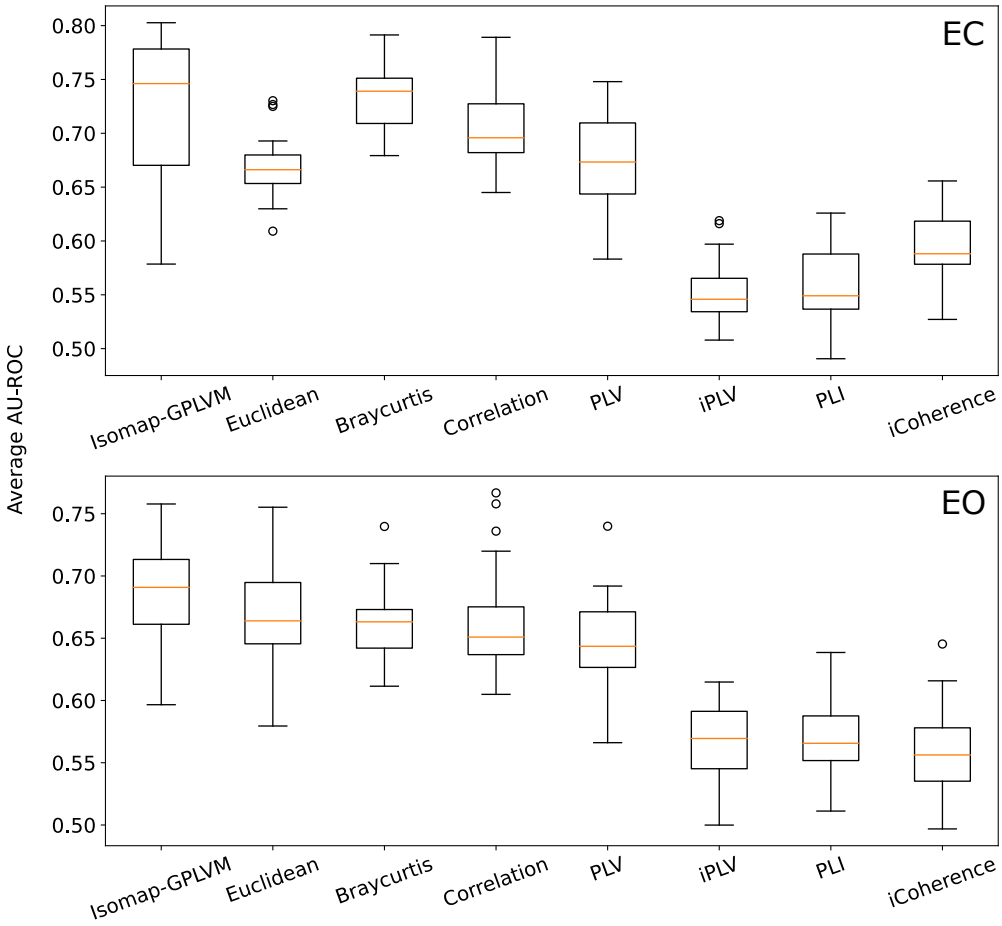


Figure 3.4: ***Comparison of the proposed Isomap-GPLVM FC measure against commonly used measures using SVM-MCV channel-specific approach.*** *The distribution of the average AU-ROC across all channel-specific feature spaces in each FC measure is shown for both EO and EC conditions.*

3.5.2 Kernel (dis)similarity matrices of HC and AD groups

The Mann–Whitney U test [169] is used for the element-wise statistical comparison of the kernel matrices between the HC and AD groups. Due to the multiple statistical comparisons done here, the p-values need to be approximately corrected [63, 209]. Also, due to the large number of comparisons (i.e. 23 EEG channels correspond to 253 channel combinations), controlling the false discovery rate (i.e. positive results that could be in fact negative) [29]) is preferred over controlling the family-wise error rate [63, 209]. Therefore, the Benjamini-Hochberg [29] false discovery rate controlling (FDR) method is used to obtain the corrected p-values. Pairwise kernel (dis)similarities (FC measures) that have statistically significant differences between the HC and AD

groups (p-values < 0.05) are denoted as 1's in significance matrix $\mathbf{S} \in \mathbb{R}^{23\times23}$, zero otherwise. The significance matrices for both the EC and EO conditions are illustrated in Fig. 3.5, where blue elements indicate the statistically significant changes in the pairwise connectivities between the HC and AD groups.

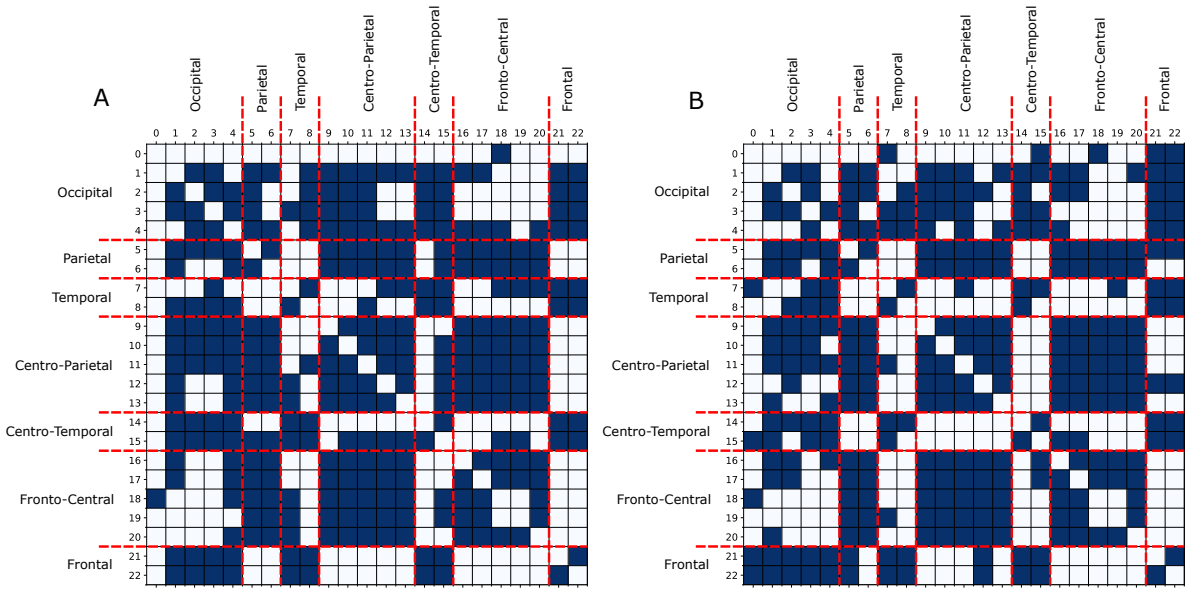


Figure 3.5: ***The significance matrices,* S*, for both EC (A) and EO (B) conditions.*** *The figures show, for both EC and EO conditions, the statistical significance of all the elements of the kernel (dis)similarity matrices between the HC and AD groups (based on all epochs of all participants). The corresponding channels are provided in Table 3.2. The significant FC changes are indicated in blue.*

From Fig. 3.5, it is evident that there are localised FC changes within certain underlying cortical regions (e.g. within the centro-parietal EEG region) and global EEG FC changes between regions (e.g. between the centro-parietal and occipital EEG regions). This can be a reflection of the specific patterns of dysfunction that have been mentioned in the literature, in which AD EEG data exhibits a specific change in FC compared to HC [242, 63, 64] and connectivity in certain regions of the EEG being affected [73, 66, 256, 6, 80, 14]. These FC changes could be linked to within-frequency and cross-frequency coupling between brain regions [20, 222].

3.5.3 Global functional connectivity changes and channel pair selections

All pairwise kernel (dis)similarities are used (Fig. 3.1B) to determine, in a global sense, which spatio-temporal FC differences between cortical regions (EEG channels) are more important in distinguishing between the HC and AD groups. Table 3.4 and 3.5 shows the top 20 channel pairs that are ranked according to the averaged normalise linear SVM weights (Section 3.4.5, Fig. 3.1B) for EO and EC conditions, respectively. EEG channel pairs in these two tables are arranged in a way so that channels related to the same underlying cortical regions can be grouped.

Table 3.4: *EO condition. Ranking of (dis)similarity features of channel-pairs–only the top 20 are shown*

Channel Pairs (Indexes and names)				Averaged normalise linear SVM weight	Ranking	Connecting regions
11	1	CZ-PZ	P4-O2	0.84	6	
13	1	C4-CZ	P4-O2	0.78	7	
9	1	C3-P3	P4-O2	0.68	12	CP - O
12	3	C3-CZ	T5-O1	0.68	13	
13	4	C4-CZ	T6-O2	0.60	19	
21	3	F7-F3	T5-O1	0.88	3	
22	3	F8-F4	T5-O1	0.72	8	F - O
22	4	F8-F4	T6-O2	0.59	20	
18	9	F3-FZ	C3-P3	0.68	11	FC - CP
18	17	F3-FZ	FZ-CZ	0.68	14	FC - FC
16	1	F3-C3	P4-O2	1.00	1	
17	1	FZ-CZ	P4-O2	0.88	4	
16	4	F3-C3	T6-O2	0.71	10	FC - O
16	2	F3-C3	P3-O1	0.64	16	
20	1	F4-C4	P4-O2	0.63	17	
5	3	P3-PZ	T5-O1	0.91	2	
5	2	P3-PZ	P3-O1	0.86	5	
5	1	P3-PZ	P4-O2	0.72	9	P - O
6	1	P4-PZ	P4-O2	0.66	15	
6	3	P4-PZ	T5-O1	0.62	18	

Tables 3.4 and 3.5 show that inter-regional FC between EEG channels from the occipital region and other regions, i.e. parietal (P–O, Fig. 3.6C), centro-parietal (CP–O, Fig. 3.6A), and fronto-central (FC–O, Fig. 3.6B), attain a considerable space among the top 20 rankings. This is observed in both EO and EC conditions. However, in the EC condition this is specific to the right occipital region (channels P4-O2 and T6-O2, Table 3.5). In particular to the EO condition, as shown in Table 3.4, FC between EEG channels from the frontal and occipital regions (F–O, Fig. 3.6D) have a significant presence within the top 20 weightings. Therefore, this suggests that connectivity between the occipital region and those regions mentioned above, respective

to each condition, can be important in identifying people with mild to moderate AD as shown in Fig 3.6.

Table 3.5: *EC condition. Ranking of (dis)similarity features of channel-pairs–only the top 20 are shown*

Channel Pairs (Indexes and names)				Averaged normalise linear SVM weight	Ranking	Connecting regions
9	4	C3-P3	T6-O2	0.99	3	CP - O
11	4	CZ-PZ	T6-O2	0.96	4	
13	4	C4-CZ	T6-O2	0.90	5	
10	4	C4-P4	T6-O2	0.90	7	
11	1	CZ-PZ	P4-O2	0.80	9	
12	4	C3-CZ	T6-O2	0.77	11	
9	1	C3-P3	P4-O2	0.72	14	
22	15	F8-F4	T4-C4	0.68	17	F - CT
21	4	F7-F3	T6-O2	0.62	19	F - O
17	4	FZ-CZ	T6-O2	1.00	1	FC - O
16	4	F3-C3	T6-O2	1.00	2	
16	1	F3-C3	P4-O2	0.80	10	
17	1	FZ-CZ	P4-O2	0.75	12	
18	4	F3-FZ	T6-O2	0.70	15	
20	4	F4-C4	T6-O2	0.62	20	
1	0	P4-O2	O1-O2	0.67	18	O - O
6	4	P4-PZ	T6-O2	0.90	6	P - O
5	4	P3-PZ	T6-O2	0.86	8	
5	1	P3-PZ	P4-O2	0.74	13	
6	1	P4-PZ	P4-O2	0.70	16	

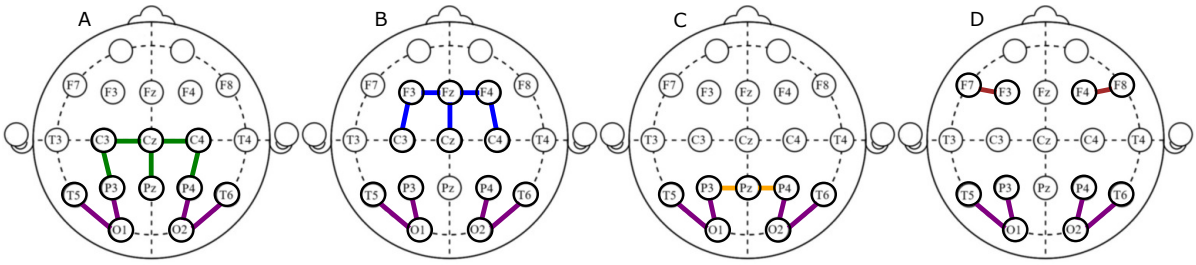


Figure 3.6: ***Inter-regional connectivity between EEG channels from the regions shown can be important in identifying people with mild to moderate AD. A) CP–O, B) FC–O, C) P–O*** *for both EO and EC conditions while* ***D) F–O*** *only for the EO condition (Tables 3.4 and 3.5). EEG channels related to the corresponding cortical regions are shown in different colours. Occipital region (O)–purple, Parietal region (P)–orange, Centro-parietal region (CP)–green, Fronto-central region (FC)–blue and Frontal region (F)–red. It shoud be noted that bipolar channels give an estimate of the instantaneous electric field along the scalp surface midway between the pair of electrodes*

3.5.4 Channel-specific functional connectivity changes and channel pair selections

In this section, the results of the channel-specific EEG FC analysis (Section 3.4.5, Fig. 3.2), using kernel (dis)similarity matrices, are presented. This approach determines, at the EEG sensor level, significant changes in FC specific to the cortical region between the HC and AD groups. The channel-specific approach provides another layer of information.

To form a channel-specific feature space, each row of all kernel matrices that correspond to a particular channel is used (Fig. 3.2). SVM-MCV is then applied to each feature space individually to identify individual EEG channels that exhibit distinguishable changes in FC with the rest of the EEG data. The average AU-ROC of the channel-specific feature space is used as an evaluation metric. The normalised average linear SVM weights (Section 3.4.4) of the channel-specific feature space are used to rank the importance of FC changes relative to the channel being considered (Section 3.4.5).

Table 3.6: *The average AU-ROC values of the channel-specific feature spaces for the EO condition. Channel-specific feature spaces with average AU-ROC* > 0.7 *are only shown*

Channel-specific feature space	Channels with normalised average weight 0.9-1	Average AU-ROC of feature space
F3-FZ	P4-O2	0.758
FZ-CZ	P4-O2, F3-FZ	0.737
F3-C3	P4-O2	0.734
C3-P3	P4-O2, T5-O1	0.733
CZ-PZ	P4-O2, F3-FZ	0.719
F4-C4	P4-O2, F3-FZ	0.714
P3-PZ	T5-O1	0.713
C4-P4	P4-O2	0.712
T5-O1	T6-O2, F7-F3	0.710
C3-CZ	P4-O2, T5-O1	0.706

Tables 3.6 and 3.7 report channel-specific feature spaces with an average AU-ROC > 0.7 for the EO and EC conditions, respectively. These tables also report the kernel (dis)similarity features that attain a high rank in the feature space being considered (i.e. normalised average linear SVM weight 0.9-1). Fig. 3.7 illustrates the channel-specific feature spaces with an average AU-ROC > 0.7 mapped to the placement of 10–20 international electrodes.

As seen from Fig. 3.7, the EEG channels associated with the channel-specific feature

Table 3.7: *The average AU-ROC values of the channel-specific feature spaces for the EC condition. Channel-specific feature spaces with average AU-ROC* > 0.7 *are only shown*

Channel-specific feature space	Channels with normalised average weight 0.9-1	Average AU-ROC of feature space
C3-P3	T6-O2	0.803
CZ-PZ	T6-O2	0.799
T6-O2	T5-O1, T3-T5, F7-F3	0.797
C4-P4	T6-O2	0.792
F3-FZ	T6-O2	0.790
P3-PZ	T6-O2	0.785
FZ-CZ	T6-O2	0.772
C3-CZ	T6-O2	0.769
F3-C3	T6-O2	0.764
P4-O2	F3-C3, F4-FZ, F8-F4	0.758
F4-C4	T6-O2	0.757
C4-CZ	T6-O2	0.746
T4-C4	F8-F4	0.721

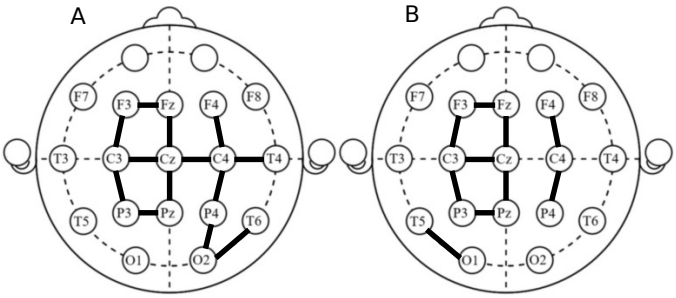


Figure 3.7: ***EC (A) and EO (B) channel-specific feature spaces with average AU-ROC*** > 0.7. *The channel-specific feature spaces with average AU-ROC* > 0.7 *are mapped into the 10–20 international electrode placement. This is illustrated for both EC and EO conditions. These channels lie mostly within the fronto-parietal regions of the cortex for both conditions.*

spaces with an average AU-ROC > 0.7 lie mostly within the fronto-parietal regions of the cortex for both conditions. However, in the EC condition (Fig. 3.7A), channel-specific feature spaces associated with the EEG channels in the right hemisphere appear to be important (average AU-ROC > 0.7). In the case of mild to moderate AD, significant FC changes between these EEG channels and the rest of the EEG is observed from the proposed FC analysis methodology.

3.6 Discussion

The results presented in the previous section indicate that certain key areas of the brain are affected by AD (Fig. 3.5, 3.6 and 3.7). In order to identify whether Isomap-GPLVM FC results are consistent with fMRI results of mild to moderate AD, first, it is necessary to determine how the bipolar channels used in this study (Section 3.3.1) relate to the functional connectivity networks [278] of the brain.

Yeo *et. al.* [278] revealed the existence of seven primary functional networks using time correlations between the fMRI of 1,075 Regions of Interest (ROI). These networks are shown to be valid for multiple participants and robust against various data processing methods. The seven functional networks are visual network (VN), somatomotor network (SN), dorsal attention network (DAN), ventral attention network (VEN), limbic network (LN), fronto-parietal network (FPN) and default mode network (DMN). Based on [278], Rojas *et. al.* [219] used the electrode positions of the international standard 10–20 EEG and the 10–10 EEG as seed positions to provide a reproducible model demonstrating the relationship between the 10–20 EEG electrode positions and the seven functional networks revealed by [278]. This is carried out by simultaneous acquisition of EEG and resting-state fMRI (rs-fMRI). Rojas *et. al.* used the Sørensen–Dice index (F1 score) to quantify the similarities between the positions of the 10–20 electrode placements and the seven functional networks mentioned above. The bipolar EEG channels used in this study estimate the electric field midway between the pair of electrodes that form the said channel [195]. Therefore, to determine approximate similarities between a bipolar channel (Fig. 3.3) and the functional networks, the average of the Sørensen–Dice indices of the two electrodes (Figure 9 and Supplementary Table 3 in [219]) that form the bipolar channel is used. An example of this is shown in the appendix (A).

Fig. 3.8 illustrates the relationship between the bipolar EEG channels used in this study and functional networks FPN, DAN, VAN and DMN. In mild to moderate AD, the connectivity changes within networks FPN, DAN, VAN and DMN have been previously reported to be significant [20, 192, 282, 88]. Therefore, the following discussion will only focus on these networks, as shown in Fig. 3.8.

For the EO condition, changes in FC between the cortical regions (EEG channels) shown in Fig. 3.6 B and D can be speculated as VAN related [282]. Considering both EO and EC conditions, changes in the inter-regional FC (as shown in Fig. 3.6 A, B, C

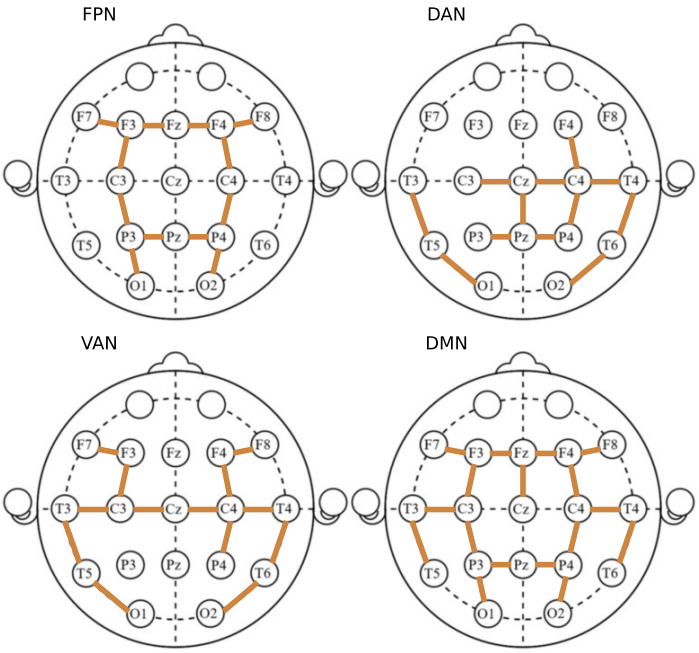


Figure 3.8: ***Relationship between the bipolar EEG channels used and the functional networks.*** *The functional networks discussed in relation to this study and the bipolar channels used is illustrated here. FPN (fronto-parietal network), DAN (dorsal attention network), VAN (ventral attention network) and DMN (default mode network).*

and D) can be linked to changes in connectivity within the FPN and DMN networks [282]. Fig. 3.6 A and C can be linked to DAN [282, 88] while Fig. 3.6 B can be linked to VAN.

The results in Section 3.5.4 show the channels that have the most significant FC changes with the rest of the EEG (Fig. 3.7). This can be a reflection of the FC changes in FPN, DAN, VAN and DMN networks. The EEG channels shown in Fig. 3.7 are mostly related to the fronto-parietal region of the cortex (fronto-central and centro-parietal regions combined, Table 3.2). This region has been reported to play an important role in the diagnosis of AD in several studies using fMRI [192] (prodromal AD), rs-fMRI [282, 88] (mild, moderate and severe AD) and EEG [20] (mild AD). Neufang *et al.* [192] pointed out, at the early stages of AD, the volume of regional grey matter is related to the reduction in the effective connectivity (through dynamic causal modelling) in the fronto-parietal region. While Babiloni *et al.* [19] found that a measure of nonlinear inter-dependence (via the synchronisation likelihood) is significantly reduced in the fronto-parietal channels of eyes-closed EEG in mild AD patients. These studies are consistent with the Isomap-GPLVM FC results in showing that the

connectivity between the EEG channels in the fronto-parietal region (Fig. 3.7) and the rest of the brain regions have significantly changed in mild to moderate AD.

3.7 Chapter Summary

A novel FC analysis and channel selection method based on kernel-based nonlinear manifold learning is presented in this work. The FC measure takes both local and global spatio-temporal (dis)similarities between EEG channels into account and ranks the pairwise FC measures that are better at distinguishing HC from patients with neurodegenerative diseases. It was demonstrated how a kernel-based (dis)similarity matrix via manifold learning can be used as a measure of spatio-temporal FC between EEG channels and to determine the important inter-relationships in characterising patients with mild to moderate AD. The methodology presented can determine changes in cortical (EEG channel) inter-relationships that are crucial in distinguishing AD patients from HCs. The chapter also demonstrates its efficacy against other commonly used FC measures. Furthermore, the results reported in this chapter are consistent with other previous studies while linking connectivity changes to functional networks.

The findings from this chapter will be used in the following chapters to investigate the detailed forms of nonlinearity using nonlinear dynamic modelling [32] and nonlinear causality measures in the time and frequency domains [286, 102, 103]. These in-depth dynamical analysis methods will be applied to the channel pairs and regions determined using the Isomap-GPLVM method. Thus, enabling the further study of the underlining dynamic processes, linear and nonlinear dynamic features, in patients with AD.

Chapter 4

Information Flow at an EEG Sensor Level Between Specific Brain Regions that are Significant to Alzheimer's Disease

4.1 Introduction

The temporal interactions between brain regions even at rest show a significant level of complexity [65, 87]. Therefore, evaluating brain connectivity during the resting-state has become a vital area of research in neuroscience as it provides insight into the functional organisation of the brain, helping to identify distinct resting-state brain networks [278] and their interactions. This type of investigation is popular not only for comprehending the mechanisms underlying the typical resting-state but also for detecting abnormalities in pathological conditions such as Alzheimer's, schizophrenia, and depression [178, 97, 98, 70]. While fMRI is commonly used to estimate connectivity, there is a growing interest in utilising electrophysiological data obtained from EEG or MEG due to their higher temporal resolution.

In Chapter 3, the FC of the resting-state EEG was analysed to determine the important changes in the pairwise statistical dependency between HC and AD groups. This Chapter aims to use these subsets of EEG channel-pairs to conduct an exploratory

analysis to investigate the directed connectivity (EFC, see Chapter 2 Section 2.3) or how the information flow between these channels varies in time between healthy cohorts and patients with mild to moderate AD. This information will be used as prior knowledge to build input-output dynamic models to analyse the significant linear and nonlinear dynamical changes in the cortex of patients with AD.

The measurement of information transfer can be accomplished through various directed information metrics, with transfer entropy being widely recognized and well researched, especially in neuroscience to investigate the flow of information and interactions between different brain regions or neuronal populations [45, 266, 258]. The model-free nature of transfer entropy, its ability to capture dynamic dependencies [131] and its sensitivity to higher order correlations [193], makes it appealing for exploratory analysis of unknown nonlinear dynamical interactions [272].

4.2 Application of Transfer Entropy in Neuroscience

Transfer entropy, introduced by Schreiber [228] and Paluš [199], has shown to be a valuable tool in numerous application scenarios across diverse fields. These fields include neuroscience [266, 258], physiology [71, 72] and complex systems theory [160, 4]. The extensive range of these application fields indicates that transfer entropy serves as a valuable and fundamental measure for comprehending complex systems, such as those that can be described as networks of interacting processes [54, 201]. However, it should be noted that the relationship between transfer entropy and the connectivity strength between interacting processes can exhibit non-monotonic behaviour. Additionally, transfer entropy is responsive to internal changes within the sub-processes. Therefore, care must be taken when interpreting transfer entropy as a measure of connectivity strength [54]. Despite this, an important aspect of transfer entropy over other methods such as DCM is that it does not require prior assumptions about data generation, i.e., it is a model-free inference on EFC. Due to this advantage, transfer entropy is widely employed in neuroscience to evaluate directed connectivity in EEG/MEG datasets where no prior assumptions are available [61, 70, 1, 200, 260, 238, 123]. Recently variants of transfer entropy such as phase transfer entropy [165] has been introduced to comprehend information flow between two processes based on phase dynamics.

Transfer entropy has been used to assess the information flow between neurons [161, 266, 258]. Despite numerous studies that have used entropy, mutual information and complexity measures such as Lempel-Ziv complexity on scalp EEG to understand cognitive deficits [241, 283, 62, 177, 179], the use of transfer entropy on the resting-state EEG is limited. In particular, at the time of writing, the study conducted by McBride *et al.* [178] is the only application of transfer entropy for diagnosing AD based on the EEG (this is based on a keyword search on Scopus). Consequently, the work presented in this chapter, which focuses on an exploratory analysis of time-varying EFC between brain regions that show significant FC changes in mild to moderate AD, using the EEG, is novel.

4.3 Transfer Entropy: Theory and methods

In comparison to the quality of prediction of future instances of a process Y, when only the past instances of Y are considered. Causality, as defined by Wiener [274] and Granger [90] is the improvement in the prediction of Y when incorporating the past instances of another process U along with the past of Y. If there is an improvement in the prediction then it is an indication of U impacting Y or information flow from process U to process Y.

In information theory, Shannon entropy [235, 214] quantifies the reduction in uncertainty of a random discrete variable when the variable is measured. Causality on the other hand is the increase of prediction power. Associating uncertainty reduction with improvement in prediction, causality can be expressed in terms of information-theoretic concepts [228]. Transfer entropy incorporates the causal principle within the framework of information theory, utilizing conditional probabilities [228, 199, 266] to infer EFC from data. Essentially, if signal U influences signal Y, then the probability of Y given its past will differ when considering the probability of Y given its past and the past of U. Another way to understand this is that the Shannon entropy of the current instance of Y decreases when past instances of U is incorporated alongside the past instances of Y [261].

4.3.1 Definitions of Transfer Entropy

Given two stochastic processes U and Y whose current instances in time are $u(t)$ and $y(t)$ respectively, where t is a time instance. Defining $\mathbf{U}_t^{n_b}$ and $\mathbf{Y}_t^{n_a}$ as the respective

time-dependent state vectors of U and Y where

$$\mathbf{Y}_t^{n_a} = [y(t-1)\ ,\ y(t-2)\ ,\ \cdots\ ,\ y(t-n_a)]\,. \tag{4.1}$$

$$\mathbf{U}_t^{n_b} = [u(t-1)\ ,\ u(t-2)\ ,\ \cdots\ ,\ u(t-n_b)] \tag{4.2}$$

n_b and n_a are the dimensions of the delay embedding space (embedding dimensions) [248], which describes the number of past instances of time used to reconstruct the phase space of the respective process [133, Chapter 3]. The time difference between each sample of time, for example between t and $t-1$, is $\triangle t$ which is the embedding delay (sampling time). Schreiber [228] proposed a measure of causality within the information theoretic framework, in which, if the dynamics of Y are completely independent of the past instances of U or if there is no information flow from U to Y, then

$$p\left(y(t) \mid \mathbf{Y}_t^{n_a}, \mathbf{U}_t^{n_b}\right) = p\left(y(t) \mid \mathbf{Y}_t^{n_a}\right). \tag{4.3}$$

To identify any deviation from this relationship, i.e. presence of information flow from U to Y. The expected Kullback-Leibler divergence between the two probability distributions in eq. (4.3) is taken to define the transfer entropy from U (source) to Y (target) as

$$TE\left(U \to Y\right) = \sum_{y(t),\ \mathbf{Y}_t^{n_a},\ \mathbf{U}_t^{n_b}} p\left(y(t), \mathbf{Y}_t^{n_a}, \mathbf{U}_t^{n_b}\right)\ \log_2\left(\frac{p\left(y(t) \mid \mathbf{Y}_t^{n_a}, \mathbf{U}_t^{n_b}\right)}{p\left(y(t) \mid \mathbf{Y}_t^{n_a}\right)}\right). \tag{4.4}$$

Transfer entropy can also be derived using Shannon entropy [235]. Let Ω_Y be the probability space of the stochastic process Y where $p(y_\omega)$ is its distribution, $y_\omega \in \Omega_Y$ are the possible instances of Y and $\sum_{y_\omega \in \Omega_Y} p(y_\omega) = 1$. The Shannon entropy of the process Y is then given by

$$H(Y) = -\sum_{y_\omega \in \Omega_Y} p(y_\omega) \log_2 p(y_\omega) \tag{4.5}$$

where the summation over y_ω is the sum of all possible instances of Y. The conditional entropy of the process Y given the process U is

$$H(Y \mid U) = -\sum_{y_\omega \in \Omega_Y} p(y_\omega) \sum_{u_\omega \in \Omega_U} p(y_\omega \mid u_\omega) \log_2 p(y_\omega \mid u_\omega) \tag{4.6}$$

where Ω_U is the probability space of the stochastic process U where $p(u_\omega)$ is its distribution. In eq. (4.6) the summation over y_ω and u_ω is the sum of all possible instances of Y and U respectively. Similarly, the joint entropy between the processes Y and U is given as

$$H(Y,U) = -\sum_{y_\omega \in \Omega_Y} \sum_{u_\omega \in \Omega_U} p(y_\omega, u_\omega) \log_2 p(y_\omega, u_\omega). \tag{4.7}$$

The relationship between $H(Y \mid U)$ and $H(Y,U)$ is

$$H(Y,U) = H(U) + H(Y \mid U) = H(Y) + H(U \mid Y). \tag{4.8}$$

Paluš showed that transfer entropy can be defined using conditional mutual information [199] which finally reduces to

$$TE\,(U \to Y) = H(y(t) \mid \mathbf{Y}_t^{n_a}) - H(y(t) \mid \mathbf{Y}_t^{n_a}, \mathbf{U}_t^{n_b}). \tag{4.9}$$

Eq. (4.9) can be re-written in terms of joint entropy using the relationship in eq. (4.8) as

$$TE\,(U \to Y) = \Big(H(y(t), \mathbf{Y}_t^{n_a}) - H(\mathbf{Y}_t^{n_a})\Big) - \Big(H(y(t), \mathbf{Y}_t^{n_a}, \mathbf{U}_t^{n_b}) - H(\mathbf{Y}_t^{n_a}, \mathbf{U}_t^{n_b})\Big). \tag{4.10}$$

From the definition given in eq. (4.9), it is clear that transfer entropy measures the decrease in uncertainty of the current instance of Y, $y(t)$, when the information about its past instances, $\mathbf{Y}_t^{n_a}$, is included alongside the past instances of U, $\mathbf{U}_t^{n_b}$. If $TE\,(U \to Y)$ has a value other than zero, it implies that past instances of U influence the future instances of Y. Therefore, transfer entropy is an asymmetric measure. This asymmetry in transfer entropy is an important characteristic for determining the direction of information flow between Y and U. If there is no information flow from U to Y then $TE\,(U \to Y) = H(y(t) \mid \mathbf{Y}_t^{n_a})$. Therefore, following [89, 106] a normalisation of $TE\,(U \to Y)$ can be given as

$$NTE\,(U \to Y) = \frac{TE\,(U \to Y)}{H(y(t) \mid \mathbf{Y}_t^{n_a})}. \tag{4.11}$$

The process of normalisation is beneficial as it allows for the comparison of information flows without considering the level of dependency between $y(t)$ and its past instances $\mathbf{Y}_t^{n_a}$ [89]. By doing so, it helps to standardise the measurement in relation to the varying

complexities of U and Y. For the interested reader to know more about transfer entropy and the derivations shown here please refer to [228, 199, 273, 39].

4.3.2 Estimation of Transfer Entropy

The estimation of transfer entropy, as seen in the previous subsection, involves the estimation of joint/conditional entropy or mutual information, which eventually involves the estimation of probability densities. One potential approach to designing an estimator involves identifying the parameters that most accurately match the sample probability densities with a recognized distribution. Although this method is computationally simple, it assumes a particular model for the probability distribution, which can be challenging to justify without additional limitations. Non-parametric methods like fixed and adaptive histogram or partition techniques are commonly employed and well-known [168, 31, 153, 265, 17, 220]. However, alternative non-parametric approaches, such as Kernel Density Estimator (KDE) or k-nearest-neighbour estimators (KNN), have demonstrated greater efficiency and accuracy in handling data while avoiding arbitrary decisions associated with binning [228, 131, 267]. Nevertheless, these methods are computationally expensive [47, 105] and necessitate a substantial volume of neural data to converge, unless the probability distributions underlying the data are adequately smooth [267, 190]. Consequently, Bullmann *et. al.* and Heer *et. al.* [47, 105] has shown that applying a recursive filter or a Gaussian filter on multivariate histograms yields good approximations to high-dimensional Probability Density Function (PDF)s and is comparable to KDE. Furthermore, this technique is computationally efficient in several orders of magnitude when compared to KDE. Therefore, in this study, the above mentioned technique for the estimation of PDFs will be used. However, it is necessary to employ suitable surrogate techniques in order to mitigate the bias and variance introduced by the approximation technique for estimating transfer entropy [31, 265].

The nonlinear nature in the electrophysiology of brain activity has been well studied [242, 42] and the EEG does indeed exhibit nonlinear dynamics such as Cross-Frequency Coupling (CFC) effects [48, 115, 42]. This study is concerned with the analysis of cortical nonlinear dynamical changes in AD using the EEG. Therefore, the random phase surrogate method [208, 142] is employed to statistically assess the significance of the transfer entropy measure to identify nonlinear dependencies [193]. Random phase surrogate testing involves the application of the Fourier transform, randomising the

phase information and then applying the inverse Fourier transform to obtain time-series with only linear dependencies. This will destroy any higher-order nonlinear correlation [208, 193, 142] within the data. Several surrogates of the source time-series, $U_{s(i)}$, are generated where $i = 1, \cdots, M$ and M is the number of surrogates. Then the transfer entropy measures, $TE_i(U_{s(i)} \rightarrow Y)$, between these surrogate sources and the target, Y, are used to make a distribution of the null hypothesis. If the transfer entropy, $TE(U \rightarrow Y)$, from the original source, U, is outside the 5^{th} percentile and the 95^{th} percentile of the null hypothesis distribution, then $TE(U \rightarrow Y)$ is considered to be significant. Thus, the presence of significant nonlinear dependencies in the information flow from U to Y. This particular method for statistical significance is used because the null hypothesis distribution can be non-Gaussian. Therefore, if $TE(U \rightarrow Y)$ is significant after accounting for bias and possible spurious nonlinear dependencies by using the above mentioned procedure, similar to [172, 106], the effective transfer entropy is given by

$$ENTE\,(U \rightarrow Y) = \frac{TE\,(U \rightarrow Y) - \mathcal{M}\Big(TE_i(U_{s(i)} \rightarrow Y)\Big)}{H(y(t) \mid \mathbf{Y}_t^{n_a})} \tag{4.12}$$

where $\mathcal{M}()$ denotes the median.

4.4 Exploratory analysis of information flow within the EEG

The aim of this chapter is to understand how the direction of information flow within the subset of pairwise channels (See Chapter 3, Subsection 3.5.3, Tables 3.4 and 3.5) that exhibit significant FC changes between HC and AD groups. The FC (statistical dependencies) between these channels were examined using the full 12 seconds of the respective EEG epochs (Section 3.3) However, the direction of information flow between these channels can change with time [196, 107, 263]. Therefore, an exploratory analysis is conducted where transfer entropy is applied within certain overlapping time windows to identify any statistically significant differences in the direction of information flow between HC and AD groups.

The choice of window size is a critical factor that can significantly impact connectivity estimates. When shorter windows are used, the efficiency of connectivity estimates

decreases and there is a possibility of magnifying the perceived variability in connectivity over time [154, 108, 149, 159]. Park *et al.* [202] showed that there is an improvement in the consistency of connectivity estimates by employing time-varying models. This highlights the significance of capturing and modeling the dynamic changes in connectivity over time. Leonardi and Van de Ville propose setting the minimum window size to be as large as the period of the lowest frequency component in resting state data [154]. One concern when using longer time windows is, the time-series within the window can be non-stationary, which can lead to spurious non-linear dependencies and thus affects the transfer entropy estimations [275, 210]. However, non-stationary effects could arise due to the physiological phenomena of the actual biophysical system rather than spurious effects and should not be avoided [153]. Furthermore, there are certain forms of CFC that are indeed non-stationary [16, 188, 58, 225]. The issues of spurious non-linear dependencies can be resolved by the random phase surrogate method that is being used in this study (see Subsection 4.3.2), which is specifically designed to mitigate this type of issue [193]. Nevertheless, it should be noted, when it comes to the EEG, there has not yet been a systematic approach to determining the appropriate window length [263]. As proposed by Van de Steen *et al.* [263], one potential strategy is to investigate the relationship between window length and the amount of information gained, in order to make a well-founded decision.

In the context of mild to moderate AD, many studies have pointed out the importance of delta, theta, alpha and beta frequency ranges in both spectral and CFC based analysis [121, 138, 80, 269, 110, 136]. Since this study is using data from mild to moderate AD patients, the transfer entropy estimates for both HC and AD EEG data will be estimated within the frequency bands $2 - 30$Hz, which covers the above mentioned bands [21]. For computational efficiency, the data will be down-sampled to 100Hz, where the Nyquist frequency is 50Hz.

Let U and Y be a channel-pair being considered. The effective normalised transfer entropy, $ENTE_w^{n_a,n_b}(U \to Y)$, for a specific combination of delays n_a and n_b within a time window w is estimated using the methodology outlined in Subsection 4.3.2. $w = 1, \cdots, W$ in which W is the total number of windows. For example, a 3-second 50% overlapping sliding-window along 12 seconds of EEG data, $W = 6$. n_a and n_b are the respective delay embedding dimensions of Y and U, eq. (4.1) and (4.2) respectively.

As previously mentioned, a range of frequencies are affected in mild to moderate

AD. Therefore, the dynamic interactions between EEG channels can happen along a broadband of frequencies. Therefore, in order to capture the complex dynamic nonlinear dependencies between the channel-pair U and Y, specific to the direction $U \rightarrow Y$, the respective delay embedding dimensions n_a and n_b need to be estimated appropriately. This is done using a grid search where the effective transfer entropy $ENTE_w^{n_a,n_b}(U \rightarrow Y)$ is evaluated for several combinations of n_a and n_b. The combination n'_a and n'_b that results in the maximum effective normalised transfer entropy, $ENTE_w^{n'_a,n'_b}$ is chosen. The grid search is done for the ranges $n_a = [1, 10]$ and $n_b = [1, 10]$. Effective normalised transfer entropy is evaluated using 50 surrogates as no difference in the results was seen when more than 50 surrogates were included. This is done for all time windows and the time window with the highest $ENTE_w^{n'_a,n'_b}$, $ENTE_{max}^{n'_a,n'_b}$ is finally selected. $ENTE_{max}^{n'_a,n'_b}(U \rightarrow Y)$ will indicate the highest level of information flow (dynamic nonlinear dependence), within the 12-second EEG, between the channel-pair U and Y in the direction $U \rightarrow Y$. $ENTE_{max}^{n'_a,n'_b}$ for the channel-pair U and Y is evaluated for both directions $U \rightarrow Y$ and $Y \rightarrow U$. This is because, in electrophysiological signals, dynamic interactions (information flow) can take place in different directions [124, 107, 69].

This study uses EEG data from 20 HC and 20 AD participants. From each participant 3 epochs of 12-second EEG data are available (Section 3.3). Considering a specific direction, $ENTE_{max}^{n'_a,n'_b}$, for a channel-pair, is evaluated individually for all the epochs of a participant. The highest $ENTE_{max}^{n'_a,n'_b}$ from all the epochs, $ENTE_{max'}^{n'_a,n'_b}$ is selected from each participant for statistical comparison between the AD and HC groups with respect to the direction of information flow. $ENTE_{max'}^{n'_a,n'_b}(U \rightarrow Y)$ will indicate the highest level of information flow (dynamic nonlinear dependence) in a participant, with regard to all 12-second EEG epochs, between the channel-pair U and Y in the direction $U \rightarrow Y$. Essentially, from a dynamic systems perspective, $ENTE_{max'}^{n'_a,n'_b}$ indicates the highest level of nonlinear dynamic interaction between a channel-pair, in a specific direction.

For a channel-pair U and Y, $ENTE_{max'}^{n'_a,n'_b}(U \rightarrow Y)$ and $ENTE_{max'}^{n'_a,n'_b}(Y \rightarrow U)$ are evaluated for all HC and AD EEG data for all three 12-second epochs (Section 3.3). The Mann–Whitney U test [169] is performed to identify any statistically significant differences (p-value ≤ 0.05) between HC and AD $ENTE_{max'}^{n'_a,n'_b}$ metrics with respect to the direction of information flow $U \rightarrow Y$ and $Y \rightarrow U$. This is done for selected channel-pairs, in both Eyes-open (EO) and Eyes-close (EC) cases (Chapter 3, Subsection 3.5.3, Tables 3.4 and 3.5 respectively). The Benjamini-Hochberg [29] false discovery rate

controlling method is used to account for multiple comparisons.

In summary, the EEG data of selected channel-pairs are filtered between $2 - 30$Hz (using the FFT method, see Section 3.3) and down-sampled to 100Hz. The effective normalised transfer entropy between the channel-pairs is estimated along a sliding time window across the 12-second epoch. The highest effective normalised transfer entropy from all time windows, $ENTE_{max}^{n'_a,n'_b}$ is selected with respect to each direction between the channel-pairs. $ENTE_{max}^{n'_a,n'_b}$ is evaluated for all HC and AD participant data. The highest value of this metric across all epochs, $ENTE_{max'}^{n'_a,n'_b}$, of each participant is then used for statistical comparisons between HC and AD groups to identify significant differences in the information flow with respect to the direction of flow. The next section will highlight these results followed by a discussion.

4.5 Results and Discussion

In the previous chapter (Chapter 3), a subset of channel pairs that have significant FC (pairwise statistical dependencies) between HC and AD groups are identified. It was shown that these channel pairs and its relation to functional networks are consistent with those obtained from previous studies using fMRI, rsfMRI and EEG. In this chapter, the information flow between these channel-pairs is assessed in different time windows using transfer entropy. This is an exploratory analysis to investigate the statistically significant changes in information flow (directed dynamic nonlinear dependencies), between the selected channel-pairs, in mild to moderate AD in comparison to HC. This information will then be used as prior knowledge in the next chapter to understand, at an EEG sensor level, the cortical nonlinear dynamical changes between HC and AD groups. This type of analysis using the resting-state EEG in the context of AD is novel.

In Section 4.4, the question of what time window to use was discussed. It is recommended to set the minimum window size to be as large as the period of the lowest frequency component in resting-state data. The lowest frequency considered in this study is 2Hz, a period of 0.5 seconds. In this study sliding time windows of lengths 1, 2, 3 and 4 seconds are used. This is to allow a long enough time window to capture any low-frequency related interactions [110] between the EEG channel-pairs considered (Chapter 3, Subsection 3.5.3, Tables 3.4 and 3.5). The sliding window with a 50% overlap is considered to reduce computational time. The grid search for the delay

embedding (eq.(4.1) and (4.2)) is done for the ranges $n_a = [1, 10]$ and $n_b = [1, 10]$. Effective normalised transfer entropy is evaluated using 50 surrogates as no difference in the results was seen when more than 50 surrogates were included.

It was found that for the frequency range considered $2 - 30\ Hz$, only the 4-second sliding time window produced any significant results considering participants. This is maybe due to the interactions of the lower frequency bands, i.e. delta ($2-4\ Hz$), theta ($4-8\ Hz$) and alpha ($8-13\ Hz$) [21], taking more precedence in the information flow in the AD case [121, 138, 80, 269, 110, 136].

Tables 4.1 and 4.2 illustrate the respective results for the EC and EO cases. These outcomes are obtained using the procedures outlined in Section 4.4 and the methodology described in Section 4.3.2 for evaluating effective normalised transfer entropy. From the subset of channels considered, the tables 4.1 and 4.2 show the statistically significant differences in information flow between the HC and AD groups. The $ENTE^{n'_a,n'_b}_{max'}$ metric, Section 4.4) is used for this purpose. In tables 4.1 and 4.2, $\Delta ENTE$ represents the median differences in $ENTE^{n'_a,n'_b}_{max'}$ between the HC and AD groups. Thus, a negative $\Delta ENTE$ indicates an increased information flow in the AD group, while a positive $\Delta ENTE$ suggests a reduction of information flow in the AD group. The cortical regions indicated in the said tables are in respect to Table 3.2 in Chapter 3, Section 3.5. These findings will be presented, accompanied by a discussion of comparable results from other studies.

Table 4.1: *EC case, effective transfer entropy*

$ENTE(U \to Y)$ (Indexes and names)				$\Delta ENTE$	Length of Time Window	Connecting regions
4	13	T6-O2	C4-CZ	-0.004335	4s	CP - O
13	4	C4-CZ	T6-O2	-0.004332	4s	

Table 4.2: *EO case, effective transfer entropy*

$ENTE(U \to Y)$ (Indexes and names)				$\Delta ENTE$	Length of Time Window	Connecting regions
3	12	T5-O1	C3-CZ	-0.004206	4s	CP - O
12	3	C3-CZ	T5-O1	-0.003354	4s	
4	13	T6-O2	C4-CZ	-0.004342	4s	
3	5	T5-O1	P3-PZ	-0.007552	4s	P - O
5	3	P3-PZ	T5-O1	-0.007575	4s	

This study reveals, in mild to moderate AD, considering the frequency range $2-30\ Hz$ and a 4-second sliding time window with 50% overlap, the statistically significant changes in information flow among brain regions (at an EEG sensor level at the resting-state) occur between the parietal and occipital regions (P–O) and also between the centro-parietal and occipital regions (CP–O). Interestingly the results obtained indicate, in both EC and EO EEG, intra-hemispheric information flow in CP–O has increased in the AD group. In the EO case it is both P–O and CP–O. Thus, indicating that, in mild to moderate AD, intra-hemispheric information flow in both P–O and CP–O has increased. This alludes to a compensatory mechanism due to the loss of connectivity in other regions [243].

From the EC EEG (Table 4.1), it is seen that information flow within the right hemisphere of CP–O has increased in the AD group. In the EO case (Table 4.2), however, indicates an increased information flow within both left and the right hemispheres of CP–O. Furthermore, in the EO case, the left hemisphere of P–O has increased information flow and comparing the $\Delta ENTE$ values, this seems more prominent overall.

In [262] the authors demonstrated an augmented neural complexity measure in AD patients in the delta and theta bands, while the multichannel correlation dimension was amplified in the beta band. As pointed out by [243], these results contradict a simplistic notion of decreased complexity in AD as it suggests a pattern of both declines and increases in connectivity across different frequency bands. Another study further clarified this pattern by revealing a reduction in mainly long-distance left hemisphere connectivity in low alpha and beta bands, and an *upsurge in parietal theta and parietal-occipital beta/gamma connectivities* in AD patients relative to HC [245]. The selective loss of long-distance left hemisphere connectivity is noteworthy, considering the outcomes of [198] and [198], which also indicate a particular susceptibility of the left hemisphere. The escalated connectivity in the parietal and occipital regions is noteworthy and comparable to this study, and it is plausible that these alterations may indicate a compensatory mechanism [243].

Similar results in the resting-state EEG of mild AD patients were shown by Frantzidis *et al.* [82]. The study found that compensatory mechanisms were evident through the formation of additional hubs (connection points in graph theory) on the left frontal and parietal regions. In resting-state MRI and fMRI, Behfar *et al.* [28] using a graph theoretical method showed that compensatory changes in AD involve an increase of

degree centrality (potential of influence) in cognition-related brain regions of the middle frontal gyrus, precentral gyrus, and superior parietal lobe despite local atrophy.

Even though the results in this study are comparable to the other work, a more comprehensive analysis using transfer entropy is needed to improve the results presented here. Shorter overlaps between time windows and consideration of a wider frequency range, $2 - 50\ Hz$ is needed. This will be done in a future study.

4.6 Chapter Summary

In the previous chapter, a subset of EEG channel-pairs that have significant FC changes between HC and AD were identified. The aim of this chapter is to identify any statistically significant differences, in mild to moderate AD, concerning the direction of information flow in the selected channel-pairs. Since the direction of information flow between channel-pairs can change with time, an exploratory analysis is conducted using transfer entropy within overlapping time windows. The effective transfer entropy, between the selected channel-pairs, is estimated for the range of delays using random phase surrogate testing method to mitigate spurious nonlinear dependencies. The effective normalised transfer entropy measure is used to compare the changes in the highest levels of information flow between HC and AD groups, respective to each direction between the channel-pairs. The results indicate that, in mild to moderate AD, intra-hemispheric information flow between parietal and occipital and between centro-parietal and occipital regions has increased within both hemispheres. Overall an increase in the intra-hemispheric information flow in the left hemisphere of the parieto-occipital is more prominent. It is plausible that the increase in information flow in AD, between those mentioned regions, may indicate a compensatory mechanism. The findings of this chapter are comparable to other studies based on resting-state EEG and rsfMRI.

Chapter 5

Nonlinear System Identification and Frequency Response Analysis for Characterising AD Using the Resting-State EEG

5.1 Introduction

Global and local processes in the brain are believed to interact through CF interactions, facilitating the integration of information across different brain regions [119]. Moreover, previous studies have reported specific alterations in CF interactions within the EEG related to neurodegenerative diseases like AD [271, 80, 117]. In this chapter, the novel application of system identification and frequency response analysis is used to characterise AD in relation to the changes in CF interactions.

In Chapter 3, the analysis focused on the FC of the resting-state EEG to identify significant changes in the pairwise statistical relationship between the HC and AD groups. Building upon this, Chapter 4 utilised these specific subsets of EEG channel-pairs to conduct an exploratory analysis, investigating the directed connectivity (EFC) or the flow of information (directed dynamic nonlinear dependencies) between these channels in individuals with mild to moderate AD compared to healthy cohorts. The findings from Chapter 4 revealed that patients with AD exhibited both increases and decreases

in information flow between specific cortical regions (EEG channels) compared to HC, suggesting the presence of compensatory mechanisms.

In this chapter, the direction of information flow, along with the specific time window and epoch where the maximum transfer of information occurs between the respective channel-pairs in each participant, will serve as prior knowledge. This information will be utilised to construct input-output dynamic models using system identification techniques. These identified models will then be analysed in the frequency-domain using frequency response analysis methods (NOFRFs) to examine the changes in CF interactions involved in the compensatory mechanisms within the cortical layers of individuals with AD.

System identification focuses on acquiring mathematical representations of dynamic systems using input-output data collected from the system. The ultimate goal of the system identification process is to develop mathematical models that can accurately link the input data to the corresponding output data of the system [32, 163]. Additionally, it is important for the model to accurately describe the behaviour of the underlying system, enabling the analysis of its dynamics [164, 143, 239]. Since the frequency response characteristics of identified models remain unchanged and unique for all local solutions. Using the frequency response analysis methods, the CF interactions within the system can be analysed [32, Chapter 6]. These techniques have also been successfully applied in neuroscience to study the interactions in different brain regions using electrophysiological data [102, 104, 37, 99]. However, the application of these methods for characterising AD is novel.

5.2 System Identification Methodology

System identification is a technique that infers and builds black-box time-series models that describe the dynamic behaviour of linear and nonlinear systems from experimental input-output data. Thus, it involves two main objectives:

- Accurately mapping the input(s) to the output(s) of the system, enabling the prediction of new and unseen data.
- Capturing the underlying dynamics of the system within the model.

The second objective is particularly crucial in the context of identifying changes in

cortical dynamics in AD compared to HC, as neurodegeneration is known to cause significant changes in brain network complexity, thus impacting nonlinear dependencies between EEG channels (ref lit rev, prev chapter). Therefore, the identification procedure must accurately capture the transmissibility dynamics between EEG channels to effectively infer underlying cortical dynamic changes.

With respect to the aforementioned objectives, system identification is used to determine a specific functional relationship that maps past inputs (input lagged terms), denoted as

$$\mathbf{U}(t) = [u(t-1)\ ,\ u(t-2)\ ,\ \cdots\ ,\ u(t-n_b)]\,, \tag{5.1}$$

and past outputs (output lagged terms), represented as

$$\mathbf{Y}(t) = [y(t-1)\ ,\ y(t-2)\ ,\ \cdots\ ,\ y(t-n_a)]\,, \tag{5.2}$$

to the present output in time $y(t)$. n_a and n_b are the number of past output and input time instances considered and relates to the Lyaponov exponents of the actual system that is being modelled [180]. The said functional mapping is described by the equation:

$$y(t) = f\big(\mathbf{Y}(t), \mathbf{U}(t)\big) + e(t) \tag{5.3}$$

where $y(t)$ and $u(t)$ refer to the output and input respectively, while $e(t)$ represents the error between the predicted output $f\big(\mathbf{Y}(t), \mathbf{U}(t)\big)$ and the actual output $y(t)$ at time instance t. Throughout the following sections, the notations defined above are carried through.

5.3 Black-box time-series model structures: ARX and NARX models

Within the field of system identification, various black-box modelling structures are available, including the Volterra series, neural networks, fuzzy models, as well as a range of linear and nonlinear auto-regressive time-series model structures [33, 162, 240, 191], among others. This overview will specifically concentrate on discrete-time black-box model structures, specifically those based on linear and nonlinear auto-regressive models with exogenous inputs.

5.3.1 Linear auto-regressive model with exogenous input: ARX model structure

One of the most common linear black-box time-series models is the Auto-Regressive model with eXogenous input (Auto-Regressive with eXogenous input (ARX)). The ARX model structure is given by

$$y(t) = \sum_{i=1}^{n_a} a_i y(t-i) + \sum_{j=1}^{n_b} b_j u(t-j) + e(t) \tag{5.4}$$

where a_i, $i = 1, \cdots, n_a$, and b_j, $j = 1, \cdots, n_j$, are the weightings on the past outputs and inputs, respectively. $e(t)$ are the model residuals. Eq.(5.4) can be written in the format shown in eq.(5.3) as

$$y(t) = \mathbf{X}(t) \times \boldsymbol{\theta} + e(t) \tag{5.5}$$

where $\mathbf{X}(t) = [\mathbf{Y}(t)\ , \mathbf{U}(t)]$ is the vector containing information of the past outputs $\mathbf{Y}(t)$ (eq.(5.2)) and inputs $\mathbf{U}(t)$ (eq.(5.1)) at a given time instance t. $\boldsymbol{\theta}$ is the vector containing the parameters of the model structure where $\boldsymbol{\theta} = [\boldsymbol{\theta_y}\ , \boldsymbol{\theta_u}]^T$. $\boldsymbol{\theta_y} = [a_1, \cdots, a_{n_a}]^T$ and $\boldsymbol{\theta_u} = [b_1, \cdots, b_{n_b}]^T$ are the weights for $\mathbf{Y}(t)$ and $\mathbf{U}(t)$ respectively. It can be seen from eq.(5.4) and (5.5), the ARX model structure models the present output $y(t)$ as a summation of weighted past inputs and outputs, inferring a linear relationship between $y(t)$, $\mathbf{U}(t)$ and $\mathbf{Y}(t)$. Thus, in relation to eq.(5.3), the ARX model structure is a linear functional mapping.

5.3.2 Nonlinear auto-regressive model with exogenous input: NARX model structure

When a system displays nonlinear properties (Section 2.5), the model structure utilised to represent it must also be nonlinear to accurately capture the system dynamics. The Nonlinear Auto-Regressive with eXogenous input (Nonlinear Auto-Regressive with eXogenous input (NARX)) model [155] is a nonlinear extension to the linear ARX model and has been extensively applied in research pertaining to model identification of complex nonlinear systems, analysis, and control of diverse nonlinear systems [287, 79, 207, 140, 101, 254, 55, 38, 52]. The NARX model is a nonlinear functional

mapping between the past outputs $\mathbf{Y}(t)$ and inputs $\mathbf{U}(t)$ where

$$y(t) = f\big(\mathbf{X}(t)\big) + e(t). \tag{5.6}$$

In eq.(5.6) above, $f(\ )$ is a nonlinear mapping function which describes the nonlinear dynamics. $\mathbf{X}(t) = [\mathbf{Y}(t)\ , \mathbf{U}(t)]$ and $e(t)$ are the model residuals. Eq.(5.7) below presents the polynomial NARX model [32], which is the widely used representation of the NARX model structure.

$$\begin{cases} y(t) = f\big(\mathbf{X}(t)\big) + e(t) \\ \\ f\big(\mathbf{X}(t)\big) = \displaystyle\sum_{n=0}^{N_p} \phi_n\big(\mathbf{X}(t)\big) \\ \\ \phi_n\big(\mathbf{X}(t)\big) = \displaystyle\sum_{p=0}^{n} \Bigg(\sum_{k_1=1}^{n_a} \sum_{k_2=1}^{n_a} \cdots \sum_{k_n=1}^{n_b} \Big(C_{p,q}(k_1, \cdots, k_{p+q}) \prod_{i=1}^{p} y(t-k_i) \prod_{i=p+1}^{p+q} u(t-k_i) \Big) \Bigg) \end{cases} \tag{5.7}$$

where n represents the order of the polynomial, where $n = 1, \cdots, N_p$, and N_p denotes the maximum polynomial order or the highest degree of polynomial nonlinearity. $q = n - p$. The term $\phi_n\big(\mathbf{X}(t)\big)$ corresponds to the n^{th} order component of the polynomial NARX model. $C_{p,q}(\cdots)$ refers to the model parameters associated with the polynomial terms of degree n. For $n = 1$, $\phi_1\big(\mathbf{X}(t)\big)$ encompasses all linear combinations of past outputs $\mathbf{Y}(t)$ and past inputs $\mathbf{U}(t)$. For $n \geq 2$, $\phi_n\big(\mathbf{X}(t)\big)$ incorporates the nonlinear terms arising from the n^{th} order polynomial combinations involving different instances of past outputs and inputs in $\mathbf{Y}(t)$ and $\mathbf{U}(t)$. Eq.(5.7) can be written in a more compact form as

$$\begin{cases} y(t) = \Big[\bar{\boldsymbol{\phi}}_1, \cdots, \bar{\boldsymbol{\phi}}_{N_p}\Big] \times \bar{\boldsymbol{\theta}} + e(t) \\ \\ \bar{\boldsymbol{\theta}} = \big[\boldsymbol{\theta}_1, \cdots, \boldsymbol{\theta}_{N_p}\big]^T \end{cases} \tag{5.8}$$

where $\bar{\boldsymbol{\phi}}_n$ is a vector containing the nonlinear terms resulting from the n^{th} order polynomial combinations of different past output and input instances, i.e. vector containing the $\prod_{i=1}^{p} y(t-k_i) \prod_{i=p+1}^{p+q} u(t-k_i)$ terms (nonlinear lagged terms) in eq.(5.7). The row vector $\boldsymbol{\theta}_n$ contains the respective parameters or weights for the elements in $\bar{\boldsymbol{\phi}}_n$, associated with $C_{p,q}(k_1, \cdots, k_{p+q})$ in eq.(5.7). The NARX structure has the ability to provide a concise portrayal of a diverse array of nonlinear complex systems. This

is accomplished without the need for any prior knowledge of the underlying physics, which is attributable to the black-box nature of the NARX representation.

5.4 Procedures in system identification

The systematic approach to system identification encompasses the solution of four key problems [32, 162]. These problems are sometimes solved concurrently or iteratively, depending on the identification algorithm and strategy employed [32, 162, 118, 76, 96, 74, 156, 34]. The following four steps summarize these problems, which will be discussed in detail in the subsequent sections:

1. **Structure detection**. Determining an appropriate structure that effectively maps the input-output variables (lagged terms) based on the type of system and the acquired data. In the nonlinear instance, for example, in regard to the polynomial NARX, this also involves determining the highest degree of polynomial nonlinearity.

2. **Parameter estimation**. Estimating the parameters that quantify the weight of each term in the given model structure.

3. **Model selection**. Selecting the most suitable model that achieves a desirable trade-off between bias and variance from a set of competing models.

4. **Model validation**. Validating the selected model using performance criteria and validation tests to establish confidence in the model based on the intended purpose of modelling.

The following subsections will discuss these procedures with appropriate methods to obtain parsimonious models.

5.4.1 Model structure detection

As seen from eq.(5.5) and eq.(5.8), (N)ARX models can be easily represented in a matrix format. Therefore, given an appropriate model structure, the relevant parameters can be evaluated using linear regression. However, determining what linear and nonlinear terms to include in the model structure is significant to obtain parsimonious models, especially in the nonlinear case [32, Chapter 1]. The Forward Regression

OLS (Forward Regression OLS (FRO)) algorithm [53, 35], based on Orthogonal Least Squares (Orthogonal Least Squares (OLS)) along with an appropriate term selection criterion [139, 270, 109], can efficiently choose model terms (regressors) in a forward selection approach (model terms are selected one at a time sequentially based on a selection criterion) to achieve a globally optimum parsimonious model. The FRO can evaluate the impact of each term on the output, independent of the influence of other terms. This evaluation depends on the criterion used for term selection. This enables the selection of the appropriate terms to be included in the final model in a sequential fashion–forward selection.

5.4.2 Orthogonal least squares method

The OLS method is a reliable means of estimating model parameters. This is particularly true when compared to traditional ordinary least squares, which may prove unreliable due to the need to compute the inverse of an information matrix that is frequently ill-conditioned. Given a linear regression model

$$\mathbf{y} = \mathbf{\Phi}\boldsymbol{\theta} + \boldsymbol{\epsilon}, \tag{5.9}$$

where $\mathbf{\Phi} \in \mathbb{R}^{L\times M}$ is the information matrix (regression matrix). $\mathbf{y} \in \mathbb{R}^{L\times 1}$ is a vector contain observations of the dependent variable. $\boldsymbol{\theta} \in \mathbb{R}^{M\times 1}$ is the vector of model parameters or weights for the respective terms in $\mathbf{\Phi}$. $\boldsymbol{\epsilon} \in \mathbb{R}^{L\times 1}$ is the vector containing the model residuals. An orthogonal decomposition of the regression matrix in eq.(5.9) is carried out where

$$\mathbf{y} = \mathbf{\Phi}\mathcal{A}^{-1}\mathcal{A}\boldsymbol{\theta} + \boldsymbol{\epsilon}, \tag{5.10}$$

$$\mathbf{y} = \mathbf{W}\mathbf{g} + \boldsymbol{\epsilon}, \tag{5.11}$$

$\mathbf{W} = \mathbf{\Phi}\mathcal{A}^{-1}$, where $\mathbf{W} = [\mathbf{w}_1, \cdots, \mathbf{w}_M] \in \mathbb{R}^{L\times M}$ in which $\mathbf{w}_i$ is the i^{th} auxiliary orthogonal regressor corresponding to the i^{th} regressor (model term) in $\mathbf{\Phi}$, where $i = 1, \cdots, M$. $\mathcal{A} \in \mathbb{R}^{M\times M}$ is an upper triangular matrix. $\mathbf{g} = [g_1, \cdots, g_M]^T$ is the associated parameter vector where $\mathbf{g} = \mathcal{A}\boldsymbol{\theta}$. The orthogonal decomposition of $\mathbf{\Phi} = \mathbf{W}\mathcal{A}$ is typically obtained by employing the modified Gram-Schmidt algorithm [53].

The columns in $\mathbf{W}$ are orthogonal with respect to each other, therefore, the corresponding parameters or weights in $\mathbf{g}$ are uncoupled. This enables the assessment of the individual contribution of each regressor (associated model term) in $\mathbf{W}$ towards reduc-

ing the error between the observed outputs $\mathbf{y}$ and the predicted outputs $\mathbf{\Phi\theta} = \mathbf{Wg}$.

5.4.3 The Forward Regression OLS

The FRO algorithm [53, 35] is a method used for model term selection in linear regression. It aims to identify the most relevant model terms that contribute to the prediction or explanation of the dependent variable according to a term selection criterion. The FRO algorithm achieves this using the OLS method along with the term selection criterion to assess the contribution of each term independently and sequentially. The FRO procedure for model structure detection [53, 35] is summarised as follows:

1. Initially, evaluate the assessment metric, of the term selection criterion used, for all the individual regressors. Then identify the best model term to add according to the criterion.
2. Orthogonalise the rest of the regressors with respect to the selected regressor(s). Then evaluate the assessment metric for the remaining regressors and select the next best term to add according to the criterion.
3. Repeat step 2 until a stopping criterion is satisfied.

5.4.4 Model term selection criteria for the Forward Regression OLS algorithm

This sub-section will highlight the term selection criteria, ERR and PRESS, that are commonly used with the FRO algorithm.

The Error Reduction Ratio (ERR)

The ERR gives a measure of how each regressor (model term), in an orthogonal sense, contributes to the variance of the observed output [32, Chapter 3]. Thus, when using the ERR, the aim is to maximise the explained variance of the model based on the contribution of each regressor to the observed output variance [139]. The ERR value (assessment metric) for each orthogonal regressor was evaluated. The regressor that is associated with the highest ERR is added to the model in step 2 of the FRO procedure shown in the previous sub-section.

The Prediction Sum of Squares (PRESS) statistic

Formulating a model construction algorithm that directly optimises the model's generalisation ability, without requiring a distinct validation dataset, is a desirable goal. One solution is the use of the PRESS-statistic, which offers the necessary capability to achieve the said objectives with leave-one-out cross-validation [270, 109]. [109] demonstrated the computation of the PRESS is simplified through the use of the orthogonalisation procedure inherent in FRO and can be easily incorporated into the algorithm for term selection purposes. Essentially, using the PRESS-statistic (assessment metric), the regressor that reduces the predicted leave-one-out cross-validation error is selected each time in step 2 of the FRO procedure. Thus a fully automated procedure without resorting to any other validation data set for model evaluation using the FRO can be achieved.

In comparison, the ERR aims to maximise the predicted explained variance while the PRESS aims to reduce the predicted leave-one-out cross-validation error. Therefore, in this work, due to the added benefit of cross-validation, the PRESS-statistic criterion is used for term selection in the iterative FRO (iFRO) algorithm. Therefore, model validation in relation to cross-validation is incorporated into the structure detection algorithm.

5.4.5 The iterative Forward Regression OLS (iFRO)

The original FRO procedure presented above follows a specific orthogonalisation path [96], which is determined by the initial term chosen in step 1. As mentioned in [96], the selection of subsequent terms may vary depending on the initial term, leading to a global solution space of potential models based on the chosen orthogonalisation path (based on the initial term that is selected). Consequently, the classical FRO approach may not always yield an optimal solution across the entire solution space of possible models. To address this limitation, the iterative-FRO (iFRO) algorithm proposed in [96] seeks the optimal solution within the global solution space while preserving the benefits of simplicity and computational efficiency. The effectiveness of the new algorithm is supported by both theoretical analysis and simulations [96].

The iFRO algorithm has the ability to generate multiple competing models by following different orthogonalisation paths [96]. This is achieved by forcing various initial terms in step 1 of the FRO procedure, generating several competing models. Among

these models, the best-performing one can be selected [96] based on a model selection criterion. The generation of competing models is efficient, and the iFRO algorithm has a higher likelihood of achieving a global optimum solution in terms of minimising predicted errors compared to the original FRO algorithm. For more detailed information on the iFRO algorithm, readers are encouraged to refer to [96]. In this study, the iFRO algorithm will be utilised for determining the model structure, with the PRESS-statistic serving as the assessment metric for term selection.

5.4.6 Model selection

Model selection and structure detection are two distinct processes. Structure detection involves selecting the appropriate regressors (terms) that should be included in the model. It often yields multiple competing models to choose from. On the other hand, model selection is the methodology used to choose the most suitable model from a set of candidate models. The selected model should be capable of predicting unseen data effectively while achieving a good balance between bias and variance.

The presence of bias and variance in model predictions arises from different sources of error in the modelling process. Bias reflects how well the model fits a specific dataset, while variance relates to the model's flexibility in capturing various aspects of a system, such as different operating conditions [162]. Increasing the complexity of a model, typically achieved by adding more terms in parametric models, enhances flexibility and reduces the error between predicted and observed values. However, excessive complexity can lead to overfitting, where the model captures noise in the data and increases variance. Conversely, reducing the number of model terms decreases the fit and increases bias but lowers variance.

When a new dataset is obtained under the same operating conditions as the identification (training) dataset, both overly complex and overly simple models perform poorly. An overly simple model fails to capture the true underlying dynamics of the system, while an overly complex model fits random noise sequences [32, Chapter 1][162]. Therefore, selecting an appropriate model structure involves a compromise, depending on the intended purpose of the model [162].

Model selection criteria

Akaike's Information Criteria (Akaike's Information Criteria (AIC)) [13], Final Prediction Error (Final Prediction Error (FPE)) [12], and Bayesian Information Criterion (Bayesian Information Criteria (BIC)) [230] are commonly used criteria for model selection. These criteria assess models based on the error between the model's predicted output and the observed output, as well as the model's complexity. They assign a score to each model, with a penalty applied for higher model complexity. The model with the lowest score is selected as the preferred choice. BIC, in particular, applies a higher penalty for increased model complexity compared to AIC and FPE [40]. Consequently, models that minimise the BIC score tend to have lower complexity compared to models obtained by minimising the AIC and FPE scores.

5.5 Application of System Identification and Frequency Response Analysis in Neuroscience

The first system identification based nonlinear GC analysis was first introduced by [286, 285]. An initial study with this type of causality analysis was done on EEG data of AD and healthy individuals by [37]. The ERR-causality test [284] was used to analyse the EEG data of both resting states, eyes open (EO) and eyes closed (EC). The study pointed out the significance of using this method as a non-invasive and economical diagnostic aid. The ERR-causality enabled for time-domain linear and nonlinear interactions between EEG channels to be analysed individually even under coloured noise. The NARX based causality tests were extended to the nonlinear frequency response by [102, 103]. In which, the GFRFs from an identified NARX model were used for the causality analysis, essentially generalising the spectral Granger causality [54] to the nonlinear instance.

The OLS-based (N)ARX modelling methods have some interesting advantages in the application of EEG analysis. As such; model structure selection under non-stationary conditions and in the presence of coloured noise [101], are the most significant features. This is because EEG signals have both these circumstances. Furthermore, as mentioned in Chapter 2 Subsection 2.5.5, the use of the NOFRFs instead of the GFRFs is novel and has an interesting outlook in relation to EEG analysis. This is because the identified model can be decomposed to the respective orders of nonlinearities in the frequency-

domain. This enables the observation of CF interactions between the input-ouptut EEG channels considered at higher-order nonlinearities (nonlinearities higher than an order of 2). This type of dynamical analysis has not yet been done in neuroscience.

5.6 Characterising Alzheimer's Disease Using System Identification and Frequency Response Analysis

As discussed in Subsection 2.5.3 and shown by eq.(2.4), the composition of the output spectrum, $Y(j\omega)$ (output frequency response), in dynamic nonlinear systems is complex. In such systems, the properties of the nonlinear dynamics are described by the GFRFs, $H_n(\cdots)$, eq.(2.6). As shown in eq.(2.4), $Y(j\omega)$ is dependent on both $H_n(\cdots)$ and how the input spectrum $U(j\omega)$ operates on $H_n(\cdots)$ to produce $Y_n(j\omega)$ for $n = 1, \cdots, N$, where N is the maximum nonlinearity considered. $Y_n(j\omega)$ is the output of the n^{th} order nonlinearity (n^{th} order output frequency response) such that $Y(j\omega) = \sum_n Y_n(j\omega)$. Therefore, to understand the CF interactions between the frequency components in the input and the output of a nonlinear dynamic system, the system needs to be probed appropriately.

As shown in eq.(2.10) and discussed in Subsection 2.5.5, in nonlinear systems, the composition of the output spectrum, $Y(j\omega)$, for input with a spectrum $U(j\omega)$ is described by the NOFRFs in a much more trivial manner than in eq.(2.4). Therefore, the NOFRFs can be used to easily decompose $Y(j\omega)$ into $Y_n(j\omega)$ for $n = 1, \cdots, N$ to observe the nonlinearities involved in specific CF interactions.

The aim of this chapter is to characterise AD in relation to the changes in CF interactions in comparison to HC. The findings from Chapter 4 are used to achieve this. The channel-pairs that exhibit significant changes in information flow (directed dynamic nonlinear dependencies), from tables 4.1 and 4.2, are used to build NARX models through system identification using the iFRO method (Subsection 5.4.5). The PRESS-statistic (Subsection 5.4.4) is used in iFRO for model term selection. The NARX models identified are specific to the EEG channel-pairs and the direction of information flow between the channels.

Cortical regions of the brain are interconnected, and this is reflected in the EEG.

Thus, when building input-output models between EEG channel-pairs, care has to be taken so that the bi-variate (between two channels) models do not overly fit the data, as the respective EEG channels will contain information related to other underlying cortical regions. Therefore, to be conservative in a bias-variance sense, the BIC [230] criterion is used to select an appropriate model from the several competing models produced using the iFRO algorithm (Subsection 5.4.5).

The identified NARX models are analysed in the freqeuncy-domain using the NOFRFs to decompose the model into the respective orders of nonlinearities to produce the n^{th} order output frequency response, $Y_n(j\omega)$ in eq.(2.13). More specifically, the identified model is probed within a specific frequency range. The resulting n^{th} order output frequency responses, $Y_n(j\omega)$ where $n = 1, \cdots, N$, is observed to understand the CF interactions between the probing input frequency range and the output frequencies produced by the model. The n^{th} order output frequency responses produced by the respective models of EEG channel-pairs will enable to characterise the changes in CF interactions that occur in patients with mild to moderate AD in comparison to HC.

The method to evaluate the NOFRFs $G_n(j\omega)$ (eq.(2.11)) and the subsequent $Y_n(j\omega)$ for $n = 1, \cdots, N$ (eq.(2.13)) is achieved using the procedures as outlined in [95]. The specifics of the probing inputs used to evaluate the respective NOFRFs and the n^{th} order output frequency responses are given in the following subsection.

5.6.1 Probing inputs to analyse the identified input-output models

Brain electrical activity can be segmented into frequency bands; delta (δ), theta (θ), alpha (α), beta (β), and gamma (γ). These frequency bands provide valuable insights into the operational dynamics of the brain [48, 233]. Therefore, to understand the CF interactions between these bands, the identified NARX model between two EEG channels needs to be probed with several appropriate inputs. Each probing input will be specific to the frequency bands mentioned earlier. The corresponding output spectrum of the NARX model to each probing input is used to observe the CF interactions between the input frequencies and the output frequencies. It should be noted in this chapter only up to the β frequency range is considered.

In Hz, the frequency ranges $f_b \in [f_1, f_2]$, where $b \in \{\delta,\ \theta,\ \alpha,\ \beta\}$, of the bands mentioned above are as follows [21]; $f_\delta \in [2, 4]$, $f_\theta \in [4, 8]$, $f_\alpha \in [8, 13]$ and $f_\beta \in [14, 30]$.

The frequency spectrum of the probing input $U^b(j\omega)$, specific to the frequency range f_b is such that

$$U^b(j\omega) = \begin{cases} U^b(j\omega) & \text{when } |\omega| = f_b \\ 0 & \text{otherwise.} \end{cases} \tag{5.12}$$

The time-domain representation of $U^b(j\omega)$ [145, 146, 144], denoted as $u_b(t)$ is

$$u_b(t) = A\frac{1}{2\pi} \times \frac{\sin(f_2 \times t) - \sin(f_1 \times t)}{t} \tag{5.13}$$

where A is the amplitude. Figure 5.1 below illustrates the magnitude spectra, $\left|U^b(j\omega)\right|$, of the probing inputs, $u_b(t)$, specific to the frequency bands mentioned previously.

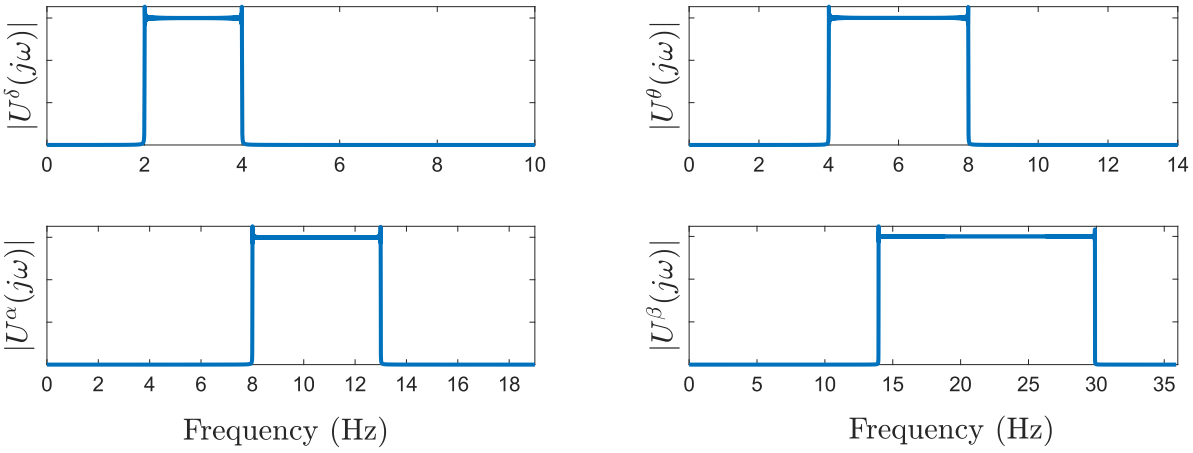


Figure 5.1: ***Magnitude spectra of the probing inputs within the respective frequency ranges.*** *The magnitude profile of the probing inputs,* $\left|U^b(j\omega)\right|$ *where* $b \in \{\delta,\ \theta,\ \alpha,\ \beta\}$*, against frequency (in Hz) is shown. The magnitude profile of these probing inputs follows the relationship shown in eq.*(5.12)*.*

5.6.2 Analysis of CF interactions from identified NARX models using output frequency responses from each order of nonlinearity

Let X and Z be two EEG channels where $x(t)$ and $z(t)$ are the respective time-series of the channels. Considering the direction of information flow (dynamic nonlinear dependencies) from $X \rightarrow Z$, a NARX model $\mathcal{M}_{X\rightarrow Z}$ is identified using the iFRO method (Subsection 5.4.5). In this instance, channel X is considered as the input and Z as the output.

To identify the NARX model that is able to reconstruct the dynamics from $X \to Z$, the number of past instances of $x(t)$ and $z(t)$, n_b and n_a respectively (Section 5.2, eq.(5.1) and (5.2)), that are used in the iFRO algorithm (Subsection 5.4.5) has to be appropriately specified [10, 11, 9, 181]. This information is taken from Chapter 4, where the delay embedding dimensions n'_a and n'_b that results in $ENTE^{n'_a,n'_b}_{max'}(X \to Z)$ (Chapter 4, Section 4.4) is used. Therefore, n'_a will be the number of past instances of $z(t)$ (output) and n'_b will be the number of past instances of $x(t)$ (input) that will be considered in the iFRO algorithm to identify the NARX model $\mathcal{M}_{X\to Z}$. This is specific to each participant EEG.

The identified model $\mathcal{M}_{X\to Z}$ is used to study the CF interactions between the channels X and Z respective to the direction $X \to Z$. This is achieved by stimulating the model $\mathcal{M}_{X\to Z}$ with the probing inputs $u_b(t)$ to produce the outputs $y_b(t)$ where $b \in \{\delta,\ \theta,\ \alpha,\ \beta,\ \gamma\}$. $u_b(t)$ has a spectrum $U^b(j\omega)$ and the corresponding output $y_b(t)$ will have a spectrum $Y^b(j\omega)$. The NOFRF (Subsection 2.5.5) method as outlined in [95] is used to decompose the model output $Y^b(j\omega)$ into the n^{th} order output frequency responses $Y^b_n(j\omega)$, in which $n = 1, \cdots, N$ where N is the maximum order of nonlinearity considered.

The magnitude profile of $Y^b_n(j\omega)$, $\left|Y^b_n(j\omega)\right|$, is used to observe the CF interactions between the frequencies of the input spectrum $U^b(j\omega)$, $f_b \in [f_1, f_2]$ (Figure 5.1), and the output $Y^b(j\omega)$ at different orders of nonlinearity. More specifically, the normalised magnitude profile of $\left|Y^b_n(j\omega)\right|$,

$$\left|\overline{Y^b_n}(j\omega)\right| = \frac{\left|Y^b_n(j\omega)\right|}{|Y^b(j\omega)|} \tag{5.14}$$

is used for this purpose. The normalised n^{th} order output frequency response, $|\overline{Y^b_n}(j\omega)|$, in eq.(5.14) above shows the percentage contribution of each n^{th} order nonlinearity to the output $Y^b(j\omega)$ at the frequency ω. Specific to the direction $X \to Z$, in $|\overline{Y^b_n}(j\omega)|$, for $\omega = f_b$ shows the level of CF interactions between the channels X and Z within the frequency range f_b at different orders of nonliterary. Similarly, for all ω outside the range f_b, $|\overline{Y^b_n}(j\omega)|$ shows the level of CF interactions between the frequencies f_b and other frequency ranges outside f_b.

NARX models, capturing the channel dynamics $X \to Z$, are identified from all participant EEGs, AD and HC. The respective normalised n^{th} order output frequency

responses, $|\overline{Y_n^b}(j\omega)|$, from all AD and HC participants are statistically compared to identify significant changes in CF interactions, specific to f_b, at different orders of nonlinearity $n = 1, \cdots, N$. It was observed that the $|\overline{Y_n^b}(j\omega)|$ in the AD group showed a considerable amount of variations (within a broader distribution) than the HC group. Therefore, the non-parametric two-sample Kolmogorov-Smirnov test [174, 183, 171] is used to identify any statistical difference between the variations in the CF interactions between AD and HC groups. The Benjamini-Hochberg [29] false discovery rate controlling method is used to account for multiple comparisons. The statistical test is applied for all HC and AD $|\overline{Y_n^b}(j\omega)|$ to identify any significant changes between the groups the at each ω.

5.7 Results and Discussion

In Chapter 3, a subset of channel pairs that have significant FC (pairwise statistical dependencies) between HC and AD groups are identified. In the previous chapter (Chapter 5), the information flow between these channel-pairs was assessed in different time windows using transfer entropy. This was an exploratory analysis to investigate the statistically significant changes in information flow (directed dynamic nonlinear dependencies), between the selected channel-pairs, in mild to moderate AD in comparison to HC. This information is used as prior knowledge in this chapter to understand, at an EEG sensor level, the cortical nonlinear dynamics in regard to changes in CF interactions between HC and AD groups. This type of analysis where system identification and frequency response analysis are applied to analyse CF interaction using the resting-state EEG in the context of AD is novel.

From the findings presented in Chapter 4, tables 4.1 and 4.2, the channel-pairs that have significant changes in information flow are analysed using system identification and frequency response analysis according to the methodology presented in Section 5.6. The statistically significant results from the frequency response analysis are reported in tables 5.1 and 5.2 for the EC and EO cases respectively. In a general sense, the significant changes in CF interactions between AD and HC groups are summarised below.

In the EC case, within the right hemisphere changes in the CF interactions between the centro-parietal (CP) and occipital regions (O) are as follows. $\mathrm{O}_\delta \rightarrow \mathrm{CP}_\theta$ interactions have increased while $\mathrm{O}_\delta \rightarrow \mathrm{CP}_\delta$ interactions have reduced. $\mathrm{O}_\theta \rightarrow \mathrm{CP}_\delta$ has

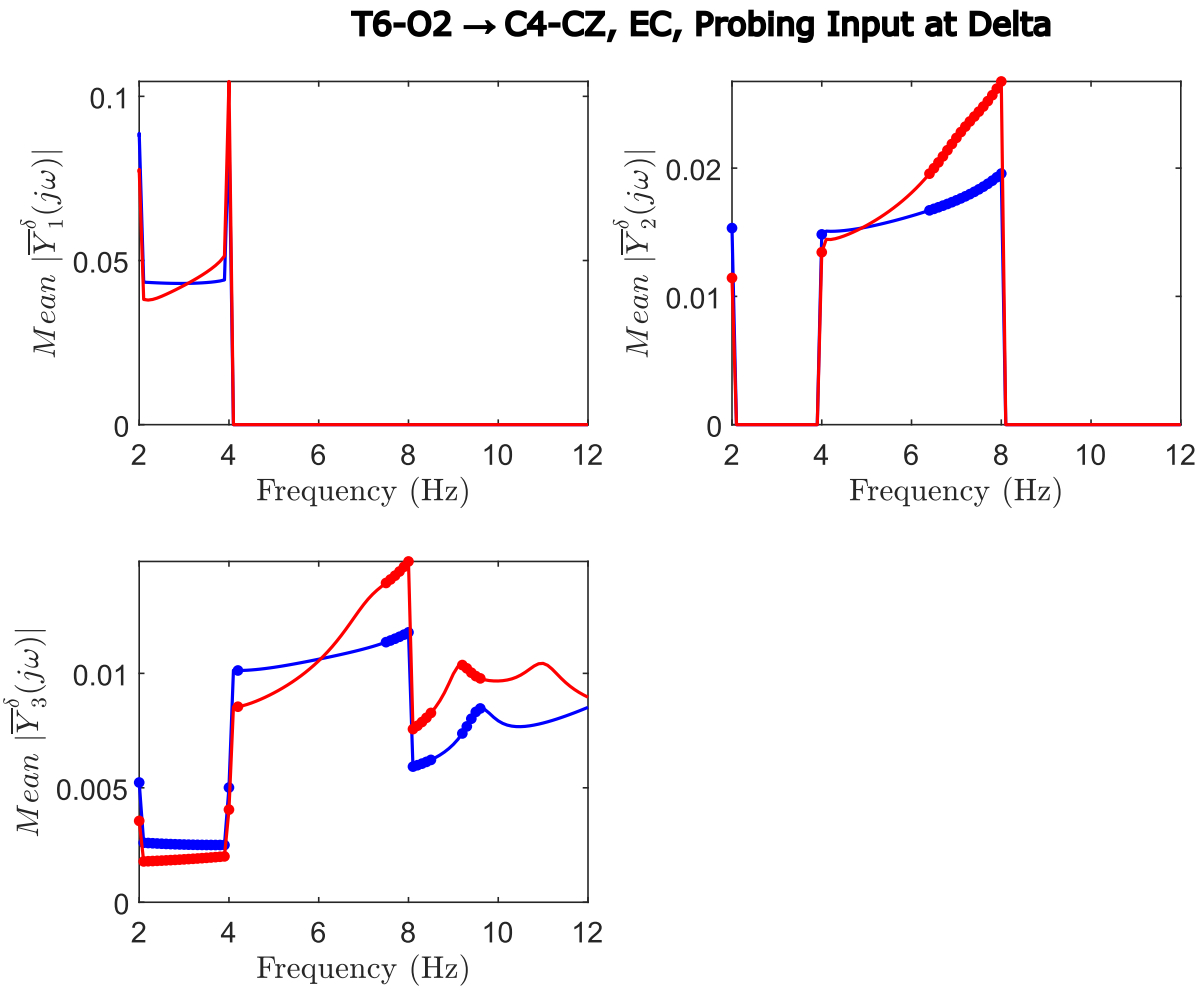


Figure 5.2: ***Frequency response analysis of NARX model T6-O2 → C4-CZ for the EC case, probed in the delta frequency range.*** *Indicated by red is the mean $|Y_n^{\delta}(j\omega)|$ for the AD group and blue for the HC group. The frequencies at which there is a significant statistical difference between the groups are indicated by the markers on the mean $|Y_n^{\delta}(j\omega)|$ plots for both HC and AD. No significant changes are observed in linear CF interactions. In the 2^{nd} order nonlinearity, $|Y_2^{\delta}(j\omega)|$, considerable differences in CF interactions between δ and the higher θ frequencies are seen where the mean $|Y_2^{\delta}(j\omega)|$ shows an average increase in the AD group. In the 3^{rd} order nonlinearity, $|Y_3^{\delta}(j\omega)|$, considerable differences in CF interactions within the δ range is seen where AD shows a decrease. While CF interactions between the δ range and the higher edge of the θ band show an average increase in the AD group. Furthermore, CF interactions between the delta band and certain frequencies in the α band indicate significant differences with an average increase in the AD group. The CF interactions concerning the δ band, between the channels in the direction T6-O2 → C4-CZ for EC case, is averagely exactly opposite to the EO case.*

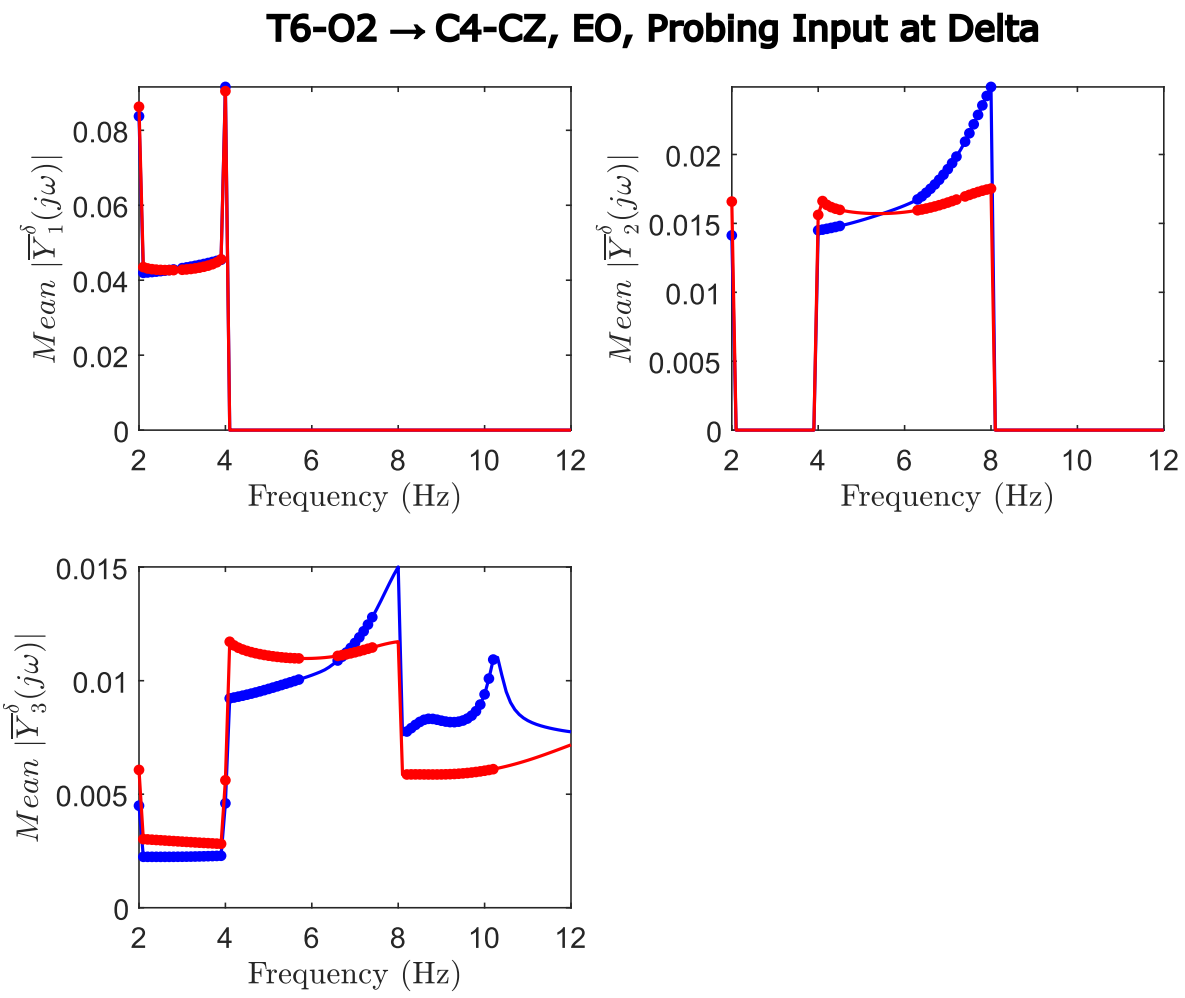


Figure 5.3: ***Frequency response analysis of NARX model T6-O2 → C4-CZ for the EO case, probed in the delta frequency range.*** *Indicated by red is mean $|Y_n^{\delta}(j\omega)|$ for the AD group and blue for the HC group. The frequencies at which there is a significant statistical difference between the groups are indicated by the markers on mean $|Y_n^{\delta}(j\omega)|$ plots for both HC and AD. Statistically significant changes are observed in linear CF interactions within the delta band. In the 2^{nd} order nonlinearity, $|Y_2^{\delta}(j\omega)|$, considerable differences in CF interactions between δ and the higher θ frequencies and certain lower θ frequencies are seen. Where the mean $|Y_2^{\delta}(j\omega)|$ shows an average increase in the AD group in the lower θ frequencies and the opposite in the higher θ frequencies. In the 3^{rd} order nonlinearity, $|Y_3^{\delta}(j\omega)|$, significant differences in CF interactions within the δ range is seen where AD shows an increase. While CF interactions between the δ range and the lower θ frequencies show an average decrease in the AD group the opposite is seen in the higher θ frequencies. Furthermore, CF interactions between the δ band and several frequencies in the lower α band indicate significant differences with an average decrease in the AD group. The CF interactions concerning the δ band, between the channels in the direction T6-O2 → C4-CZ for EO case, is averagely exactly opposite to the EC case.*

Table 5.1: *EC case, statistically significant changes in CF interactions*

$\mathcal{M}_{U\to Y}$ (Indexes and names)				Significant Frequencies in $\lvert Y_n^b(j\omega)\rvert$ – (b,n)	Average Change of $\lvert Y_n^b(j\omega)\rvert$ in AD	Connecting regions
4	13	T6-O2	C4-CZ	(δ, 2): 6.4–8.0 Hz	Increased	CP - O
				(δ, 3): 2.0–4.2 Hz	Decreased	
				(δ, 3): 7.5–9.6 Hz	Increased	
				(θ, 1): 4.0–6.1 Hz	Decreased	
				(θ, 1): 6.8–8.0 Hz	Increased	
				(θ, 2): 8.0–9.3 Hz	Increased	
				(θ, 3): 10.3–10.6 Hz	Increased	
				(θ, 3): 15.2–15.7 Hz	Decreased	

Table 5.2: *EO case, statistically significant changes in CF interactions*

$\mathcal{M}_{U\to Y}$ (Indexes and names)				Significant Frequencies in $\lvert Y_n^b(j\omega)\rvert$ – (b,n)	Average Change of $\lvert Y_n^b(j\omega)\rvert$ in AD	Connecting regions
12	3	C3-CZ	T5-O1	(δ, 1): 2.0–3.0 Hz	Decreased	CP - O
				(δ, 1): 3.3–4.0 Hz	Increased	
				(δ, 2): 6.2–6.9 Hz	Increased	
4	13	T6-O2	C4-CZ	(δ, 1): 2.0–2.6 Hz	Increased	
				(δ, 1): 2.7–4.0 Hz	Decreased	
				(δ, 2): 4.1–4.5 Hz	Increased	
				(δ, 2): 6.3–8.0 Hz	Decreased	
				(δ, 3): 2.0–6.7 Hz	Increased	
				(δ, 3): 6.8–10.2 Hz	Decreased	

reduced while $O_\theta \to CP_\theta$ has increased. $O_\theta \to CP_\beta$ interactions had decreased within a narrow band around $15Hz$.

In the EO case. Changes in CF interactions between AD and HC are seen in both the left and the right hemisphere in the CP–O. Within the left hemisphere, $CP_\delta \to O_\delta$ has both increased and decreased while $CP_\delta \to O_\theta$ has increased within a narrow band of $6-7Hz$. Within the right hemisphere, $O_\delta \to CP_\delta$ interactions have both increased and decreased while $O_\delta \to CP_\delta$ in higher orders have increased in interactions throughout the δ band. $O_\delta \to CP_\theta$ interactions more prominently have reduced. This is the exact opposite of the EC case as seen in Figures 5.2 and 5.3.

The use of system identification and frequency response analysis in identifying significant changes in directed CF interactions have been demonstrated with more prominent changes in higher-order nonlinearities. Thus, indicating that in mild to moderate AD directed CF interactions have complex changes. These changes are both increases and decreases in CF interactions further alluding to the compensation mechanisms involved in AD. This chapter analyses directed broadband CF interactions between cortical regions at an EEG sensor level. However, most of the channel-pairs that showed significant changes in information flow from Chapter 4 did not show up on these results.

5.8 Chapter Summary

In this chapter, a novel application of system identification and frequency response analysis methods from control systems engineering is presented. Building upon the findings of Chapter 4, data-driven models are constructed to capture the nonlinear dynamics between pairs of channels. These models are then examined in the frequency domain to identify noteworthy alterations in CF interactions between the AD and HC groups. This chapter analyses directed broadband CF interactions between cortical regions at an EEG sensor level. It was observed that significant changes in directed CF interactions with more prominent changes in higher-order nonlinearities. Intra-hemispheric directed CF interactions are statistically significant in both hemispheres in opposite directions in the EO case while in the EC case only within the right hemisphere was observed. This is preliminary work in applying system identification and frequency response analysis to examine directed broadband CF interactions. Further improvements are required which depend upon the improvement of the work presented in Chapter 4.

Chapter 6

Conclusions and Future Work

The primary objective of the research presented in this thesis is to utilise nonlinear analysis methods to characterise mild to moderate AD in comparison to HC (healthy controls). Specifically, the focus is on novel applications and developing data-driven techniques for examining FC, EFC and CF interactions for understanding neurodegenerative diseases using the EEG. To achieve this methods from manifold learning, information theory, and control systems engineering are employed. Consequently, this research aims to contribute towards a comprehensive data-driven framework for analysing nonlinear dynamics in cortical neural activity using the EEG. Furthermore, the research work undertaken seeks to contribute to early diagnosis and characterisation of disease progression in neurodegenerative conditions like AD.

A novel FC measure that generalises (dis)similarity using Kernel-based manifold learning was introduced to mitigate issues in using and interpreting various (dis)similarity measures used for FC. This was used to identify EEG channel-pairs that showed significant differences in FC between mild to moderate AD and HC groups. An exploratory analysis on the above mentioned EEG channel-pairs was conducted using transfer entropy to identify significant changes in the level of information flow (directed dynamic nonlinear dependencies) between these channel-pairs. This type of analysis of the resting-state EEG in the context of AD using transfer entropy is novel. The findings from the exploratory analysis were then used to produce data-driven dynamic nonlinear models using system identification. The data-driven models were then analysed in the frequency-domain using frequency response analysis for the novel characterisation

of broadband CF interactions in mild to moderate AD.

Chapter 3 of this thesis introduces a new method for analysing FC and selecting channels based on kernel-based nonlinear manifold learning. Considering local and global spatio-temporal similarities and dissimilarities between EEG channels, the FC measure ranks pairwise measures that distinguish patients with neurodegenerative diseases from healthy controls. The presented methodology can identify changes in cortical inter-relationships that are important in characterising patients with AD, and the results are consistent with previous studies while linking connectivity changes to functional networks. The main purpose of Chapter 3 is to introduce this novel FC analysis and channel selection methodology and its computational procedure. The chapter also demonstrates its efficacy against other commonly used FC measures. The robustness of the method presented against volume conduction effects in the EEG could be further assessed [43]. The Isomap-GPLVM method can be made to explicitly control the compromise between local similarity and global dissimilarity information being learnt. This can be achieved by including appropriate prior probabilities of the latent positions, in the GPLVM log-likelihood [259, 46]. This development can be used to improve the classification performance further. With the above considerations, the proposed methodology in Chapter 3 can be developed into a diagnostic tool not only for detecting neurodegenerative diseases but also for determining the important FC changes related to the disease.

From Chapter 3, the subset of channel-pairs that are identified as having significant changes in FC in mild to moderate AD was used in Chapter 4 to ascertain any statistically significant variations in the direction of information flow within the selected channel-pairs in mild to moderate AD. An exploratory analysis utilising transfer entropy in overlapping time windows is conducted since the direction of information flow between channel-pairs may vary over time. To reduce spurious non-linear dependencies, a random phase surrogate testing method is used to estimate the effective transfer entropy between the chosen channel-pairs for a range of delays. The effective transfer entropy measurements for the range of delays are then combined to create a metric that will be employed to compare the changes in information flow between AD groups with respect to each time window and direction between the channel-pairs. The findings suggest that in individuals with mild to moderate AD, there is an enhanced intra-hemispheric information flow between parietal and occipital regions, as well as between centro-parietal and occipital regions, in both hemispheres. Particu-

larly, there is a more noticeable increase in intra-hemispheric information flow in the left hemisphere in the parieto-occipital region. This increase in information flow between these specific regions in AD may potentially signify a compensatory mechanism. The results obtained in this chapter align with other studies that utilize resting-state EEG and rsfMRI. To further improve the results presented in this chapter, a broader range of frequencies should be considered and transfer entropy should be applied at longer delays to accommodate interactions between lower and higher frequency bands. Furthermore, sliding windows with more overlaps should also be considered.

Chapter 5 is a novel application of system identification and frequency response analysis for the characterisation of broadband CF interactions in mild to moderate AD. Findings from Chapter 4 are used to produce data-driven models that capture the nonlinear dynamics between channel-pairs. These models are analysed in the frequency-domain to identify significant changes in broadband CF interactions between AD and HC groups. It was observed that these changes are both increase and decrease CF interactions, more prominent in higher order nonlinearities, further alluding to the compensation mechanisms involved in AD. The work represented is an initial exploration of applying system identification and frequency response analysis to study directed broadband CF interactions. Further enhancements depend on improving the work presented in Chapter 4.

With the improvements to the work presented in Chapter 4 which would improve the analysis in Chapter 5, future work should focus on formulating a formal data-driven framework for cortical brain activity and the characterisation of neurodegenerative diseases based on the methods used in this thesis. Subsequently investigating the potential development of EEG biomarkers using such a framework, employing larger sample sizes. Additionally, further research should explore whether the identified EEG biomarkers exhibit longitudinal changes over relatively short time frames at the individual level.

Bibliography

[1] A. K. Abbas, G. Azemi, S. Amiri, S. Ravanshadi, and A. Omidvarnia. Effective connectivity in brain networks estimated using eeg signals is altered in children with adhd. *Computers in Biology and Medicine*, 134:104515, 2021.

[2] A. Abbott. Dementia: A problem for our age. *Nature*, 475(7355):S2–S4, Jul 2011.

[3] A. B. Abdessalem, N. Dervilis, D. J. Wagg, and K. Worden. Automatic kernel selection for gaussian processes regression with approximate bayesian computation and sequential monte carlo. *Frontiers in Built Environment*, 3, 2017. ISSN 2297-3362.

[4] F. Abdul Razak and H. J. Jensen. Quantifying 'causality' in complex systems: Understanding transfer entropy. *PLOS ONE*, 9(6):1–14, 06 2014.

[5] D. Abásolo, R. Hornero, C. Gómez, M. García, and M. López. Analysis of EEG background activity in Alzheimer's disease patients with lempel-ziv complexity and central tendency measure. *Medical Engineering and Physics*, 28(4):315–322, 2006.

[6] D. Abásolo, R. Hornero, and P. Espino. Approximate entropy of EEG background activity in Alzheimer's disease patients. *Intelligent Automation & Soft Computing*, 15(4):591–603, 2009.

[7] A. T. Adebisi and K. C. Veluvolu. Brain network analysis for the discrimination of dementia disorders using electrophysiology signals: A systematic review. *Frontiers in Aging Neuroscience*, 15, 2023.

[8] H. Adeli, S. Ghosh-Dastidar, and N. Dadmehr. Alzheimer's disease: Models of computation and analysis of eegs. *Clinical EEG and Neuroscience*, 36(3):131–140, 2005.

[9] L. Aguirre. A nonlinear correlation function for selecting the delay time in dynamical reconstructions. *Phys. Lett. A*, 203(2–3):88–94,, 1995.

[10] L. AGUIRRE and S. BILLINGS. Discrete reconstruction of strange attractors of chua's circuit. *Int. J. Bifurc. Chaos*, 04(04):853–864,, 1994.

[11] L. Aguirre and S. Billings. Dynamical effects of over parametrization in nonlinear models. *Phys*, D, vol. 80:26–40,, 1995.

[12] H. Akaike. Fitting autoregressive models for prediction. *Annals of the Institute of Statistical Mathematics*, 21(1):243–247, Dec 1969.

[13] H. Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723, 1974.

[14] N. K. Al-Qazzaz, S. H. B. M. Ali, S. A. Ahmad, K. Chellappan, M. S. Islam, and J. Escudero. Role of EEG as biomarker in the early detection and classification of dementia. *The Scientific World Journal*, 2014:906038, Jun 2014.

[15] T. Alotaiby, F. E. A. El-Samie, S. A. Alshebeili, and I. Ahmad. A review of channel selection algorithms for EEG signal processing. *EURASIP Journal on Advances in Signal Processing*, 2015(1):66, Aug 2015. ISSN 1687-6180.

[16] J. Aru, J. Aru, V. Priesemann, M. Wibral, L. Lana, G. Pipa, W. Singer, and R. Vicente. Untangling cross-frequency coupling in neuroscience. *Current Opinion in Neurobiology*, 31:51–61, 2015.

[17] S. Assecondi, D. Ostwald, and A. P. Bagshaw. Reliability of Information-Based Integration of EEG and fMRI Data: A Simulation Study. *Neural Computation*, 27(2):281–305, 02 2015.

[18] C. Babiloni, R. Ferri, D. V. Moretti, A. Strambi, G. Binetti, G. Dal Forno, F. Ferreri, B. Lanuzza, C. Bonato, F. Nobili, G. Rodriguez, S. Salinari, S. Passero, R. Rocchi, C. J. Stam, and P. M. Rossini. Abnormal fronto-parietal coupling of brain rhythms in mild Alzheimer's disease: a multicentric EEG study. *European Journal of Neuroscience*, 19(9):2583–2590, 2004.

[19] C. Babiloni, R. Ferri, G. Binetti, A. Cassarino, G. D. Forno, M. Ercolani, F. Ferreri, G. B. Frisoni, B. Lanuzza, C. Miniussi, F. Nobili, G. Rodriguez, F. Rundo, C. J. Stam, T. Musha, F. Vecchio, and P. M. Rossini. Fronto-parietal coupling of brain rhythms in mild cognitive impairment: A multicentric EEG study. *Brain Research Bulletin*, 69(1):63–73, 2006.

[20] C. Babiloni, R. Lizio, N. Marzano, P. Capotosto, A. Soricelli, A. I. Triggiani, S. Cordone, L. Gesualdo, and C. Del Percio. Brain neural synchronization and functional coupling in Alzheimer's disease as revealed by resting state EEG rhythms. *International Journal of Psychophysiology*, 103:88–102, 2016. Research on Brain Oscillations and Connectivity in A New Take-Off State.

[21] C. Babiloni, R. J. Barry, E. Başar, K. J. Blinowska, A. Cichocki, W. H. Drinkenburg, W. Klimesch, R. T. Knight, F. Lopes da Silva, P. Nunez, R. Oostenveld, J. Jeong, R. Pascual-Marqui, P. Valdes-Sosa, and M. Hallett. International federation of clinical neurophysiology (ifcn) – eeg research workgroup: Recommendations on frequency and topographic analysis of resting state eeg rhythms. part 1: Applications in clinical research studies. *Clinical Neurophysiology*, 131(1): 285–307, 2020.

[22] C. Babiloni, K. Blinowska, L. Bonanni, A. Cichocki, W. De Haan, C. Del Percio, B. Dubois, J. Escudero, A. Fernández, G. Frisoni, B. Guntekin, M. Hajos, H. Hampel, E. Ifeachor, K. Kilborn, S. Kumar, K. Johnsen, M. Johannsson, J. Jeong, F. LeBeau, R. Lizio, F. Lopes da Silva, F. Maestú, W. J. McGeown, I. McKeith, D. V. Moretti, F. Nobili, J. Olichney, M. Onofrj, J. J. Palop, M. Rowan, F. Stocchi, Z. M. Struzik, H. Tanila, S. Teipel, J. P. Taylor, M. Weiergräber, G. Yener, T. Young-Pearse, W. H. Drinkenburg, and F. Randall. What electrophysiology tells us about alzheimer's disease: a window into the synchronization and connectivity of brain neurons. *Neurobiology of Aging*, 85:58–73, 2020.

[23] C. Babiloni, X. Arakaki, H. Azami, K. Bennys, K. Blinowska, L. Bonanni, A. Bujan, M. C. Carrillo, A. Cichocki, J. de Frutos-Lucas, C. Del Percio, B. Dubois, R. Edelmayer, G. Egan, S. Epelbaum, J. Escudero, A. Evans, F. Farina, K. Fargo, A. Fernández, R. Ferri, G. Frisoni, H. Hampel, M. G. Harrington, V. Jelic, J. Jeong, Y. Jiang, M. Kaminski, V. Kavcic, K. Kilborn, S. Kumar, A. Lam,

L. Lim, R. Lizio, D. Lopez, S. Lopez, B. Lucey, F. Maestú, W. J. McGeown, I. McKeith, D. V. Moretti, F. Nobili, G. Noce, J. Olichney, M. Onofrj, R. Osorio, M. Parra-Rodriguez, T. Rajji, P. Ritter, A. Soricelli, F. Stocchi, I. Tarnanas, J. P. Taylor, S. Teipel, F. Tucci, M. Valdes-Sosa, P. Valdes-Sosa, M. Weiergräber, G. Yener, and B. Guntekin. Measures of resting state eeg rhythms for clinical trials in alzheimer's disease: Recommendations of an expert panel. *Alzheimer's & Dementia*, 17(9):1528–1553, 2021.

[24] H. Bakhshayesh, S. P. Fitzgibbon, A. S. Janani, T. S. Grummett, and K. J. Pope. Detecting connectivity in eeg: A comparative study of data-driven effective connectivity measures. *Computers in Biology and Medicine*, 111:103329, 2019.

[25] C. Ballard, S. Gauthier, A. Corbett, C. Brayne, D. Aarsland, and E. Jones. Alzheimer's disease. *The Lancet*, 377(9770):1019–1031, 2011.

[26] D. S. Bassett, P. Zurn, and J. I. Gold. On the nature and use of models in network neuroscience. *Nature Reviews Neuroscience*, 19(9):566–578, Sep 2018.

[27] A.-K. Beck, G. Lütjens, K. Schwabe, R. Dengler, J. K. Krauss, and P. Sandmann. Thalamic and basal ganglia regions are involved in attentional processing of behaviorally significant events: evidence from simultaneous depth and scalp eeg. *Brain Structure and Function*, 223(1):461–474, Jan 2018.

[28] Q. Behfar, S. K. Behfar, B. von Reutern, N. Richter, J. Dronse, R. Fassbender, G. R. Fink, and O. A. Onur. Graph theory analysis reveals resting-state compensatory mechanisms in healthy aging and prodromal alzheimer's disease. *Frontiers in Aging Neuroscience*, 12, 2020.

[29] Y. Benjamini and Y. Hochberg. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1):289–300, 1995.

[30] D. Berg. Biomarkers for the early detection of parkinson's and Alzheimer's disease. *Neurodegenerative Diseases*, 5(3-4):133–136, 2008.

[31] M. Besserve, B. Schölkopf, N. K. Logothetis, and S. Panzeri. Causal relationships between frequency bands of extracellular signals in visual cortex revealed by an information theoretic analysis. *Journal of Computational Neuroscience*, 29(3): 547–566, Dec 2010.

[32] S. Billings. *Nonlinear System Identification: NARMAX METHODS IN THE TIME, FREQUENCY, AND SPATIO-TEMPORAL DOMAINS*, volume 13. John Wiley & Sons, Ltd, Chichester, UK, 2013.

[33] S. BILLINGS and O. BOAGHE. The response spectrum map, a frequency domain equivalent to the bifurcation diagram. *Int. J. Bifurc. Chaos*, 11(07): 1961–1975,, 2001.

[34] S. Billings and K. Mao. Model identification and assessment based on model predicted output, 1998.

[35] S. BILLINGS, M. Korenberg, and S. CHEN. Identification of non-linear output-affine systems using an orthogonal least-squares algorithm, 1987.

[36] S. Bitzer and C. K. I. Williams. Kick-starting GPLVM optimization via a connection to metric MDS. In *NIPS 2010 Workshop on Challenges of Data Visualization*, 2010.

[37] D. Blackburn, Y. Zhao, M. D. Marco, S. Bell, F. He, H.-L. Wei, S. Lawrence, Z. Unwin, M. Blyth, J. Angel, K. Baster, T. Farrow, I. Wilkinson, S. Billings, A. Venneri, and P. Sarrigiannis. A pilot study investigating a novel non-linear measure of eyes open versus eyes closed EEG synchronization in people with Alzheimer's disease and healthy controls. *Brain Sciences*, 8(7):134, jul 2018.

[38] O. Boaghe, S. Billings, L. Li, P. Fleming, and J. Liu. Time and frequency domain identification and analysis of a gas turbine engine. *Control Eng. Pract*, 10(12): 1347–1356,, 2002.

[39] T. Bossomaier, L. Barnett, M. Harré, and J. T. Lizier. *Transfer Entropy*, pages 65–95. Springer International Publishing, Cham, 2016. ISBN 978-3-319-43222-9.

[40] G. Box, G. Jenkins, G. Reinsel, and G. Ljung. *Time series analysis: forecasting and control*. John Wiley & Sons, Inc, Hoboken, New Jersey, 5th edition, 2015.

[41] M. Breakspear. Dynamic connectivity in neural systems. *Neuroinformatics*, 2 (2):205–224, Jun 2004.

[42] M. Breakspear. Dynamic models of large-scale brain activity. *Nature Neuroscience*, 20(3):340–352, Mar 2017.

[43] C. T. Briels, D. N. Schoonhoven, C. J. Stam, H. de Waal, P. Scheltens, and A. A. Gouw. Reproducibility of EEG functional connectivity in Alzheimer's disease. *Alzheimer's Research & Therapy*, 12(1):68, Jun 2020.

[44] S. L. Brunton, J. L. Proctor, and J. N. Kutz. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences*, 113(15):3932–3937, 2016.

[45] A. Buehlmann and G. Deco. Optimal information transfer in the cortex through synchronization. *PLOS Computational Biology*, 6(9):1–10, 09 2010.

[46] F. Buettner and F. J. Theis. A novel approach for resolving differences in single-cell gene expression patterns from zygote to blastocyst. *Bioinformatics*, 28(18): i626–i632, 09 2012.

[47] M. Bullmann, T. Fetzer, F. Ebner, F. Deinzer, and M. Grzegorzek. Fast kernel density estimation using gaussian filter approximation. In *2018 21st International Conference on Information Fusion (FUSION)*, pages 1233–1240, 2018.

[48] G. Buzsáki. *Rhythms of the Brain*. Oxford University Press, 10 2006. ISBN 9780195301069.

[49] J. Cao, Y. Zhao, X. Shan, H.-l. Wei, Y. Guo, L. Chen, J. A. Erkoyuncu, and P. G. Sarrigiannis. Brain functional and effective connectivity based on electroencephalography recordings: A review. *Human Brain Mapping*, 43(2):860–879, 2022.

[50] J. V. Carter, J. Pan, S. N. Rai, and S. Galandiuk. Roc-ing along: Evaluation and interpretation of receiver operating characteristic curves. *Surgery*, 159(6): 1638–1645, 2016.

[51] M. Chehelcheraghi, C. van Leeuwen, E. Steur, and C. Nakatani. A neural mass model of cross frequency coupling. *PLOS ONE*, 12(4):1–19, 04 2017.

[52] S. Chen and S. Billings. Representations of non-linear systems: the narmax model. *Int. J. Control*, 49(3):1013–1032,, 1989.

[53] S. Chen, S. Billings, and W. Luo. Orthogonal least squares methods and their application to non-linear system identification. *Int. J. Control*, 50(5):1873–1896,, 1989.

[54] D. Chicharro and A. Ledberg. Framework to study dynamic dependencies in networks of interacting processes. *Phys. Rev. E*, 86:041901, Oct 2012.

[55] N. Chiras, C. Evans, and D. Rees. Global nonlinear modeling of gas turbine dynamics using narmax structures. *J. Eng. Gas Turbines Power*, 124(4):817–826,, 2002.

[56] H. Choi and S. Choi. Robust kernel isomap. *Pattern Recognition*, 40(3):853–862, 2007. ISSN 0031-3203.

[57] M. X. Cohen. Assessing transient cross-frequency coupling in eeg data. *Journal of Neuroscience Methods*, 168(2):494–499, 2008.

[58] S. R. Cole, R. van der Meij, E. J. Peterson, C. de Hemptinne, P. A. Starr, and B. Voytek. Nonsinusoidal beta oscillations reflect cortical pathophysiology in parkinson's disease. *Journal of Neuroscience*, 37(18):4830–4840, 2017.

[59] R. A. Cooper and M. Ritchey. Cortico-hippocampal network connections support the multidimensional quality of episodic memory. *eLife*, 8:e45591, mar 2019.

[60] J. S. Damoiseaux and M. D. Greicius. Greater than the sum of its parts: a review of studies combining structural connectivity and resting-state functional connectivity. *Brain Structure and Function*, 213(6):525–533, Oct 2009.

[61] A. Das and V. Menon. Causal dynamics and information flow in parietal-temporal-hippocampal circuits during mental arithmetic revealed by high-temporal resolution human intracranial eeg. *Cortex*, 147:24–40, 2022.

[62] J. Dauwels, F. Vialatte, and A. Cichocki. Diagnosis of Alzheimer's disease from EEG signals: where are we standing? *Current Alzheimer Research*, 7(6):487–505, 2010.

[63] J. Dauwels, F. Vialatte, T. Musha, and A. Cichocki. A comparative study of synchrony measures for the early diagnosis of Alzheimer's disease based on EEG. *NeuroImage*, 49(1):668–693, 2010. ISSN 1053-8119.

[64] J. Dauwels, F.-B. Vialatte, and A. Cichocki. On the early diagnosis of Alzheimer's disease from EEG signals: A mini-review. In R. Wang and F. Gu, editors, *Advances in Cognitive Neurodynamics (II)*, pages 709–716, Dordrecht, 2011. Springer Netherlands.

[65] G. Deco, V. K. Jirsa, P. A. Robinson, M. Breakspear, and K. Friston. The dynamic brain: From spiking neurons to neural masses and cortical fields. *PLOS Computational Biology*, 4(8):1–35, 08 2008.

[66] B. Deng, L. Cai, S. Li, R. Wang, H. Yu, Y. Chen, and J. Wang. Multivariate multi-scale weighted permutation entropy analysis of EEG complexity for Alzheimer's disease. *Cognitive Neurodynamics*, 11(3):217–231, Jun 2017.

[67] N. Deschle, J. Ignacio Gossn, P. Tewarie, B. Schelter, and A. Daffertshofer. On the validity of neural mass models. *Frontiers in Computational Neuroscience*, 14, 2021.

[68] R. Dorf and R. Bishop. *Modern Control Systems*. Pearson Education, Inc, Upper Saddle River, New Jersey, 12th edition, 2011.

[69] T. Dupré la Tour, L. Tallot, L. Grabot, V. Doyère, V. van Wassenhove, Y. Grenier, and A. Gramfort. Non-linear auto-regressive models for cross-frequency coupling in neural time series. *PLOS Computational Biology*, 13(12):1–32, 12 2017.

[70] A. Ekhlasi, A. M. Nasrabadi, and M. R. Mohammadi. Direction of information flow between brain regions in adhd and healthy children based on eeg by using directed phase transfer entropy. *Cognitive Neurodynamics*, 15(6):975–986, Dec 2021.

[71] L. Faes and G. Nollo. Bivariate nonlinear prediction to quantify the strength of complex dynamical interactions in short-term cardiovascular variability. *Medical and Biological Engineering and Computing*, 44(5):383–392, May 2006.

[72] L. Faes, G. Nollo, and A. Porta. Compensated transfer entropy as a tool for reliably estimating information transfer in physiological time series. *Entropy*, 15 (1):198–219, 2013.

[73] T. H. Falk, F. J. Fraga, L. Trambaiolli, and R. Anghinah. EEG amplitude modulation analysis for semi-automated diagnosis of Alzheimer's disease. *EURASIP Journal on Advances in Signal Processing*, 2012(1):192, Aug 2012.

[74] M. Farina and L. Piroddi. An iterative algorithm for simulation error based identification of polynomial input–output models using multi-step prediction. *Int. J. Control*, 83(7):1442–1456,, 2010.

[75] G. W. Fenton. Electrophysiology of Alzheimer's disease. *British Medical Bulletin*, 42(1):29–33, 03 1986.

[76] P. Fleming and R. Purshouse. Evolutionary algorithms in control systems engineering: A survey. *Control Eng. Pract*, 10(11):1223–1241,, 2002.

[77] E. Florin and S. Baillet. The brain's resting-state activity is shaped by synchronized cross-frequency coupling of neural oscillations. *NeuroImage*, 111:26–35, 2015.

[78] A. Fornito, A. Arnatkevičiūtė, and B. D. Fulcher. Bridging the gap between connectome and transcriptome. *Trends in Cognitive Sciences*, 23(1):34–50, 2019.

[79] B. Foss and T. Johansen. Parallel nonlinear decoupling for process control — a narmax approach. *Annu. Rev. Autom. Program*, 17:49–54,, 1992.

[80] F. Fraga, T. Falk, P. Kanda, and R. Anghinah. Characterizing Alzheimer's disease severity via resting-awake EEG amplitude modulation analysis. *PloS one*, 8:e72240, 08 2013.

[81] M. J. Frank, B. Loughry, and R. C. O'Reilly. Interactions between frontal cortex and basal ganglia in working memory: A computational model. *Cognitive, Affective, & Behavioral Neuroscience*, 1(2):137–160, Jun 2001.

[82] C. A. Frantzidis, A. B. Vivas, A. Tsolaki, M. A. Klados, M. Tsolaki, and P. D. Bamidis. Functional disorganization of small-world brain networks in mild alzheimer's disease and amnestic mild cognitive impairment: an eeg study using relative wavelet entropy (rwe). *Frontiers in Aging Neuroscience*, 6, 2014.

[83] U. Friese, M. Köster, U. Hassler, U. Martens, N. Trujillo-Barreto, and T. Gruber. Successful memory encoding is associated with increased cross-frequency coupling between frontal theta and posterior gamma oscillations in human scalp-recorded eeg. *NeuroImage*, 66:642–647, 2013.

[84] K. J. Friston. Functional and effective connectivity: A review. *Brain Connectivity*, 1(1):13–36, 2011.

[85] E. Gallego-Jutglà, J. Solé-Casals, F.-B. Vialatte, J. Dauwels, and A. Cichocki. A theta-band eeg based index for early diagnosis of alzheimer's disease. *Journal of Alzheimer's Disease*, 43:1175–1184, 2015. 4.

[86] D. George. Continuous nonlinear systems, 1959.

[87] A. Ghosh, Y. Rho, A. R. McIntosh, R. Kötter, and V. K. Jirsa. Cortical network dynamics with time delays reveals functional connectivity in the resting brain. *Cognitive Neurodynamics*, 2(2):115–120, Jun 2008.

[88] N. Gour, O. Felician, M. Didic, L. Koric, C. Gueriot, V. Chanoine, S. Confort-Gouny, M. Guye, M. Ceccaldi, and J. P. Ranjeva. Functional connectivity changes differ in early and late-onset Alzheimer's disease. *Human Brain Mapping*, 35(7): 2978–2994, 2014.

[89] B. Gourévitch and J. J. Eggermont. Evaluating information transfer between auditory cortical neurons. *Journal of Neurophysiology*, 97(3):2533–2543, 2007.

[90] C. W. J. Granger. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3):424–438, 1969.

[91] T. D. Griffith and J. E. Hubbard. System identification methods for dynamic models of brain activity. *Biomedical Signal Processing and Control*, 68:102765, 2021.

[92] S. Gu, F. Pasqualetti, M. Cieslak, Q. K. Telesford, A. B. Yu, A. E. Kahn, J. D. Medaglia, J. M. Vettel, M. B. Miller, S. T. Grafton, and D. S. Bassett. Controllability of structural brain networks. *Nature Communications*, 6(1):8414, Oct 2015.

[93] R. Gunawardena, P. G. Sarrigiannis, D. J. Blackburn, and F. He. Kernel-based nonlinear manifold learning for eeg functional connectivity analysis with application to alzheimer's disease. In *2022 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, pages 1–5, 2022.

[94] R. Gunawardena, P. G. Sarrigiannis, D. J. Blackburn, and F. He. Kernel-based nonlinear manifold learning for eeg-based functional connectivity analysis and channel selection with application to alzheimer's disease. *Neuroscience*, 523:140–156, 2023.

[95] S. R. A. S. Gunawardena and Z. Q. Lang. An effective method for evaluation of the nonlinear output frequency response functions from system input-output data. In *2018 UKACC 12th International Conference on Control (CONTROL)*, pages 134–139, 2018.

[96] Y. Guo, L. Guo, S. Billings, and H.-L. Wei. An iterative orthogonal forward regression algorithm. *Int. J. Syst. Sci*, 46(5):776–789,, 2015.

[97] D. J. Harmah, C. Li, F. Li, Y. Liao, J. Wang, W. M. A. Ayedh, J. C. Bore, D. Yao, W. Dong, and P. Xu. Measuring the non-linear directed information flow in schizophrenia by multivariate transfer entropy. *Frontiers in Computational Neuroscience*, 13, 2020.

[98] F. Hasanzadeh, M. Mohebbi, and R. Rostami. Graph theory analysis of directed functional brain networks in major depressive disorder based on eeg signal. *Journal of Neural Engineering*, 17(2):026010, mar 2020.

[99] F. He and Y. Yang. Nonlinear system identification of neural systems from neurophysiological signals. *Neuroscience*, 458:213–228, 2021.

[100] F. He, S. A. Billings, H.-L. Wei, P. G. Sarrigiannis, and Y. Zhao. Spectral analysis for nonstationary and nonlinear systems: A discrete-time-model-based approach. *IEEE Transactions on Biomedical Engineering*, 60(8):2233–2241, aug 2013.

[101] F. He, H.-L. Wei, and S. A. Billings. Identification and frequency domain analysis of non-stationary and nonlinear systems using time-varying NARMAX models. *International Journal of Systems Science*, 46(11):2087–2100, nov 2013.

[102] F. He, S. A. Billings, H.-L. Wei, and P. G. Sarrigiannis. A nonlinear causality measure in the frequency domain: Nonlinear partial directed coherence with applications to EEG. *Journal of Neuroscience Methods*, 225:71–80, mar 2014.

[103] F. He, H.-L. Wei, S. A. Billings, and P. G. Sarrigiannis. A nonlinear generalization of spectral granger causality. *IEEE Transactions on Biomedical Engineering*, 61 (6):1693–1701, jun 2014.

[104] F. He, P. Sarrigiannis, S. Billings, H. Wei, J. Rowe, C. Romanowski, N. Hoggard, M. Hadjivassilliou, D. Rao, R. Grünewald, A. Khan, and J. Yianni. Nonlinear

interactions in the thalamocortical loop in essential tremor: A model-based frequency domain analysis. *Neuroscience*, 324:377–389, 2016. ISSN 0306-4522.

[105] J. Heer. Fast & accurate gaussian kernel density estimation. In *2021 IEEE Visualization Conference (VIS)*, pages 11–15, 2021.

[106] C. Heeyoul. Localization and regularization of normalized transfer entropy. *Neurocomputing*, 139:408–414, 2014.

[107] A. Hillebrand, P. Tewarie, E. van Dellen, M. Yu, E. W. S. Carbo, L. Douw, A. A. Gouw, E. C. W. van Straaten, and C. J. Stam. Direction of information flow in large-scale resting-state networks is frequency-dependent. *Proceedings of the National Academy of Sciences*, 113(14):3867–3872, 2016.

[108] J. Hlinka and M. Hadrava. On the danger of detecting network states in white noise. *Frontiers in Computational Neuroscience*, 9, 2015.

[109] X. Hong, P. Sharkey, and K. Warwick. Automatic nonlinear predictive model-construction algorithm using forward regression and the press statistic. *IEE Proceedings-Control Theory and Applications*, 150(3):245–254, 2003.

[110] A. Horvath, A. Szucs, G. Csukly, A. Sakovics, G. Stefanics, and A. Kamondi. EEG and ERP biomarkers of alzheimer's disease: a critical review. *Front Biosci (Landmark Ed)*, 23(2):183–220, Jan. 2018.

[111] M. Hossin and M. N. Sulaiman. A review on evaluation metrics for data classification evaluations. *International Journal of Data Mining & Knowledge Management Process*, 5:01–11, 2015.

[112] J. C. Houk and S. P. Wise. Feature article: Distributed modular architectures linking basal ganglia, cerebellum, and cerebral cortex: Their role in planning and controlling action. *Cerebral Cortex*, 5(2):95–110, 03 1995.

[113] A. J. Hughes, S. E. Daniel, L. Kilford, and A. J. Lees. Accuracy of clinical diagnosis of idiopathic parkinson's disease: a clinico-pathological study of 100 cases. *Journal of Neurology, Neurosurgery & Psychiatry*, 55(3):181–184, mar 1992.

[114] M. M. Husain and R. K. Garrett. Clinical diagnosis and management of alzheimer's disease. *Neuroimaging Clinics of North America*, 15(4):767–777, 2005.

[115] A. Hyafil, A.-L. Giraud, L. Fontolan, and B. Gutkin. Neural cross-frequency coupling: Connecting architectures, mechanisms, and functions. *Trends in Neurosciences*, 38(11):725–740, 2015.

[116] I. A. Illán, J. M. Górriz, J. Ramírez, E. W. Lang, D. Salas-Gonzalez, and C. G. Puntonet. Bilateral symmetry aspects in computer-aided alzheimer's disease diagnosis by single-photon emission-computed tomography imaging. *Artificial Intelligence in Medicine*, 56(3):191–198, 2012.

[117] N. Jackson, S. R. Cole, B. Voytek, and N. C. Swann. Characteristics of waveform shape in parkinson's disease detected with scalp electroencephalography. *eNeuro*, 6(3), 2019.

[118] W. Jacobs. Control-focused, nonlinear and time-varying modelling of dielectric elastomer actuators with frequency response analysis. *Smart Mater. Struct*, 24 (5):055002,, 2015.

[119] O. Jensen and L. L. Colgin. Cross-frequency coupling between neuronal oscillations. *Trends in Cognitive Sciences*, 11(7):267–269, 2007.

[120] O. Jensen, E. Spaak, and J. M. Zumer. Human brain oscillations: From physiological mechanisms to analysis and cognition. In S. Supek and C. J. Aine, editors, *Magnetoencephalography: From Signals to Dynamic Cortical Networks*, pages 359–403. Springer Berlin Heidelberg, Berlin, Heidelberg, 2014.

[121] J. Jeong. Eeg dynamics in patients with alzheimer's disease. *Clinical Neurophysiology*, 115(7):1490–1505, 2004.

[122] J. Jeong. EEG dynamics in patients with Alzheimer's disease. *Clinical Neurophysiology*, 115(7):1490–1505, jul 2004.

[123] Y. Jia, K. Kudo, L. B. N. Hinkley, M. Fisher, S. Vinogradov, S. Nagarajan, and K. Subramaniam. Abnormal Information Flow in Schizophrenia Is Linked to Psychosis. *Schizophrenia Bulletin*, 48(6):1384–1393, 09 2022.

[124] H. Jiang, A. Bahramisharif, M. A. van Gerven, and O. Jensen. Measuring directionality between neuronal oscillations of different frequencies. *NeuroImage*, 118: 359–367, 2015.

[125] B. Jiao, R. Li, H. Zhou, K. Qing, H. Liu, H. Pan, Y. Lei, W. Fu, X. Wang, X. Xiao, X. Liu, Q. Yang, X. Liao, Y. Zhou, L. Fang, Y. Dong, Y. Yang, H. Jiang, S. Huang, and L. Shen. Neural biomarker diagnosis and prediction to mild cognitive impairment and alzheimer's disease using eeg technology. *Alzheimer's Research & Therapy*, 15(1):32, Feb 2023.

[126] V. Jirsa and V. Müller. Cross-frequency coupling in real and virtual brain networks. *Frontiers in Computational Neuroscience*, 7, 2013.

[127] T. Joachims. Making large-scale SVM learning practical. In B. Schölkopf, C. Burges, and A. Smola, editors, *Advances in Kernel Methods - Support Vector Learning*, chapter 11, pages 169–184. MIT Press, Cambridge, MA, 1999.

[128] A. Jobbagy, E. Furnee, P. Harcos, and M. Tarczy. Early detection of parkinson's disease through automatic movement evaluation. *IEEE Engineering in Medicine and Biology Magazine*, 17(2):81–88, 1998.

[129] J. JONES and S. BILLINGS. Interpretation of non-linear frequency response functions. *Int. J. Control*, 52(2):319–346,, 1990.

[130] V. Jurcak, D. Tsuzuki, and I. Dan. 10/20, 10/10, and 10/5 systems revisited: Their validity as relative head-surface-based positioning systems. *NeuroImage*, 34(4):1600–1611, 2007.

[131] A. Kaiser and T. Schreiber. Information transfer in continuous processes. *Physica D: Nonlinear Phenomena*, 166(1):43–62, 2002.

[132] M. J. Kaminski and K. J. Blinowska. A new method of the description of the information flow in the brain structures. *Biological Cybernetics*, 65(3):203–210, Jul 1991.

[133] H. Kantz and T. Schreiber. *Nonlinear Time Series Analysis*. Cambridge University Press, 2 edition, 2003.

[134] D. L. Kauffman. *Systems one: An introduction to systems thinking*. Future Systems, Incorporated, 1980.

[135] D. Klepl, F. He, M. Wu, D. J. Blackburn, and P. Sarrigiannis. Eeg-based graph neural network classification of alzheimer's disease: An empirical evaluation of functional connectivity methods. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 30:2651–2660, 2022.

[136] D. Klepl, F. He, M. Wu, D. J. Blackburn, and P. G. Sarrigiannis. Cross-frequency multilayer network analysis with bispectrum-based functional connectivity: A study of alzheimer's disease. *Neuroscience*, 521:77–88, 2023.

[137] W. Koch, S. Teipel, S. Mueller, J. Benninghoff, M. Wagner, A. L. Bokde, H. Hampel, U. Coates, M. Reiser, and T. Meindl. Diagnostic power of default mode network resting state fmri in the detection of alzheimer's disease. *Neurobiology of Aging*, 33(3):466–478, 2012.

[138] T. Koenig, L. Prichep, T. Dierks, D. Hubl, L. Wahlund, E. John, and V. Jelic. Decreased eeg synchronization in alzheimer's disease and mild cognitive impairment. *Neurobiology of Aging*, 26(2):165–171, 2005.

[139] M. Korenberg, S. Billings, Y. Liu, and P. McIlroy. Orthogonal parameter estimation algorithm for non-linear stochastic systems. *Int. J. Control*, 48(1):193–210,, 1988.

[140] S. Kukreja, H. Galiana, and R. Kearney. Narmax representation and identification of ankle dynamics. *IEEE Trans. Biomed. Eng*, 50(1):70–81,, 2003.

[141] C. Lainscsek, M. E. Hernandez, J. Weyhenmeyer, T. J. Sejnowski, and H. Poizner. Non-linear dynamical analysis of EEG time series distinguishes patients with parkinson's disease from healthy individuals. *Frontiers in Neurology*, 4, 2013.

[142] G. Lancaster, D. Iatsenko, A. Pidde, V. Ticcinelli, and A. Stefanovska. Surrogate data for hypothesis testing of physical systems. *Physics Reports*, 748:1–60, 2018.

[143] I. D. Landau and G. Zito. *Digital control systems: design, identification and implementation*, volume 130. Springer, 2006.

[144] Z. Lang and S. Billings. Energy transfer properties of non-linear systems in the frequency domain. *International Journal of Control*, 78(5):345–362, 2005.

[145] Z.-Q. Lang and S. Billings. Output frequency characteristics of nonlinear systems. *Int. J. Control*, 64(6):1049–1067,, 1996.

[146] Z.-Q. Lang and S. Billings. Output frequencies of nonlinear systems. *Int. J. Control*, 67(5):713–730,, 1997.

[147] C. Laske, H. R. Sohrabi, S. M. Frost, K. L. de Ipiña, P. Garrard, M. Buscema, J. Dauwels, S. R. Soekadar, S. Mueller, C. Linnemann, S. A. Bridenbaugh, Y. Kanagasingam, R. N. Martins, and S. E. O'Bryant. Innovative diagnostic tools for early detection of Alzheimer's disease. *Alzheimer's & Dementia*, 11(5): 561–578, nov 2015.

[148] J. Laub, V. Roth, J. M. Buhmann, and K.-R. Müller. On the information and representation of non-euclidean pairwise data. *Pattern Recognition*, 39(10):1815–1826, 2006.

[149] T. O. Laumann, A. Z. Snyder, A. Mitra, E. M. Gordon, C. Gratton, B. Adeyemo, A. W. Gilmore, S. M. Nelson, J. J. Berg, D. J. Greene, J. E. McCarthy, E. Tagliazucchi, H. Laufs, B. L. Schlaggar, N. U. F. Dosenbach, and S. E. Petersen. On the Stability of BOLD fMRI Correlations. *Cerebral Cortex*, 27(10):4719–4732, 09 2016.

[150] N. Lawrence. Probabilistic non-linear principal component analysis with gaussian process latent variable models. *Journal of Machine Learning Research*, 6(60): 1783–1816, 2005.

[151] N. D. Lawrence. Gaussian process latent variable models for visualisation of high dimensional data. In *Proceedings of the 16th International Conference on Neural Information Processing Systems*, NIPS'03, page 329–336, Cambridge, MA, USA, 2003. MIT Press.

[152] N. D. Lawrence and J. Quiñonero Candela. Local distance preservation in the GPLVM through back constraints. In *Proceedings of the 23rd International Conference on Machine Learning*, ICML '06, page 513–520, New York, NY, USA, 2006. Association for Computing Machinery.

[153] J. Lee, S. Nemati, I. Silva, B. A. Edwards, J. P. Butler, and A. Malhotra. Transfer entropy estimation and directional coupling change detection in biomedical time series. *BioMedical Engineering OnLine*, 11(1):19, Apr 2012.

[154] N. Leonardi and D. Van De Ville. On spurious and real fluctuations of dynamic functional connectivity during rest. *NeuroImage*, 104:430–436, 2015.

[155] I. J. Leontaritis and S. A. Billings. Input-output parametric models for non-linear systems part i: deterministic non-linear systems. *Int. J. Control*, 41(2):303–328,, 1985.

[156] K. Li, J. Peng, and G. Irwin. A fast nonlinear model identification method. *IEEE Trans. Automat. Contr*, 50(8):1211–1216,, 2005.

[157] P. Li and S. Chen. A review on gaussian process latent variable models. *CAAI Transactions on Intelligence Technology*, 1(4):366–376, oct 2016.

[158] R. Li, T. Nguyen, T. Potter, and Y. Zhang. Dynamic cortical connectivity alterations associated with alzheimer's disease: An eeg and fnirs integration study. *NeuroImage: Clinical*, 21:101622, 2019.

[159] M. A. Lindquist, Y. Xu, M. B. Nebel, and B. S. Caffo. Evaluating dynamic bivariate correlations in resting-state fmri: A comparison study and a new approach. *NeuroImage*, 101:531–546, 2014.

[160] J. T. Lizier, M. Prokopenko, and A. Y. Zomaya. Local information transfer as a spatiotemporal filter for complex systems. *Phys. Rev. E*, 77:026110, Feb 2008.

[161] J. T. Lizier, J. Heinzle, A. Horstmann, J.-D. Haynes, and M. Prokopenko. Multivariate information-theoretic measures reveal directed information structure and task relevant changes in fmri connectivity. *Journal of Computational Neuroscience*, 30(1):85–107, Feb 2011.

[162] L. Ljung. *System Identification: Theory for the User*. Prentice Hall, 2nd edition, 1997.

[163] L. Ljung. *System identification*. Springer, 1998.

[164] L. Ljung. Perspectives on system identification. *Annual Reviews in Control*, 34 (1):1–12, 2010.

[165] M. Lobier, F. Siebenhühner, S. Palva, and J. M. Palva. Phase transfer entropy: A novel phase-based measure for directed connectivity in networks coupled by oscillatory interactions. *NeuroImage*, 85:853–872, 2014. New Horizons for Neural Oscillations.

[166] F. Lotte, L. Bougrain, A. Cichocki, M. Clerc, M. Congedo, A. Rakotomamonjy, and F. Yger. A review of classification algorithms for EEG-based brain–computer interfaces: a 10 year update. *Journal of Neural Engineering*, 15(3):031005, apr 2018.

[167] S. J. Luck. *An introduction to the event-related potential technique*. MIT press, 2014.

[168] C. Magri, K. Whittingstall, V. Singh, N. K. Logothetis, and S. Panzeri. A toolbox for the fast information analysis of multiple-site lfp, eeg and spike train recordings. *BMC Neuroscience*, 10(1):81, Jul 2009.

[169] H. B. Mann and D. R. Whitney. On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. *The Annals of Mathematical Statistics*, 18(1):50 – 60, 1947.

[170] A. D. Marimpis, S. I. Dimitriadis, and R. Goebel. Dyconnmap: Dynamic connectome mapping—a neuroimaging python module. *Human Brain Mapping*, 42 (15):4909–4939, 2021.

[171] G. Marsaglia, W. W. Tsang, and J. Wang. Evaluating kolmogorov's distribution. *Journal of Statistical Software*, 8(18):1–4, 2003.

[172] R. Marschinski and H. Kantz. Analysing the information flow between financial time series. *The European Physical Journal B - Condensed Matter and Complex Systems*, 30(2):275–281, Nov 2002.

[173] L. Marzetti, A. Basti, F. Chella, A. D'Andrea, J. Syrjälä, and V. Pizzella. Brain functional connectivity through phase coupling of neuronal oscillations: A perspective from magnetoencephalography. *Frontiers in Neuroscience*, 13, 2019.

[174] F. J. Massey. The kolmogorov-smirnov test for goodness of fit. *Journal of the American Statistical Association*, 46(253):68–78, 1951.

[175] A. G. d. G. Matthews, M. van der Wilk, T. Nickson, K. Fujii, A. Boukouvalas, P. León-Villagrá, Z. Ghahramani, and J. Hensman. GPflow: A Gaussian process library using TensorFlow. *Journal of Machine Learning Research*, 18(40):1–6, apr 2017.

[176] A. Maturana-Candelas, C. Gómez, J. Poza, N. Pinto, and R. Hornero. Eeg characterization of the alzheimer's disease continuum by means of multiscale entropies. *Entropy*, 21(6), 2019.

[177] J. McBride, X. Zhao, T. Nichols, V. Vagnini, N. Munro, D. Berry, and Y. Jiang. Scalp EEG-based discrimination of cognitive deficits after traumatic brain injury using event-related tsallis entropy analysis. *IEEE Transactions on Biomedical Engineering*, 60(1):90–96, 2013.

[178] J. McBride, X. Zhao, N. Munro, G. Jicha, C. Smith, and Y. Jiang. Discrimination of mild cognitive impairment and alzheimer's disease using transfer entropy measures of scalp eeg. *Journal of Healthcare Engineering*, 6:316746, 2015.

[179] J. C. McBride, X. Zhao, N. B. Munro, G. A. Jicha, F. A. Schmitt, R. J. Kryscio, C. D. Smith, and Y. Jiang. Sugihara causality analysis of scalp eeg for detection of early alzheimer's disease. *NeuroImage: Clinical*, 7:258–265, 2015.

[180] E. Mendes and S. Billings. On overparametrization of nonlinear discrete systems. *Int. J. Bifurc. Chaos*, 08(03):535–556,, 1998.

[181] E. M. A. M. Mendes and S. A. Billings. On overparametrization of nonlinear discrete systems. *International Journal of Bifurcation and Chaos*, 08(03):535–556, 1998.

[182] C. A. Micchelli, Y. Xu, and H. Zhang. Universal kernels. *J. Mach. Learn. Res.*, 7:2651–2667, 2006.

[183] L. H. Miller. Table of percentage points of kolmogorov statistics. *Journal of the American Statistical Association*, 51(273):111–121, 1956.

[184] L. A. Moctezuma and M. Molinas. EEG channel-selection method for epileptic-seizure classification based on multi-objective optimization. *Frontiers in Neuroscience*, 14, 2020.

[185] R. Mohanty, W. A. Sethares, V. A. Nair, and V. Prabhakaran. Rethinking measures of functional connectivity via feature extraction. *Scientific Reports*, 10 (1):1298, Jan 2020.

[186] R. Moran, K. Stephan, T. Seidenbecher, H.-C. Pape, R. Dolan, and K. Friston. Dynamic causal models of steady-state responses. *NeuroImage*, 44(3):796–811, 2009.

[187] S. Murakami and Y. Okada. Contributions of principal neocortical neurons to magnetoencephalography and electroencephalography signals. *The Journal of Physiology*, 575(3):925–936, 2006.

[188] A. Nakhnikian, S. Ito, L. Dwiel, L. Grasse, G. Rebec, L. Lauridsen, and J. Beggs. A novel cross-frequency coupling detection method using the generalized morse wavelets. *Journal of Neuroscience Methods*, 269:61–73, 2016.

[189] A. Nayfeh and D. Mook. *Nonlinear Oscillations*. Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts. Wiley, 1979. ISBN 9780471035558.

[190] I. Nelken, G. Chechik, T. D. Mrsic-Flogel, A. J. King, and J. W. H. Schnupp. Encoding stimulus information by spike numbers and mean response time in primary auditory cortex. *Journal of Computational Neuroscience*, 19(2):199–221, Oct 2005.

[191] O. Nelles. *Nonlinear System Identification*. Springer, Berlin, Heidelberg, 2001.

[192] S. Neufang, A. Akhrif, V. Riedl, H. Förstl, A. Kurz, C. Zimmer, C. Sorg, and A. Wohlschlaeger. Disconnection of frontal and parietal areas contributes to impaired attention in very early Alzheimer's disease. *Journal of Alzheimer's disease : JAD*, 25:309–21, 03 2011.

[193] J. M. Nichols, M. Seaver, S. T. Trickey, M. D. Todd, C. Olson, and L. Overbey. Detecting nonlinearity in structural systems using the transfer entropy. *Phys. Rev. E*, 72:046217, Oct 2005.

[194] A. Nordberg, J. O. Rinne, A. Kadir, and B. Långström. The use of pet in alzheimer disease. *Nature Reviews Neurology*, 6(2):78–87, Feb 2010.

[195] P. L. Nunez and R. Srinivasan. *Electric Fields of the Brain: The neurophysics of EEG*. Oxford University Press, 2 edition, 01 2006.

[196] A. Omidvarnia, G. Azemi, B. Boashash, J. M. O'Toole, P. B. Colditz, and S. Vanhatalo. Measuring time-varying information flow in scalp eeg signals: Orthogonalized partial directed coherence. *IEEE Transactions on Biomedical Engineering*, 61(3):680–693, 2014.

[197] G. C. O'Neill, P. Tewarie, D. Vidaurre, L. Liuzzi, M. W. Woolrich, and M. J. Brookes. Dynamics of large-scale electrophysiological networks: A technical review. *NeuroImage*, 180:559–576, 2018.

[198] D. Osipova, J. Ahveninen, S. Kaakkola, I. P. Jääskeläinen, J. Huttunen, and E. Pekkonen. Effects of scopolamine on meg spectral power and coherence in elderly subjects. *Clinical Neurophysiology*, 114(10):1902–1907, 2003.

[199] M. Paluš, V. Komárek, Z. c. v. Hrnčíř, and K. Štěrbová. Synchronization as adjustment of information rates: Detection from bivariate time series. *Phys. Rev. E*, 63:046211, Mar 2001.

[200] Y. Pan, J. Yang, T. Zhang, J. Wen, F. Pang, and Y. Luo. Characterization of the abnormal cortical effective connectivity in patients with sleep apnea hypopnea syndrome during sleep. *Computer Methods and Programs in Biomedicine*, 204: 106060, 2021.

[201] F. Parente and A. Colosimo. Modelling a multiplex brain network by local transfer entropy. *Scientific Reports*, 11(1):15525, Jul 2021.

[202] H.-J. Park, K. J. Friston, C. Pae, B. Park, and A. Razi. Dynamic effective connectivity in resting state fmri. *NeuroImage*, 180:594–608, 2018. Brain Connectivity Dynamics.

[203] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[204] Z. Peng, Z. Lang, and S. Billings. Resonances and resonant frequencies for a class of nonlinear systems. *J. Sound Vib*, 300(3–5):993–1014,, 2007.

[205] L. Pezard, R. Jech, and E. Růžička. Investigation of non-linear properties of multichannel EEG in the early stages of parkinson's disease. *Clinical Neurophysiology*, 112(1):38–45, 2001.

[206] E. Powers. Application of higher order spectral analysis to cubically nonlinear system identification. *IEEE Trans. Signal Process*, 42(7):1746–1765,, 1994.

[207] S. K. Pradhan and B. Subudhi. Narmax modeling of a two-link flexible robot. In *2011 Annual IEEE India Conference*, pages 1–5, 2011.

[208] D. Prichard and J. Theiler. Generating surrogate data for time series with several simultaneously measured variables. *Phys. Rev. Lett.*, 73:951–954, Aug 1994.

[209] T. Puoliväli, S. Palva, and J. M. Palva. Influence of multiple hypothesis testing on reproducibility in neuroimaging research: A simulation study and python-based software. *Journal of Neuroscience Methods*, 337:108654, 2020.

[210] A. M. T. Ramos and E. E. N. Macau. Minimum sample size for reliable causal inference using transfer entropy. *Entropy*, 19(4), 2017.

[211] C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning*, chapter 4, pages 79–104. The MIT Press, 2006. ISBN 0-262-18253-X. URL `http://gaussianprocess.org/gpml/`.

[212] R. Rastgoufard and A. Alsamman. GPLVM and lava floor distance for label-deficient semi-supervised learning: Case study. In *2016 19th International Conference on Information Fusion (FUSION)*, pages 84–90, 2016.

[213] P. A. Reuter-Lorenz, C. L. Grady, R. Cabeza, and N. A. Dennis. Frontal Lobes and Aging: Deterioration and Compensation. In *Principles of Frontal Lobe Function*. Oxford University Press, 02 2013. ISBN 9780199837755.

[214] F. Reza. *An Introduction to Information Theory*. Dover Books on Mathematics Series. Dover, 1994. ISBN 9780486682105.

[215] P. Risold, R. Thompson, and L. Swanson. The structural organization of connections between hypothalamus and cerebral cortex. *Brain Research Reviews*, 24 (2):197–254, 1997.

[216] R. C. Robinson. *Dynamical Systems: Stability, Symbolic Dynamics, and Chaos*. CRC press, 1998.

[217] R. C. Robinson. *An Introduction to Dynamical Systems: Continuous and Discrete*, volume 19. American Mathematical Soc., 2012.

[218] G. Rodriguez-Bermudez and P. J. Garcia-Laencina. Analysis of EEG signals using nonlinear dynamics and chaos: a review. *Applied mathematics & information sciences*, 9(5):2309, 2015.

[219] G. M. Rojas, C. Alvarez, C. E. Montoya, M. de la Iglesia-Vayá, J. E. Cisternas, and M. Gálvez. Study of resting-state functional connectivity networks using EEG electrodes position as seed. *Frontiers in Neuroscience*, 12, 2018.

[220] A. Rozo, J. Morales, J. Moeyersons, R. Joshi, E. G. Caiani, P. Borzée, B. Buyse, D. Testelmans, S. Van Huffel, and C. Varon. Benchmarking transfer entropy methods for the study of linear and nonlinear cardio-respiratory interactions. *Entropy*, 23(8), 2021.

[221] W. J. Rugh. *Nonlinear system theory The Volterra/Wiener Approach*. Johns Hopkins University Press Baltimore, 1981.

[222] S. Sadaghiani, M. J. Brookes, and S. Baillet. Connectomics of human electrophysiology. *NeuroImage*, 247:118788, 2022.

[223] V. Sakkalis. Review of advanced techniques for the estimation of brain connectivity measured with eeg/meg. *Computers in Biology and Medicine*, 41(12): 1110–1117, 2011. Special Issue on Techniques for Measuring Brain Connectivity.

[224] K. Sameshima and L. A. Baccalá. Using partial directed coherence to describe neuronal ensemble interactions. *Journal of Neuroscience Methods*, 94(1):93–103, 1999.

[225] S. Samiee and S. Baillet. Time-resolved phase-amplitude coupling in neural oscillations. *NeuroImage*, 159:270–279, 2017.

[226] B. Schölkopf. The kernel trick for distances. In *Proceedings of the 13th International Conference on Neural Information Processing Systems*, NIPS'00, page 283–289, Cambridge, MA, USA, 2000. MIT Press.

[227] J. Schoukens and L. Ljung. Nonlinear system identification: A user-oriented road map. *IEEE Control Systems Magazine*, 39(6):28–99, 2019.

[228] T. Schreiber. Measuring information transfer. *Phys. Rev. Lett.*, 85:461–464, Jul 2000.

[229] E. Schulz, M. Speekenbrink, and A. Krause. A tutorial on gaussian process regression: Modelling, exploring, and exploiting functions. *Journal of Mathematical Psychology*, 85:1–16, aug 2018.

[230] G. Schwarz. Estimating the dimension of a model. *The Annals of Statistics*, 6 (2):461–464, 1978.

[231] B. Schönheit, R. Zarski, and T. G. Ohm. Spatial and temporal relationships between plaques and tangles in alzheimer-pathology. *Neurobiology of Aging*, 25 (6):697–711, 2004.

[232] A. K. Seth, A. B. Barrett, and L. Barnett. Granger causality analysis in neuroscience and neuroimaging. *Journal of Neuroscience*, 35(8):3293–3297, 2015.

[233] A. J. Shackman, B. W. McMenamin, J. S. Maxwell, L. L. Greischar, and R. J. Davidson. Identifying robust and sensitive frequency bands for interrogating neural oscillations. *NeuroImage*, 51(4):1319–1333, 2010.

[234] G. Shan. Monte carlo cross-validation for a study with binary outcome and limited sample size. *BMC Medical Informatics and Decision Making*, 22(1):270, Oct 2022.

[235] C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27(3):379–423, 1948.

[236] J. Shao. Linear model selection by cross-validation. *Journal of the American Statistical Association*, 88(422):486–494, 1993.

[237] A. S. Shirkhorshidi, S. Aghabozorgi, and T. Y. Wah. A comparison study on similarity and dissimilarity measures in clustering continuous data. *PLOS ONE*, 10(12):1–20, 12 2015.

[238] E. Shumbayawonda, A. Fernández, M. P. Hughes, and D. Abásolo. Investigation of changes in causality throughout life—a magnetoencephalogram study using

granger causality and transfer entropy. In L. Lhotska, L. Sukupova, I. Lacković, and G. S. Ibbott, editors, *World Congress on Medical Physics and Biomedical Engineering 2018*, pages 233–236, Singapore, 2019. Springer Singapore.

[239] S. Simani, C. Fantuzzi, R. J. Patton, S. Simani, C. Fantuzzi, and R. J. Patton. *Model-based fault diagnosis techniques*. Springer, 2003.

[240] J. Sjöberg. Nonlinear black-box modeling in system identification: a unified overview. *Automatica*, 31(12):1691–1724,, 1995.

[241] R. Sneddon, W. R. Shankle, J. Hara, A. Rodriquez, D. Hoffman, and U. Saha. Eeg detection of early alzheimer's disease using psychophysical tasks. *Clinical EEG and Neuroscience*, 36(3):141–150, 2005.

[242] C. Stam. Nonlinear dynamical analysis of EEG and MEG: Review of an emerging field. *Clinical Neurophysiology*, 116(10):2266–2301, oct 2005.

[243] C. Stam. Use of magnetoencephalography (MEG) to study functional brain networks in neurodegenerative disorders. *Journal of the Neurological Sciences*, 289(1):128–134, 2010.

[244] C. Stam, B. Jelles, H. Achtereekte, S. Rombouts, J. Slaets, and R. Keunen. Investigation of EEG non-linearity in dementia and parkinson's disease. *Electroencephalography and Clinical Neurophysiology*, 95(5):309–317, nov 1995.

[245] C. Stam, B. Jones, I. Manshanden, A. van Cappellen van Walsum, T. Montez, J. Verbunt, J. de Munck, B. van Dijk, H. Berendse, and P. Scheltens. Magnetoencephalographic evaluation of resting-state functional connectivity in alzheimer's disease. *NeuroImage*, 32(3):1335–1344, 2006.

[246] K. J. Stam, D. L. J. Tavy, B. Jelles, H. A. M. Achtereekte, J. P. J. Slaets, and R. W. M. Keunen. Non-linear dynamical analysis of multichannel EEG: Clinical applications in dementia and parkinson's disease. *Brain Topography*, 7(2):141–150, dec 1994.

[247] K. E. Stephan. On the role of general system theory for functional neuroimaging. *Journal of Anatomy*, 205(6):443–470, 2004.

[248] F. Takens. Detecting strange attractors in turbulence. In D. Rand and L.-S. Young, editors, *Dynamical Systems and Turbulence, Warwick 1980*, pages 366–381, Berlin, Heidelberg, 1981. Springer Berlin Heidelberg. ISBN 978-3-540-38945-3.

[249] J. B. Tenenbaum, V. de Silva, and J. C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500):2319–2323, 2000.

[250] R. D. Terry, E. Masliah, D. P. Salmon, N. Butters, R. DeTeresa, R. Hill, L. A. Hansen, and R. Katzman. Physical basis of cognitive alterations in alzheimer's disease: Synapse loss is the major correlate of cognitive impairment. *Annals of Neurology*, 30(4):572–580, 1991.

[251] M. E. Tipping and C. M. Bishop. Probabilistic principal component analysis. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 61 (3):611–622, aug 1999.

[252] J. Touchon and K. Ritchie. Prodromal cognitive disorder in alzheimer's disease. *International Journal of Geriatric Psychiatry*, 14(7):556–563, 1999.

[253] L. R. Trambaiolli, A. C. Lorena, F. J. Fraga, P. A. Kanda, R. Anghinah, and R. Nitrini. Improving Alzheimer's disease diagnosis with machine learning techniques. *Clinical EEG and Neuroscience*, 42(3):160–165, 2011.

[254] J.-H. Tsai, W.-T. Hsu, L.-G. Lin, S.-M. Guo, and J. Tann. A modified narmax model-based self-tuner with fault tolerance for unknown nonlinear stochastic hybrid systems with an input–output direct feed-through term. *ISA Trans*, 53(1): 56–75,, 2014.

[255] A. Tsoli and O. C. Jenkins. Neighborhood denoising for learning high-dimensional grasping manifolds. In *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3680–3685, 2008.

[256] L. Tylová, J. Kukal, V. Hubata-Vacek, and O. Vyšata. Unbiased estimation of permutation entropy in EEG analysis for Alzheimer's disease classification. *Biomedical Signal Processing and Control*, 39:424–430, 2018.

[257] K. Tzimourta, T. Afrantou, P. Ioannidis, L. Astrakas, and N. Giannakeas. Analysis of electroencephalographic signals complexity regarding Alzheimer's disease. *Computers & Electrical Engineering*, 76:198–212, 04 2019.

[258] M. Ursino, G. Ricci, and E. Magosso. Transfer entropy as a measure of brain connectivity: A critical analysis with the help of neural mass models. *Frontiers in Computational Neuroscience*, 14, 2020.

[259] R. Urtasun, D. J. Fleet, A. Geiger, J. Popović, T. J. Darrell, and N. D. Lawrence. Topologically-constrained latent variable models. In *Proceedings of the 25th International Conference on Machine Learning*, ICML '08, page 1080–1087, New York, NY, USA, 2008. Association for Computing Machinery. ISBN 9781605582054.

[260] V. Vakorin, B. Misic, O. Krakovska, and A. McIntosh. Empirical and theoretical aspects of generation and transfer of information in a neuromagnetic source network. *Frontiers in Systems Neuroscience*, 5, 2011.

[261] V. A. Vakorin, O. A. Krakovska, and A. R. McIntosh. Confounding effects of indirect connections on causality estimation. *Journal of Neuroscience Methods*, 184(1):152–160, 2009.

[262] A.-M. van Cappellen van Walsum, Y. Pijnenburg, H. Berendse, B. van Dijk, D. Knol, P. Scheltens, and C. Stam. A neural complexity measure applied to meg data in alzheimer's disease. *Clinical Neurophysiology*, 114(6):1034–1040, 2003.

[263] F. Van de Steen, H. Almgren, A. Razi, K. Friston, and D. Marinazzo. Dynamic causal modelling of fluctuating connectivity in resting-state eeg. *NeuroImage*, 189:476–484, 2019.

[264] M. P. van den Heuvel and O. Sporns. A cross-disorder connectome landscape of brain dysconnectivity. *Nature Reviews Neuroscience*, 20(7):435–446, Jul 2019.

[265] R. Vicente and M. Wibral. *Efficient Estimation of Information Transfer*, pages 37–58. Springer Berlin Heidelberg, Berlin, Heidelberg, 2014. ISBN 978-3-642-54474-3.

[266] R. Vicente, M. Wibral, M. Lindner, and G. Pipa. Transfer entropy—a model-free measure of effective connectivity for the neurosciences. *Journal of Computational Neuroscience*, 30(1):45–67, Feb 2011.

[267] J. D. Victor. Binless strategies for estimation of information from neural data. *Phys. Rev. E*, 66:051903, Nov 2002.

[268] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020.

[269] J. Wang, Y. Fang, X. Wang, H. Yang, X. Yu, and H. Wang. Enhanced gamma activity and cross-frequency interaction of resting-state electroencephalographic oscillations in patients with alzheimer's disease. *Frontiers in Aging Neuroscience*, 9, 2017.

[270] L. Wang and W. R. Cluett. Use of press residuals in dynamic system identification. *Automatica*, 32(5):781–784, 1996.

[271] X. Wang, X. Cui, C. Ding, D. Li, C. Cheng, B. Wang, and J. Xiang. Deficit of cross-frequency integration in mild cognitive impairment and alzheimer's disease: A multilayer network approach. *Journal of Magnetic Resonance Imaging*, 53(5): 1387–1398, 2021.

[272] M. Wibral, B. Rahm, M. Rieder, M. Lindner, R. Vicente, and J. Kaiser. Transfer entropy in magnetoencephalographic data: Quantifying information flow in cortical and cerebellar networks. *Progress in Biophysics and Molecular Biology*, 105(1):80–97, 2011. BrainModes: The role of neuronal oscillations in health and disease.

[273] M. Wibral, R. Vicente, and M. Lindner. *Transfer Entropy in Neuroscience*, pages 3–36. Springer Berlin Heidelberg, Berlin, Heidelberg, 2014. ISBN 978-3-642-54474-3.

[274] N. Wiener. The theory of prediction. *Modern mathematics for engineers*, 1956.

[275] P. Wollstadt, M. Martínez-Zarzuela, R. Vicente, F. J. Díaz-Pernas, and M. Wibral. Efficient transfer entropy analysis of non-stationary neural time series. *PLOS ONE*, 9:1–21, 07 2014.

[276] Q.-S. Xu and Y.-Z. Liang. Monte carlo cross validation. *Chemometrics and Intelligent Laboratory Systems*, 56(1):1–11, 2001.

[277] Q.-S. Xu, Y.-Z. Liang, and Y.-P. Du. Monte carlo cross-validation for selecting a model and estimating the prediction error in multivariate calibration. *Journal of Chemometrics*, 18(2):112–120, 2004.

[278] B. T. T. Yeo, F. M. Krienen, J. Sepulcre, M. R. Sabuncu, D. Lashkari, M. Hollinshead, J. L. Roffman, J. W. Smoller, L. Zöllei, J. R. Polimeni, B. Fischl, H. Liu, and R. L. Buckner. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J. Neurophysiol*, 106(3):1125–1165, Jun 2011.

[279] R. Yuvaraj, M. Murugappan, U. R. Acharya, H. Adeli, N. M. Ibrahim, and E. Mesquita. Brain functional connectivity patterns for emotional state classification in parkinson's disease patients without dementia. *Behavioural Brain Research*, 298:248–260, 2016.

[280] P. Zerzucha and B. Walczak. Concept of (dis)similarity in data analysis. *TrAC Trends in Analytical Chemistry*, 38:116–128, 2012. ISSN 0165-9936.

[281] C. Zhang, C. Liu, X. Zhang, and G. Almpanidis. An up-to-date comparison of state-of-the-art classification algorithms. *Expert Systems with Applications*, 82: 128–150, 2017.

[282] H.-Y. Zhang, S.-J. Wang, B. Liu, Z.-L. Ma, M. Yang, Z.-J. Zhang, and G.-J. Teng. Resting brain connectivity: changes during the progress of Alzheimer disease. *Radiology*, 256(2):598–606, Aug. 2010.

[283] P. Zhao, P. Van-Eetvelt, C. Goh, N. Hudson, S. Wimalaratna, and E. Ifeachor. Characterization of eegs in alzheimer's disease using information theoretic methods. In *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 5127–5131, 2007.

[284] Y. Zhao, S. A. Billings, H. Wei, and P. G. Sarrigiannis. Tracking time-varying causality and directionality of information flow using an error reduction ratio test with applications to electroencephalography data. *Phys. Rev. E*, 86:051919, Nov 2012.

[285] Y. Zhao, S. A. Billings, H. Wei, F. He, and P. G. Sarrigiannis. A new narx-based granger linear and nonlinear casual influence detection method with applications to eeg data. *Journal of Neuroscience Methods*, 212(1):79–86, 2013.

[286] Y. Zhao, S. A. Billings, H. Wei, F. He, and P. G. Sarrigiannis. A new NARX-based granger linear and nonlinear casual influence detection method with applications to EEG data. *Journal of Neuroscience Methods*, 212(1):79–86, jan 2013.

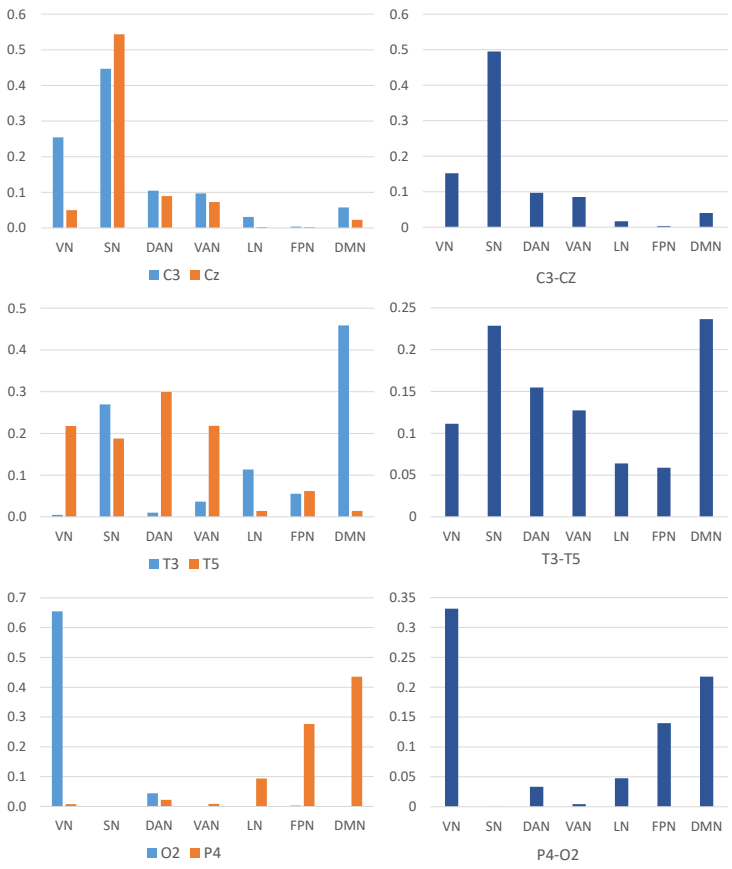
[287] G. Zito and I. Landau. Narmax model identification of a variable geometry turbocharged diesel engine. In *Proceedings of the 2005, American Control Conference*, page 1021–1026, 2005.

[288] M. H. Zweig and G. Campbell. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clinical Chemistry*, 39(4): 561–577, 04 1993.

Appendices

Appendix A

Examples: Evaluating the approximate similarity between a bipolar channel and the functional networks


0.6
0.5
0.4
0.3
0.2
0.1
0.0
VN SN DAN VAN LN FPN DMN
C3 Cz
C3-CZ
0.25
0.15
0.05
T3 T5
T3-T5
0.7
0.35
O2 P4
P4-O2