

CRANFIELD UNIVERSITY

JUN CAO

Estimation and visualisation of Brain Functional and Effective  
Connectivity

School of Aerospace, Transport and Manufacturing  
PhD in Manufacturing

PhD

Academic Year: 2020 - 2023

Supervisor: Professor Yifan Zhao  
Associate Supervisor: Professor John Ahmet Erkoyuncu  
February 2023

CRANFIELD UNIVERSITY

School of Aerospace, Transport and Manufacturing  
PhD in Manufacturing

PhD  
Academic Year: 2020 - 2023

JUN CAO

Estimation and visualisation of Brain Functional and Effective  
Connectivity

Supervisor: Professor Yifan Zhao  
Associate Supervisor: Professor John Ahmet Erkoyuncu  
February 2023

This thesis is submitted in partial fulfilment of the requirements for  
the degree of PhD in Manufacturing.

© Cranfield University 2023. All rights reserved. No part of this  
publication may be reproduced without the written permission of the  
copyright owner.

## ABSTRACT

Functional and effective connectivity are two important concepts in the field of neuroscience that describe how different regions of the brain communicate and work together to support various cognitive and behavioural functions. Despite the many advances in functional and effective connectivity research, there are still several important research gaps that need to be addressed. This thesis explores the novel estimation and visualisation of brain functional and effective connectivity using electroencephalography recordings, with a particular focus on its potential impact on the diagnosis and monitoring of neurological disorders. This thesis proposes two novel methods for estimating brain functional connectivity and effective connectivity. The first method, Revised Hilbert-Huang Transformation, outperforms wavelet-based methods in terms of promising features and time-frequency resolution, providing a potential biomarker and diagnostic tool for Alzheimer's disease. The second method, causality detection attention-based convolutional neural networks, effectively estimates effective connectivity networks and identifies disrupted connectivity in Alzheimer's disease patients. These methods contribute to the growing literature on connectivity estimation and offer valuable insights into the neural mechanisms underlying cognitive processes and neurodegenerative diseases, providing potential diagnostic and monitoring tools for healthcare professionals. This thesis also introduces a novel directed structure learning GNN (DSL-GNN) to leverage several EBC estimations to extract discriminative biomarkers for dementia classification. In studies of Alzheimer's disease, epilepsy, Parkinson's disease, and workload classification, the thesis demonstrates that the proposed brain connectivity methods have better performance compared with traditional methods based on individual channel. It suggests that functional and effective connectivity may track more changes from healthy people to patients to a certain extent, providing the possibility for earlier and more accurate diagnoses. Specifically, the thesis finds that specific regions of the brain can contribute to the diagnosis of epilepsy and dementia disease as well as workload classification based on brain connectivity. By advising the appropriate placement of electroencephalography sensors based on these identified

regions, doctors and researchers can more efficiently and accurately diagnose and classify these neurological disorders, reducing the burden on healthcare systems.

Keywords:

Electroencephalography (EEG), dementia, workload, machine learning, biomarker, classification

## ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to all those who have contributed to the completion of this thesis.

First and foremost, I would like to thank my supervisors, Professor Yifan Zhao and Professor John Ahmet Erkoyuncu, for their unwavering support, invaluable guidance, and insightful comments throughout the entire research process. Their dedication and encouragement have been instrumental in shaping my research ideas and helping me develop the necessary skills to conduct high-quality research. They have challenged me to think critically, provided timely feedback on my work, and offered invaluable advice on how to navigate the complex landscape of academia. Their guidance and mentorship have been invaluable to me, and I am truly grateful for their support.

I would like to thank my colleagues at Cranfield University, who have provided me with a stimulating academic environment and helped me in various ways, especially Dr Robin Du, Dr Lichao Yang, Dr Youdao Wang, Dr Xiaocai Shan, Dr Feiyang He and Kailun Deng. Their encouragement, support, and collaboration have made my time at Cranfield University a truly rewarding experience, and I am grateful for their friendship and camaraderie.

I also want to express my deep appreciation to my family and friends for their unconditional love, support, and encouragement during my PhD journey. Their belief in me has been a constant source of motivation and inspiration. I would like to thank my parents for their unwavering support and encouragement, and for instilling in me the values of hard work and perseverance. I would also like to thank my partner, Taojie Wang, for her unwavering love and support, and for being a constant source of motivation and encouragement throughout my PhD journey.

Thank you all for your contributions to this thesis and for being part of this journey. I am truly grateful for your support and encouragement, and I look forward to continuing our collaborations in the future.

# TABLE OF CONTENTS

|                                                                                                                           |      |
|---------------------------------------------------------------------------------------------------------------------------|------|
| ABSTRACT .....                                                                                                            | i    |
| ACKNOWLEDGEMENTS.....                                                                                                     | iii  |
| LIST OF FIGURES.....                                                                                                      | viii |
| LIST OF TABLES .....                                                                                                      | xvii |
| LIST OF ABBREVIATIONS.....                                                                                                | xix  |
| 1 Introduction.....                                                                                                       | 1    |
| 1.1 Background.....                                                                                                       | 1    |
| 1.2 Research Gaps.....                                                                                                    | 3    |
| 1.3 Aim and Objectives .....                                                                                              | 4    |
| 1.4 Research Methodology adopted to achieve the stated objectives. ....                                                   | 6    |
| 1.5 List of journal papers.....                                                                                           | 8    |
| 1.6 Reference .....                                                                                                       | 9    |
| 2 Literature Review .....                                                                                                 | 11   |
| 2.1 Introduction .....                                                                                                    | 11   |
| 2.2 Estimation .....                                                                                                      | 15   |
| 2.2.1 Non-parametric methods.....                                                                                         | 17   |
| 2.2.2 Parametric methods .....                                                                                            | 18   |
| 2.3 Visualisation of brain connectivity .....                                                                             | 21   |
| 2.4 Applications and Machine Learning .....                                                                               | 23   |
| 2.4.1 Practical application .....                                                                                         | 23   |
| 2.4.2 Hand-crafted vs. Imaging features .....                                                                             | 24   |
| 2.5 Discussion .....                                                                                                      | 25   |
| 2.5.1 Connectivity estimation .....                                                                                       | 25   |
| 2.5.2 Functional Connectivity vs. Effective Connectivity .....                                                            | 26   |
| 2.5.3 Phase-based Connectivity vs. Amplitude-based Connectivity.....                                                      | 26   |
| 2.5.4 Precautions during analysis .....                                                                                   | 27   |
| 2.5.5 Role of Visualisation.....                                                                                          | 27   |
| 2.5.6 Research Gaps and Future Direction.....                                                                             | 28   |
| 2.6 Conclusions .....                                                                                                     | 31   |
| 2.7 Appendix The conceptual and mathematical details for calculating<br>brain connectivity.....                           | 32   |
| 2.7.1 Functional Connectivity .....                                                                                       | 32   |
| 2.7.2 Effective Connectivity .....                                                                                        | 38   |
| 2.8 Reference .....                                                                                                       | 46   |
| 3 Time-frequency Functional Connectivity Estimation.....                                                                  | 76   |
| 3.1 A Revised Hilbert-Huang Transformation to Track Non-stationary<br>Association of Electroencephalography Signals ..... | 76   |
| 3.1.1 Introduction .....                                                                                                  | 76   |
| 3.1.2 Method .....                                                                                                        | 80   |
| 3.1.3 Simulations.....                                                                                                    | 87   |

|                                                                                                                                                                    |     |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 3.1.4 Application To EEG Data .....                                                                                                                                | 96  |
| 3.1.5 Conclusion .....                                                                                                                                             | 104 |
| 3.1.6 Supplementary .....                                                                                                                                          | 105 |
| 3.1.7 References .....                                                                                                                                             | 111 |
| 3.2 Ultra-high-resolution time-frequency analysis of EEG to characterise<br>brain functional connectivity with the application in Alzheimer's disease. ...         | 116 |
| 3.2.1 Introduction .....                                                                                                                                           | 116 |
| 3.2.2 Materials and Methodology .....                                                                                                                              | 121 |
| 3.2.3 Results .....                                                                                                                                                | 133 |
| 3.2.4 Discussion .....                                                                                                                                             | 140 |
| 3.2.5 Conclusion .....                                                                                                                                             | 145 |
| 3.2.6 Supplementary .....                                                                                                                                          | 146 |
| 3.2.7 References .....                                                                                                                                             | 148 |
| 4 The Application of Functional Connectivity on Brain Disease Classification                                                                                       | 160 |
| 4.1 Using interictal seizure-free EEG data to recognise patients with<br>epilepsy based on machine learning of brain functional connectivity.....                  | 160 |
| 4.1.1 INTRODUCTION.....                                                                                                                                            | 160 |
| 4.1.2 METHODOLOGY .....                                                                                                                                            | 166 |
| 4.1.3 RESULTS.....                                                                                                                                                 | 175 |
| 4.1.4 DISCUSSIONS .....                                                                                                                                            | 182 |
| 4.1.5 CONCLUSIONS.....                                                                                                                                             | 184 |
| 4.1.6 Supplementary .....                                                                                                                                          | 186 |
| 4.1.7 REFERENCES.....                                                                                                                                              | 198 |
| 4.2 Spatial – temporal graph convolutional network for Alzheimer<br>classification based on brain functional connectivity imaging of<br>electroencephalogram. .... | 205 |
| 4.2.1 Introduction .....                                                                                                                                           | 205 |
| 4.2.2 Methods .....                                                                                                                                                | 210 |
| 4.2.3 Results .....                                                                                                                                                | 219 |
| 4.2.4 Conclusions.....                                                                                                                                             | 226 |
| 4.2.5 Supplementary .....                                                                                                                                          | 230 |
| 4.2.6 References.....                                                                                                                                              | 232 |
| 5 The Application of Functional Connectivity on Mental workload<br>Classification .....                                                                            | 240 |
| 5.1 Introduction .....                                                                                                                                             | 240 |
| 5.2 Materials and Methods.....                                                                                                                                     | 243 |
| 5.2.1 Dataset.....                                                                                                                                                 | 244 |
| 5.2.2 Signal pre-processing and Feature extraction.....                                                                                                            | 245 |
| 5.2.3 Feature Selection and Fusion .....                                                                                                                           | 248 |
| 5.2.4 Machine-Learning Classification .....                                                                                                                        | 249 |
| 5.3 Results.....                                                                                                                                                   | 250 |
| 5.3.1 Time Interval Selection.....                                                                                                                                 | 250 |

|                                                                                                                   |     |
|-------------------------------------------------------------------------------------------------------------------|-----|
| 5.3.2 Machine-Learning Classification Performance .....                                                           | 251 |
| 5.3.3 Visualisation .....                                                                                         | 255 |
| 5.4 Discussion .....                                                                                              | 260 |
| 5.4.1 EEG vs. fNIRS .....                                                                                         | 261 |
| 5.4.2 Univariate vs. Multivariate Features .....                                                                  | 262 |
| 5.4.3 Independent vs. Hybrid Feature .....                                                                        | 262 |
| 5.5 Conclusions .....                                                                                             | 263 |
| 5.6 Supplementary.....                                                                                            | 264 |
| 5.7 References .....                                                                                              | 266 |
| 6 Effective Connectivity Estimation Using Attention-based Convolutional<br>Neural Networks.....                   | 273 |
| 6.1 Introduction .....                                                                                            | 273 |
| 6.2 Methods.....                                                                                                  | 276 |
| 6.2.1 Causality Detection Attention-based Convolutional Neural<br>Networks (CD-ABCNN).....                        | 276 |
| 6.2.2 MVAR-model Simulation .....                                                                                 | 279 |
| 6.2.3 Real EEG data .....                                                                                         | 280 |
| 6.3 Results.....                                                                                                  | 282 |
| 6.3.1 Study 1: Simulated Data.....                                                                                | 282 |
| 6.3.2 Study 2: AD and HC EEG .....                                                                                | 283 |
| 6.4 Discussion .....                                                                                              | 289 |
| 6.5 References .....                                                                                              | 291 |
| 7 The Application of Effective Connectivity on Dementia Classification<br>Supported by Graph Neural Networks..... | 298 |
| 7.1 Introduction .....                                                                                            | 298 |
| 7.2 Methods and materials.....                                                                                    | 302 |
| 7.2.1 Framework .....                                                                                             | 303 |
| 7.2.2 Pre-processing .....                                                                                        | 303 |
| 7.2.3 Directed Structure Learning Graph Neural Network.....                                                       | 303 |
| 7.2.4 Effective brain connectivity estimation.....                                                                | 307 |
| 7.2.5 Dataset.....                                                                                                | 311 |
| 7.2.6 Classification .....                                                                                        | 312 |
| 7.3 Results.....                                                                                                  | 314 |
| 7.3.1 Classification Performance.....                                                                             | 314 |
| 7.3.2 Band Contribution .....                                                                                     | 315 |
| 7.3.3 Feature Visualisation.....                                                                                  | 317 |
| 7.4 Discussion .....                                                                                              | 319 |
| 7.5 Conclusion .....                                                                                              | 323 |
| 7.6 References .....                                                                                              | 323 |
| 8 Conclusion.....                                                                                                 | 333 |
| 8.1 Contribution to Knowledge.....                                                                                | 333 |
| 8.2 Impact.....                                                                                                   | 337 |



|                                     |     |
|-------------------------------------|-----|
| 8.3 Limitation and Future Work..... | 338 |
| REFERENCES.....                     | 341 |

## LIST OF FIGURES

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |    |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figure 1-1 The structure of PhD Thesis .....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 5  |
| Figure 2-1 Number of Publications per year from PubMed search using keywords "EEG and Machine learning" or "EEG and AI" (Trend1) and "brain connectivity and EEG" (Trend2) in the period 2005–2020 .....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | 12 |
| Figure 2-2 Examples of the visualisation of brain connectivity (A) Structural and functional networks are topologically similar. Examples of structural and functional adjacency matrices from one patient. Similarity between structural network architecture and cross-correlation and coherence functional networks is visually evident.(Chu et al., 2015) (B) Contribution of non-normalized directed transfer function (NDTF) pairs variables to PC1 in terms of Principal component analysis (PCA) loadings. Only these NDTF pairs which showed statistical differences between Nold and AD groups on the level $p < 0.0005$ contributed. It provides information about the importance of a particular parameter in the classification procedures (Blinowska et al., 2017) (C) Constructed functional connectivity map (The threshold of is applied as the connection strength) for subject diagnosed with left frontal region epilepsy and subject diagnosed with generalized epilepsy. (Sargolzaei et al., 2015) (D) Effective brain networks (averaged over all participants) for responses to different emotional music (Shahabi and Moghimi, 2016).(E) A revised Circular Graph plot overlaid with EEG electrode locations to highlight the real electrode locations and their corresponding locations in the plot. (Zhao et al., 2020) ..... | 20 |
| Figure 3-1 The flow chart of the proposed framework for tracking non-stationary association based on Revised Hilvert-Huang Transform (RHHT).....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | 81 |
| Figure 3-2 Example 1 (a) Instantaneous frequency of Y1, Y2 and their components; b) RHHT of Y1, Y2 and associated IMFs; (c) CWT of Y1, Y2 and associated bands.....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | 90 |
| Figure 3-3 Example 1 (a) RHHT cross-spectrum (CS) of Y1, Y2, and IMFs before and after significant test; (b) CWT cross-spectrum of Y1, Y2, and associated bands before and after significant test.....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | 91 |
| Figure 3-4 Example 2 (a) Instantaneous frequency of Y1, Y2 and their components; (b) RHHT of Y1, Y2 and associated IMFs; (c) CWT of Y1, Y2 and associated bands.....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | 93 |
| Figure 3-5 Example 2 (a) RHHT cross-spectrum (CS) of Y1, Y2 and IMFs before and after significant test; (b) CWT cross-spectrum of Y1, Y2, and bands before and after significant test. ....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | 94 |
| Figure 3-6 (a) RHHT of Y1 and Y2 and the cross-spectrum significance result with 95% confidence interval (b) CWT of Y1 and Y2 and the cross-spectrum significance result with 95% confidence interval.....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 97 |

|                                                                                                                                                                                                                                                                                                                                                                                                  |     |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Figure 3-7 The boxplots of top 10 smallest p-values of all 207 pairs for each frequency band of EC (a) and EO (b) data by RHHT and CWT cross-spectrum.....                                                                                                                                                                                                                                       | 98  |
| Figure 3-8 The boxplots of top 10 highest classification accuracies of all 207 pairs for each frequency band of EC (a) and EO (b) data by RHHT and CWT cross-spectrum.....                                                                                                                                                                                                                       | 103 |
| Figure 3-9 Rows from top to bottom show the results for EO data calculated by the frequency bands of 1-50Hz, Gamma, Beta, Alpha, Theta, and Delta respectively. The left two columns (EC data) and the right two columns (EO data) show the p-values of each cross-bipolar channel based on the cross-spectrums by RHHT and CWT respectively. ....                                               | 107 |
| Figure 3-10 The results of synthetic Examples 1 with SNR being 30dB.....                                                                                                                                                                                                                                                                                                                         | 108 |
| Figure 3-11 The results of synthetic Examples 2 with SNR being 30dB.....                                                                                                                                                                                                                                                                                                                         | 108 |
| Figure 3-12 The results of synthetic Examples 1 with SNR being 20dB.....                                                                                                                                                                                                                                                                                                                         | 109 |
| Figure 3-13 The results of synthetic Examples 2 with SNR being 20dB.....                                                                                                                                                                                                                                                                                                                         | 109 |
| Figure 3-14 The results of synthetic Examples 1 with SNR being 20dB with filtering.....                                                                                                                                                                                                                                                                                                          | 110 |
| Figure 3-15 The results of synthetic Examples 2 with SNR being 20dB with filtering.....                                                                                                                                                                                                                                                                                                          | 110 |
| Figure 3-16 (a) Bipolar channel locations. The numbers 1, 2, 3, 4, 5, 6 and 7 represent the left frontal, right frontal, left temporal, right temporal, parietal, left occipital and right occipital regions, respectively (b) An example of 23-bipolar-channel EEG recordings.....                                                                                                              | 122 |
| Figure 3-17 207 possible channel pair combinations and 46 discarded pairs marked with X (symmetry) .....                                                                                                                                                                                                                                                                                         | 123 |
| Figure 3-18 Flow Diagram of the proposed framework .....                                                                                                                                                                                                                                                                                                                                         | 124 |
| Figure 3-19 flowchart of CEEMDAN algorism .....                                                                                                                                                                                                                                                                                                                                                  | 128 |
| Figure 3-20 Top ten classification accuracies in EO (A, B), EC (C, D) and EO&EC (E, F) using features from Wavelet (left column) and RHHT cross-spectrum (right column) .....                                                                                                                                                                                                                    | 134 |
| Figure 3-21 Time-frequency brain connectivity analysis of channel pair Fz-Cz:O1-O2. (a) Wavelet cross-spectrum for a HC participant. (b) Wavelet cross-spectrum for an AD participant. (c) RHHT cross-spectrum for a HC participant. (d) RHHT cross-spectrum for an AD participant. (e) A violin plot to present peak frequency values for channel pair Fz-Cz:O1-O2 ( $p < 0.0001$ , ANOVA)..... | 136 |
| Figure 3-22 Topographic plots representing Full-band connnectivity PFoCS for each participant (Left) and average (Right) extrated by Wavelet and RHHT                                                                                                                                                                                                                                            |     |

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |     |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| techniques , by using each available derivation against the occipital O1-O2 channel.....                                                                                                                                                                                                                                                                                                                                                                                                                                 | 138 |
| Figure 3-23 Topographic map of the classification accuracy using PSD.....                                                                                                                                                                                                                                                                                                                                                                                                                                                | 139 |
| Figure 3-24 Comparison of the RHHT and previous brain connectivity methods.<br>* Mann-Whitney Test, $p < 0.001$ , it means significant outperformance in a band.....                                                                                                                                                                                                                                                                                                                                                     | 141 |
| Figure 3-25 An example of the resulting Revised Hilbert-Huang Transformation (RHHT) based on complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) from the EEG recordings of channel Fz-Cz:O1-O2.....                                                                                                                                                                                                                                                                                            | 146 |
| Figure 3-26 An example of the resulting Revised Hilbert-Huang Transformation (RHHT) cross-spectrum the EEG recordings of channel Fz-Cz:O1-O2 ...                                                                                                                                                                                                                                                                                                                                                                         | 147 |
| Figure 3-27 Topographic plots representing mean coherence connectivity for each participant (Left) and average (Right) extracted by Wavelet techniques, by using each available derivation against the parietal Cz-Pz channel.....                                                                                                                                                                                                                                                                                       | 148 |
| Figure 4-1 A typical EEG recording before during and after a generalised epileptic seizure (typical absence) recorded from the left centrottemporal area. The ictal segment (i.e., seizure)) contains large magnitude epileptic discharges, which can be easily distinguished from the seizure-free recording (interictal).....                                                                                                                                                                                          | 162 |
| Figure 4-2 EEG recording from one of the patients of the epileptic cohort showing an interictal epileptiform discharge (IED). This demonstrates a clear generalised distribution in keeping with the patient's diagnosis of generalised epilepsy. Of note, the EEG is "normal-looking", i.e. Empirically classified as normal, prior to the generalised paroxysmal discharge. "Normal-looking" EEG epochs were the dominant feature on this interictal EEG recording, a common occurrence in this group of patients..... | 164 |
| Figure 4-3 Flow Diagram for EG vs. HC and EG vs. NEAD classification .....                                                                                                                                                                                                                                                                                                                                                                                                                                               | 167 |
| Figure 4-4 The box charts of the classification accuracy of HC vs EG against various frequency bands for the 9 selected measures, where only the top 10 pairs in terms of classification accuracy were considered. Each box shows highest, lowest, and median accuracy.....                                                                                                                                                                                                                                              | 176 |
| Figure 4-5 The top 10 bipolar pairs based on the percentage of appearance of top 3 classification performance between HC and EG for different connectivity measures and bands; (a) EO; (b) EC.....                                                                                                                                                                                                                                                                                                                       | 177 |
| Figure 4-6 The heatmap for the HC vs EG classification performance based on CohMean in Beta band for the EC state. ....                                                                                                                                                                                                                                                                                                                                                                                                  | 178 |

|                                                                                                                                                                                                                                                                                                                                   |     |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Figure 4-8 The distribution of mean coherence values in the Beta band of 4 channel pairs corresponding to the highest classification accuracy (>90%) for HC vs EG in the EC state. ....                                                                                                                                           | 179 |
| Figure 4-7 (a) The locations of 6 channel pairs corresponding to the highest classification accuracy (>90%) for the HC vs EG based on CohMean in the Beta band in the EC state (Red area is better than Yellow area); (b) The ROC and AUC of the 6 pairs. ....                                                                    | 179 |
| Figure 4-9 The heatmap for the NEAD vs EG classification performance based on MI in Delta band for the EO state. ....                                                                                                                                                                                                             | 180 |
| Figure 4-10 The box charts of the classification accuracy of NEAD and EG against various frequency bands for the 9 selected measures, where only the top 10 pairs in terms of classification accuracy were considered. Each box shows highest, lowest, and median accuracy. ....                                                  | 181 |
| Figure 4-11 The distribution of MI values in Delta band of two channel pairs corresponding highest classification accuracy (>70%) for the NEAD vs EG for the EO state. ....                                                                                                                                                       | 182 |
| Figure 4-12 The spatial distribution of classification accuracy using the univariate method, where the same colormap scale is used. ....                                                                                                                                                                                          | 183 |
| Figure 4-13 The distribution of Mean coherence in the Beta band of two channel pairs with different p-value ....                                                                                                                                                                                                                  | 187 |
| Figure 4-14 The best classification accuracy of HC vs EG using F4C4-FZCZ for each connectivity feature, frequency band and eye state. ....                                                                                                                                                                                        | 189 |
| Figure 4-15 The best classification accuracy of NEAD vs EG for each connectivity feature, frequency band and eye state. ....                                                                                                                                                                                                      | 190 |
| Figure 4-16 The evaluation of best classifier using features providing highest accuracy classification between NEAD and EG (a) SVM (b) optimized SVM (c) KNN, k=15 (d) decision tree ....                                                                                                                                         | 191 |
| Figure 4-17 The evaluation of the best K value of KNN based on ROC and AUC ....                                                                                                                                                                                                                                                   | 192 |
| Figure 4-18 The classification results using two features for EG vs HC. ....                                                                                                                                                                                                                                                      | 192 |
| Figure 4-19 Spatial distribution of frequency response of each channel for Delta band, where the subjects are divided into 5 groups: EG with medication (row 1), EG without medication (row 2), NEAD with medication (row 3), NEAD without medication (row 4), HC (row 5). The same colormap scale is used across all plots. .... | 193 |
| Figure 4-20 Spatial distribution of frequency response of each channel for Theta band, where the subjects are divided into 5 groups: EG with medication (row 1), EG without medication (row 2), NEAD with medication (row 3), NEAD without medication (row 4), HC (row 5). The same colormap scale is used across all plots ....  | 194 |

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |     |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Figure 4-21 Spatial distribution of frequency response of each channel for Alpha band, where the subjects are divided into 5 groups: EG with medication (row 1), EG without medication (row 2), NEAD with medication (row 3), NEAD without medication (row 4), HC (row 5). The same colormap scale is used across all plots. ....                                                                                                                                                                                                                                                                                                                                                                                                                         | 195 |
| Figure 4-22 Spatial distribution of frequency response of each channel for Beta band, where the subjects are divided into 5 groups: EG with medication (row 1), EG without medication (row 2), NEAD with medication (row 3), NEAD without medication (row 4), HC (row 5). The same colormap scale is used across all plots. ....                                                                                                                                                                                                                                                                                                                                                                                                                          | 196 |
| Figure 4-23 Spatial distribution of frequency response of each channel for Gamma band, where the subjects are divided into 5 groups: EG with medication (row 1), EG without medication (row 2), NEAD with medication (row 3), NEAD without medication (row 4), HC (row 5). The same colormap scale is used across all plots. ....                                                                                                                                                                                                                                                                                                                                                                                                                         | 197 |
| Figure 4-24 The flowchart of (a) the proposed ST-GCN framework and (b) the T-CNN framework for comparison. $X$ is sized $M \times N$ ( $M = 25$ , representing the length of the mini-epoch channel; $N = 23$ , representing the 23 channels) and the input $W$ is sized $N \times N$ . The $(i, j)$ th entry of the adjacency matrix $W$ denotes spatial coupling correlation strength between the $i$ th and $j$ th of all the 23 different channels, the detailed calculation of which is presented in Section 2.2. $Kt$ is the size of the temporal 1D-Conv filter, set as 3 here. $Cl$ ( $l = 1, 2, \dots, 6$ ) is the number of filters in each layer. The estimated computational complexity of ST-GCN and T-CNN are 54M and 52M respectively..... | 210 |
| Figure 4-25 The adjacency matrices of one mini-epoch of AD from six FC methods.....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 220 |
| Figure 4-26 The adjacency matrices of one mini-epoch of HC from six FC methods.....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 221 |
| Figure 4-27 Training and convergence process for Full band data with WC connectivity by the ST-GCN model.....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 223 |
| Figure 4-28 The averaged WC adjacency matrices of ADs and HCs.....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | 224 |
| Figure 4-29 Threshold of averaged WC adjacency matrices of ADs and HCs.....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 225 |
| Figure 4-31 The 3D brain mapping of averaged WC adjacency matrices for EO state of HCs.....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 227 |
| Figure 4-30The 3D brain mapping of averaged WC adjacency matrices for EO state of ADs .....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 227 |
| Figure 4-32 The 3D brain mapping of averaged WC adjacency matrices for EC state of ADs .....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | 228 |

|                                                                                                                                                                                                                                                                                                                                                                             |     |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Figure 4-33 The 3D brain mapping of averaged WC adjacency matrices for EC state of HCs.....                                                                                                                                                                                                                                                                                 | 228 |
| Figure 5-1 Flowchart of the proposed framework. The pipeline contains four main steps: pre-processing, feature extraction, feature selection and machine-learning classification.....                                                                                                                                                                                       | 244 |
| Figure 5-2 Layout of a set in the experiment. A single task consisted of a 2 s instruction indicating the type of task (0-, 2-, or 3-back), a 40 s task period that consisted of 20 trials, a 1 s stop period, and a 20 s rest period. Each participant completed nine sets of n-back tasks. ....                                                                           | 245 |
| Figure 5-3 Time window analysis (A) Time interval analysis for fNIRS features; Time window-size evaluation for EEG and fNIRS features for (B) 0-back vs. 2-back, (C) 0-back vs. 3-back, (D) 2-back vs. 3-back. ....                                                                                                                                                         | 250 |
| Figure 5-4 Comparison of four FBC estimations (MI, PCC, MSC and PLV) in terms of the average of the Top 10 classification accuracies along with maximum and minimum value. ....                                                                                                                                                                                             | 252 |
| Figure 5-5 The receiver operating characteristic (ROC) curves for three binary classification tasks: 0-back vs. 2-back, 0-back vs. 3-back, and 2-back vs. 3-back. ....                                                                                                                                                                                                      | 255 |
| Figure 5-6 Topographic map of the EEG alpha-band PSD. Left: average; Middle: each participant; Right: Accuracy using each-channel PSD as the input. The area that provides the highest accuracy is highlighted. ....                                                                                                                                                        | 257 |
| Figure 5-7 Topographic map of the fNIRS HbR features. Left: average; Middle: each participant; Right: Accuracy using each-channel HbR feature as the input. The area that provided the highest accuracy is highlighted. ....                                                                                                                                                | 259 |
| Figure 5-8 The heat map of MI FBC features and the accuracy results. (A–C) shows the MI value for each participant in 0-back, 2-back, and 3-back task. (D–F) represents the classification accuracy for 0-back vs. 2-back, 0-back vs.3-back and 2-back vs. 3-back, respectively, using each pair of EEG channels as the input where the FBC value was estimated by MI. .... | 259 |
| Figure 5-9 Channels and locations for the EEG (Left) and fNIRS (Right) recordings.....                                                                                                                                                                                                                                                                                      | 264 |
| Figure 5-10 fNIRS average HbO value of each of 26 participants in three levels of workload .....                                                                                                                                                                                                                                                                            | 265 |
| Figure 5-11 fNIRS average HbR value of each of 26 participants in three levels of workload .....                                                                                                                                                                                                                                                                            | 265 |
| Figure 5-12 A sample of significant test to represent the difference among three-level workload with the purpose of selecting limited numbers of features. (p-value<0.0001) .....                                                                                                                                                                                           | 266 |
| Figure 6-1 The architecture of the proposed Causality Detection Attention-based Convolutional Neural Networks (CD-ABCNN). Here is an example to                                                                                                                                                                                                                             |     |

predict  $\mathbf{x}_t$ , with 2 hidden layers ( $L=2$ ), kernel size  $K = 2$  (shown as arrows) and dilation coefficient  $d = 2$ , leading to a receptive field  $R = 8$ . It consists of  $N$  channels, one for each input time series. A PReLU activation function is applied after each convolution. Weights  $\mathbf{W}$  are shared across layers, indicated by the identical colours. The  $N$  channels have  $T$  steps,  $L = 1$  hidden layer in the depthwise convolution and  $N \times 2$  kernels with kernel size  $K = 2$ . The attention scores  $\mathbf{a}$  are multiplied elementwise with the input time series, followed by an elementwise multiplication with the kernel. In the pointwise convolution, all channel outputs are combined to construct the prediction  $\mathbf{x}_t$ ..... 276

Figure 6-2 Methodology of CD-ABCNN performed on practical 23 channel EEG recordings. CD-ABCNN employs 23 independent CNNs, denoted as  $\mathbf{N1} \dots \mathbf{N23}$ , where each network takes each channel of 23-channel EEGs of length  $T$  as input ( $N$  is equivalent to the number of time series in the input data set). The CD-ABCNN constructs a temporal causal graph, namely effective connectivity networks. .... 278

Figure 6-3 Performance of the proposed CD-ABCNN using the MVAR-model simulated data. (a) **i** the structure of the designed 5-node networks; **ii** the simulated data based on 5-node MVAR-simulated networks; **iii** and **iv** the ground truth of the causal relationship of the designed 5-node networks; **v** and **vi** the learnt of the causal relationship of the designed 5-node networks; . (b) **i** the structure of the designed 9-node networks; **ii** the simulated data based on 9-node MVAR-simulated networks; **iii** and **iv** the ground truth of the causal relationship of the designed 9-node networks; **v** and **vi** the learnt of the causal relationship of the designed 9-node networks; (c) the performance of the proposed CD-ABCNN on 5-node and 9-node simulated time series..... 281

Figure 6-4 The visualisation of the learnt ECNs of one AD (left) and one HC (right). (a) and (d) the heat map of the learnt ECNs; (b) and (e) the chord diagram of the learnt ECNs. The order and locations of channels are designed to correspond the brain areas; (c) and (f) the learnt ECNs with corresponding brain regions. 23 bipolar derivations were created from 8 different brain regions labelled by distinct colour: the left frontal (F7–F3, F3–FZ), right frontal (F8–F4, F4–FZ), left frontocentral, temporo- central, temporal and centroparietal (F3–C3, T3– C3, T3–T5, C3–P3), right frontocentral, temporo- central, temporal and centroparietal (F4–C4, T4–C4, T4–T6, C4–P4), midline frontocentral, centroparietal and left and right parasagittal central regions (Fz– Cz, Cz–Pz, C3–Cz, C4–Cz), left parietal, parieto-occipital and temporo-occipital areas (P3–Pz, P3– O1, T5–O1), right parietal, parieto-occipital and temporo-occipital areas (P4–Pz, P4–O2, T6–O2) and midline occipital region (O1–O2)). .... 284



Figure 6-5 The statistical comparison of AD and HC in terms of the learned ECNs under EO and EC condition, using the proposed CD-ABCNN framework. (a) the sum of ECNs of HC under EC condition; (b) the sum of ECNs of AD under EC condition. The yellow circles indicate specific pairs of channels show learned effective connectivity for all AD patients, but not for all HCs. (c) the heat map of difference between HC and AD under EC condition, namely (a)-(b); (d) the sum of ECNs of HC under EO condition; (e) the sum of ECNs of AD under EO condition. The green circles indicate specific pairs of channels show learned effective connectivity for all HCs, but not for all AD patients. The red circles indicate specific pairs of channels that imply the AD and HC are affected by the eyes' conditions (EO/EC) to varying degrees. (f) the heat map of difference between HC and AD under EO condition, namely (d)-(e); (g) the comparison of AD and HC in terms of top 15 pairs of channels that shows the most significant difference under EC condition; (h) the comparison of AD and HC in terms of top 15 pairs of channels that shows the most significant difference under EO condition. Some combinations have common EEG channels because a bipolar approach is employed (such as F4–C4 and C4–P4), which may result in a misleading high false connectivity between some of the bipolar channels. To avoid this issue, the authors neglected any pair with common EEG electrodes. Therefore, the 46 channel pairs that have this characteristic were not considered. .... 286

Figure 6-6 The regional comparison of AD and HC in terms of the learned ECNs under EO and EC condition, using the proposed CD-ABCNN framework. (a) the sum of learnt ECNs of HC under EC condition; (b) the sum of learnt ECNs of AD under EC condition. The arrow points out the bi-directed causal relationship that represent the most significant difference between HC and AD under the EC condition. (c) the sum of learnt ECNs of HC under EO condition; (d) the sum of learnt ECNs of AD under EO condition. . The arrow points out the causal relationship that represent the most significant difference between HC and AD under the EC condition. The width of the lines represents variant numbers of participants. Significant difference between EO and EC for HCs are highlighted by the green rectangles.... 288

Figure 7-1 (A) Effective brain connectivity visualised by a heat map marked with the corresponding brain regions; (B) the proposed framework includes four main steps: pre-processing, feature extraction (EBC and PSD), graph construction, and DSL-GNN classification. The proposed DSL-GNN employs brain networks as the input, which is constructed by a graph comprised of PSD as nodes and EBC as directional edges. .... 302

Figure 7-2 The Flowchart of the proposed DSL-GNN. The input is constructed by a graph comprised of PSD as nodes and EBC as directional edges. . 304

Figure 7-4 Classification accuracy in Full band using DSL-GNN, employing five EBC estimations: PDC, GPDC, GGC, DTF and ffDTF. .... 313

Figure 7-3 Classification accuracy in Full band using SVM, employing five EBC estimations: PDC, GPDC, GGC, DTF and ffDTF..... 313

|                                                                                                                                                             |     |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Figure 7-5 The normalised average ffDTF in different band for HC, AD and PD .....                                                                           | 318 |
| Figure 7-6 The normalised average PSD in different bands for HC, AD and PD. The PSD distribution of PD shows more difference, compared with HC and AD. .... | 319 |

## LIST OF TABLES

|                                                                                                                                                                                                                                                                                                                           |     |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Table 2-1 Comparison of methods for quantifying brain connectivity using EEG .....                                                                                                                                                                                                                                        | 16  |
| Table 2-2 Recent applications combining brain connectivity estimations with machine learning methods.....                                                                                                                                                                                                                 | 24  |
| Table 2-3 The potential directions of EEG brain connectivity research .....                                                                                                                                                                                                                                               | 30  |
| Table 3-1 THE DETAILS OF PARTICIPANTS AND AVAILABLE DATA.....                                                                                                                                                                                                                                                             | 105 |
| Table 3-2 THE NUMBERS OF EFFECTIVE CROSS-CHANNELS FOR CLASSIFICATION WITH RHHT- AND CWT-BASED CROSS-SPECTRUM .....                                                                                                                                                                                                        | 106 |
| Table 3-3 Extracted features from the cross-curves.um and two average curves .....                                                                                                                                                                                                                                        | 131 |
| Table 3-4 A comparison of performances of the various functional connectivity methods and univariate methods (maximum accuracy achieved by each method) .....                                                                                                                                                             | 144 |
| Table 4-1 The properties of the estimated functional connectivity.....                                                                                                                                                                                                                                                    | 169 |
| Table 4-2 THE COMPARISON OF THE HIGHEST CLASSIFICATION USING A SINGLE FEATURE BETWEEN THE UNIVARIATE APPROACH AND THE PROPOSED APPROACH .....                                                                                                                                                                             | 184 |
| Table 4-3 THE DETAILS OF PARTICIPANTS AND AVAILABLE DATA.....                                                                                                                                                                                                                                                             | 188 |
| Table 4-4 THE LIST OF 23 BIPOLAR CHANNELS.....                                                                                                                                                                                                                                                                            | 189 |
| Table 4-5 Classification accuracy for HC and AD in EC and EO states within 6 bands by ST-GCN with different adjacency matrices. ....                                                                                                                                                                                      | 222 |
| Table 4-6 Classification accuracy for HC and AD in EC and EO states within 6 bands by T-CCN with different adjacency matrices. ....                                                                                                                                                                                       | 222 |
| Table 4-7 ST-GCN Model performance in a hand-out validation, where 12 ADs and 13 HCs were used for training and validation and the remaining 7 ADs, and 7 HCs were used for testing. Overall, the classification accuracy of the full-band data is still the best, with an average of 70.5% for EC and 68.9% for EO. .... | 230 |
| Table 4-8 T-CCN Model performance in a hand-out validation, where 12 ADs and 13 HCs were used for training and validation and the remaining 7 ADs and 7 HCs were used for testing. Overall, the classification accuracy of the full-band data is lower than the ST-GCN model. ....                                        | 231 |
| Table 5-1 Comparison of four neuroimaging techniques.....                                                                                                                                                                                                                                                                 | 241 |

|                                                                                                                                                 |     |
|-------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Table 5-2 SVM classification accuracy of 0-back vs. 2-back using different features.....                                                        | 253 |
| Table 5-3 SVM classification accuracy of 0-back vs. 3-back using different features.....                                                        | 253 |
| Table 5-4 SVM classification accuracy of 2-back vs. 3-back using different features.....                                                        | 253 |
| Table 5-5 Performance of classification for 3 binary classification tasks. ....                                                                 | 254 |
| Table 5-6 A comparative analysis of the previous research and the proposed work. ....                                                           | 260 |
| Table 6-1 The Number of subjects in four groups.....                                                                                            | 280 |
| Table 7-1 Classification accuracy using SVM in five different bands using five EBC estimations and PSD (%) .....                                | 315 |
| Table 7-2 Classification accuracy using DSL-GNN in five different bands using five EBC estimations and PSD (%).....                             | 316 |
| Table 7-3 Performance of popular features and classifiers that are implemented on the same dataset in the discrimination task of AD or PD ..... | 320 |

## LIST OF ABBREVIATIONS

|          |                                                                   |
|----------|-------------------------------------------------------------------|
| AAFT     | Amplitude Adjusted Fourier Transform                              |
| AAR      | Automatic Artifact Removal                                        |
| AD       | Alzheimer's Disease                                               |
| ADHD     | Attention-Deficit/Hyperactivity Disorder                          |
| AED      | Anti-Epileptic Drugs                                              |
| ANOVA    | One-way Analysis of Variance                                      |
| AR       | Autoregressive                                                    |
| ARMAX    | Autoregressive Moving Average with Exogenous Input                |
| ARX      | Autoregressive with Exogenous                                     |
| AUC      | the Area Under the ROC Curve                                      |
| BiFEC    | Brain Functional and Effective Connectivity                       |
| CD-ABCNN | Causality Detection Attention-Based Convolutional Neural Networks |
| CEEMDAN  | Complete Ensemble EMD with Adaptive Noise                         |
| CNN      | Convolutional Neural Networks                                     |
| CPSD     | Cross-Power Spectral Density                                      |
| CWT      | Continuous Wavelet Transform                                      |
| DCM      | Dynamic Causal Model                                              |
| DCM      | Dynamic Causal Modelling                                          |
| DCM      | Dynamic Causal Modelling                                          |
| DSL-GNN  | directed structure learning GNN                                   |
| DTF      | Directed Transfer Function                                        |
| DTF      | Directed Transfer Function                                        |
| DWT      | Discrete Wavelet Transform                                        |
| EBC      | Effective Brain Connectivity                                      |
| EBC      | Effective Brain Connectivity                                      |
| EC       | Eyes Close                                                        |
| ECN      | Effective Connectivity Network                                    |
| EEG      | Electroencephalography                                            |
| EEMD     | ensemble EMD                                                      |
| EG       | Epilepsy Group                                                    |
| EMD      | Empirical Mode Decomposition                                      |
| EO       | Eyes Open                                                         |
| ERR      | Error Reduction Ratio                                             |
| FBC      | Functional Brain Connectivity                                     |

|         |                                                              |
|---------|--------------------------------------------------------------|
| ffDTF   | full frequency Direct Transfer Function                      |
| fMRI    | Functional Magnetic Resonance Imaging                        |
| fNIRS   | Functional Near-Infrared Spectroscopy                        |
| FXS     | Fragile X Syndrome                                           |
| GC      | Granger Causality                                            |
| GCNs    | Graph Convolutional Networks                                 |
| GGC     | Geweke-Granger Causality                                     |
| GNN     | Graph Neural Network                                         |
| GNN     | Graph Neural Network                                         |
| GPDC    | Generalized Partial Directed Coherence                       |
| HbO     | Oxygenated Hemoglobin                                        |
| HbR     | Deoxygenated Hemoglobin                                      |
| HC      | Healthy Control                                              |
| IEDs    | Interictal Epileptiform Discharges                           |
| IMF     | Intrinsic Mode Functions                                     |
| IPC     | Imaginary Part of Coherence                                  |
| iWASOBI | Improved Weight-Adjusted Second-Order Blind Identification   |
| KNN     | K-Nearest Neighbour                                          |
| LDA     | Linear Discriminant Analysis                                 |
| LR      | Logistic Regression                                          |
| mBLL    | modified Beer–Lambert equation                               |
| MEG     | Magnetoencephalography                                       |
| MI      | Mutual Information                                           |
| ML      | Machine Learning                                             |
| MRI     | Magnetic Resonance Imaging                                   |
| MSC     | Magnitude Squared Coherence                                  |
| MVAR    | Multivariate Autoregressive                                  |
| NARMAX  | Nonlinear Autoregressive Moving Average with Exogenous Input |
| NDTF    | Non-normalized Directed Transfer Function                    |
| NDTF    | Non-normalised Directed Transfer Function                    |
| NEAD    | Non-Epileptic Attack Disorder                                |
| NMF     | Nonnegative Matrix Factorization                             |
| PCA     | Principal Component Analysis                                 |
| PCC     | Pearson Correlation Coefficient                              |
| PD      | Parkinson's disease                                          |
| PDC     | Partial Directed Coherence                                   |

|        |                                         |
|--------|-----------------------------------------|
| PDC    | Partial Directed Coherence              |
| PESR   | Penalised Error-to-Signal Ratio         |
| PET    | Position Emission Tomography            |
| PFoCS  | Peak Frequency of Cross Spectrum        |
| PLI    | Phase Lag Index                         |
| PLV    | Phase Locking Value                     |
| PS     | Phase Synchronization                   |
| PSD    | Power Spectral Density                  |
| PSD    | Power Spectrum Density                  |
| PSDD   | Power Spectral Density Difference       |
| RBF    | Radial Basis Function                   |
| RF     | Random Forest                           |
| RHHT   | Revised Hilbert-Huang Transformation    |
| ROC    | Receiver Operating Characteristic       |
| ROIs   | Regions of Interests                    |
| SEM    | Structural Equation Model               |
| SL     | Synchronization Likelihood              |
| SL     | Synchronization Likelihood              |
| STFC   | Short-Time Fourier Coherence            |
| STFT   | Short-Time Fourier Transform            |
| ST-GCN | Spatial-Temporal GCN                    |
| SVM    | Support Vector Machine                  |
| tDCS   | transcranial Direct Current Stimulation |
| TE     | Transfer Entropy                        |
| TE     | Transfer Entropy                        |
| TMS    | Transcranial Magnetic Stimulation       |
| WC     | Wavelet Coherence                       |
| WC     | Wavelet Coherence                       |
| WT     | Wavelet Transform                       |

# **1 Introduction**

## **1.1 Background**

Analysing non-invasive brain signals is a growing trend in the field of neuroscience and medical research. Non-invasive brain signal analysis refers to the use of techniques that allow scientists to study brain activity without requiring invasive procedures like surgery or implantation of electrodes (Li et al., 2020). The trend towards analysing non-invasive brain signals is driven by a desire to understand the brain better and develop new treatments for neurological and psychiatric disorders. Non-invasive techniques are preferred over invasive ones because they are less risky, less expensive, and more accessible to a wider range of patients (Van de Steen et al., 2019; Yasin et al., 2021). In this case, a type of neurophysiological signal, so-called scalp electroencephalography (EEG), has been playing an increasingly important role in this field since EEG has the advantages of low cost, high temporal sampling rate, and improved accessibility at hospitals. Meanwhile, growing attention was paid to brain connectivity estimated from EEGs due to increasing evidence that brain connectivity can reveal the function of different brain regions and complex cortical intercommunications among them. Brain connectivity can be subdivided into neuroanatomical (or structural), functional and effective connectivity (Cao et al., 2022). Neuroanatomical connectivity refers to structural links such as synapses or fibre pathways at the microscopic scale of neurons. Functional connectivity is defined as statistical interdependence among remote neurophysiological regions (Qiao et al., 2018; Sadaghiani and Kleinschmidt, 2016). In terms of effective connectivity, it indicates the directed influence of one neural region over others (Hasanzadeh et al., 2022; Saeedi et al., 2020; Schumacher et al., 2015). Selecting the most appropriate connectivity measure is crucial, as widely distributed complex brain networks generate diverse topological signal processing and intercommunications in distinct conditions. Therefore, many novel methods have been developed from various aspects. This includes consideration of linear or nonlinear behaviours, the information in



time, frequency or time-frequency domain, parametric or non-parametric measures, and directed or undirected information (Cao et al., 2022).

Visualisation plays a crucial role in the research of EEG brain connectivity. Specifically, it establishes the base for further practical and clinical uses of novel methods. It could not only improve the efficiency for other researchers to understand and evaluate the proposed approaches and produced results but also benefit experts in presenting and explaining the finding obtained from various advanced methods to people without basic knowledge of EEG signals and brain connectivity. Besides, EEG study could be combined with appropriate visualisation methods to study potential functions, oscillations and intercommunications within the dynamic architecture of the human brain. The current medical instrument related to EEG focuses on oscillations and biomarkers within a signal channel. Therefore, from my point of view, development and innovations of visualisation can promote the upgrading of EEG-related equipment, as well as explore more secrets within the human brain. For instance, the utilization of appropriate visualizations can significantly enhance researchers' comprehension of the distinctive contributions made by various brain regions. Moreover, the resultant insights can offer valuable perspectives for the refinement of EEG sensor design, encompassing considerations such as configuration and layout.

Increasing evidence shows that brain connectivity analysis, quantified by correlation and coherence, information theory and causality analysis, can reflect on the changes in neurological patients. It has provided a significant number of promising results relevant to different neurological disorders such as Alzheimer's disease, epilepsy, and Parkinson's disease. In addition to neurological and neurophysiological studies and associated clinical applications, there is a wide range of applications where the research objectives are to assist humans in understanding brain behaviour, for example, emotion recognition, object recognition, mental assessment and biometric identification. However, there are still a lot of limitations and challenges.

## 1.2 Research Gaps

First of all, lack of interpretability is a main issue in this field. The volume of EEG recordings is usually massive, and their interpretability is highly related to clinical practice needing specific knowledge to understand. Therefore, the practical use of EEG is extremely dependent on doctors' experience, and the process is usually time-consuming.

Secondly, changes in brain connectivity may occur in the early stage of neurological disorders, while this sometimes cannot be diagnosed by traditional digital signal processing methods relying on a single channel. In this case, potential features may be explored from brain functional and effective connectivity to support early diagnosis and progression evaluation of neurological diseases.

Thirdly, it has been proved that brain connectivity can reveal the underlying function of different brain regions and complex cortical intercommunications among them, which helps better understand many neurological conditions, including disorders and emotions. However, the potential of brain connectivity in cognitive research (e.g., multi-class workload classification problem) has not been fully explored.

Fourthly, EEG study could be combined with appropriate visualisation methods to study potential functions, oscillations, and intercommunications within the dynamic architecture of the human brain. The current medical instrument related to EEG focuses more on oscillations and biomarkers within a single channel, but there are limited approaches to interpret the EEG brain connectivity. The development and innovations of visualisation may promote the upgrading of EEG-related equipment, as well as explore more secrets within the human brain. In this case, there is high demand for satisfactory visualisation techniques for brain connectivity.

Fifthly, the literature suggests that effective brain connectivity (EBC) has the potential to reveal abnormal cortical changes in the patients' brain caused by neurological disorders to an extent. In the meantime, to deal with complex brain

networks, a graph neural network (GNN) has been increasingly popular in recent years. However, the effectiveness of the combination of the mentioned EBC and GNN techniques has been largely unexplored in the field of dementia diagnosis.

Finally, brain network interactions are dynamic and may be time-varying (the associated signals are non-stationary), as phase synchronisation and phase scattering occur within the millisecond range. Transient associations have been observed, usually highly nonlinear, among different brain regions. Even though recent years have witnessed the development of brain connectivity techniques, limited methods can analyse the non-stationary and non-linear behaviours of brain networks. Exploring hidden information within EEG signals is far beyond the capabilities of commonly available methods.

### **1.3 Aim and Objectives**

This PhD project aims to develop an EEG-based brain functional and effective connectivity (BiFEC) estimation and visualisation technique to identify new sensitive, non-invasive and reproducible biomarkers for diagnosing and monitoring neurological diseases. The aim results in the following four objectives:

- 1) To review state-of-the-art brain connectivity quantification methods and well-accepted visualisation approaches to represent brain functional and effective connectivity (BiFEC). This objective links to Chapter 2.
- 2) To develop novel methods to estimate BiFEC in time, frequency and time-frequency domains, primarily focusing on exploring the non-linear and highly dynamic intercommunications with the associated spatial distribution. This objective links to Chapter 3 and Chapter 6.
- 3) To design brain connectivity visualisation approaches to reveal brain activity and disorder evaluated by EEG recordings. This objective links to Chapters 3-7.
- 4) To evaluate the proposed BiFEC methods by applying them to real EEG data from patients with various neurological disorders (e.g., Alzheimer's

disease, epilepsy, and Parkinson's disease) or from participants in mental workload classification tasks. This objective links to Chapters 4, 5 and 7.

The structure of PhD thesis is illustrated in Figure 1-1. The first chapter introduces the background of EEG brain connectivity, summarises the relevant research gap, clarifies the main aim and individual objectives as well as the corresponding methodology, and provides the structure of the thesis. Chapter 2 is a literature review of the brain functional and effective connectivity based on EEG recordings to achieve Objective 1. Chapter 3 and Chapter 6 introduce several novel functional and effective connectivity estimation approaches proposed by this PhD project to achieve Objective 2. Chapter 4, Chapter 5, and Chapter 7 introduce practical applications of functional connectivity and effectivity connectivity, respectively, to achieve Objective 4. In the meantime, various visualisation methods are designed and adopted to assist in developing and utilising functional and effective connectivity to achieve Objective 3.

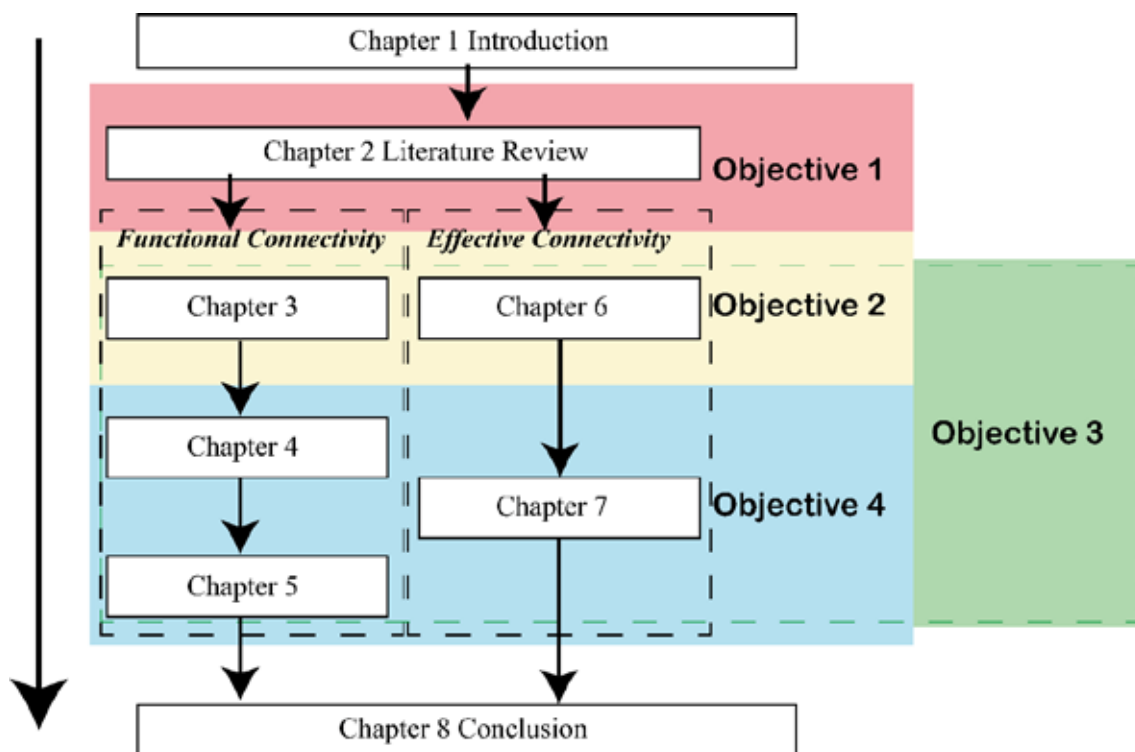


Figure 1-1 The structure of PhD Thesis

## **1.4 Research Methodology adopted to achieve the stated objectives.**

- 1) To systematically review brain functional and effective connectivity estimation and analysis methods in accordance with their properties, such as linear or nonlinear, parametric or non-parametric, time, frequency or time-frequency domain and directed or undirected. Besides, various distinct approaches to connectivity visualisation are reviewed. This review plays a crucial role in my thesis to identify the research gaps. (Objective 1)
- 2) There are three typical categories of functional connectivity measures: (1) Correlation and coherence, (2) Phase synchronization measures, and (3) Information-based measures. Slightly different from functional connectivity, effective connectivity is event-dependent and relies more on a model, for example, the structural equation model (SEM), dynamic causal model (DCM) and directed transfer function (DTF) based on the autoregressive model. After reviewing and evaluating previous methods, this project explores several novel approaches to measure functional and effective connectivity, especially targeting to estimate nonlinear and dynamic information from non-stationary brain networks. The study is expected to be involved in the time, frequency, and time-frequency domains and explore information individually and conjunctively. (Objective 2)
- 3) Data-driven methods gradually represents the power of calculation causal intercommunication between the multiple components of brain networks. Unlike parametric-model-based methods, non-parametric approaches do not need to build a model before calculating connections. In other words, it could avoid assumptions of an unknown system, improving the efficiency of algorithms to a certain extent. In this case, some novel deep learning methods will be applied to learn brain effective connectivity, revealing whole-brain causal information flows. (Objective 2)
- 4) With the purpose of increasing the interpretability of results obtained from EEG recordings, many researchers developed or employed various distinguished visualisation approaches. The visualisation methods can be categorised into three groups: heat map, data statistics and head map. The

continuous advancement of visualisation methods applied in brain connectivity has been successful in effectively capturing distinguishing information from EEG recordings. Depending on the cases, an appropriate visualisation pattern could be designed to reveal more valuable brain connectivity information from different aspects. (Objective 3)

- 5) To evaluate the validation of proposed methods, those methods will be realised and verified in the case of Alzheimer's disease, epilepsy, Parkinson's disease and mental workload classification. The new approaches are expected to reflect the difference between healthy controls and patients or patients in different stages. (Objective 4)
- 6) One-way analysis of variance (ANOVA) is employed to evaluate the significance of differences in the investigated features of different groups. In this case, it guides to select significant channels for each frequency band based on the p-value, preparing for optimising the input features of machine learning. ANOVA is widely used to test the significant difference between the different groups of participants. ANOVA tests the null hypothesis, i.e., the means of the tested groups are equal, and the p-value indicates the statistical significance. The rejection of the null hypothesis leads to the conclusion that the two groups are statistically different. To select promising features that offer significant differences between different groups and to reduce the computational burden of machine learning, the estimation is undertaken using a determined p-value. (Objective 4)
- 7) Moreover, to evaluate the effectiveness of a variety of features and biomarkers extracted from signal processing and brain connectivity analysis, machine learning (ML) methods tend to be used, such as support vector machine (SVM), random forest (RF), K-Nearest Neighbour (KNN), convolutional neural networks (CNN) and graph neural network (GNN). With the purpose of further evaluating the machine learning algorithm performance, accuracy, sensitivity, and specificity will be calculated. Moreover, the receiver operating characteristic (ROC) curve and the area under the ROC curve (AUC) can be further calculated to assess the goodness of classification. (Objective 4)

- 8) In terms of EEG data resources, this project is cooperating with a local NHS hospital so that clinical EEGs can be obtained and used with a Data Management Plan in place. Furthermore, supported by the doctors' knowledge and experience, some medical and neuroscience findings can be explored and summed up. (Objective 4)

### **1.5 List of journal papers**

1. Cao, J., Garro, E.M., Zhao, Y., 2022a. EEG/fNIRS Based Workload Classification Using Functional Brain Connectivity and Machine Learning. *Sensors* 22, 7623. <https://doi.org/10.3390/s22197623>
2. Cao, J., Grajcar, K., Shan, X., Zhao, Y., Zou, J., Chen, L., Li, Z., Grunewald, R., Zis, P., Marco, M. De, Unwin, Z., Blackburn, D., Sarrigiannis, P.G., 2021. Using interictal seizure-free EEG data to recognise patients with epilepsy based on machine learning of brain functional connectivity. *Biomed. Signal Process. Control* 67, 102554. <https://doi.org/10.1016/j.bspc.2021.102554>
3. Cao, J., Zhao, Y., Shan, X., Blackburn, D., Wei, J., Erkoyuncu, J.A., Chen, L., Sarrigiannis, P.G., 2022b. Ultra-high-resolution time-frequency analysis of EEG to characterise brain functional connectivity with the application in Alzheimer's disease. *J. Neural Eng.* 19, 046034. <https://doi.org/10.1088/1741-2552/ac84ac>
4. Cao, J., Zhao, Y., Shan, X., Wei, H., Guo, Y., Chen, L., Erkoyuncu, J.A., Sarrigiannis, P.G., 2022c. Brain functional and effective connectivity based on electroencephalography recordings: A review. *Hum. Brain Mapp.* 43, 860–879. <https://doi.org/10.1002/hbm.25683>
5. Shan, X., Cao, J., Huo, S., Chen, L., Sarrigiannis, P.G., Zhao, Y., 2022. Spatial-temporal graph convolutional network for Alzheimer classification based on brain functional connectivity imaging of electroencephalogram. *Hum. Brain Mapp.* 43, 5194–5209. <https://doi.org/10.1002/hbm.25994>
6. Shan, X., Huo, S., Yang, L., Cao, J., Zou, J., Chen, L., Sarrigiannis, P.G., Zhao, Y., 2021. A Revised Hilbert-Huang Transformation to Track Non-Stationary Association of Electroencephalography Signals. *IEEE Trans. Neural Syst. Rehabil. Eng.* 29, 841–851. <https://doi.org/10.1109/TNSRE.2021.3076311>

7. (Submitted to Pattern Recognition) Dementia Classification Using a Graph Neural Network on Imaging of Effective Brain Connectivity
8. Imaging of Effective Connectivity Network Using Causality Detection Attention-based Convolutional Neural Networks (CD-ABCNN)

## 1.6 Reference

Cao, J., Zhao, Y., Shan, X., Wei, H., Guo, Y., Chen, L., Erkoyuncu, J.A., Sarrigiannis, P.G., 2022. Brain functional and effective connectivity based on electroencephalography recordings: A review. *Hum. Brain Mapp.* 43, 860–879. <https://doi.org/10.1002/hbm.25683>

Hasanzadeh, F., Mohebbi, M., Rostami, R., 2022. A Nonlinear Effective Connectivity Measure Based on Granger Causality and Volterra Series. *IEEE J. Biomed. Heal. Informatics* 26, 2299–2307. <https://doi.org/10.1109/JBHI.2021.3138199>

Li, J., Chan, S.C., Liu, Z., Chang, C., 2020. A Novel Adaptive Fading Kalman Filter-Based Approach to Time-Varying Brain Spectral/Connectivity Analyses of Event-Related EEG Signals. *IEEE Access* 8, 51230–51245. <https://doi.org/10.1109/ACCESS.2020.2979551>

Qiao, J., Lv, Y., Cao, C., Wang, Z., Li, A., 2018. Multivariate Deep Learning Classification of Alzheimer's Disease Based on Hierarchical Partner Matching Independent Component Analysis. *Front. Aging Neurosci.* 10, 1–12. <https://doi.org/10.3389/fnagi.2018.00417>

Sadaghiani



Schumacher, E.M., Stiris, T.A., Larsson, P.G., 2015. Clinical Neurophysiology Effective connectivity in long-term EEG monitoring in preterm infants. *Clin. Neurophysiol.* 126, 2261–2268. <https://doi.org/10.1016/j.clinph.2015.01.017>

Van de Steen, F., Almgren, H., Razi, A., Friston, K., Marinazzo, D., 2019. Dynamic causal modelling of fluctuating connectivity in resting-state EEG. *Neuroimage* 189, 476–484. <https://doi.org/10.1016/j.neuroimage.2019.01.055>

Yasin, S., Hussain, S.A., Aslan, S., Raza, I., Muzammel, M., Othmani, A., 2021. EEG based Major Depressive disorder and Bipolar disorder detection using Neural Networks:A review. *Comput. Methods Programs Biomed.* 202. <https://doi.org/10.1016/j.cmpb.2021.106007>

## 2 Literature Review

This chapter is based a journal review paper, 'Brain functional and effective connectivity based on EEG recordings: a review', published in **Human Brain Mapping**.

Cao, J., Zhao, Y., Shan, X., Wei, H., Guo, Y., Chen, L., Erkoyuncu, J.A., Sarrigiannis, P.G., 2021b. Brain functional and effective connectivity based on electroencephalography recordings: A review. *Hum. Brain Mapp.* n/a, 1–20. <https://doi.org/https://doi.org/10.1002/hbm.25683>

### 2.1 Introduction

In the past decades, plenty of prior researchers showed their interest in understanding brain activities through analysing non-invasive brain signals. Scalp electroencephalography (EEG), has been yearly increasingly attractive resulting in a growing number of publications (Diessen ., 2015). On one hand, to extract hidden information from EEG recordings, various traditional digital signal processing methods were employed, such as independent component analysis (ICA) (Mierlo ., 2014; Tafreshi ., 2019a; van Mierlo ., 2014; Vecchio ., 2014), power spectral density (PSD) (Erra ., 2017; Ko ., 2009; Lias ., 2010; Saito ., 2006) and discrete wavelet transform (DWT) (Amin ., 2020; Cvetkovic ., 2008; Dhiman ., 2014; Ibrahim ., 2018; Vijay Anand and Shantha Selvakumari, 2019), etc. They usually focus on either a single EEG channel, which are univariate methods essentially, or blind signal separation. Those methods may not consider the intercommunications between different EEG signals. On the other hand, growing attention was paid to brain connectivity. There is increasing evidence that brain connectivity can reveal the function of different brain regions and complex cortical intercommunications among them (Babiloni ., 2005; V. Sakkalis, 2011; Tafreshi ., 2019a; van Mierlo ., 2014), which helps better understand many neurological conditions including disorders and emotions. As shown in Figure 2-1, the number of publications relevant to “EEG and brain connectivity” grew sustainably in recent years, especially after 2012.

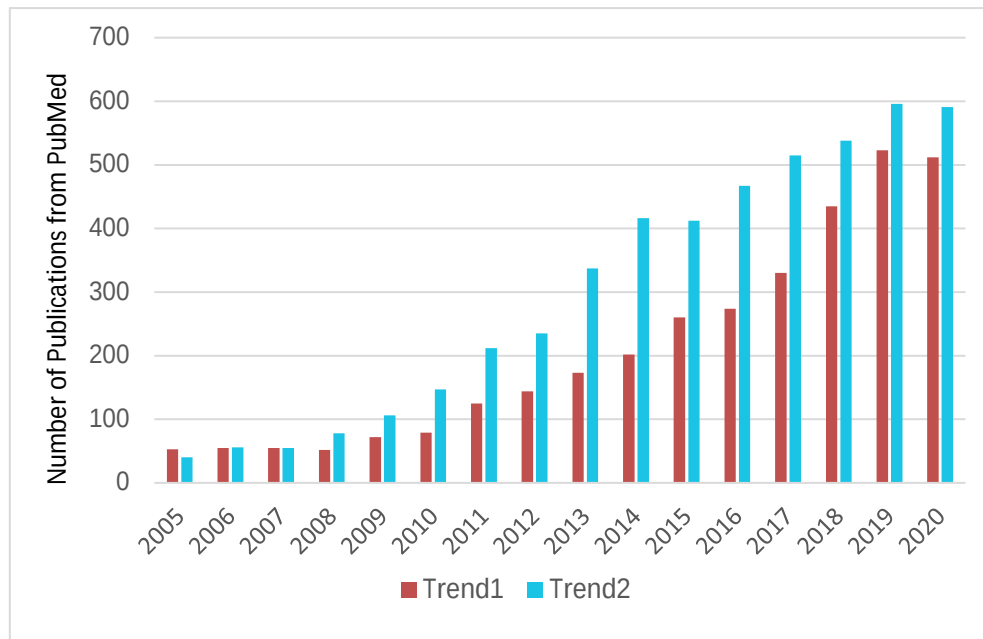


**Figure 2-1 Number of Publications per year from PubMed search using keywords "EEG and Machine learning" or "EEG and AI" (Trend1) and "brain connectivity and EEG" (Trend2) in the period 2005–2020**

Brain connectivity can be subdivided into neuroanatomical (or structural), functional and effective connectivity. Neuroanatomical connectivity refers to structural links such as synapses or fibre pathways at the microscopic scale of neurons (Sakkalis, 2011). Magnetic resonance imaging (MRI) and diffusion tensor imaging (DTI) are well-accepted techniques to reveal the brain structural connectivity with a relatively high spatial resolution (Clayden, 2013; Fan et al., 2016; Mohanty et al., 2020), while EEG cannot directly reveal structural connections and it is applied to estimate functional and effective connectivity. Compared with MRI, EEG offers higher temporal resolution, thus connectivity can be estimated at shorter time scales. Meanwhile, EEG has the capacity to enable early detection of anomalies in pathophysiological processes affecting brain networks at a lower cost before clinical symptoms emerge and before structural alterations are visible in MRI (Marzetti et al., 2019; Sadaghiani et al., 2021). Functional connectivity is defined as statistical interdependence among spatially distant neurophysiological regions (Friston, 2011), usually measured

by correlation, coherence and information theory (Cao et al., 2021a; Colclough et al., 2016; Marzetti et al., 2019; Sarrigiannis et al., 2015; Shan et al., 2021; Wang et al., 2014; Zhao et al., 2018). In terms of effective connectivity, it indicates the causal influence of one neural region over others (Astolfi et al., 2004; Tafreshi et al., 2019a) by combining imaging techniques like EEG and magnetoencephalography (MEG) with mathematical models of interconnected brain regions (Friston, 2001; Ponten et al., 2010). Different from the functional connectivity that only addresses statistical relationships, effective connectivity tends to reveal underlying mechanisms of interaction among neural regions and it is dynamic (event-dependent) and depends on a model of connections (Friston, 2011; He et al., 2013; Seth et al., 2015). It is believed that brain connectivity contains enormous potentials which could aid our understanding of brain networks. There are debates if the volume conduction effect can be avoided from sensor-level estimates of brain connectivity, due to the transmission from neural excitation to the surface measurement (He et al., 2019). Some researchers prefer to calculate the connectivity in source-space after solving the inverse problem (Moezzi and Goldsworthy, 2018; Palva et al., 2018; Supp et al., 2007).

Moreover, based on a variety of features and biomarkers extracted from signal processing and brain connectivity analysis, recent years have witnessed a considerable increase of published papers using machine learning (ML) methods for EEG classification (Figure 1), such as support vector machine (SVM), random forest (RF) and K-Nearest Neighbour (KNN) (Akbarian and Erfanian, 2020; Blinowska et al., 2017; Lehmann et al., 2007; Rajendra Acharya et al., 2012), as well as deep learning approaches (Ball et al., 2016; Chen et al., 2019; Saeedi et al., 2020).

Visualisation is indispensable in brain connectivity analysis and highly promoting to interpret brain activity and intercommunications. Zhao et al. (2020) suggest that the imaging and study of brain functional connectivity can effectively revolutionise our understanding of brain degradation or dysfunction in a user friendly and systematic way. Besides, some researchers believe that

connectivity results are supposed to be presented using appropriate visualisations that would in reality be interpreted by neurophysiologists (Haufe et al., 2013). On the other hand, the visualisation of brain connectivity generates novel and promising input images for some deep learning methods. Literature shows that the Convolution Neural Network (CNN) has the ability to identify the complexity present in EEG signals using advanced visualising techniques of brain connectivity (Gao et al., 2019; Mammone et al., 2020; Wang et al., 2019).

Although over recent years an incessant flurry of numerous promising approaches related to brain functional and effective connectivity has contributed to neural research in understanding brain network function and some reviews tried to summarise the methods and discuss the limitations (Bastos and Schoffelen, 2016; Diessen et al., 2015; V. Sakkalis, 2011; van Mierlo et al., 2014), there is a lack of comprehensive review in the last few years in terms of estimation, applications, and particularly their associated machine learning approaches. Moreover, there is no review available on a variety of visualisation methods to represent the variation of functional and effective connectivity.

Section 2 of this Chapter systematically reviews brain functional and effective connectivity estimation and analysis methods in accordance with their properties and applied areas. Those popular methods to quantify brain connectivity are grouped as parametric and non-parametric techniques and their conceptual and mathematical details are represented in the Appendix. Various distinct approaches of connectivity visualisation, grouped into three categories, are reviewed in Section 3. Section 4 provides a summary and critical evaluation of brain connectivity methods and their applications in the fields of neural disease and brain activity research. We then discuss the role of visualisation for brain connectivity analysis in detail. Besides, a discussion on the significance of machine learning and its various effective combination with brain connectivity estimations are represented in Section 4. This section closes with the identified research gaps and future direction. The conclusions are given in Section 5.

## 2.2 Estimation

A review of high-frequently used methods to estimate brain connectivity based on EEG is given in

Table 2-1 associated with their properties. These methods are either linear or nonlinear, parametric or non-parametric, belong to functional or effective

|      | Linearity |           | Signal processing |                | Brain connectivity |    | Domain |           |                |
|------|-----------|-----------|-------------------|----------------|--------------------|----|--------|-----------|----------------|
|      | Linear    | Nonlinear | Parametric        | non-Parametric | FC                 | EC | Time   | Frequency | Time-Frequency |
| DCM  |           |           |                   |                |                    |    |        |           |                |
| MSC  |           |           |                   |                |                    |    |        |           |                |
| STFC |           |           |                   |                |                    |    |        |           |                |
| WC   |           |           |                   |                |                    |    |        |           |                |
| PLV  |           |           |                   |                |                    |    |        |           |                |
| GS   |           |           |                   |                |                    |    |        |           |                |
| GC   |           |           |                   |                |                    |    |        |           |                |
| PDC  |           |           |                   |                |                    |    |        |           |                |
| Corr |           |           |                   |                |                    |    |        |           |                |
| SL   |           |           |                   |                |                    |    |        |           |                |
| TE   |           |           |                   |                |                    |    |        |           |                |
| MI   |           |           |                   |                |                    |    |        |           |                |
| DTF  |           |           |                   |                |                    |    |        |           |                |
| PS   |           |           |                   |                |                    |    |        |           |                |
| SEM  |           |           |                   |                |                    |    |        |           |                |
| IPC  |           |           |                   |                |                    |    |        |           |                |
| PLI  |           |           |                   |                |                    |    |        |           |                |
| ERR  |           |           |                   |                |                    |    |        |           |                |

connectivity, in time, frequency or time-frequency domain. The conceptual and mathematical details of these approaches are represented individually in the Appendix. It can be observed from this table that the functional connectivity is usually estimated by non-parametric methods while the effective connectivity estimates are based on parametric techniques. In this review, parametric and non-parametric methods are grouped separately.

|      | Linear | Nonlinear | Parametric | non-<br>Parametric | FC | EC | Time | Frequency | Time-<br>Frequency |
|------|--------|-----------|------------|--------------------|----|----|------|-----------|--------------------|
| DCM  |        |           |            |                    |    |    |      |           |                    |
| MSC  |        |           |            |                    |    |    |      |           |                    |
| STFC |        |           |            |                    |    |    |      |           |                    |
| WC   |        |           |            |                    |    |    |      |           |                    |
| PLV  |        |           |            |                    |    |    |      |           |                    |
| GS   |        |           |            |                    |    |    |      |           |                    |
| GC   |        |           |            |                    |    |    |      |           |                    |
| PDC  |        |           |            |                    |    |    |      |           |                    |
| Corr |        |           |            |                    |    |    |      |           |                    |
| SL   |        |           |            |                    |    |    |      |           |                    |
| TE   |        |           |            |                    |    |    |      |           |                    |
| MI   |        |           |            |                    |    |    |      |           |                    |
| DTF  |        |           |            |                    |    |    |      |           |                    |
| PS   |        |           |            |                    |    |    |      |           |                    |
| SEM  |        |           |            |                    |    |    |      |           |                    |
| IPC  |        |           |            |                    |    |    |      |           |                    |
| PLI  |        |           |            |                    |    |    |      |           |                    |
| ERR  |        |           |            |                    |    |    |      |           |                    |

**Table 2-1 Comparison of methods for quantifying brain connectivity using EEG**

---

## **2.2.1 Non-parametric methods**

### **2.2.1.1 Linear methods**

Linear intercommunication of the brain has been studied for several decades and is relatively straightforward in terms of computation and interpretation (Diessen et al., 2015). Correlation measures are well-accepted to represent linear connections. Pearson correlation coefficient (PCC) and cross-correlation were applied in the time domain to estimate functional connectivity from multi-channel EEG recordings (Fagerholm et al., 2020; Handojoseno et al., 2013; Ibrahim et al., 2018; Lee and Hsieh, 2014; Tarokh et al., 2010; Vortmann et al., 2019). It should be noted that PCC is the cross-correlation value at the zero time lag. In other words, if there is any time delay between the two EEG signals, PCC cannot accurately represent the strength of linear correlation. The counterpart of the cross-correlation in the frequency domain is the coherence. Coherence is sensitive to both changes in power and phase relationships between EEG signals (V. Sakkalis, 2011). The Magnitude Squared Coherence (MSC) (Battaglia and Brovelli, 2019) and the imaginary part of coherency (Ewald et al., 2012; Haufe and Ewald, 2019; Nolte et al., 2004; Stam et al., 2007) were widely utilised and the latter one is less sensitive to volume conduction (Nentwich et al., 2020; Nolte et al., 2004). The limitation of coherence and correlation is that only either spectral components or temporal information can be observed while no information about brain dynamics is provided. Time-frequency analysis methods are popular to study the changes in cortical connectivity, simultaneously extracting spectral and temporal information (Sankari et al., 2012). The Short-Time Fourier Coherence (STFC) (Chen et al., 2013; Wendling et al., 2009) and Wavelet Coherence (WC) (Ieracitano et al., 2017; Lachaux et al., 2002; Qassim et al., 2017; Sankari et al.,



2012; Sankari and Adeli, 2011) are utilised by several studies to produce EEG functional connectivity in the time-frequency domain. STFC employs a fixed sliding window to achieve spectral analysis within the time window, while WC optimises and adjusts the wavelet base for different signal frequencies to characterise time-varying coherence (V. Sakkalis, 2011).

### **2.2.1.2 Nonlinear methods**

Phase synchronization is a category that focuses on the phase coupling of oscillation systems. The Phase Locking Value (PLV) (Bajo et al., 2015; Bedo et al., 2020; Delgado-Restituto et al., 2019; Mheich et al., 2015; Sadaghiani and Kleinschmidt, 2016) and the Phase Lag Index (PLI) (Chaturvedi et al., 2019; Fraga González et al., 2018; Liao et al., 2019; Stam et al., 2007) are high-frequently used to obtain the strength of phase synchronization. Information theory is regarded as another efficient method in the case of extracting nonlinear interactions among EEG signals. To be more specific, Mutual Information (MI) (Melia et al., 2015; Meng et al., 2015; Piho and Tjahjadi, 2020; Rajendra Acharya et al., 2012; Yin et al., 2017) and Synchronization Likelihood (SL) (Altenburg et al., 2003; Chriskos et al., 2018; Mumtaz et al., 2018; Pijnenburg et al., 2004; Yu et al., 2017) are used to estimate the undirected functional connectivity from EEG recordings, while the directed transfer of information can be quantified by Transfer Entropy (TE). Choosing SL as a connectivity measure might lead to biased results since SL is sensitive to volume conduction effects (Boersma et al., 2013). In contrast, TE has demonstrated its robustness against volume conduction. (Harmah et al., 2020; Huang et al., 2015; McBride et al., 2015; Yang et al., 2013; Yao and Wang, 2017). It should be noticed that information-based approaches are fully model-free. That is to say, relatively fewer assumptions tend to be imposed on the data, but it trades in the need for larger data sets (Seth et al., 2015; Zhao et al., 2020).

### **2.2.2 Parametric methods**

The parametric methods refer to a framework or procedure where there is a specification of a model that requires a set of fixed parameters to be fitted to the

observed signals (V. Sakkalis, 2011; Salman ., 2018; Zhao ., 2012). Compared with model-free techniques, parametric-model-based methods are more commonly accepted to estimate the effective connectivity of multi-channel EEGs. In this case, parametric effective connectivity is based on theoretical models that describe how brain areas interact and influence each other (V. Sakkalis, 2011). Dynamic Causal Modelling (DCM) (Brown and Friston, 2012; Herz ., 2012; Lee ., 2020; Van de Steen ., 2019) applies a Bayesian framework to assess model performance, and Structural Equation Modelling (SEM) (Astolfi ., 2004; Babiloni ., 2003; Sartori ., 2012) is a generalized linear modelling framework combining factor analysis with path modelling (Kaur ., 2019). DCM and SEM treat the brain as a deterministic nonlinear and linear system respectively (Astolfi ., 2004; David ., 2006; Friston ., 2003).

Unlike DCM and SEM that require a certain amount of a priori knowledge about connectivity, many researchers have developed various data-driven methods for quantifying effective connectivity based on Granger causality (Salman ., 2018). Granger causality-related methods occupy a useful middle ground between fully model-free and highly model-dependent approaches (Seth ., 2015). Most of the Granger causality measures are constructed based on autoregressive models, in which the present samples of the signals are predicted using a linear or nonlinear combination of the past samples (Omidvarnia ., 2011; van Mierlo ., 2014; Zhao ., 2013, 2020). According to the considered coefficients of the autoregressive models, Partial Directed Coherence (PDC) (Ghumare ., 2018; He ., 2014a; Mazher ., 2017; Silfverhuth ., 2012; Varotto ., 2012, 2014) and Directed Transfer Function (DTF) (Haufe ., 2011; Omidvarnia ., 2011; Omidvarnia ., 2011) were applied in the field of EEG-based neuroscience. PDC is computationally more efficient and more robust than DTF since it does not involve any matrix inversion (He ., 2014a). Granger causality was originally developed in the context of linear autoregressive with exogenous (ARX) input models (Geweke, 1982), while some researchers focused on nonlinear causality, which is generated by nonlinear ARX model in time and



2013b, 2013a). Besides the traditional Granger method, the Error Reduction Ratio–causality (ERR-causality) test was proposed to estimate the time-varying direction and strength of linear or nonlinear causality between two signals as well as their relative time shift (Sarrigiannis et al., 2014; Zhao et al., 2012, 2020)

## **2.3 Visualisation of brain connectivity**

To effectively interpret results obtained from EEG recordings, many researchers have developed or employed various distinguished visualisation approaches. This section aims to review the typical visualisation methods for functional and effective connectivity and provide corresponding critical comparisons. The visualisation methods can be categorised into three groups: heat map, data statistics and head map.

Heat map generally employs an adjacency matrix to represent the quantification of brain connectivity, which performs well in showing the overall relationship between all pairs of available EEG channels. Chu et al. (2015) found a significant similarity between functional and structural networks within dysplasia patients. Regions of Interests (ROIs) for structural network analysis were selected to overlap with ROIs used for constructing the functional networks. The functional connectivity was quantified by cross-correlation and coherence and visualised by a heat map, as shown in Figure 2-2(a). This visualisation method was also used in pediatric epilepsy studies (Sargolzaei et al., 2015) and the analysis of spontaneous blood oxygen level-dependent signals (Chang et al., 2013).

The data statistic method tends to quantitatively compare targeted pairs and visualise significant ones by setting a threshold. Blinowska et al. (2017) demonstrated a statistical difference among a series of directed pairs in the contribution of Non-normalised Directed Transfer Function (NDTF) pairs values (Figure 2-2 (B)), indicating the most important connections for the discrimination between healthy elderly and Alzheimer's disease individuals (Blinowska et al., 2017). Similarly, PLI was estimated and visualised by a statistic method for patients with Fragile X Syndrome (FXS) (Van Der Molen et al., 2014), where the

healthy control group shows the brain connectivity difference in typical frequency ranges.

With a combination of brain connectivity and a head map, researchers were able to represent the distinction between health and disease and responses to external stimuli while demonstrating the importance of specific brain regions. Figure 2-2 (C) and (D) visualise brain functional connectivity for epilepsy diagnosis (Sargolzaei et al., 2015) and effective connectivity for human emotional responses to various types of music (Shahabi and Moghimi, 2016) respectively. Besides, Zhao et al. (2020) proposed a revised circular graph to visualise the functional connectivity quantification for the classification of Alzheimer's disease (Figure 2-2(E)), which demonstrates the potential contribution of specific regions to disease diagnosis.

Both heat map and head map methods can represent connectivity distributions reflecting an extent involvement of specific brain regions. The heat map employs an adjacency matrix to show interconnections between each electrode pair, while the head map helps isolate and visualise brain areas of interest that can be used in developing specific applications, such as the representation of connectivity differences for diagnostic purposes (i.e., brain network function in health and various neurological disorders). It is worth noting that the determination of an appropriated threshold is important for the implementation of an appropriate head map method, since too much or too little information may limit connectivity interpretation. Moreover, a heat map is usually used for visualising functional connectivity but not effective connectivity while the head map can reveal both types of connectivity. Another advantage of heat map is that it can generate appropriate inputs for deep learning models. For instance, Chen et al. (2019) used mutual information to build the heat map, which is then used as the inputs to the CNNs. The data statistics approach, compared with the other two methods, focuses more on quantification. To be more specific, this method tends to provide numerical comparison and select the most valuable connections by setting thresholds. Therefore, it is more suitable for studies that aim to reflect on data difference among a limited number of electrode pairs.

However, the overall topological connectivity characteristics cannot be represented, which limits the ability to infer the contribution of structural connectivity to the findings.

## **2.4 Applications and Machine Learning**

### **2.4.1 Practical application**

Increasing evidence exists that brain connectivity analysis, quantified by statistical dependence (e.g. correlation and coherence), information theory, and parametric and non-parametric causality analysis, can reveal the changes of the brain signal behaviour or patterns of neurological patients. Over the past few years, there was substantial developments and emergences of a great number of promising results for analysing and understanding different types of neurological disorders such as Alzheimer's disease (Bajo et al., 2015; delEtoile and Adeli, 2017; Engels et al., 2015; Sankari and Adeli, 2011), epilepsy (Clemens et al., 2013; Douw et al., 2010; Visani et al., 2010; Xie and Krishnan, 2013) and Parkinson's disease (Chaturvedi et al., 2019; Evangelisti et al., 2019; Handojoseno et al., 2013; Yuvaraj et al., 2016). In addition to neurological and neurophysiological studies and associated clinical applications, there is a wide range of applications where the research objectives are to assist human in understanding brain behaviour, for example, emotion recognition (Khosrowabadi et al., 2010; Lee and Hsieh, 2014; Shahabi and Moghimi, 2016), object recognition (Kaur et al., 2019; Supp et al., 2007; Tafreshi et al., 2019a), mental assessment (Al-Shargie et al., 2019; Cattai et al., 2018; Rathee et al., 2017) and biometric identification (Fraschini et al., 2019; Rocca et al., 2014; Wang et al., 2019).

In the meantime, the rapid progression of machine learning technology, applied in the field of brain connectivity has led to very significant developments, aiming to achieve a deeper and better understanding of brain network behaviours for the health group and patients with various neurological conditions. Table 2-2 Recent applications combining brain connectivity estimations with machine learning methods. shows a depiction of some recent applications of brain connectivity estimations supported by machine learning methods.

## 2.4.2 Hand-crafted vs. Imaging features

On one hand, state-of-the-art methods for EEG-based applications mostly apply a procedure that comprises hand-crafted features and traditional machine learning classifiers. For example, 'PCC, WC, MSC, PS and MI + SVM' is applied to object recognition (Tafreshi et al., 2019a), 'WC+LDA' is used for the diagnosis of Parkinson's disease-related dementia and Alzheimer's disease (Jeong et al., 2016), 'DTF + SVM' is used for the detection of brain responses to emotional music (Shahabi and Moghimi, 2016) and 'SL + SVM, LR and NB' is used for the diagnosis of major depressive disorder (Mumtaz et al., 2018). On the other hand, various visualisation methods provide robust features for deep learning in recent years. For instance, Saeedi et al. (2020) reconstructed a connectivity image using 2 connectivity measures (PDC and DTF) and eight frequency bands for the input of deep learning networks. The proposed achieved an accuracy of 99.24% in the case of major depressive disorder diagnosis.

**Table 2-2 Recent applications combining brain connectivity estimations with machine learning methods.**

| Applications                                                                                   | Estimation + ML method                                                       |
|------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------|
| Object recognition (Tafreshi et al., 2019a)                                                    | PCC, WC, MSC, PS and MI. + SVM                                               |
| Diagnosis of Parkinson's disease-related dementia and Alzheimer's disease (Jeong et al., 2016) | WC+ linear discriminant analysis (LDA)                                       |
| Prediction of Freezing of Gait in Parkinson's disease patients (Handojoseno et al., 2013)      | PCC +Multilayer Perceptron Neural Network and k-Nearest Neighbor classifier  |
| Emotion Recognition (Piho and Tjahjadi, 2020)                                                  | MI + SVM, naive Bayes (NB) classifier, and K-nearest neighbours (KNN)        |
| Detection of brain responses to emotional music (Shahabi and Moghimi, 2016)                    | DTF + SVM                                                                    |
| Discrimination between Alzheimer's patients and healthy individuals (Blinowska et al., 2017)   | DTF + artificial neural networks (ANNs)                                      |
| Depression diagnosis (Saeedi et al., 2020)                                                     | PDC and DTF + Long short-term memory and convolutional neural networks (CNN) |
| Attention-deficit/hyperactivity disorder identification (Chen et al., 2018)                    | MI + CNN                                                                     |

---

|                                                                   |                                                                                                            |
|-------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------|
| .., 2019)                                                         |                                                                                                            |
| Diagnosis of Alzheimer's disease (Zhao .., 2020)                  | ERR + KNN                                                                                                  |
| Diagnosis of major depressive disorder (Mumtaz 2018)              | .. SL + SVM, logistic regression (LR) and NB                                                               |
| Classification of autism spectrum disorder (Jamal 2014)           | .. PS + LDA and SVM                                                                                        |
| Speech categorization decisions (Al-Fahad .., 2019)               | PCC and graph network + SVM and LDA                                                                        |
| Transcranial magnetic stimulation monitoring (Gupta 2019)         | .. Coherence + principal component analysis (PCA) along sparse nonnegative matrix factorization (NMF)      |
| Detecting Disorders of Consciousness (Wang .., 2020)              | ensemble of SVMs + Power Spectral Density Difference (PSDD) incorporating with a recursive Cosine function |
| Sedation scale estimation (Sanz-García .., 2019)                  | PS + SVM                                                                                                   |
| Detecting Psycho-Physiological Insomnia (2015)                    | .. MI, PCC and MSC + NB, random forest, regression methods and nearest neighbour based methods             |
| Investigation of the effect of Clozapine therapy (Ravan .., 2015) | cross-power spectral density (CPSD) + fuzzy c-mean                                                         |
| Face perception tasks (Jamal .., 2015)                            | PLV + LDA and KNN                                                                                          |

---

## 2.5 Discussion

### 2.5.1 Connectivity estimation

Brain connectivity has shown significant potential in assessing brain function in patients with different neurological disorders and tracking various cognitive and emotional brain states for healthy participants. It is crucial to select the most appropriate connectivity measure, as widely distributed complex brain networks generate diverse topological signal processing and intercommunications in distinct conditions. Therefore, many novel methods have been developed from manifold aspects. This includes consideration of linear or nonlinear behaviors, the information in time, frequency or time-frequency domain, parametric or non-parametric measures, directed or undirected information.



### **2.5.2 Functional Connectivity vs. Effective Connectivity**

In the studies of effective connectivity, parametric methods have been extensively used to quantify directional interconnections among separated brain regions. The main advantage of parametric modelling of effective analysis is that the model structure is usually transparent and compact, with lagged signals being the model variables. The estimation of parametric models usually does not need a large number of samples, therefore parametric modelling allows to perform a transient or time-dependent connectivity analysis (Li *et al.*, 2019; Zhao *et al.*, 2012). On the other hand, functional connectivity captures statistical independence between distributed and spatially separated neuronal areas, using e.g. a bi-variate measure (Wang *et al.*, 2019). Functional connectivity is generally model-free, which decreases the need of setting strict assumptions on signals. However, the amount of data needs to be increased to meet the calculation demand of functional connectivity estimation.

### **2.5.3 Phase-based Connectivity vs. Amplitude-based Connectivity**

Brain connectivity can also be divided into phase-based connectivity and amplitude-based connectivity. Connectivity of the amplitude is usually estimated with correlation (Brookes *et al.*, 2011; Chang *et al.*, 2013; Hipp and Siegel, 2015; Wang *et al.*, 2020), while phase coupling is generally estimated with coherence-based methods and phase synchronisation (Bastos and Schoffelen, 2016; Chaturvedi *et al.*, 2019; Fagerholm *et al.*, 2020; Nolte *et al.*, 2004; Stam *et al.*, 2007). These aspects further capture different neural processes (He *et al.*, 2019; Siems and Siegel, 2020). As a result, the phase- and amplitude-based approaches give partially overlapping, partially differing results. At the same time, it has been shown that in noisy signals, phase and amplitude dynamics influence each other and the reliability of phase estimation inherently depends on the signal-to-noise ratio (SNR) and may generally be more accurate in the presence of higher signal amplitudes (He *et al.*, 2019).

#### **2.5.4 Precautions during analysis**

Volume conduction, an important issue when analyzing EEG recordings, has been shown to influence sensor-space connectivity analyses (Haufe et al., 2013), for example, phase coupling (Palva et al., 2018), Granger causality (Haufe et al., 2013), and correlation methods (Hipp and Siegel, 2015). To avoid the effect of volume conduction, some methods have been proposed, such as imaginary part of the coherency (van Mierlo et al., 2014), DTF (Kamifiski and Bfinowska, 1991) and PDC (Baccalá and Sameshima, 2001). Besides, calculating source-level connectivity from sensor-level EEGs could get rid of the influence caused by volume conduction or field propagation (He et al., 2019). This process is the so-called inverse problem (Diessen et al., 2015). In the meantime, some researchers suggest the interpretation of connectivity measures from sensor-level EEG recordings is not straightforward. Instead, source-level EEG is believed to be a reliable tool for measuring connectivity and it can be reconstructed from the scalp EEG (Moezzi and Goldsworthy, 2018).

#### **2.5.5 Role of Visualisation**

Visualisation plays a crucial role in the research of EEG brain connectivity. Specifically, it establishes the base for further practical and clinical uses of novel methods. It could not only improve the efficiency for other researchers to understand and evaluate the proposed approaches and produced results but also benefit experts in presenting and explaining the finding obtained from various advanced methods to people without basic knowledge of EEG signals and brain connectivity. The heat map, data statistics and head map generally conclude various popular visualisation methods and meet the distinct requirements and aims of researchers.

Besides, EEG study could be combined with appropriate visualisation methods to study potential functions, oscillations and intercommunications within a dynamic architecture of the human brain (Chen et al., 2013). The current medical instrument related to EEG focuses on oscillations and biomarkers within a signal channel (Ratti et al., 2017). Therefore, from the authors' point of view, development and innovations of visualisation have the potential to

promote the upgrading of EEG-related equipment, as well as explore more secrets within the human brain.

Another opportunity for brain connectivity visualisation is the combination with deep learning methods which has attracted increasing investigation for brain study. There is a significant amount of works to use a variety of deep learning methods in the understanding of EEG recording, but most of them use the raw data directly. The limitation of such approaches is the lack of transparency because they usually cannot reveal which brain region has the anomaly and how it leads to neurological disorders. There is increasing evidence that Convolution Neural Network (CNN) has superior performance to classify images by learning the spatial pattern of raw images automatically, instead of using handcrafted features. Therefore, it is promising to use the brain connectivity visualisation, in a form of images, to be fed into CNN in the classification. Essentially, in this approach, brain connectivity acts as handcrafted features, but in an image format instead of singular values. Chen proposed a general framework for the attention-deficit/hyperactivity disorder (ADHD) identification problem by combining an EEG-based brain network, estimated by MI, with CNN techniques (2019). The proposed framework achieved a convincing performance with an accuracy of 94.67% with the heat map visualisation. Besides functional connectivity, effective connectivity visualisation can also provide valuable inputs for CNN. For example, Saeedi (2020) constructed a connectivity image with 2 connectivity measures (PDC and DTF) and eight frequency bands as the input of deep learning networks. The experiments show that the CNN applied on the constructed image of effective connectivity achieves the best results with an accuracy of 99.24% in the case of major depressive disorder diagnosis.

### **2.5.6 Research Gaps and Future Direction**

Although significant outcomes in this area have been achieved, limitations and significant challenges remain. Firstly, brain network interactions are dynamic and may be time-varying (the associated signals are non-stationary), as phase synchronisation and phase scattering occur within the millisecond range (Varela

., 2001). Transient associations, usually highly nonlinear, among different brain regions have been observed (Sarrigiannis ., 2018, 2014). Even though recent years witnessed the development of brain connectivity techniques, there are limited methods that can analyse non-stationary and non-linear behaviours of brain networks. Exploring the hidden information within EEG signals is far beyond the capabilities of commonly available methods. Secondly, simplifying the interpretation of data-driven methods is a challenge. While traditional parametric models extended from AR models, e.g. autoregressive moving average with exogenous input (ARMAX) and nonlinear autoregressive moving average with exogenous input (NARMAX) models (Gu ., 2021) and dynamic causal model (David ., 2006) are parsimonious and transparent, whose individual model terms have a clear explanation and can be linked back to the original neural system, other types of parametric models e.g. nonlinear kernel models (Shen ., 2016) tend to be complicated and it becomes difficult to relate the models back to the underlying system due to the lack of transparency. Interpretation of the outcomes of non-parametric methods e.g. neural networks (Abbasvandi and Nasrabadi, 2019; Saeedi ., 2020) require sufficient mathematical knowledge and probably some specific and professional modelling skills. To be more specific, understanding neural networks requires a grasp of key mathematical concepts like calculus, linear algebra, and probability theory. This knowledge helps in comprehending how these networks process data and optimize their parameters. Effective interpretation also relies on skills in designing, implementing, and fine-tuning neural network models. This involves selecting the right network architecture, optimizing hyperparameters, and addressing issues like overfitting. Therefore, the overall performance of such methods heavily depends on the experience and knowledge of the end-users, typically clinicians or doctors if the tools are developed for clinical applications. Thirdly, there is no real-time feedback of brain functional connectivity during data acquisition in current EEG systems. As a result, scientists (e.g., neurologists or neurophysiologists) cannot flexibly focus on the examination of specific network areas in real-time or determine where the most significant abnormalities state or behaviour takes place. Some

researchers have started to integrate EEG with augmented reality (Mercier-Ganady et al., 2014; Vortmann et al., 2019), but they only focused on the information obtained from a single channel. There is a lack of applications to track and visualise time-varying brain connectivity in real-time. Fourthly, from the authors' point of view, there are a lot of visualisation and imaging techniques in neuroscience and some of them used for other imaging modalities also have the potential to visualise EEG brain connectivity. For instance, BiImage Suite provides an attractive and interpretative visualisation for fMRI functional connectivity (Finn et al., 2015; Shen et al., 2018).

Therefore, even though brain connectivity techniques have been significantly developed over the past years, more validation and improvements/developments are still required. Table 2-3 suggests potential directions and possible strategies, which will help deepen our understanding of brain functional and effective connectivity from a variety of novel perspectives, as well as how to enhance their availability and reliability.

**Table 2-3 The potential directions of EEG brain connectivity research**

| Future Direction | Purpose or Strategy                                                                                                                                                                                                 |
|------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Novel Estimation | Extract more valuable information from EEG signals by robust brain connectivity methods, especially non-stationary and non-linear intercommunications                                                               |
| Interpretability | Design appropriate visualisation methods to reduce the difficulty of understanding the actual implication of brain connectivity estimation and its outcomes, such as disease diagnosis and brain activity analysis. |
| Universality     | Build a large dataset covering people with different ages, genders and diseases conditions to develop and evaluate universal brain connectivity methods                                                             |

|                       |                                                                                                                                                                                                  |
|-----------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Real-time Research    | Establish a real-time sensor and monitoring system based on advanced brain connectivity estimation and visualisation approaches, capturing dynamic neuro-connectivity and assisting observation. |
| Improved diagnosis    | Using the visualisation of estimated brain functional or effective connectivity as the input of deep learning method to maintain the transparency and improve the classification accuracy        |
| Application Extension | Pursue a deeper understanding of the brain network and explore potential fields where EEG brain connectivity can be used.                                                                        |

## 2.6 Conclusions

This Chapter reviewed recent studies in estimating brain functional and effective connectivity using EEG with associated visualisation as well as their applications. It is observed that functional connectivity is usually measured using non-parametric methods while effective connectivity is measured by parametric methods. In comparison to parametric methods, non-parametric methods require relatively fewer assumptions, but they trade in the need for larger data sets. Nonlinear dynamics and the spatiotemporal characteristics of brain networks hidden within EEG recordings have the potential to be identified using appropriate functional and effective connectivity methods. Although progress has been made the potential of those methods is far from being fully explored. It is expected that future research in this area will focus on better integrating different methods in particular machine learning approaches to improve the accuracy of disease diagnosis meanwhile increasing transparency. Real-time visualisation of brain connectivity during the data acquisition stage can be introduced to improve the data quality, and help scientists better identify which areas of the brain exhibit connectivity deficits under various neurological conditions.

## 2.7 Appendix The conceptual and mathematical details for calculating brain connectivity.

### 2.7.1 Functional Connectivity

#### 2.7.1.1 Correlation

Pearson correlation coefficient is able to evaluate the linear interdependency between two signals in the time domain and it ranges from -1 to +1. The correlation coefficient between signal  $X$  and  $Y$  can be defined as follows (van Mierlo et al., 2014):

$$\rho_{xy} = \frac{E[(x - \mu_x)(y - \mu_y)]}{\sigma_x \sigma_y} \quad (2-1)$$

where  $E$  is the expected value,  $\mu_x$  and  $\mu_y$  are the mean values and  $\sigma_x$  and  $\sigma_y$  are the standard deviations of  $X$  and  $Y$  time series.

Cross-correlation differs from Pearson correlation coefficient since it is a function with respect to time lag (Ibrahim et al., 2018; van Mierlo et al., 2014)

$$\rho_{xy}(\tau) = \frac{E[(x_n - \mu_x)(y_{n+\tau} - \mu_y)]}{\sigma_x \sigma_y} \quad (2-2)$$

#### 2.7.1.2 Magnitude Squared Coherence

Magnitude squared coherence (MSC) is a linear method to estimate the interconnections between two signals in the frequency domain calculated by PSD. The MSC of signal  $X$  and  $Y$  can be written as (Mierlo et al., 2014; van Mierlo et al., 2014; Wendling et al., 2009)

$$MSC_{xy}(f) = C_{xy}^2 = \frac{S_{xy}(f)^2}{|S_{xx}(f)| \times |S_{yy}(f)|} \quad (2-3)$$

where  $S_{xx}(f)$  and  $S_{yy}(f)$  are the PSD of signal  $X$  and  $Y$  respectively, and  $S_{xy}(f)$  is the cross PSD at frequency  $f$ .

To avoid the volume conduction effects, the imaginary part of the coherency is employed by a few studies instead of looking at the magnitude and phase of the coherency (Diessen et al., 2015; Ewald et al., 2012; van Mierlo et al., 2014).

$$C_{xy} = RC_{xy} + IC_{xy} \quad (2-4)$$

### 2.7.1.3 Wavelet Coherence

Wavelet coherence is generally acknowledged as a qualitative estimator that can represent the time-varying relations in the time-frequency domain between signals with a specifically determined scale (Labat, 2005; Tafreshi et al., 2019a). To be more specific, the wavelet transforms of a signal  $x$  is a function of both time and frequency, which is defined as the convolution of the input with a Wavelet family  $\theta(u)$ :

$$W_x(t, f) = \int_{-\infty}^{+\infty} x(u) \theta_{t,f}(u) du \quad (2-5)$$

Wavelet cross-spectrum around time  $t$  and frequency  $f$ , given input signals  $x$  and  $y$ , can be derived by the Wavelet transforms of  $x$  and  $y$ :

$$CW_{xy}(t, f) = \int_{t-\delta/2}^{t+\delta/2} W_x(\tau, f) W_y(\tau, f) d\tau \quad (2-6)$$



where  $*$  defines the complex conjugate and  $\delta$  is assumed as a frequency-depending scalar.

Wavelet Coherence (WC) at the time  $t$  and frequency  $f$  is derived as:

$$WC_{xy}(t, f) = \frac{|CW_{xy}(t, f)|}{|CW_{xx}(t, f) \times CW_{yy}(t, f)|^{1/2}} \quad (2-7)$$

#### 2.7.1.4 Short Time Fourier Coherence

Short-Time Fourier Transform (STFT) is also implemented in the time-frequency domain instead of the classical fast Fourier transform approach to calculate coherence, and coherence may be calculated around a number of time instants. This technique produces the so-called “coherogram”, which forms a three-dimensional matrix of time and frequency against coherence. However, stationarity is still required within each time interval for which coherence is calculated, meaning that in practice one should carefully decide on the optimal section length over which each coherence estimate is measured. Window length and overlapping within each coherence estimate affect the resolution of the measure (Delgado-Restituto et al., 2019; V. Sakkalis, 2011).

#### 2.7.1.5 Phase synchronization

Phase synchronization (PS) assumes two oscillation systems without amplitude synchronization can have phase synchronization. The phase locking value (PLV) and the phase lag index (PLI) are high-frequently used to obtain the strength of phase synchronization (van Mierlo et al., 2014). The instantaneous phase of a signal  $x$  is given by:

$$\phi_x(t) = \arctan \frac{\tilde{x}(t)}{x(t)} \quad (2-8)$$

where  $\tilde{x}(t)$  is the Hilbert transform of  $x(t)$  which is defined as:

$$\tilde{x}(t) = \frac{1}{\pi} PV \int_{-\infty}^{+\infty} \frac{x(\tau)}{t - \tau} d\tau \quad (2-9)$$

where  $PV$  refers to the Cauchy principal value. The Phase locking value for two signals is defined as:

$$PLV = \left| \frac{1}{N} \sum_{j=0}^{N-1} e^{j(x(j \cdot t) - y(j \cdot t))} \right| \quad (2-10)$$

where  $t$  defines the sampling period and  $N$  indicates the sample number of each signal (Mormann et al., 2000). The range of PLV is from 0 to 1, where 0 shows a lack of synchronization and 1 indicates strict phase synchronization.

The phase lag index (PLI) captures the asymmetry of the distribution of phase differences between two signals and is calculated based on the relative phase difference between the two signals (Borrego-Salcido et al., 2019; Wendling et al., 2009).

$$PLI = |E[\text{sign}(\phi(t))]| \quad (2-11)$$

The resulting value lies in the interval  $[0, 1]$ , where a higher value indicates more phase synchrony.

### 2.7.1.6 Synchronization likelihood

Synchronization likelihood (SL) is a measure of multivariate synchronization to investigate the interactions between nonlinear dynamical systems without any knowledge about the governing equations (C.J. Stam and Dijk, 2002). For an M-channel time series  $x_{k,n}$  ( $k$  denotes channel number ( $k = 1, 2, \dots, M$ )), embedded vectors  $X_{k,n}$  are reconstructed with time-delay embedding:

$$X_{k,n} = (x_{k,n}, x_{k,n+\tau}, \dots, x_{k,n+(m-1)\tau}) \quad (2-12)$$

where  $m$  is the embedding dimension and  $\tau$  denotes time lag. The estimated probability that embedded vectors  $X_n$  are closer to each other than a distance  $\varepsilon$  is (Wendling et al., 2009):

$$P_{k,n}^{\varepsilon} = \frac{1}{2(w_2 - w_1)} \sum_{\substack{j=1 \\ w_1 < |i-j| < w_2}}^N \theta(\varepsilon - |X_{k,i} - X_{k,j}|) \quad (2-13)$$

where  $| \cdot |$  is the Euclidean distance;  $\theta$  stands for the Heaviside step function,  $w_1$  is the Theiler correction and  $w_2$  determines the length of the sliding window. Letting  $P_{x,n}^{\varepsilon} = P_{ref}$ , where  $P_{ref} \geq 1$ , the number of channels  $H_{i,j}$  where the embedded vectors  $X_{k,i}$  and  $X_{k,j}$  will be closer together than this critical distance  $\varepsilon_{k,i}$ :

$$H_{i,j} = \sum_{k=1}^M \theta(\varepsilon_{k,i} - |X_{k,i} - X_{k,j}|) \quad (2-14)$$

Synchronization likelihood at time  $n$  can be obtained by averaging over all values of  $j$

$$SL_n = \frac{1}{2P_{ref}(w_2 - w_1)} \sum_{\substack{j=1 \\ w_1 < |i-j| < w_2}}^N \left( \frac{H_{i,j}}{M} - \frac{1}{1} \right) \quad (2-15)$$

All aforementioned measures are normalized between 0 and 1; the 0 value means that the two signals are completely independent. On the opposite, the 1 value means that the two signals are completely synchronized(Wendling et al., 2009).

### 2.7.1.7 Mutual information

According to information theory, Mutual information (MI) of two random variables  $X$  and  $Y$  shows how a random variable is informative for the other one. Let,  $P(x)$  and  $P(y)$  be the probability distributions of random variables  $X$  and  $Y$ , respectively. The entropy of  $X$  and  $Y$  is defined as(Tafreshi et al., 2019a):

$$H(X) = - \sum_{j=1}^N P(x_j) \log_b(P(x_j)) \quad (2-16)$$

$$H(Y) = - \sum_{j=1}^N P(y_j) \log_b(P(y_j)) \quad (2-17)$$

where  $N$  defines window length.

$H(Y|X)$  and  $H(X, Y)$  are conditional entropy and joint entropy between  $X$  and  $Y$ , defined respectively as:

$$H(X, Y) = - E_X[E_Y[\log_b P(X, Y)]] \quad (2-18)$$

where  $E_Y$  is the expected value function.

$$H(Y|X) = -E_X[E_Y[\log_b P(Y|X)]] \quad (2-19)$$

MI of two random variables  $X$  and  $Y$  is computed as follows (Khosrowabadi et al., 2010):

$$MI(X, Y) = H(X) + H(Y) - H(X, Y) = H(X) - H(X|Y) = H(Y) - H(Y|X) \quad (2-20)$$

$MI(X, Y) = 0$  if and only if random variables  $X$  and  $Y$  are statistically independent.

## 2.7.2 Effective Connectivity

### 2.7.2.1 Granger Causality

Granger causality can estimate effective interactions from time-series data. Granger causality implements a statistical, predictive notion of causality whereby causes precede and help predict their effects (Seth et al., 2015). It is defined in both the time and frequency domains, and it allows for the conditioning out of common causal influences. One time series is said to Granger cause a second one if the inclusion of the past values of the first into the modelling of the second significantly reduces the variance of the modelling error. According to Granger causality, if the past values of  $X$  contain information that helps to predict  $Y$  above and beyond the information contained in past values of  $Y$  alone then signal  $X$  “Granger-causes” signal  $Y$ . The Granger causality from signal  $X$  to  $Y$  and the one from signal  $Y$  to  $X$  can be investigated separately. (van Mierlo et al., 2014).

### 2.7.2.2 Partial directed coherence

PDC is a method that quantifies the relation between two among  $N$  signals, avoiding volume conduction by estimating influences of all other signals (Tafreshi et al., 2019a). PDC improves the concept of Partial Coherence by estimating causal influences. This method is estimated on multivariate autoregressive (MVAR).

$X(t)$  is a set of estimated signals from  $N$  recording channels:

$$X = [x_1(t), x_2(t), \dots, x_N(t)]^T \quad (2-21)$$

The MVAR process is an expressive description of the data set  $X$ :

$$\sum_{r=0}^P A(r)X(t-r) = E(t) \quad (2-22)$$

In this model  $E(t)$  is a zero-mean multivariate uncorrelated white noise vector.  $A(r)$  are the autoregressive coefficients matrix and its elements  $a_{ij}(r)$  indicates the influence of  $X_j(t-r)$  on  $X_i(t)$  and  $p$  represents the border of the model.

PDC from the  $i$ th channel to the  $j$ th channel at frequency  $f$  is defined as follows:

$$PDC_{ij}(f) = \frac{\bar{A}_{ij}(f)}{\sqrt{\sum_{m=1}^N \bar{A}_{mj}(f)\bar{A}_{mj}(f)}} \quad (2-23)$$

where  $A_{ij}(f)$  is a frequency-domain description of  $a_{ij}(r)$  (Baccalá and Sameshima, 2001; Granger and Aug, 2007)

$$\bar{A}_{ij}(f) = \begin{cases} 1 - \sum_{r=1}^P a_{ij}(r) e^{-j2\pi fr}, & \text{if } i = j \\ \sum_{r=1}^P a_{ij}(r) e^{-j2\pi fr}, & \text{otherwise} \end{cases} \quad (2-24)$$

### 2.7.2.3 Directed transfer function

The directed transfer function (DTF) is also based on the concept of Granger causality estimated on MVAR, which models all signals simultaneously (Kamifiski and Bfinowska, 1991).  $X(t)$  is a set of estimated signals from  $N$  recording channels:

$$X = [x_1(t), x_2(t), \dots, x_N(t)]^T \quad (2-25)$$

The MVAR process is an expressive description of the data set  $X$ :

$$\sum_{r=0}^p A(r)X(t-r) = E(t) \quad (2-26)$$

In this model  $E(t)$  is a zero-mean multivariate uncorrelated white noise vector with  $A(0) = 1$ .  $A(1), A(2), \dots, A(p)$  is coefficients matrix and  $p$  represent the border of the model. Equation (20) can be transformed into the frequency domain which is defined as:

$$A(f)X(f) = E(f) \quad (2-27)$$

where

$$A(f) = \sum_{r=0}^p A(r)e^{-j2\pi f \cdot r} \quad (2-28)$$

So,  $X(f)$  can be obtained by

$$X(f) = A^{-1}(f)E(f) = H(f)E(f) \quad (2-29)$$

$H(f)$  is the transfer function of the system and its elements  $H_{ij}(f)$  indicate the causal influence from the  $j$ th input to the  $i$ th output at frequency  $f$ .

The DTF is defined as:

$$\gamma_{ij}^2 = |H_{ij}(f)|^2 \quad (2-30)$$

Normalized DTF is defined as:

$$\gamma_{ij}^2(f) = \frac{|H_{ij}(f)|^2}{\sum_{m=1}^N |H_{im}(f)|^2} \quad (2-31)$$

where  $\gamma_{ij}^2(f)$  describes the influence ratio of the  $j$ th channel-related cortical area on the  $i$ th channel-related cortical area with respect to the influence of all estimated cortical signals (Babiloni et al., 2005). This is an important difference between DTF and PDC since DTF is normalized for the structure that receives the signal, whereas PDC is normalized for the structure that sends the signal (Baccalá and Sameshima, 2001).

#### 2.7.2.4 Structural equation modelling

A set of linear structural equations are employed by the structural equation modelling (SEM), exploring causal intercommunications among the observed variables and parameters. In terms of neural systems, a measure of covariance represents the degree to which the activities of two or more regions are related. The SEM technique estimates the parameters by minimizing the difference between the observed covariances and those implied by a structural or path model (Astolfi et al., 2004; Babiloni et al., 2003).

#### 2.7.2.5 Dynamic Causal Modelling

The aim of dynamic causal modelling (Friston et al., 2003) is to make inferences about the coupling among brain regions or sources and how that coupling is



influenced by experimental factors. DCM uses the notion of effective connectivity, defined as the influence one neuronal system exerts over another. The central idea behind DCM is to treat the brain as a deterministic nonlinear dynamical system that is subject to inputs and produces outputs. In neuroimaging, DCM starts with a reasonably realistic neuronal model of interacting cortical regions. This model is then supplemented with a forward model of how neuronal activity is transformed into measured responses. This enables effective connectivity of the neuronal model to be estimated from observed data (David et al., 2006; Yvert et al., 2012).

#### 2.7.2.6 Transfer Entropy

Transfer entropy (TE) is an alternative measure of effective connectivity based on information theory. TE does not require a model of the interaction and is inherently nonlinear (Vicente et al., 2011a). TE for two observed time series  $x_t$  and  $y_t$  can be written as

$$TE(X \rightarrow Y) = \sum_{y_{t+u}, y_t^{d_y}, x_t^{d_x}} p(y_{t+u}, y_t^{d_y}, x_t^{d_x}) \log \frac{p(y_{t+u} | y_t^{d_y}, x_t^{d_x})}{p(y_{t+u} | y_t^{d_y})}$$

where  $t$  is a discrete-valued time-index and  $u$  denotes the prediction time, a discrete-valued time-interval.  $y_t^{d_y}$  and  $x_t^{d_x}$  are  $d_y$ - and  $d_x$ -dimensional delay vectors (Vicente et al., 2011a).

#### 2.7.2.7 Error Reduction Ratio (ERR)

Error Reduction Ratio (ERR) is a parametric method in the time domain, which indicates how much of the change in the system response, in percentage, can be accounted for by including the relevant model terms. This capability allows the computing of the contribution of linear and nonlinear terms independently without fitting the full nonlinear model, where parameter estimation and model validation are required. Consider a linear regression

$$Y = P \quad (2-33)$$

where

$$Y = \begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(M) \end{bmatrix}, P = \begin{bmatrix} P^T(1) \\ P^T(2) \\ \vdots \\ P^T(M) \end{bmatrix}, \quad = \begin{bmatrix} (1) \\ (2) \\ \vdots \\ (M) \end{bmatrix} \quad (2-34)$$

and  $P^T(k) = (p_1(k), p_2(k), \dots, p_N(k))$ .  $p_1, p_2, \dots, p_N$  are the model terms selected from the candidate term set. Matrix  $P$  can be decomposed as  $P = W \times A$  where

$$W = \begin{bmatrix} w_1(1) & w_2(1) & \dots & w_N(1) \\ w_1(2) & w_2(2) & \dots & w_N(2) \\ \vdots & \vdots & \ddots & \vdots \\ w_1(M) & w_2(M) & \dots & w_N(M) \end{bmatrix} \quad (2-35)$$

is an orthogonal matrix because

$$W^T W = \text{Diag} \left[ \sum_{t=1}^M w_1^2(t), \dots, \sum_{t=1}^M w_N^2(t) \right] \quad (2-36)$$

and  $A = \{a_{ij}\}$  is an upper triangular matrix with unity diagonal elements

$$A = \begin{bmatrix} 1 & a_{12} & \dots & a_{1N} \\ & 1 & & \dots \\ & & 1 & a_{N-1N} \\ & & & 1 \end{bmatrix} \quad (2-37)$$

Equation (2-33) is then rewritten as

$$Y = WG, \quad (2-38)$$

where  $G = A^{-1} = [g_1 \ g_2 \ \dots \ g_N]^T$ . Equation (2-38) is now ready to represent the relation between  $Y$  and  $G$ .

The importance of each model term to the variation of the system output is then estimated. Initially, set values  $a_{ij} = 0$  for  $i \neq j$  ( $A$  then becomes an identity matrix), so  $w_1(k) = p_1(k)$ , and calculate  $g_1$  as

$$g_1 = \frac{\sum_{k=1}^M w_1(k)y(k)}{\sum_{k=1}^M w_1^2(k)}. \quad (2-39)$$

For  $j = 2, 3, \dots, M$ , set  $a_{jj} = 1$  and then calculate

$$a_{ij} = \frac{\sum_{k=1}^M w_i(k)p_j(k)}{\sum_{k=1}^M w_i^2(k)}, \quad (2-40)$$

where  $i = 1, 2, \dots, j-1$ . Next, the algorithm calculates

$$w_j(k) = p_j(k) + \sum_{i=1}^{j-1} a_{ij}w_i(k) \quad (2-41)$$

and

$$g_j = \frac{\sum_{k=1}^M w_j(k)y(k)}{\sum_{k=1}^M w_j^2(k)}. \quad (2-42)$$

The ERR value for each term  $p_i$  is finally defined as

$$ERR_i = \frac{\hat{g}_i^2 \sum_{k=1}^H x_i^2(k)}{\sum_{k=1}^H y^2(k)}. \quad (2-43)$$

Values of ERR range from 0% to 100%. The larger the value of ERR, the higher the dependence between this term and the output. To stop the search procedure and determine the number of significant terms  $N$ , a criterion called Penalised Error-to-Signal Ratio (PESR) is used (Zhao et al., 2012). It can be written as

$$PESR_m = \frac{1}{\left(1 - \frac{\lambda m}{H}\right)^2} \left(1 + \sum_{i=1}^m ERR_i\right) \quad (2-44)$$

This criterion is introduced to monitor the search procedure, where  $m$  denotes the index of the selected terms. The search procedure stops when  $PESR_m$  achieves a local minimum. In this Chapter, the value of  $\lambda$  was chosen as 8.

To calculate the contribution of the input to the output, the sum of ERR of all selected terms, denoted by  $SERR$ , is calculated by

$$SERR = \sum_{i=1}^N ERR_i. \quad (2-45)$$

Note  $N$  is the number of the selected terms, not the number of total candidate terms. The value of  $SERR$  ( $0 \leq SERR \leq 1$ ) describes the percentage explained by the identified model to the system output. If the considered inputs can fully explain the variation of the system output, the value of  $SERR$  is equal to 100%.

The linearity of connectivity is represented by the sum of ERR of the linear terms, and it can be computed by

$$ERR_L(y, u) = \frac{\sum_{b=1}^N (ERR_b | p_b, C_1)}{SERR}. \quad (2-46)$$

The nonlinearity of connectivity is represented by the sum of ERR of the nonlinear terms, and it can be computed by

$$ERR_N(y, u) = \frac{\sum_{b=1}^N (ERR_b | p_b, C_1)}{SERR}. \quad (2-47)$$

## 2.8 Reference

Abbasvandi, Z., Nasrabadi, A.M., 2019.

Comput. Biol. Med. 110, 93–107.  
<https://doi.org/10.1016/j.compbiomed.2019.05.012>

Akbarian, B., Erfanian, A., 2020.

. Biomed. Signal Process. Control 59, 101878.  
<https://doi.org/10.1016/j.bspc.2020.101878>

Al-Fahad, R., Yeasin, M., Bidelman, G.M., 2019.

. bioRxiv. <https://doi.org/10.1101/686048>

Al-Shargie, F., Tariq, U., Hassanin, O., Mir, H., Babiloni, F., Al-Nashash, H., 2019.

. Brain Sci. 9. <https://doi.org/10.3390/brainsci9120363>

Altenburg, J., Vermeulen, R.J., Strijers, R.L.M., Fetter, W.P.F., Stam, C.J., 2003.

Clin. Neurophysiol. 114, 50–55. [https://doi.org/10.1016/S1388-2457\(02\)00322-X](https://doi.org/10.1016/S1388-2457(02)00322-X)

Amin, H.U., Yusoff, M.Z., Ahmad, R.F., 2020.

. Biomed. Signal Process. Control 56. <https://doi.org/10.1016/j.bspc.2019.101707>

Angelakis, E., Stathopoulou, S., Frymiare, J.L., Green, D.L., Lubar, J.F., Kounios, J., 2007.

. Clin. Neuropsychol. 21, 110–129. <https://doi.org/10.1080/13854040600744839>

Astolfi, L., Cincotti, F., Mattia, D., Salinari, S., Babiloni, C., Basilisco, A., Maria, P., Ding, L., Ni, Y., He, B., 2004.

22, 1457–1470.

<https://doi.org/10.1016/j.mri.2004.10.006>

. J. Med. Syst.

<https://doi.org/10.1007/s10916-015-0219-1>

Aydore, S., Pantazis, D., Leahy, R.M., 2013.

. Neuroimage 74, 231–244.

Babiloni, C., Del Percio, C., Bordet, R., Bourriez, J.L., Bentivoglio, M., Payoux, P., Derambure, P., Dix, S., Infarinato, F., Lizio, R., Triggiani, A.I., Richardson, J.C., Rossini, P.M., 2013.

. Clin. Neurophysiol. 124, 837–850.

<https://doi.org/10.1016/j.clinph.2012.09.017>

Babiloni, C., Lizio, R., Marzano, N., Capotosto, P., Soricelli, A., Triggiani, A.I., Cordone, S., Gesualdo, L., Del Percio, C., 2016.

. Int. J. Psychophysiol. 103, 88–102.

<https://doi.org/10.1016/j.ijpsycho.2015.02.008>

Babiloni, F., Cincotti, F., Babiloni, C., Carducci, F., Mattia, D., Astolfi, L., Basilisco, A., Rossini, P.M., Ding, L., Ni, Y., Cheng, J., Christine, K., Sweeney, J., He, B., 2005.

24, 118–131. <https://doi.org/10.1016/j.neuroimage.2004.09.036>

Babiloni, F., Cincotti, F., Basilisco, A., Maso, E., Bufano, M., Babiloni, C., Carducci, F., Rossini, P., Cerutti, S., Rubin, D.B.D., 2003.

. Int. IEEE/EMBS Conf. Neural Eng. NER

2003-Janua, 79–82. <https://doi.org/10.1109/CNE.2003.1196760>

- Baccalá, L.A., Sameshima, K., 2001.  
*Psychol. Biol. Cybern.* 84, 463–474.  
<https://doi.org/10.1007/PL00007990>
- Bajo, R., Pusil, S., López, M.E., Canuet, L., Pereda, E., Osipova, D., Maestú, F., Pekkonen, E., 2015.  
*Sci. Rep.* 5, 1–6.  
<https://doi.org/10.1038/srep09748>
- Ball, G., Aljabar, P., Arichi, T., Tusor, N., Cox, D., Merchant, N., Nongena, P., Hajnal, J. V., Edwards, A.D., Counsell, S.J., 2016.  
*Neuroimage* 124, 267–275. <https://doi.org/10.1016/j.neuroimage.2015.08.055>
- Bastos, A.M., Schoffelen, J.M., 2016.  
*Front. Syst. Neurosci.* 9, 1–23. <https://doi.org/10.3389/fnsys.2015.00175>
- Battaglia, D., Brovelli, A., 2019.  
*Cogn. Neurosci.*
- Bedo, N., Ribary, U., Ward, L.M., 2020.  
*Pone* 9, 1–24.  
<https://doi.org/10.1371/journal.pone.0088940>
- Benwell, C.S.Y., Davila-Pérez, P., Fried, P.J., Jones, R.N., Trivison, T.G., Santarnecchi, E., Pascual-Leone, A., Shafi, M.M., 2020.  
*Neurobiol. Aging* 85, 83–95.  
<https://doi.org/10.1016/j.neurobiolaging.2019.10.004>
- Blinowska, K.J., Rakowski, F., Kaminski, M., De Vico Fallani, F., Del Percio, C., Lizio, R., Babiloni, C., 2017.  
*Clin. Neurophysiol.* 128, 667–680.  
<https://doi.org/10.1016/j.clinph.2016.10.002>

Boersma, M., Smit, D.J. a., Boomsma, D.I., de Geus, E.J. c., delemarre-Van de Waal, H.A., Stam, C.J., 2013.

. Brain Connect. 3, 50–60.

<https://doi.org/10.1089/brain.2012.0106>

Borrego-Salcido, C., Juárez-Del-Toro, R., Cruz.Aké, S., 2019.

. Rev. Mex.

Econ. y Finanz. 15, 37–55. <https://doi.org/10.21919/remef.v15i1.413>

Brookes, M.J., Hale, J.R., Zumer, J.M., Stevenson, C.M., Francis, S.T., Barnes, G.R., Owen, J.P., Morris, P.G., Nagarajan, S.S., 2011.

. Neuroimage

56, 1082–1104. <https://doi.org/10.1016/j.neuroimage.2011.02.054>

Brown, H.R., Friston, K.J., 2012.

. Neuroimage 63, 223–231.

<https://doi.org/10.1016/j.neuroimage.2012.06.044>

C.J. Stam, Dijk, B.W. van, 2002.

. Phys. D

Nonlinear Phenom. 20, 94–97.

Cao, J., Grajcar, K., Shan, X., Zhao, Y., Zou, J., Chen, L., Li, Z., Grunewald, R., Zis, P., Marco, M. De, Unwin, Z., Blackburn, D., Sarrigiannis, P.G., 2021a.

. Biomed. Signal Process.

Control 67, 102554. <https://doi.org/10.1016/j.bspc.2021.102554>

Cao, J., Grajcar, K., Shan, X., Zhao, Y., Zou, J., Chen, L., Li, Z., Grunewald, R., Zis, P., Marco, M. De, Unwin, Z., Blackburn, D., Sarrigiannis, P.G., 2021b.

. Biomed. Signal Process.

Control 67, 102554. <https://doi.org/10.1016/j.bspc.2021.102554>



Cao, J., Zhao, Y., Shan, X., Wei, H., Guo, Y., Chen, L., Erkoyuncu, J.A., Sarrigiannis, P.G., 2022.

. Hum. Brain Mapp. 43, 860–879.

<https://doi.org/10.1002/hbm.25683>

Cattai, T., Colonnese, S., Corsi, M.C., Bassett, D.S., Scarano, G., De Vico Fallani, F., 2018.

. Eur. Signal Process. Conf. 2018-Septe, 1377–1381.

<https://doi.org/10.23919/EUSIPCO.2018.8553000>

Chang, C., Liu, Z., Chen, M.C., Liu, X., Duyn, J.H., 2013.

. Neuroimage 72, 227–236.

<https://doi.org/10.1016/j.neuroimage.2013.01.049>

Chaturvedi, M., Bogaarts, J.G., Kozak (Cozac), V. V., Hatz, F., Gschwandtner, U., Meyer, A., Fuhr, P., Roth, V., 2019.

. Clin. Neurophysiol. 130, 1937–1944.

<https://doi.org/10.1016/j.clinph.2019.07.017>

Chen, G., Chen, M., Hong, G., Lu, Y., Zhou, B., Gao, Y., 2020.

. Energies 13.

<https://doi.org/10.3390/en13040888>

Chen, H., Song, Y., Xiaoli, L., 2019.

. Neurocomputing 356,

83–96. <https://doi.org/10.1016/j.neucom.2019.04.058>

Chen, J., Ros, T., Gruzelier, J.H., 2013.

868, 852–868.

<https://doi.org/10.1002/hbm.21475>

Chiang, J., Wang, Z.J., McKeown, M.J., 2011.

IEEE Trans. Signal Process. 60, 453–465.

Chriskos, P., Frantzidis, C.A., Gkivogkli, P.T., Bamidis, P.D., Kourtidou-Papadeli, C., 2018.

. Front. Hum. Neurosci. 12, 1–13.  
<https://doi.org/10.3389/fnhum.2018.00110>

Clayden, J.D., 2013.

. Funct. Neurol. 28, 197–203.  
<https://doi.org/10.11138/FNeur/2013.28.3.197>

Clemens, B., Puskás, S., Besenyei, M., Spisák, T., Opposits, G., Hollódy, K., Fogarasi, A., Fekete, I., Emri, M., 2013.

. Epilepsy Res. 106, 357–369.  
<https://doi.org/10.1016/j.epilepsyres.2013.06.017>

Cohen, M.X., 2017a.

Trends Neurosci. 40, 208–218.

Cohen, M.X., 2017b.

. Elife 6, e21792.

Cohen, M.X., 2014. Analyzing neural time series data: theory and practice. MIT press.

Colclough, G.L., Woolrich, M.W., Tewarie, P.K., Brookes, M.J., Quinn, A.J., Smith, S.M., 2016.

Neuroimage 138, 284–293. <https://doi.org/10.1016/j.neuroimage.2016.05.070>

Cvetkovic, D., Derya, E., Cosic, I., 2008.

18, 861–874. <https://doi.org/10.1016/j.dsp.2007.05.009>

David, O., Kiebel, S.J., Harrison, L.M., Mattout, J., Kilner, J.M., Friston, K.J., 2006.

Neuroimage 30, 1255–1272. <https://doi.org/10.1016/j.neuroimage.2005.10.045>

de Cheveigné, A., 2010. . J. Neurosci. Methods 189, 113–120.

delEtoile, J., Adeli, H., 2017. . Neuroscientist 23, 616–626.  
<https://doi.org/10.1177/1073858417702621>

Delgado-Restituto, M., Romaine, J.B., Rodríguez-Vázquez, Á., 2019.

. IEEE Trans. Biomed. Circuits Syst. 13, 957–970.  
<https://doi.org/10.1109/TBCAS.2019.2931799>

Dhiman, R., Saini, J.S., Priyanka, 2014. . Appl. Soft Comput. J. 19, 8–17. <https://doi.org/10.1016/j.asoc.2014.01.029>

Diessen, E. Van, Numan, T., Dellen, E. Van, Kooi, A.W. Van Der, Boersma, M., Hofman, D., Lutterveld, R. Van, Dijk, B.W. Van, Straaten, E.C.W. Van, Hillebrand, A., Stam, C.J., 2015.

. Clin. Neurophysiol. 126, 1468–1481.  
<https://doi.org/10.1016/j.clinph.2014.11.018>

Douw, L., van Dellen, E., de Groot, M., Heimans, J.J., Klein, M., Stam, C.J., Reijneveld, J.C., 2010.

. BMC Neurosci. 11.  
<https://doi.org/10.1186/1471-2202-11-103>

Durongbhan, Pholpat, Zhao, Yifan, Chen, L., Zis, Panagiotis, De Marco, Matteo, Unwin, Zoe C, Venneri, Annalena, He, X., Li, S., Zhao, Yitian, Blackburn, Daniel J, Sarrigiannis, Ptolemaios G, Durongbhan, P, Zhao, Y, Zis, P, De Marco, M, Unwin, Z C, Venneri, A, Blackburn, D J, Sarrigiannis, P G, 2019.

. IEEE Trans. NEURAL Syst. Rehabil. Eng. 27.  
<https://doi.org/10.1109/TNSRE.2019.2909100>

Engels, M.M.A., Stam, C.J., Flier, W.M. Van Der, Scheltens, P., Waal, H. De, Straaten, E.C.W. Van, 2015.

. BMC Neurol. 1–8.

<https://doi.org/10.1186/s12883-015-0400-7>

Erra, R.G., Velazquez, J.L.P., Rosenblum, M., 2017.

. Front. Comput. Neurosci. 11, 98.

<https://doi.org/10.3389/fncom.2017.00098>

Evangelisti, S., Pittau, F., Testa, C., Rizzo, G., Gramegna, L.L., Ferri, L., Coito, A., Cortelli, P., Calandra-Buonaura, G., Bisquoli, F., Bianchini, C., Manners, D.N., Talozzi, L., Tonon, C., Lodi, R., Tinuper, P., 2019.

. Front.

Neurosci. 13, 1–10. <https://doi.org/10.3389/fnins.2019.00611>

Ewald, A., Marzetti, L., Zappasodi, F., Meinecke, F.C., Nolte, G., 2012.

. Neuroimage 60, 476–488.

<https://doi.org/10.1016/j.neuroimage.2011.11.084>

Fagerholm, E.D., Moran, R.J., Violante, I.R., Leech, R., Friston, K.J., 2020.

. Neuroimage 208.

<https://doi.org/10.1016/j.neuroimage.2019.116452>

Fan, L., Li, H., Zhuo, J., Zhang, Y., Wang, J., Chen, L., Yang, Z., Chu, C., Xie, S., Laird, A.R., Fox, P.T., Eickhoff, S.B., Yu, C., Jiang, T., 2016.

.

Cereb. Cortex 26, 3508–3526. <https://doi.org/10.1093/cercor/bhw157>

Faust, O., Acharya, U.R., Adeli, H., Adeli, A., 2015.

.

Seizure 26, 56–64.

Ferreri, F., Vecchio, F., Vollero, L., Guerra, A., Petrichella, S., Ponzo, D., Määttä, S., Mervaala, E., Könönen, M., Ursini, F., Pasqualetti, P., Iannello, G., Rossini, P.M., Di Lazzaro, V., 2016.

. Hum. Brain Mapp. 37, 2083–2096. <https://doi.org/10.1002/hbm.23158>

Finn, E.S., Shen, X., Scheinost, D., Rosenberg, M.D., Huang, J., Chun, M.M., Papademetris, X., Constable, R.T., 2015. . Nat. Neurosci. 18, 1664–1671. <https://doi.org/10.1038/nn.4135>

Fraga González, G., Smit, D.J.A., van der Molen, M.J.W., Tijms, J., Stam, C.J., de Geus, E.J.C., van der Molen, M.W., 2018. .

Front. Hum. Neurosci. 12, 1–12. <https://doi.org/10.3389/fnhum.2018.00341>

Fraschini, M., Pani, S.M., Didaci, L., Marcialis, G.L., 2019. . Pattern Recognit. Lett. 125, 49–54. <https://doi.org/10.1016/j.patrec.2019.03.025>

Friston, K.J., 2011. . Brain Connect. 1, 13–36. <https://doi.org/10.1089/brain.2011.0008>

Friston, K.J., 2001. . Neuroscientist 7, 406–418. <https://doi.org/10.1177/107385840100700510>

Friston, K.J., Harrison, L., Penny, W., 2003. . Neuroimage 19, 1273–1302. [https://doi.org/10.1016/S1053-8119\(03\)00202-7](https://doi.org/10.1016/S1053-8119(03)00202-7)

Fu, K., Qu, J., Chai, Y., Dong, Y., 2014. . Biomed. Signal Process. Control 13, 15–22. <https://doi.org/10.1016/j.bspc.2014.03.007>

Gao, Z., Wang, X., Yang, Y., Mu, C., Cai, Q., Dang, W., Zuo, S., 2019. . IEEE Trans. neural networks Learn. Syst. 30, 2755–2763. <https://doi.org/10.1109/TNNLS.2018.2886414>

Geweke, J., 1982.

. J. Am. Stat. Assoc. 77, 304–313.  
<https://doi.org/10.1080/01621459.1982.10477803>

Ghumare, E.G., Schrooten, M., Vandenberghe, R., Dupont, P., 2018.

. Brain Topogr. 31, 721–737. <https://doi.org/10.1007/s10548-018-0621-3>

Grandy, T.H., Werkle-Bergner, M., Chicherio, C., Schmiedek, F., Lövdén, M., Lindenberger, U., 2013.

. Psychophysiology 50, 570–582. <https://doi.org/10.1111/psyp.12043>

Granger, C.W.J., Aug, N., 2007.

37, 424–438.

Gu, Y., Yang, Y., Dewald, J.P.A., Van Der Helm, F.C.T., Schouten, A.C., Wei, H.L., 2021.

. IEEE Trans. Biomed. Eng. 68, 948–958. <https://doi.org/10.1109/TBME.2020.3013545>

Guggisberg, A.G., Honma, S.M., Findlay, A.M., Dalal, S.S., Kirsch, H.E., Berger, M.S., Nagarajan, S.S., 2008.

. Ann. Neurol. 63, 193–203.

Gupta, D., Du, X., Hong, E., Choa, F. Sen, 2019.

. Int. IEEE/EMBS Conf. Neural Eng. NER 2019-March, 1216–1219.  
<https://doi.org/10.1109/NER.2019.8716962>

Hamed, M., Salleh, S.-H., Noor, A.M., 2016.

. Neural Comput. 28, 999–1041.

Handojoseno, A.M.A., Shine, J.M., Nguyen, T.N., Tran, Y., Lewis, S.J.G., Nguyen, H.T., 2013.

. Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS 4263–4266. <https://doi.org/10.1109/EMBC.2013.6610487>

Harmah, D.J., Li, C., Li, F., Liao, Y., Wang, J., Ayedh, W.M.A., Bore, J.C., Yao, D., Dong, W., Xu, P., 2020.

. Front. Comput. Neurosci. 13, 1–15. <https://doi.org/10.3389/fncom.2019.00085>

Hassan, M., Chaton, L., Benquet, P., Delval, A., Leroy, C., Plomhause, L., Moonen, A.J.H., Duits, A.A., Leentjens, A.F.G., van Kranen-Mastenbroek, V., Defebvre, L., Derambure, P., Wendling, F., Dujardin, K., 2017.

. NeuroImage Clin. 14, 591–601. <https://doi.org/10.1016/j.nicl.2017.03.002>

Haufe, S., Ewald, A., 2019.

. Brain Topogr. 32, 625–642. <https://doi.org/10.1007/s10548-016-0498-y>

Haufe, S., Nikulin, V., Nolte, G., 2011.

. BMC Neurosci. 12, 1–2. <https://doi.org/10.1186/1471-2202-12-s1-p141>

Haufe, S., Nikulin, V. V., Müller, K.R., Nolte, G., 2013.

. Neuroimage 64, 120–133. <https://doi.org/10.1016/j.neuroimage.2012.09.036>

He, B., Astolfi, L., Valdes-Sosa, P.A., Marinazzo, D., Palva, S.O., Benar, C.G., Michel, C.M., Koenig, T., 2019.

. IEEE Trans. Biomed. Eng. 66, 2115–2137. <https://doi.org/10.1109/TBME.2019.2913928>

He, F., Billings, S.A., Wei, H.L., Sarrigiannis, P.G., 2014a.

. J. Neurosci. Methods 225, 71–80.  
<https://doi.org/10.1016/j.jneumeth.2014.01.013>

He, F., Billings, S.A., Wei, H.L., Sarrigiannis, P.G., Zhao, Y., 2013.

. IEEE Trans. Biomed. Eng. 60, 2233–2241.  
<https://doi.org/10.1109/TBME.2013.2252347>

He, F., Wei, H.L., Billings, S.A., Sarrigiannis, P.G., 2014b.

. IEEE Trans. Biomed. Eng. 61,  
1693–1701. <https://doi.org/10.1109/TBME.2014.2300636>

Herz, D.M., Christensen, M.S., Reck, C., Florin, E., Barbe, M.T., Stahlhut, C.,  
Pauls, A.K.M., Tittgemeyer, M., Siebner, H.R., Timmermann, L., 2012.

. Neuroimage 59, 3187–3193.  
<https://doi.org/10.1016/j.neuroimage.2011.11.042>

Hipp, J.F., Siegel, M., 2015.

. Curr. Biol. 25, 1368–1374.  
<https://doi.org/10.1016/j.cub.2015.03.049>

Huang, C.S., Pal, N.R., Chuang, C.H., Lin, C.T., 2015.

. Front. Hum.  
Neurosci. 9, 1–12. <https://doi.org/10.3389/fnhum.2015.00570>

Huang, Norden E, Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen,  
N.-C., Tung, C.C., Liu, H.H., 1998.

. Proc. R.  
Soc. London. Ser. A Math. Phys. Eng. Sci. 454, 903–995.

Huang, Norden E., Shen, Z., Long, S.R., Wu, M.C., Snin, H.H., Zheng, Q., Yen,  
N.C., Tung, C.C., Liu, H.H., 1998.

. Proc. R.  
Soc. A Math. Phys. Eng. Sci. 454, 903–995.  
<https://doi.org/10.1098/rspa.1998.0193>



Ibrahim, S., Djemal, R., Alsuwailem, A., 2018.

Biocybern. Biomed. Eng. 38, 16–26. <https://doi.org/10.1016/j.bbe.2017.08.006>

Ieracitano, C., Duun-Henriksen, J., Mammone, N., Foresta, F. La, Morabito, F.C., 2017.

. Proc. Int. Jt. Conf. Neural Networks 2017-May, 1297–1304. <https://doi.org/10.1109/IJCNN.2017.7966002>

Jamal, W., Das, S., Maharatna, K., Pan, I., Kuyucu, D., 2015.

. Phys. A Stat. Mech. its Appl. 434, 273–295. <https://doi.org/10.1016/j.physa.2015.03.087>

Jamal, W., Das, S., Oprescu, I.A., Maharatna, K., Apicella, F., Sicca, F., 2014.

. J. Neural Eng. 11. <https://doi.org/10.1088/1741-2560/11/4/046019>

Jeong, D.H., Kim, Y. Do, Song, I.U., Chung, Y.A., Jeong, J., 2016.

. Entropy 18. <https://doi.org/10.3390/e18010008>

Jones, S.R., 2016.

. Curr. Opin. Neurobiol. 40, 72–80.

Kamifiski, M.J., Bfinowska, K.J., 1991.

. Biol. Cybern. 210, 203–210.

Kaur, Y., Ouyang, G., Junge, M., Sommer, W., Liu, M., Zhou, C., Hildebrandt, A., 2019.

. J. Neurosci. Methods 326, 108343. <https://doi.org/10.1016/j.jneumeth.2019.108343>

Keijzer, H.M., Tjepkema-Cloostermans, M.C., Klijn, C.J.M., Blans, M., van Putten, M.J.A.M., Hofmeijer, J., 2021.

. Clin. Neurophysiol. 132, 157–164. <https://doi.org/10.1016/j.clinph.2020.10.024>

Kent, B.A., Feldman, H.H., Nygaard, H.B., 2021.

. Prog. Neurobiol. 197, 101902. <https://doi.org/10.1016/j.pneurobio.2020.101902>

Khosrowabadi, R., Heijnen, M., Wahab, A., Quek, H.C., 2010.

, in: 2010 IEEE Intelligent Vehicles Symposium. pp. 377–381. <https://doi.org/10.1109/IVS.2010.5548102>

Ko, K.E., Yang, H.C., Sim, K.B., 2009.

. Int. J. Control. Autom. Syst. 7, 865–870. <https://doi.org/10.1007/s12555-009-0521-0>

Korats, G., Le Cam, S., Ranta, R., Louis-Dorr, V., 2015.

. IEEE Trans. Biomed. Eng. 63, 1966–1973.

Kulkarni, N., 2019.

, in: Kulkarni, A.J., Satapathy, S.C., Kang, T., Kashan, A.H. (Eds.), Proceedings of the 2nd International Conference on Data Engineering and Communication Technology. Springer Singapore, Singapore, pp. 395–403.

Labat, D., 2005.

314, 275–288. <https://doi.org/10.1016/j.jhydrol.2005.04.003>

Lachaux, J.P., Lutz, A., Rudrauf, D., Cosmelli, D., Le Van Quyen, M., Martinerie, J., Varela, F., 2002.

Neurophysiol. Clin. 32, 157–174. [https://doi.org/10.1016/S0987-7053\(02\)00301-5](https://doi.org/10.1016/S0987-7053(02)00301-5)

Lee, M., Yoon, J.G., Lee, S.W., 2020.

. Front. Hum. Neurosci. 14, 1–15. <https://doi.org/10.3389/fnhum.2020.00321>

Lee, Y., Hsieh, S., 2014.

9.

<https://doi.org/10.1371/journal.pone.0095415>

Lehmann, C., Koenig, T., Jelic, V., Prichep, L., John, R.E., Wahlund, L.O., Dodge, Y., Dierks, T., 2007.

. J. Neurosci. Methods 161, 342–350. <https://doi.org/10.1016/j.jneumeth.2006.10.023>

Li, Y., Lei, M., Cui, W., Guo, Y., Wei, H.L., 2019.

. IEEE

Trans. Biomed. Eng. 66, 3509–3525.

Liao, Y., Zhou, G., Luo, Y.X., 2019.

. Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS 648–651. <https://doi.org/10.1109/EMBC.2019.8857296>

Lias, S., Sulaiman, N., Murat, Z.H., Taib, M.N., 2010.

. ISIEA 2010 - 2010 IEEE Symp. Ind. Electron. Appl. 612–616. <https://doi.org/10.1109/ISIEA.2010.5679391>

Mammone, N., Ieracitano, C., Morabito, F.C., 2020.

. Neural Networks 124, 357–372. <https://doi.org/10.1016/j.neunet.2020.01.027>

Marrelec, G., Daunizeau, J., Pélégriani-Issac, M., Doyon, J., Benali, H., 2005.

. IEEE Trans. Signal Process. 53, 3503–3516.

Marzetti, L., Basti, A., Chella, F., D'Andrea, A., Syrjäla, J., Pizzella, V., 2019.

. Front. Neurosci. 13, 1–21.

<https://doi.org/10.3389/fnins.2019.00964>

Mazaheri, A., Jensen, O., 2008.

. J. Neurosci. 28, 7781–7787.

Mazher, M., Abd Aziz, A., Malik, A.S., Ullah Amin, H., 2017.

. IEEE Access 5, 14819–14829.

<https://doi.org/10.1109/ACCESS.2017.2731784>

McBride, J., Zhao, X., Munro, N., Jicha, G., Smith, C., Jiang, Y., 2015.

. J. Healthc. Eng. 6, 55–70.

<https://doi.org/10.1260/2040-2295.6.1.55>

Melia, U., Guaita, M., Vallverdú, M., Embid, C., Vilaseca, I., Salamero, M., Santamaria, J., 2015.

. Med. Eng. Phys. 37, 297–308.

<https://doi.org/10.1016/j.medengphy.2015.01.002>

Meng, J., Yao, L., Sheng, X., Zhang, D., Zhu, X., 2015.

. IEEE Trans. Biomed. Eng. 62, 227–240.

<https://doi.org/10.1109/TBME.2014.2345458>

Mercier-Ganady, J., Lotte, F., Loup-Escande, E., Marchal, M., Lecuyer, A., 2014.

. Proc. - IEEE Virtual Real. 33–38.

<https://doi.org/10.1109/VR.2014.6802047>

Mheich, A., Hassan, M., Khalil, M., Berrou, C., Wendling, F., 2015. . J. Neurosci. Methods 242, 77–81. <https://doi.org/10.1016/j.jneumeth.2015.01.002>

Mierlo, P. Van, Papadopoulou, M., Carrette, E., Boon, P., Vandenberghe, S., Vonck, K., Marinazzo, D., 2014. Prog. Neurobiol. 121, 19–35. <https://doi.org/10.1016/j.pneurobio.2014.06.004>

Miraglia, F., Vecchio, F., Bramanti, P., Rossini, P.M., 2016. . Clin. Neurophysiol. 127, 1261–1268. <https://doi.org/10.1016/j.clinph.2015.07.040>

Moca, V. V., Bârzan, H., Nagy- . Nat. Commun. 12, 1–18. <https://doi.org/10.1038/s41467-020-20539-9>

Moezzi, B., Goldsworthy, M.R., 2018. . Front. Neurosci. 12, 10–12. <https://doi.org/10.3389/fnins.2018.00147>

Mohanty, R., Sethares, W.A., Nair, V.A., Prabhakaran, V., 2020. . Sci. Rep. 10, 1–17. <https://doi.org/10.1038/s41598-020-57915-w>

Mormann, F., Lehnertz, K., David, P., Elger, C.E., 2000. 144, 358–369.

Mousavi, A.A., Zhang, C., Masri, S.F., Gholipour, G., 2020. . Sensors (Switzerland) 20. <https://doi.org/10.3390/s20051271>

Mumtaz, W., Ali, S.S.A., Yasin, M.A.M., Malik, A.S., 2018.  
. Med. Biol. Eng. Comput. 56, 233–246.  
<https://doi.org/10.1007/s11517-017-1685-z>

Navarro, X., Porée, F., Beuchée, A., Carrault, G., 2015.  
. Med. Eng. Phys. 37, 315–320. <https://doi.org/10.1016/j.medengphy.2015.01.006>

Nentwich, M., Ai, L., Madsen, J., Telesford, Q.K., Haufe, S., Milham, M.P., Parra, L.C., 2020.  
. Neuroimage 218, 117001.  
<https://doi.org/10.1016/j.neuroimage.2020.117001>

Nolte, G., Bai, O., Wheaton, L., Mari, Z., Vorbach, S., Hallett, M., 2004.  
. Clin. Neurophysiol. 115, 2292–2307.  
<https://doi.org/10.1016/j.clinph.2004.04.029>

Omidvarnia, A., Mesbah, M., O'Toole, J.M., Colditz, P., Boashash, B., 2011.  
. 7th Int. Work. Syst. Signal Process. their Appl. WoSSPA 2011 179–182. <https://doi.org/10.1109/WOSSPA.2011.5931445>

Omidvarnia, A.H., Mesbah, M., Khlif, M.S., O'Toole, J.M., Colditz, P.B., Boashash, B., 2011.  
. Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS 1423–1426. <https://doi.org/10.1109/IEMBS.2011.6090335>

Palva, J.M., Wang, S.H., Palva, S., Zhigalov, A., Monto, S., Brookes, M.J., Schoffelen, J.M., Jerbi, K., 2018.  
. Neuroimage 173, 632–643.  
<https://doi.org/10.1016/j.neuroimage.2018.02.032>

- Petrantonakis, P.C., Hadjileontiadis, L.J., 2012. . IEEE Trans. Signal Process. 60, 2604–2616.
- Piho, L., Tjahjadi, T., 2020. 11, 722–736.
- Pijnenburg, Y.A.L., Vd Made, Y., Van Cappellen Van Walsum, A.M., Knol, D.L., Scheltens, P., Stam, C.J., 2004. . Clin. Neurophysiol. 115, 1332–1339. <https://doi.org/10.1016/j.clinph.2003.12.029>
- Poil, S.S., de Haan, W., van der Flier, W.M., Mansvelder, H.D., Scheltens, P., Linkenkaer-Hansen, K., 2013. . Front. Aging Neurosci. 5, 1–12. <https://doi.org/10.3389/fnagi.2013.00058>
- Ponten, S.C., Daffertshofer, A., Hillebrand, A., Stam, C.J., 2010. . Neuroimage 52, 985–994. <https://doi.org/10.1016/j.neuroimage.2009.10.049>
- Qassim, Y.T., Cutmore, T.R.H., Rowlands, D.D., 2017. . Microprocess. Microsyst. 51, 356–365. <https://doi.org/10.1016/j.micpro.2017.01.001>
- Qiao, J., Lv, Y., Cao, C., Wang, Z., Li, A., 2018. . Front. Aging Neurosci. 10, 1–12. <https://doi.org/10.3389/fnagi.2018.00417>
- Rajendra Acharya, U., Vinitha Sree, S., Alvin, A.P.C., Suri, J.S., 2012. . Expert Syst. Appl. 39, 9072–9078. <https://doi.org/10.1016/j.eswa.2012.02.040>

Rajintha S Gunawardena, S.A., Sarrigiannis, P.G., Blackburn, D.J., He, F., 2021.

<https://doi.org/10.1101/2021.10.15.464451>

Rathee, D., Cecotti, H., Prasad, G., 2017.

. J. Neural Eng. 14.

<https://doi.org/10.1088/1741-2552/aa785c>

Ratti, E., Waninger, S., Berka, C., Ruffini, G., Verma, A., 2017.

. Front.

Hum. Neurosci. 11, 1–7. <https://doi.org/10.3389/fnhum.2017.00398>

Ravan, M., Hasey, G., Reilly, J.P., MacCrimmon, D., Khodayari-Rostamabad, A., 2015.

. Clin. Neurophysiol. 126, 721–730.

<https://doi.org/10.1016/j.clinph.2014.07.017>

Rehman, N., Mandic, D.P., 2010.

Proc. R. Soc. A Math. Phys. Eng. Sci. 466, 1291–1302.

Rilling, G., Flandrin, P., Gonçalves, P., Lilly, J.M., 2007.

. IEEE Signal Process. Lett. 14, 936–939.

Rocca, D. La, Campisi, P., Vegso, B., Cserti, P., Kozmann, G., Babiloni, F., De Vico Fallani, F., 2014.

. IEEE Trans. Biomed. Eng. 61, 2406–2412.

<https://doi.org/10.1109/TBME.2014.2317881>

Sadaghiani, S., Brookes, M.J., Baillet, S., 2021.

. arXiv 1–39.

Sadaghiani, S., Kleinschmidt, A., 2016.

. Trends Cogn. Sci.

20, 805–817. <https://doi.org/10.1016/j.tics.2016.09.004>



Saeedi, A., Saeedi, M., Maghsoudi, A., Shalbaf, A., 2020.  
. Cogn.  
Neurodyn. 0. <https://doi.org/10.1007/s11571-020-09619-0>

Saito, K., Suzuki, H., Kawakami, Y., 2006.  
. J.  
Physiol. Sci. 56, 313–316. <https://doi.org/10.2170/physiolsci.SC006106>

Sakkalis, V., 2011.  
. Comput. Biol. Med. 41, 1110–1117.  
<https://doi.org/10.1016/j.compbiomed.2011.06.020>

Sakkalis, Vangelis, 2011.  
. Comput. Biol. Med. 41, 1110–  
1117.

Salman, H., Grover, J., Shankar, T., 2018.  
2733, 2709–2733.  
<https://doi.org/10.1162/NECO>

Sankari, Z., Adeli, H., 2011.  
. J. Neurosci.  
Methods 197, 165–170. <https://doi.org/10.1016/j.jneumeth.2011.01.027>

Sankari, Z., Adeli, H., Adeli, A., 2012.  
. Clin. EEG Neurosci. 43, 268–278.  
<https://doi.org/10.1177/1550059412444970>

Sankari, Z., Adeli, H., Adeli, A., 2011.  
. Clin. Neurophysiol. 122, 897–  
906. <https://doi.org/10.1016/j.clinph.2010.09.008>

Sanz-García, A., Pérez-Romero, M., Pastor, J., Sola, R.G., Vega-Zelaya, L.,  
Vega, G., Monasterio, F., Torrecilla, C., Pulido, P., Ortega, G.J., 2019.

. J. Neural Eng. 16. <https://doi.org/10.1088/1741-2552/ab039f>

Sargolzaei, S., Cabrerizo, M., Goryawala, M., Salah, A., Adjouadi, M., 2015. . Comput. Biol. Med. 56, 158–166. <https://doi.org/10.1016/j.compbimed.2014.10.018>

Sarrigiannis, P.G., Zhao, Y., He, F., Billings, S.A., Baster, K., Rittey, C., Yianni, J., Zis, P., Wei, H., Hadjivassiliou, M., Grünewald, R., 2018.

. Clin. Neurophysiol. 129, 602–617. <https://doi.org/10.1016/j.clinph.2017.11.029>

Sarrigiannis, P.G., Zhao, Y., He, F., Wei, H., Billings, S.A., Lawrence, S., Rowe, J., Romanowski, C., Hoggard, N., Rao, D.G., Grünewald, R., Khan, A., Hadjivassiliou, M., Yianni, J., 2015.

. Brain Stimul. <https://doi.org/10.1016/j.brs.2015.08.006>

Sarrigiannis, Ptolemaios G, Zhao, Y., Wei, H.-L., Billings, S.A., Fotheringham, J., Hadjivassiliou, M., 2014.

. Clin. Neurophysiol. 125, 32–46. <https://doi.org/10.1016/j.clinph.2013.06.012>

Sarrigiannis, Ptolemaios G., Zhao, Y., Wei, H.L., Billings, S.A., Fotheringham, J., Hadjivassiliou, M., 2014.

. Clin. Neurophysiol. <https://doi.org/10.1016/j.clinph.2013.06.012>

Sartori, E., Bertoldo, A., Zavaglia, M., Ursino, M., Toffolo, G.M., 2012.

14, 96–99.

Schmidt, M.T., Kanda, P.A.M., Basile, L.F.H., da Silva Lopes, H.F., Baratho, R., Demario, J.L.C., Jorge, M.S., Nardi, A.E., Machado, S., Ianof, J.N., Nitrini, R.,

Anghinah, R., 2013.

. Front. Aging Neurosci. 5, 1–6.  
<https://doi.org/10.3389/fnagi.2013.00060>

Schreiber, T., Schmitz, A., 2000. . Phys. D Nonlinear Phenom. 142, 346–382.

Schumacher, E.M., Stiris, T.A., Larsson, P.G., 2015. . Clin. Neurophysiol. 126, 2261–2268. <https://doi.org/10.1016/j.clinph.2015.01.017>

Schutte, N.M., Hansell, N.K., de Geus, E.J.C., Martin, N.G., Wright, M.J., Smit, D.J.A., 2013. . Twin Res. Hum. Genet. 16, 962–969. <https://doi.org/10.1017/thg.2013.55>

Seth, A.K., Barrett, A.B., Barnett, L., 2015. . J. Neurosci. 35, 3293–3297. <https://doi.org/10.1523/JNEUROSCI.4399-14.2015>

Shahabi, H., Moghimi, S., 2016.

. Comput. Human Behav. 58, 231–239. <https://doi.org/10.1016/j.chb.2016.01.005>

Shan, X., Huo, S., Yang, L., Cao, J., Zou, J., Chen, L., Sarrigiannis, P.G., Zhao, Y., 2021.

. IEEE Trans. Neural Syst. Rehabil. Eng. 29, 841–851. <https://doi.org/10.1109/TNSRE.2021.3076311>

Shen, X., Finn, E.S., Scheinost, D., Rosenberg, M.D., Chun, M.M., Papademetris, X., Constable, R.T., 2018.

. Physiol. Behav. 176, 139–148. <https://doi.org/10.1038/nprot.2016.178>.Using

Shen, Y., Baingana, B., Giannakis, G.B., 2016.

- Siems, M., Siegel, M., 2020. *Neuroimage* 209, 116538. <https://doi.org/10.1016/j.neuroimage.2020.116538>
- Silfverhuth, M.J., Hintsala, H., Kortelainen, J., Seppänen, T., 2012. *Med. Biol. Eng. Comput.* 50, 683–688. <https://doi.org/10.1007/s11517-012-0911-y>
- Stam, C.J., Nolte, G., Daffertshofer, A., 2007. *Neuroimage* 1193, 1178–1193. <https://doi.org/10.1002/hbm.20346>
- Supp, G.G., Schlögl, A., Trujillo-Barreto, N., Müller, M.M., Gruber, T., 2007. *PLoS One* 2. <https://doi.org/10.1371/journal.pone.0000684>
- Tafreshi, T.F., Daliri, M.R., Ghodousi, M., 2019a. *Cogn. Neurodyn.* 13, 555–566. <https://doi.org/10.1007/s11571-019-09556-7>
- Tafreshi, T.F., Daliri, M.R., Ghodousi, M., 2019b. *Cogn. Neurodyn.* 13, 555–566.
- Tarokh, L., Carskadon, M.A., Achermann, P., 2010. *Neuroscience* NSC 171, 622–634. <https://doi.org/10.1016/j.neuroscience.2010.08.071>
- Theiler, J., Eubank, S., Longtin, A., Galdrikian, B., Farmer, J.D., 1992. *Phys. D Nonlinear Phenom.* 58, 77–94.
- Torres, María E, Colominas, M.A., Schlotthauer, G., Flandrin, P., 2011. *Entropy*, in: 2011

IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). pp. 4144–4147.

Torres, María E., Colominas, M.A., Schlotthauer, G., Flandrin, P., 2011.  
. 2011  
IEEE Int. Conf. Acoust. speech signal Process. 4144–4147.

Uhlhaas, P.J., Singer, W., 2006.  
. Neuron 52, 155–  
168. <https://doi.org/10.1016/j.neuron.2006.09.020>

Ursino, M., Zavaglia, M., Astolfi, L., 2007.  
351–365. <https://doi.org/10.1007/s00422-006-0122-4>

Van de Steen, F., Almgren, H., Razi, A., Friston, K., Marinazzo, D., 2019.  
.   
Neuroimage 189, 476–484. <https://doi.org/10.1016/j.neuroimage.2019.01.055>

Van Der Molen, M.J.W., Stam, C.J., Van Der Molen, M.W., 2014.  
. PLoS One 9.  
<https://doi.org/10.1371/journal.pone.0088451>

van der Zande, J.J., Gouw, A.A., Steenoven, I. van, Scheltens, P., Stam, C.J.,  
Lemstra, A.W., 2018.  
. Front. Aging Neurosci. 10, 1–10.  
<https://doi.org/10.3389/fnagi.2018.00190>

van Mierlo, P., Papadopoulou, M., Carrette, E., Boon, P., Vandenberghe, S.,  
Vonck, K., Marinazzo, D., 2014.  
. Prog.  
Neurobiol. 121, 19–35. <https://doi.org/10.1016/j.pneurobio.2014.06.004>

Varela, F., Lachaux, J.P., Rodriguez, E., Martinerie, J., 2001.  
. Nat. Rev. Neurosci. 2, 229–  
239. <https://doi.org/10.1038/35067550>

Varotto, G., Fazio, P., Rossi Sebastiano, D., Duran, D., D'Incerti, L., Parati, E., Sattin, D., Leonardi, M., Franceschetti, S., Panzica, F., 2014.

. Clin. Neurophysiol. 125, 63–68.  
<https://doi.org/10.1016/j.clinph.2013.06.016>

Varotto, G., Visani, E., Canafoglia, L., Franceschetti, S., Avanzini, G., Panzica, F., 2012.

. Epilepsia 53, 359–367. <https://doi.org/10.1111/j.1528-1167.2011.03352.x>

Vecchio, F., Miraglia, F., Alù, F., Menna, M., Judica, E., Cotelli, M., Rossini, P.M., 2020.

. J. Alzheimer's Dis. 75, 1253–1261. <https://doi.org/10.3233/JAD-200171>

Vecchio, F., Miraglia, F., Bramanti, P., Rossini, P.M., 2014.

. J. Alzheimer's Dis. 41, 1239–1249.  
<https://doi.org/10.3233/JAD-140090>

Vecchio, F., Miraglia, F., Piludu, F., Granata, G., Romanello, R., Caulo, M., Onofri, V., Bramanti, P., Colosimo, C., Rossini, P.M., 2017.

. Brain Imaging Behav. 11, 473–485. <https://doi.org/10.1007/s11682-016-9528-3>

Vecchio, F., Miraglia, F., Quaranta, D., Granata, G., Romanello, R., Marra, C., Bramanti, P., Rossini, P.M., 2016.

. Neuroscience 316, 143–150.

Vicente, R., Wibral, M., Lindner, M., 2011a.

. <https://doi.org/10.1007/s10827-010-0262-3>

- Vicente, R., Wibral, M., Lindner, M., Pipa, G., 2011b. . J. Comput. Neurosci. 30, 45–67.
- Vijay Anand, S., Shantha Selvakumari, R., 2019. . Soft Comput. 23, 2645–2653. <https://doi.org/10.1007/s00500-018-3630-y>
- Visani, E., Varotto, G., Binelli, S., Fratello, L., Franceschetti, S., Avanzini, G., Panzica, F., 2010. . Clin. Neurophysiol. 121, 318–324. <https://doi.org/10.1016/j.clinph.2009.12.003>
- Vortmann, L.M., Kroll, F., Putze, F., 2019. . Front. Hum. Neurosci. 13, 1–14. <https://doi.org/10.3389/fnhum.2019.00348>
- Wang, F., Tian, Y.C., Zhang, X., Hu, F., 2020. . IEEE Trans. Emerg. Top. Comput. Intell. 1–11. <https://doi.org/10.1109/TETCI.2020.3032662>
- Wang, H.E., Bénar, C.G., Quilichini, P.P., Friston, K.J., Jirsa, V.K., Bernard, C., 2014. . Front. Neurosci. 8, 1–22. <https://doi.org/10.3389/fnins.2014.00405>
- Wang, M., El-Fiqi, H., Hu, J., Abbass, H.A., 2019. . IEEE Trans. Inf. Forensics Secur. 14, 3359–3372. <https://doi.org/10.1109/TIFS.2019.2916403>
- Wendling, F., Ansari-asl, K., Bartolomei, F., Senhadji, L., 2009. 183, 9–18. <https://doi.org/10.1016/j.jneumeth.2009.04.021>
- Wendling, F., Chauvel, P., Biraben, A., Bartolomei, F., 2010.

. Front. Syst. Neurosci. 4, 1–13.  
<https://doi.org/10.3389/fnsys.2010.00154>

Wu, Z., Huang, N.E., 2009.  
. Adv. Adapt. Data Anal. 1, 1–41.

Xie, S., Krishnan, S., 2013.  
. Med. Biol. Eng. Comput. 51, 49–60. <https://doi.org/10.1007/s11517-012-0967-8>

Yang, C., Le Bouquin Jeannes, R., Bellanger, J.J., Shu, H., 2013.  
. IEEE Trans. Biomed. Eng. 60, 1318–1327.  
<https://doi.org/10.1109/TBME.2012.2234125>

Yao, W., Wang, J., 2017.  
. Phys. A Stat. Mech. its Appl. 484, 276–281.  
<https://doi.org/10.1016/j.physa.2017.04.181>

Yildirim, O., Talo, M., Ay, B., Baloglu, U.B., Aydin, G., Acharya, U.R., 2019.  
. Comput. Biol. Med. 113, 103387. <https://doi.org/10.1016/j.compbiomed.2019.103387>

Yin, Z., Li, J., Zhang, Y., Ren, A., Von Meneen, K.M., Huang, L., 2017.  
. Biomed. Signal Process. Control 31, 331–338.  
<https://doi.org/10.1016/j.bspc.2016.08.013>

Yu, H., Lei, X., Song, Z., Wang, J., Wei, X., Yu, B., 2018.  
. Phys. A Stat. Mech. its Appl. 506, 1093–1103.  
<https://doi.org/10.1016/j.physa.2018.05.009>

Yu, H., Liu, J., Cai, L., Wang, J., Cao, Y., Hao, C., 2017.



. Phys. A Stat. Mech. its Appl. 468, 566–577. <https://doi.org/10.1016/j.physa.2016.10.068>

Yuvaraj, R., Murugappan, M., Acharya, U.R., Adeli, H., Ibrahim, N.M., Mesquita, E., 2016.

. Behav. Brain Res. 298, 248–260. <https://doi.org/10.1016/j.bbr.2015.10.036>

Yvert, G., Perrone-Bertolotti, M., Baciú, M., David, O., 2012.

. J. Neurosci. 32, 4297–4306. <https://doi.org/10.1523/JNEUROSCI.6434-11.2012>

Zhang, Y., Liu, B., Ji, X., Huang, D., 2017.

. Neural Process. Lett. 45, 365–378.

Zhao, Y., Billings, S.A., Wei, H., He, F., Sarrigiannis, P.G., 2013a.

. J. Neurosci. Methods 212, 79–86. <https://doi.org/10.1016/j.jneumeth.2012.09.019>

Zhao, Y., Billings, S.A., Wei, H., Sarrigiannis, P.G., 2012.

. Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys. 86. <https://doi.org/10.1103/PhysRevE.86.051919>

Zhao, Y., Billings, S.A., Wei, H.L., Sarrigiannis, P.G., 2013b.

. IEEE Trans. Biomed. Eng. 60, 3141–3148. <https://doi.org/10.1109/TBME.2013.2269766>

Zhao, Yifan, Laguna, R.C., Zhao, Yitian, Liu, J.J., He, X., Yianni, J., Sarrigiannis, P.G., 2018.

. Complexity 2018. <https://doi.org/10.1155/2018/7269494>

Zhao, Yifan, Zhao, Yitian, Durongbhan, P., Chen, L., Liu, J., Billings, S.A., Zis, P., Unwin, Z.C., De Marco, M., Venneri, A., Blackburn, D.J., Sarrigiannis, P.G., 2020.

. IEEE Trans. Med. Imaging 39, 1571–1581.  
<https://doi.org/10.1109/TMI.2019.2953584>

### 3 Time-frequency Functional Connectivity Estimation

This chapter is based on two published journal paper. ‘Tracking non-stationary association of two electroencephalography signals using a Revised Hilbert-Huang Transformation’ is published in **IEEE Transactions on Neural Systems and Rehabilitation Engineering**. ‘Ultra-high-resolution time-frequency analysis of EEG to characterise brain functional connectivity with the application in Alzheimer’s disease’ is published in **Journal of Neural Engineering**.

Shan, X., Huo, S., Cao, J., Yang, L., Zou, J., Chen, L., Li, Z., 2021.

. IEEE Trans. Neural Syst. Rehabil. Eng.

Shan. <https://doi.org/10.1109/TNSRE.2021.3076311>

Cao, J., Zhao, Y., Shan, X., Blackburn, D., Wei, J., Erkoyuncu, J.A., Chen, L., Sarrigiannis, P.G., 2022b.

. J. Neural Eng. 19. <https://doi.org/10.1088/1741-2552/ac84ac>

#### 3.1 A Revised Hilbert-Huang Transformation to Track Non-stationary Association of Electroencephalography Signals

##### 3.1.1 Introduction

ELECTROENCEPHALOGRAPHY (EEG) is a powerful technique that can noninvasively study the electrophysiological brain dynamics with high temporal accuracy. EEG has been instrumental in making discoveries about cognition, brain function, and dysfunction (Cohen, 2017a). Cognitive, perceptual, linguistic, emotional, and motor processes are fast and take place from tens and hundreds of milliseconds to a few seconds (Cohen, 2014). Neuronal oscillations are embedded in multiplexed and noisy signals, which have been observed with multiple spatial and temporal scales. The non-stationary, asymmetry and non-linear characteristics of EEG have challenged researchers to develop advanced methods to reveal more hidden information.

Time-frequency-based methods are suitable for analysing various fast, dynamic and spatial-temporal cognitive events. Most of the univariate-based time-frequency analysis involves matching EEG timing data to waveform templates. Wavelet Transform (WT) offers superior time and frequency resolution in comparison to other methods, like the Short-Time Fourier Transform (STFT), and is a commonly used tool for finely identifying hidden information in non-stationary signals. Recent studies have used WT in signal decomposition and adaptive filtering (Navarro et al., 2015), EEG source localization (Koratzopoulos et al., 2015), computer-aided seizure detection and epilepsy diagnosis (Faust et al., 2015), and classification of EEG signals (Zhang et al., 2017). Because the wavelet transform is based on Fourier analysis, using sine waves as templates, it has the inherent shortcoming of Fourier spectrum analysis: it can only detect linear phenomena. For example, it can solve inter-wave frequency modulation, but because the basic wavelet has a certain wavelength, it cannot solve intra-wave frequency modulation (Norden & Huang, 1998). Furthermore, WT is limited by the spectral leakage caused by the finite length of the basic wavelet function, the counterintuitive interpretation of fake high-frequency information when analysing true local low-frequency events, and the non-adaptivity of the once set basis wavelet (Norden & Huang, 1998).

It is highly debated if neural oscillations are sinusoidal (Cohen, 2017b; Jones, 2016; Mazaheri and Jensen, 2008). Jones et al. (Jones, 2016) demonstrated that peaks at specific frequencies in the spectrum can be created by dominant waveform features in the time domain rather than 'sinusoidal' type oscillations. Mazaheri and Jensen (Mazaheri and Jensen, 2008) proposed a physiological explanation for the theory that peaks and troughs of oscillatory activity fluctuate asymmetrically, peaks being more strongly modulated than troughs in response to stimuli. Sinusoid-like activity will be observed, even in noise, when signal processing involves sinusoidal filters. Thus, it is difficult for common WT-based approaches to distinguish between sinusoidal versus non-sinusoidal oscillations (Cohen, 2017a).

Instead of a sine wave template, some researchers use "physiologically defined" waveforms as an analytical regression template to explicitly incorporate neurophysiological principles into EEG data analysis. Cheveigné (2010) automatically synthesised multichannel finite impulse response templates by delaying the original data based on the criterion of reproducibility over stimulus repetitions. Cohen (2017b) proposed a generalized eigen-decomposed cross-frequency coupling framework based on multi-channel covariance matrix and pointed out that the non-sinusoidal waveform of neural oscillations may provide important insights into biophysical processes. Empirical templates are data-driven and defined from the spatio-temporal filters or the output of computational models. However, the effectiveness of this method depends on the accuracy of the template. Physiologically defined waveform templates need to be targeted to each brain region or cognitive process, and even to each individual. In addition, the data used to estimate physiological waveforms always contain noise, and the dependent biophysical hypothesis models may not be verified (Cohen, 2017a).

Unlike WT with sinusoidal or physiologically defined wavelets, EMD (Empirical Mode Decomposition) is more intuitive and adaptive without any template assumption of the analysed signal (Norden E Huang *et al.*, 1998), which makes it more suitable for describing non-stationary asymmetric and non-linear characteristics of EEG. EMD analyses the behaviour of non-stationary and nonlinear signals by decomposing them into several intrinsic mode functions (IMFs) that admit well-behaved Hilbert transforms. There are many variants of EMD, such as ensemble EMD (EEMD) (Wu and Huang, 2009), bivariate EMD (BEMD) (Rilling *et al.*, 2007), multivariate EMD (MEMD) (Rehman and Mandic, 2010), and complete ensemble EMD with adaptive noise (CEEMDAN) (Torres *et al.*, 2011). CEEMDAN can provide an exact reconstruction of the original signal and a better spectral separation of the modes with a lower computational cost, overcoming the "mode mixing" problem of EMD and the lack of ability to reconstruct the original signal of EEMD. Since the decomposition is based on the characteristics of the local time scale, with the Hilbert Transform, the IMFs generate instantaneous frequencies as functions of time that separately

estimates dynamic structures of different transient information. The final presentations are several energy-frequency-time distributions, designated as the Hilbert spectrum, and phase-time curves.

Detecting and quantifying the interactions between different EEG channel signals can lead to important insights into the dynamic networks that underpin human brain function (Aydore et al., 2013). Brain functional connectivity is defined as a measurement to evaluate statistic dependence between separated neuronal regions (Akbarian and Erfanian, 2020; Friston, 2011; Hamed et al., 2016; Schumacher et al., 2015). Brain synchronization can provide a mechanism for functionally related but spatially distinct neurons to act synergistically (Chiang et al., 2011; Marrelec et al., 2005). Abnormalities of brain functional connectivity have been reported for numerous brain disorders (Erra et al., 2017; Uhlhaas and Singer, 2006). By examining the magnitude of linear or nonlinear correlation in the domain of time, frequency, and time-frequency, studying brain functional connectivity between multiple channels has been increasingly considered as a promising approach (Diessen et al., 2015). In the time domain, the method of synchronization likelihood (SL) is thought to have the ability to detect both linear and non-linear functional interactions between EEG time series (Babiloni et al., 2016; Boersma et al., 2013; Schutte et al., 2013). Seizure prediction and epileptogenic-focus localisation have been investigated using techniques like correlation coefficient (Wendling et al., 2010) and cross-correlation (van Mierlo et al., 2014). The correlation between memory performance and quantified brain functional connectivity from EEG has been used to study Alzheimer's diseases (Vecchio et al., 2016). Babiloni et al. (2016) suggested that the technique of mutual information (MI) is useful for quantifying information transmitted between EEG rhythms in Alzheimer's diseases; MI can also be estimated in the time-frequency domain. In addition, in the time domain, Babiloni et al. (2016) and Vicente et al. (2011b) introduced a transfer entropy approach, which presented good performance in detecting unknown non-linear interactions. In the frequency domain, magnitude squared coherence (MSC) is the most widely used technique (Vangelis Sakkalis, 2011). Based on phase-locking value (PLV), researchers developed an algorithm to track functional

connectivity dynamics (Mheich et al., 2015) and found biomarkers with the purpose of seizure prediction in epilepsy (van Mierlo et al., 2014) as well as object recognition (Tafreshi et al., 2019b) paradigms in healthy subjects. In the time-frequency domain, Chen (Chen et al., 2013) and Wendling (2009) used STFT to evaluate the dynamic change of brain connectivity. Furthermore, instead of applying the constant size windows as in the STFT, wavelet coherence (WC) also performs well in quantifying time-varying connectivity (Vangelis Sakkalis, 2011). Petrantonakis and Hadjileontiadis (2012) developed adaptive methods for EEG signal segmentation in the time-frequency domain to effectively retrieve the emotion-related information. Higher time-frequency resolution means better capturing the dynamic non-stationarity and non-linearity of EEG signals (Cohen, 2014), which translates to a more reliable estimation of time-varying brain functional connectivity. Therefore, EMD with corresponding frequency analysis could be a promising analysis tool in the study of functional connectivity.

This Chapter proposes a new time-frequency brain functional connectivity analysis framework, to track the non-stationary association of two EEG signals, based on a Revised Hilbert-Huang Transform (RHHT). One of the highlights of this framework is a novel method of estimating the cross-spectrum, based on the decomposed components. Compared with other time-frequency methods, such as CWT, the proposed method provides a more accurate dynamic time-frequency feature recognition and can be used to dissect the details of dynamic connectivity between brain regions. The effectiveness of this proposed method is validated through simulated and clinical EEG data.

### **3.1.2 Method**

The proposed framework, as illustrated in Figure 3-1, introduces a CEEMDAN-based Hilbert Transform, namely RHHT, to track the non-stationary association of two EEG signals. Firstly, the two signals are decomposed by CEEMDAN to obtain the corresponding IMFs. Secondly, the Hilbert Transform is performed on the IMFs of the two signals to obtain the corresponding time-frequency spectrums respectively. The Cross-Spectrum of the two signals is then

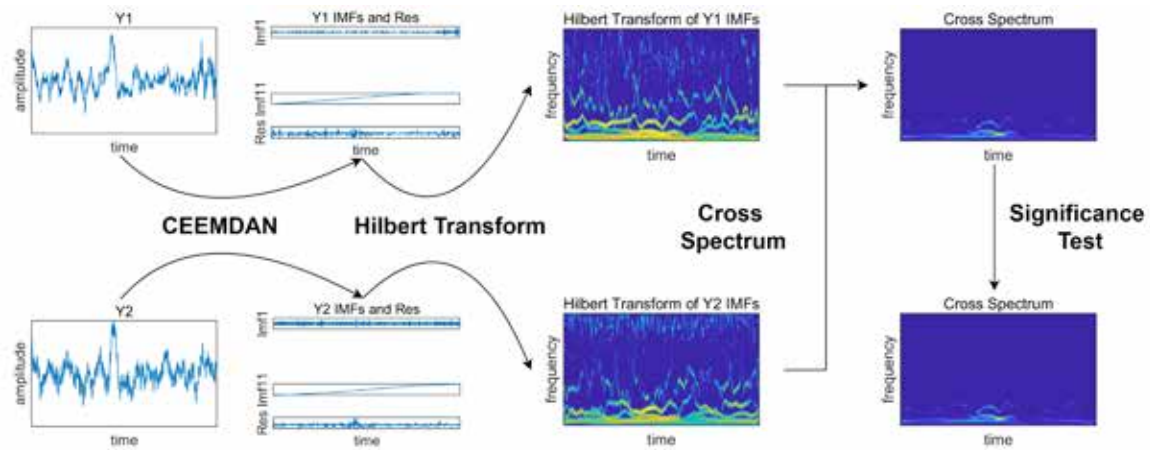


**Figure 3-1 The flow chart of the proposed framework for tracking non-stationary association based on Revised Hilbert-Huang Transform (RHHT).**

estimated and validated through a surrogate significance test. The detail of each step is presented below.

### 3.1.2.1 CEEMDAN

Empirical Mode Decomposition (EMD) can decompose complicated data into a finite and often small number of 'intrinsic mode functions' that admit well-behaved Hilbert transforms (Huang et al., 1998). An intrinsic mode function (IMF) must satisfy two conditions: (1) in the whole data set, the number of extrema and the number of zero crossings must be either equal or differ at most by one; and (2) at any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero (Huang et al., 1998). EMD separates the signal into fast oscillations and slow oscillations. EMD has the problem of "mode mixing", that is, oscillations with very disparate amplitudes exist in one mode or very similar oscillations exist in different modes. To overcome this, Ensemble EMD (EEMD) (Wu and Huang, 2009) performs the EMD over an ensemble of the signal plus Gaussian white noise. By filling the entire time-frequency space to take advantage of the dyadic filter bank behaviour of EMD, the addition of white Gaussian noise solves the mode mixing problem. But it creates a new problem that the reconstructed signal includes residual noise, and different realisations of signal plus noise may produce a different number of modes. Complete Ensemble EMD with Adaptive Noise (CEEMDAN) (Torres et al., 2011), based on EEMD, only needs less than



half of the EEMD screening iterations and can accurately reconstruct the original signal by summing the modes.

Given a signal  $x[n]$ ,  $n = 1, \dots, N$ , where  $N$  is the length of the signal  $x[n]$ , the defined operator  $E_j(\cdot)$  produces the  $j$ -th mode of  $x[n]$  by EMD. Let  $w^i \in [0, 1]$ ,  $i = 1, \dots, I$  be white noise where  $I$  is the realizations times of adding noise. Coefficient  $\varepsilon_i$  allows selecting the signal to noise ratio (SNR) at each stage. CEEMDAN method can be described by the following algorithm:

Decompose  $I$  realizations  $x[n] + \varepsilon_0 w^i[n]$  by EMD to obtain their first modes and compute the average as the first mode  $\widetilde{IMF}_1[n]$ :

$$\widetilde{IMF}_1[n] = \frac{1}{I} \sum_{i=1}^I IMF_1^i[n] = \overline{IMF}_1[n] \quad (3-1)$$

At the first stage ( $k=1$ ) calculate the first residue  $r_1[n]$ :

$$r_1[n] = x[n] - \widetilde{IMF}_1[n] \quad (3-2)$$

Decompose  $I$  realizations  $r_1[n] + \varepsilon_1 E_1(w^i[n])$  by EMD to obtain their first modes and compute the average as the second mode  $\widetilde{IMF}_2[n]$ :

$$\widetilde{IMF}_2[n] = \frac{1}{I} \sum_{i=1}^I E_1 \left( r_1[n] + \varepsilon_1 E_1(w^i[n]) \right) \quad (3-3)$$

For  $k = 2$  calculate the  $k$ -th residue:

$$r_k[n] = r_{k-1}[n] - \widetilde{IMF}_k[n] \quad (3-4)$$

Decompose realizations  $r_k[n] + \varepsilon_k E_k(w^i[n])$ , by EMD to obtain their first modes and compute the average as the  $(k+1)^{th}$  mode  $\widetilde{IMF}_{k+1}[n]$ :

$$\widetilde{IMF}_{k+1}[n] = \frac{1}{I} \sum_{i=1}^I E_k \left( r_k[n] + \varepsilon_k E_k(w^i[n]) \right) \quad (3-5)$$

Go to step 4 for the next  $k$  and perform steps 4 and 5 until the residue is no longer feasible to be decomposed (the residue does not have at least two extrema).

The final residual  $R[n]$ :

$$R[n] = x[n] - \sum_{k=1}^K \widetilde{IMF}_k[n] \quad (3-6)$$

with  $K$  being the total number of modes. The given signal  $x[n]$  can be expressed as:

$$x[n] = \sum_{k=1}^K \widetilde{IMF}_k[n] + R[n] \quad (3-7)$$

Equation (3-7) makes the proposed decomposition complete and provides an exact reconstruction of the original data. Wu and Huang (Wu and Huang, 2009) suggested to use small amplitudes of the added noise for data dominated by high-frequency signals, and vice versa. Following then, this Chapter uses a few hundred realisations and fixed the same SNR for all the stages.

### Hilbert Transform

The original signal  $x(t)$  and its Hilbert transform  $y(t)$  constitute its analytic signal  $z(t)$ :

$$y(t) = HT(x(t)) = \frac{1}{\pi} P.V. \int_{-\infty}^{\infty} \frac{x(\tau)}{t - \tau} d\tau \quad (3-8)$$

$$z(t) = x(t) + iy(t) = a(t)e^{i\theta(t)} \quad (3-9)$$

where P.V. denotes Cauchy principal value,  $a(t)$  and  $\theta(t)$  denote the instantaneous amplitude and instantaneous phase, respectively.  $a(t)$  is the trace envelope and defined as:

$$a(t) = \sqrt{x^2(t) + y^2(t)} \quad (3-10)$$

$\theta(t)$  is defined as:

$$\theta(t) = \arctan\left(\frac{y(t)}{x(t)}\right) \quad (3-11)$$

Instantaneous frequency  $f(t)$  is defined as the first derivative of  $\theta(t)$ . Thus,

$$f(t) = \frac{1}{2\pi} \frac{d\theta(t)}{dt} \quad (3-12)$$

To prevent ambiguities due to phase unwrapping in Equation (12),  $f(t)$  can be calculated instead from:

$$f(t) = \frac{1}{2\pi} \frac{x(t)y'(t) - x'(t)y(t)}{x^2(t) + y^2(t)} \quad (3-13)$$

where the prime denotes derivative with respect to time. The time-frequency spectrum of signal  $x(t)$  is defined as:

$$RHHT_x(t, f(t)) = a(t)e^{i\theta(t)} \quad (3-14)$$

The local symmetry property of the IMFs ensures that instantaneous frequencies are always positive, which admits well-behaved Hilbert transforms (Norden E Huang, 1998). The time-frequency spectrum of  $\widetilde{IMF}_k(t)$ ,  $t = 1, \dots, N$ , and  $k = 1, \dots, K$ , and  $R(t)$  are defined as:

$$RHHT_{\widetilde{IMF}_k}(t, f_k(t)) = a_k(t)e^{i\theta_k(t)} \quad (3-15)$$

$$RHHT_R(t, f_R(t)) = a_R(t)e^{i\theta_R(t)} \quad (3-16)$$

The time-frequency spectrum of  $x(t)$  is:

$$RHHT_x(t, f(t)) = \sum_{k=1}^K a_k(t)e^{i\theta_k(t)} + a_R(t)e^{i\theta_R(t)} \quad (3-17)$$

This procedure produces a multitude of instantaneous frequencies,  $f$  including  $f_k(t)$  and  $f_R(t)$ , namely one for each IMF and Residual, allowing for more in-

depth signal analysis. At each time point, the instantaneous frequencies are as many as the number of IMFs, and most applications produce up to a dozen IMFs, creating very sparse time-frequency representations. The result is a time-frequency distribution that is uniformly sampled in time but not in frequency, different from the wavelet transform.

### 3.1.2.2 RHHT Cross-Spectrum

Cross-spectrum can provide information about linear synchronization (Zhao et al., 2018). The RHHT cross-spectrum of  $x_1(t)$  and  $x_2(t)$  is defined as follows:

$$SH_{x_1x_2}(t, f) = \int_{t-\delta/2}^{t+\delta/2} RHHT_{x_1}(\tau, f) \overline{RHHT_{x_2}(\tau, f)} d\tau \quad (3-18)$$

where,  $RHHT_{x_1}(\tau, f)$  and  $RHHT_{x_2}(\tau, f)$  are the RHHT coefficients  $x_1(t)$  and  $x_2(t)$  and;  $\overline{\phantom{x}}$  means complex conjugate;  $\delta$  is the length of the integrating range.

The Continuous Wavelet Transform (CWT) cross-spectrum of  $x_1(t)$  and  $x_2(t)$  is defined as follows:

$$SW_{x_1x_2}(t, f) = \int_{t-\delta/2}^{t+\delta/2} CWT_{x_1}(\tau, f) \overline{CWT_{x_2}(\tau, f)} d\tau \quad (3-19)$$

where,  $CWT_{x_1}(\tau, f)$  and  $CWT_{x_2}(\tau, f)$  is the CWT coefficients  $x_1(t)$  and  $x_2(t)$ ;  $\overline{\phantom{x}}$  means complex conjugate;  $\delta$  is the length of the integrating range.

For CWT cross-spectrum, the length of wavelet transforms integral window  $\delta$  is defined as  $\delta = n_{cy}/f$ , where  $n_{cy}$  is the number of integral cycles in the wavelet window, independent from frequency  $f$ , and the length of each cycle is  $1/f$ . When  $n_{cy}$  is fixed, there is an inverse relationship between  $\delta$  and  $f$ . Low frequencies need a greater and a longer data set to average over, as the time resolution is low, and the frequency resolution is relatively high. High frequencies need a small value of  $\delta$ , as the data need to be superimposed are shorter on average, with time resolution being relatively high and frequency resolution being relatively low. In practical applications, if the data length is short,  $\delta$  should be a small value, and long data need a big  $\delta$ . For RHHT cross-spectrum, the selection of  $\delta$  is also important (Guggisberg et al., 2008).

$\delta$  is entirely empirical, independent of frequency. In this Chapter, the length of the integral window of RHHT coherence was set to 4.

Given two signals  $x_1(t)$  and  $x_2(t)$ , For all IMFs of  $x_1(t)$  and  $x_2(t)$ , the RHHT cross-spectrums between  $\widetilde{IMF}_{x_{1k}}(t)$  and  $\widetilde{IMF}_{x_{2k}}(t)$  can be calculated and analysed separately.

### 3.1.2.3 Significance Test

A significance test is introduced to assure the data analysis does not reflect our prejudice about the underlying system and that represents a fair account of the structures presented in the data (Schreiber and Schmitz, 2000). Consequently, for the data-driven time-frequency analysis by RHHT, the application of this method has to be justified by establishing the statistical significance of the connectivity estimated by cross-spectrum. The method of surrogate data has been a popular tool to address this problem. We first specify the “no connectivity” as our null hypothesis, then generate surrogate data sets that are consistent with this null hypothesis by Amplitude Adjusted Fourier Transform (AAFT) algorithm (Theiler et al., 1992), and finally compute a discriminating statistic for the original and each of the surrogate data sets. If the cross-spectrum value computed for the original data is significantly different than the ensemble of values computed for the surrogate data, then the null hypothesis is rejected at a given significance level and significant connectivity is detected.

Denote the original time series by  $x[n]$ , with  $n = 1, \dots, N$ . The AAFT algorithm can be summarised below:

The first step is to make a Gaussian time series  $s[n]$ , where each element is generated independently from a Gaussian pseudorandom number generator.

Re-order the time sequence of the Gaussian time series so that the ranks of both time series agree, that is, if  $x[n]$  is the  $m^{th}$  smallest of  $x$ , then  $s[n]$  will be the  $m^{th}$  smallest of  $s$ . The re-ordered  $s[n]$  is a time series that “follows” the original time series  $x[n]$  and has a Gaussian amplitude distribution.

Using the windowed or unwindowed Fourier Transform algorithm, a surrogate, call it  $s'[n]$ , of the Gaussian time series can be created. If the original time series  $x[n]$  is time re-ordered so that it follows  $s'[n]$  in the sense that the ranks agree, then the time-re-ordered time series provides a surrogate of the original time series which matches its amplitude distribution. Further, the "underlying" time series ( $s[n]$  and  $s'[n]$ ) are Gaussian and have the same Fourier power spectrum.

This procedure was repeated 100 times and then the 95% quantile (z-score = 1.96) was determined as the threshold for each time and frequency point in the time-frequency spectrum.

### 3.1.3 Simulations

This section aims at demonstrating the performance of the proposed method against the wavelet-based method using simulation examples. Each of the two simulated EEG signals has two non-linear signal components, namely the linear frequency increase Chirp signal and the frequency-modulated Cos signal. The sampling frequency is 500Hz, with a length of 1000 (equal to 2s).

The frequency of the first Chirp signal  $u_1$  increases from 30Hz to 40Hz, and the second one  $u_2$  increases from 30Hz to 46Hz, being defined as:

$$u_1 = \cos(2\pi(f_s + c_1 t) \quad t) \quad (3-20)$$

$$u_2 = \cos(2\pi(f_s + c_2 t) \quad t) \quad (3-21)$$

where  $f_s$  is the starting frequency,  $f_s = 30\text{Hz}$ ;  $c_1$  and  $c_2$  are chirpiness,  $c_1 = 5\text{Hz/s}$  and  $c_2 = 8\text{Hz/s}$ . The instantaneous frequency of  $u_1$  and  $u_2$  are:

$$f_{u_1}(t) = \frac{1}{2\pi} \frac{d\theta_{11}}{dt} = f_s + c_1 t \quad (3-22)$$

$$f_{u_2}(t) = \frac{1}{2\pi} \frac{d\theta_{12}}{dt} = f_s + c_2 t \quad (3-23)$$

where  $\theta_{11} = 2\pi(f_s + c_1 t) \quad t$  and  $\theta_{12} = 2\pi(f_s + c_2 t) \quad t$ .

The two intrawave frequency-modulated Cos signals  $v_1$  and  $v_2$ , showing different harmonic distortion, are defined as:

$$v_1 = \cos(2\pi f_a t + 0.5 \sin(2\pi f_b t)) \quad (3-24)$$

$$v_2 = \cos(2\pi f_a t + 0.5 \sin(2\pi f_c t)) \quad (3-25)$$

where  $f_a$  is the intrinsic frequency, being 6Hz;  $f_b$  and  $f_c$  are the intrawave modulation frequencies, being 6Hz and 8Hz, respectively. The main frequency of  $v_1$  or  $v_2$  exhibits intrawave frequency modulation of nonlinear systems: the further from this equilibrium point, the stronger the nonlinearity (Huang et al., 1998). The instantaneous frequency of  $v_1$  and  $v_2$  are:

$$f_{v_1}(t) = \frac{1}{2\pi} \frac{d\theta_{21}}{dt} = f_a + 0.5 f_b \cos(2\pi f_b t) \quad (3-26)$$

$$f_{v_2}(t) = \frac{1}{2\pi} \frac{d\theta_{22}}{dt} = f_a + 0.5 f_c \cos(2\pi f_c t) \quad (3-27)$$

where  $\theta_{21} = 2\pi f_a t + 0.5 \sin(2\pi f_b t)$  and  $\theta_{22} = 2\pi f_a t + 0.5 \sin(2\pi f_c t)$ .

The two simulated EEG channel signals  $Y_1$  and  $Y_2$  are defined as:

$$Y_1 = 0.8 u_1 + 0.2 v_1 + n_1 \quad (3-28)$$

$$Y_2 = 0.8 u_2 + 0.2 v_2 + n_2 \quad (3-29)$$

White Gaussian noise  $n_1$  and  $n_2$  were added to make the SNR of the observed  $Y_1$  and  $Y_2$  being 40dB. We used the function 'awgn' in Matlab to add white Gaussian noise to the signals. This function assumes that the power of the input signal is 0 dBW and noisePower is calculated by  $\frac{1}{10^{(SNR/10)}}$  dBW. The noise is produced by 'sqrt(noisePower)\*randn(size(signal))' in Matlab.

CWT is obtained using the analytic Morlet wavelet. Morlet is one of the most popular wavelet functions for biomedical signals due to its suitability and simplicity (Qassim et al., 2012). It has several advantages for the time-frequency analysis (Cohen, 2019). Firstly, Morlet convolution keeps the temporal resolution the same as the original signal. Secondly, the Morlet wavelet has a Gaussian distribution in both the time and frequency domain (Mouraux and Iannetti, 2008), non-sharp edges of which can minimise ripple effects that may be misinterpreted as oscillations. Thirdly, compared to other

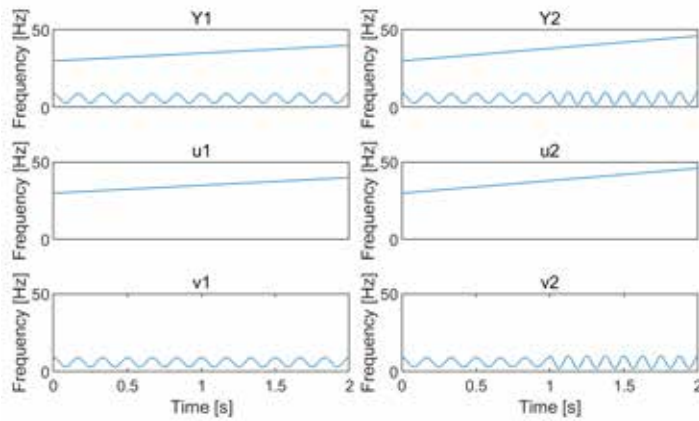
wavelets, Morlet convolution is more computationally efficient by involving the fast Fourier transform.

Two fine-tuned examples are presented to illustrate that RHHT has the following advantages over CWT: characterizing non-linear interwave with high time-frequency resolution, and capturing sudden changes in frequency and amplitude, which make the RHHT cross-spectrum superior in analyzing the linear connectivity of two-channel signals. The  $IMF_s$  from CEEMDAN are compared with the  $Band_s$  from Wavelet Convolutional Filter. In order to maintain the consistency of the Wavelet Convolutional Filter, we obtain the corresponding  $Band_s$  components by setting the centre frequency and bandwidth parameters as the dominant frequency and bandwidth of the first and second  $IMF$ . The RHHT cross-spectrums of  $Y_1$ ,  $Y_2$  and their corresponding  $IMF_s$  are compared with CWT cross-spectrums of  $Y_1$ ,  $Y_2$  and their corresponding  $Band_s$ .

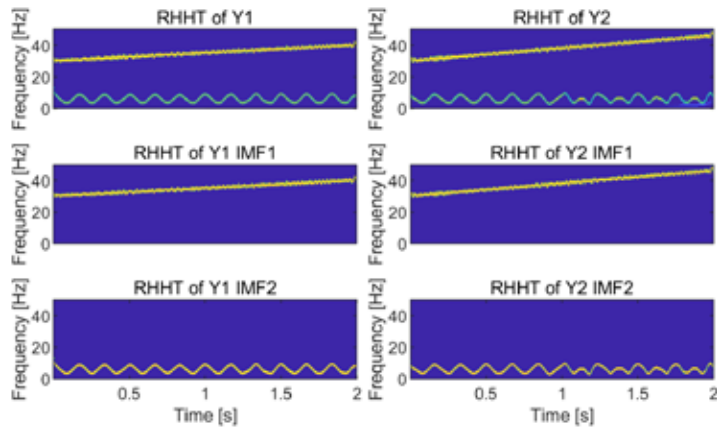
### 3.1.3.1 Example 1

In order to evaluate the performance of the proposed method in capturing similar and dissimilar oscillations within two signals, based on Eq. (2-24) and Eq. (2-25), the first half of the  $v_2$  component is adjusted to be the same as the first half of the  $v_1$  component. Figure 3-2(a) shows the theoretical instantaneous frequency of  $Y_1$ ,  $Y_2$  and their components  $u_1$ ,  $u_2$  and  $v_1$ ,  $v_2$  respectively. It can be observed from the second row of Figure 3-2(a) that  $u_1$  and  $u_2$  have a relatively similar instantaneous frequency in the beginning stage and the difference between them increases with time. The instantaneous frequencies of  $v_1$  and  $v_2$  are same in the first half, and the second half is quite different.





(a)



(b)

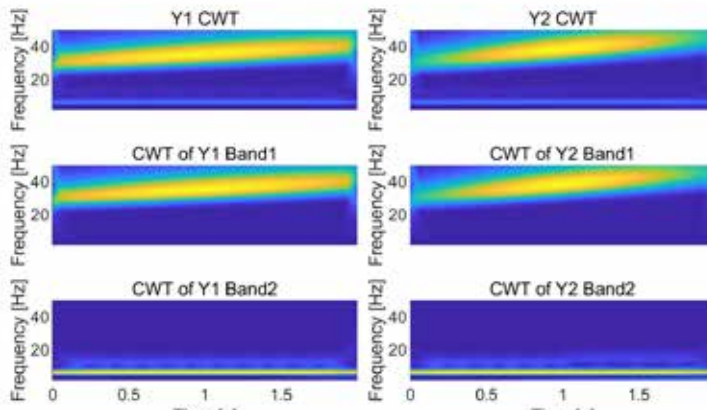


**Figure 3-2 Example 1 (a) Instantaneous frequency of Y1, Y2 and their components; b) RHHT of Y1, Y2 and associated IMFs; (c) CWT of Y1, Y2 and associated bands.**

Figure 3-2(b) and Figure 3-2(c) illustrate the time-frequency results of RHHT

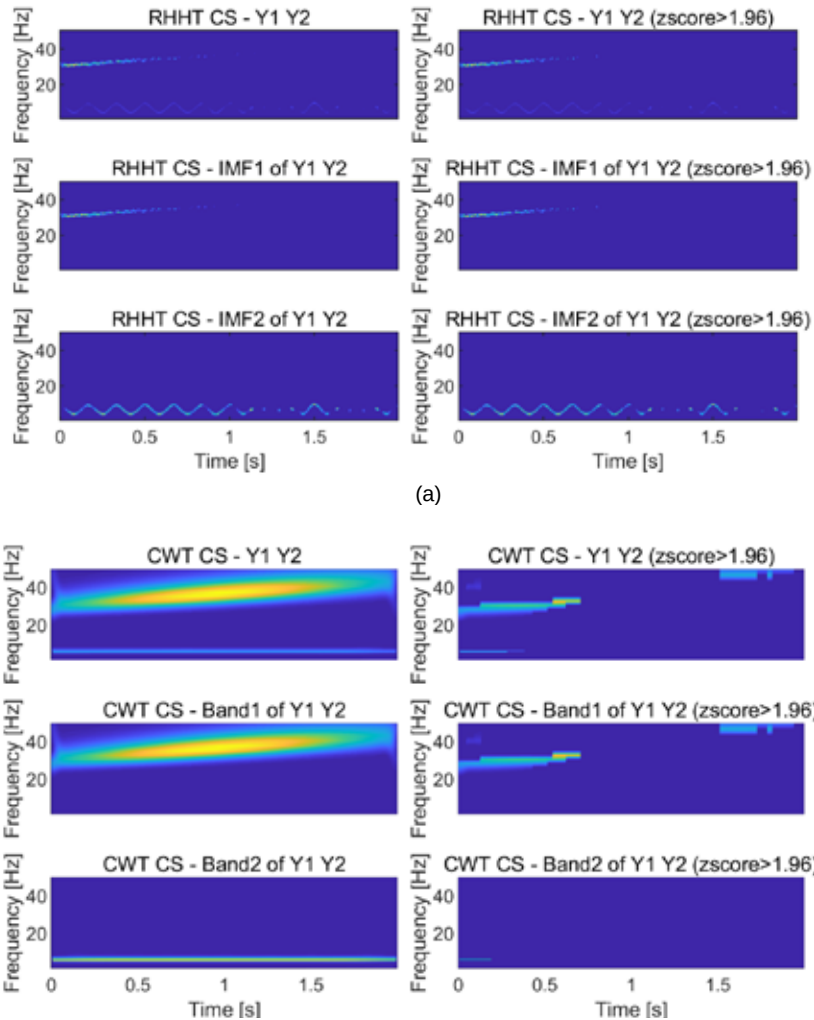


**Figure 3-3 Example 1 (a) RHHT cross-spectrum (CS) of  $Y_1$ ,  $Y_2$ , and IMFs before and after significant test; (b) CWT cross-spectrum of  $Y_1$ ,  $Y_2$ , and associated bands before and after significant test.**

and CWT. The first row shows the time-frequency maps of  $Y_1$  and  $Y_2$ ; the second and third rows show the time-frequency maps of  $Y_1$ 's and  $Y_2$ 's first and second IMF or band. The estimated instantaneous frequency energy of RHHT in Figure 3-2(b) is accurately localised in both frequency and time domains. RHHT spectrums of  $Y_1$  and  $Y_2$  contain two prominent components: the high-frequency component fluctuating increasingly from 30Hz to 40 or 46Hz, and the strongly intrawave frequency-modulated oscillations with intrinsic frequency 6Hz, in agreement with the theoretical instantaneous frequency shown in Figure 3-2(a). The observed high time-frequency resolution leads to better

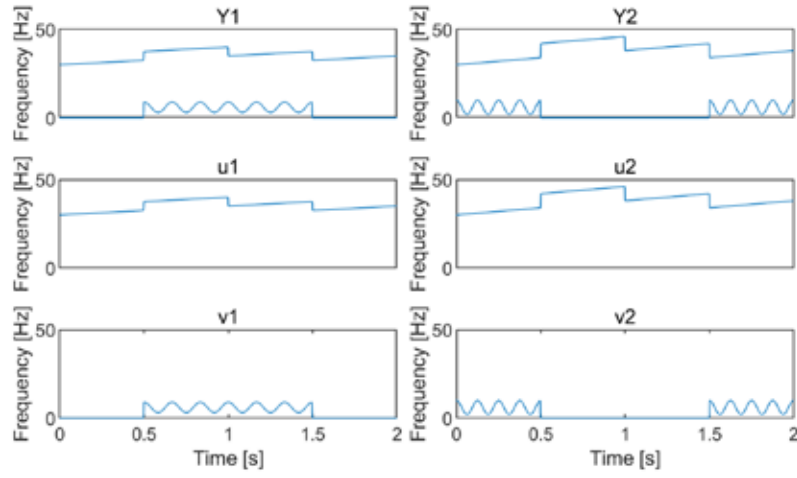
characterization of oscillations pattern of  $Y_1$ ,  $Y_2$  and their components  $u_1$ ,  $u_2$  and  $v_1$ ,  $v_2$ . For the CWT results shown in Figure 3-2 (c), neither the energy density nor the frequency is well localized. Corresponding to  $u_1$  and  $u_2$ , due to the low resolution of CWT in the high-frequency part, the results of two signals'  $Band_1$  spread energy over a much wider frequency range but fail to detect any truly detailed frequency variations. CWT spectrums of two signals'  $Band_2$  show a rich distribution of harmonics but no details of the intrawave frequency modulations. The second half period of the two signals'  $Band_2$  seems almost the same, which is different from the true theoretical state. This indicates the effective property of the RHHT spectrum in capturing the distortion harmonic components to represent the dynamic signal with modulated frequency.

The RHHT cross-spectrums of  $Y_1$ ,  $Y_2$  and their corresponding IMFs are shown in Figure 3-3(a), where it can be observed from the second column that energy density is well localized within 95% statistical significance. The RHHT cross-spectrum of  $IMF_1$  of  $Y_1$ ,  $Y_2$  in the second row shows similar time-frequency connectivity in the beginning stage and this connectivity disappears in the second half period, which matches the ground truth shown in Figure 3-2 (a). The  $IMF_2$  of  $Y_1$ ,  $Y_2$  in the third row of Figure 3-3 (a) further confirms this conclusion by successfully picking up the high and consistent association of  $v_1$  and  $v_2$  in the first half period and a less consistent association in the second half. The results of CWT cross-spectrum shown in Figure 3-3 (b), particularly the ones after the significant test, cannot fully capture the dynamic time-frequency association of the two signals. For example, the limited significant association is detected in  $Band_2$  in the first half which should be high and consistent. A significant association is detected in  $Band_1$  in the second half period, which is not true.

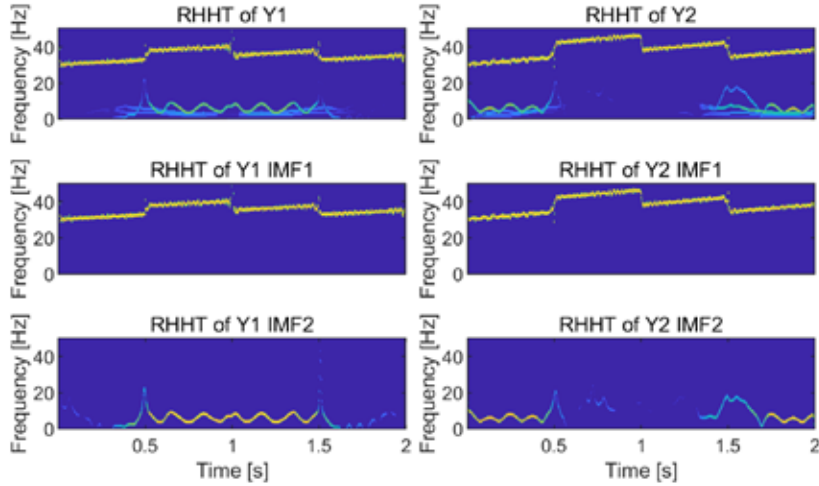
### 3.1.3.2 Example 2

In order to simulate the transient property of the non-stationary signal, based on Eq. (2-24) and Eq. (2-25), the Chirp and Cos components of the two synthetic EEG signals are adjusted by suddenly changing the frequency and amplitude. The second quarter and the fourth quarter of  $u_1$  are swapped, as is  $u_2$ . The

amplitude of the first and last quarter of  $v_1$  component is set to zero, while the



(a)



(b)

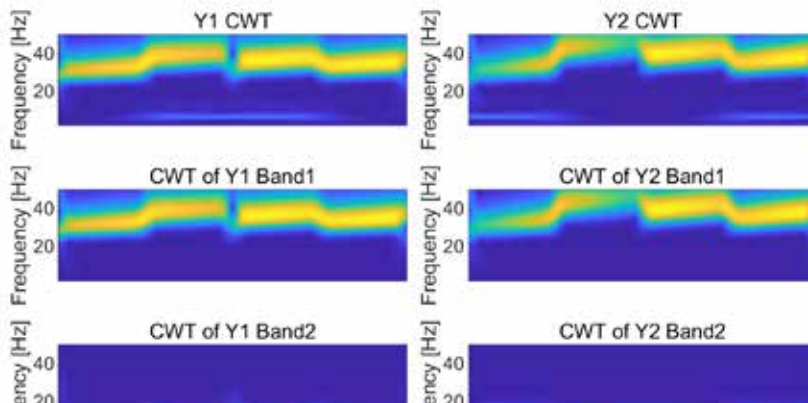


**Figure 3-4 Example 2 (a) Instantaneous frequency of Y1, Y2 and their components; (b) RHHT of Y1, Y2 and associated IMFs; (c) CWT of Y1, Y2 and associated bands.**

middle half of  $v_2$  component is set to zero. Figure 3-4(a) shows the theoretical instantaneous frequency of  $Y_1$ ,  $Y_2$  with their components  $u_1$ ,  $u_2$ , and  $v_1$ ,  $v_2$ .

Figure 3-4(b) and (c) illustrate the time-frequency results of RHHT and CWT. The first row shows the time-frequency maps of  $Y_1$  and  $Y_2$ ; the second and third rows show the time-frequency maps of the IMFs or Bands of  $Y_1$  and  $Y_2$ . It can be observed that the instantaneous frequency results of RHHT are more consistent

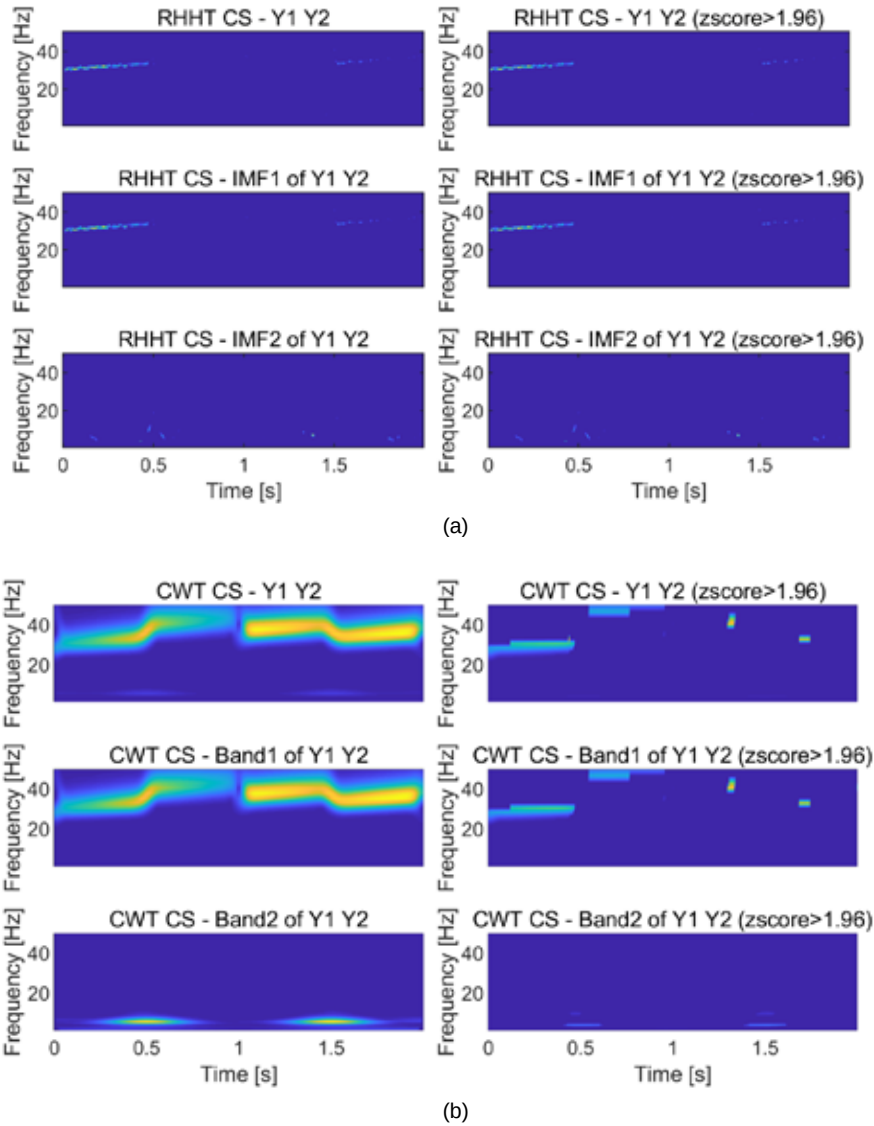


**Figure 3-5 Example 2 (a) RHHT cross-spectrum (CS) of Y1, Y2 and IMFs before and after significant test; (b) CWT cross-spectrum of Y1, Y2, and bands before and after significant test.**

with Figure 3-4(a) than the CWT results. In Figure 3-4(b), the time-frequency spectrum of  $IMF_1$  strictly reveals sudden changes of frequency at 0.5s, 1s and 1.5s, and the time-frequency spectrum of  $IMF_2$  undergoes intrawave frequency modulation and captures the sudden changes in amplitude energy at 0.5s, and 1.5s. In Figure 3-4(c), the time-frequency results of CWT reveal the transient property of the signal but are encumbered by harmonic components and the edge effects. There is energy dispersion at the sudden changes in frequency and amplitude of the two components. This indicates that RHHT spectrum can better capture non-stationarity and reveals more details of the system with a higher resolution.

In the second row of Figure 3-5(a), the RHHT cross-spectrum of  $Y_1$  and  $Y_2$ 's  $IMF_1$  shows similar time-frequency connectivity in the beginning and ending stage, which does not exist in the middle half period, in agreement with the ground truth shown in Figure 3-4(a). In the third row of Figure 3-5(a), the RHHT cross-spectrum of  $Y_1$  and  $Y_2$ 's  $IMF_2$  shows low energy of discontinuous connectivity at time 0.5s and 1.5s. However, in Figure 3-5(b), CWT cross-spectrum of  $Y_1$  and  $Y_2$ 's  $Band_1$  shows that the connectivity energy exists at some high-frequency areas with statistical significance, which may be caused by the low-frequency resolution. CWT cross-spectrum of  $Y_1$  and  $Y_2$ 's  $Band_2$  show divergent connectivity energy at time 0.5s and 1.5s within a certain time range.

To investigate the effect of and sensitivity to noise, we tested the results of the two synthetic examples with lower SNRs being as 30dB and 20dB, shown in Figure 3-11, Figure 3-10, Figure 3-12, Figure 3-13, Figure 3-14, Figure 3-15 in the Supplementary Material. For the data with the SNR being 30dB, a similar conclusion as that of 40dB can be drawn. For the data with the SNR being 20dB, compared with CWT, RHHT seems to be more sensitive to noise. However, this kind of problem can be addressed by applying a Gaussian filter before the analysis, as shown in Figure 3-14 and Figure 3-15. The superior performance of RHHT is observed again after the Gaussian filter processing.

### 3.1.4 Application To EEG Data

This section introduces a clinical application of the proposed technique to assist the diagnosis of generalized epilepsy. In this work we retrospectively selected video EEGs (vEEGs) with occasional generalised interictal epileptiform discharges (IEDs), providing electrophysiological evidence for the diagnosis of epilepsy. All selected epilepsy patients were isolated from the Royal Hallamshire Hospital (Sheffield, UK) Department of Neurophysiology database with the following inclusion criteria: standard interictal EEG available containing at least one well defined generalised IED; age between late teens to 61 years (based on a cohort of healthy control (HC) available from previous work to ensure no significant age differences between groups occurred); previous history of at least one witnessed generalised tonic-clonic seizure without any other known type of seizures. The following exclusion criteria were also applied: learning difficulties; sleep deprivation the night before the EEG was recorded; the known history of drug addiction; refractory epilepsy; any other known neurological disorder other than epilepsy.

The investigated EEG data comes from two groups: epileptic participants and HC. The data of 10 HC (6 females and 4 males, mean age  $37 \pm 15$ y) and 14 epilepsy participants (9 females and 5 males, mean age  $33 \pm 12$ y) were collected, details of which can be seen in Table 3-1 in the Supplementary Material. The epilepsy participants were patients diagnosed with generalized epilepsy and otherwise neurologically normal, without other underlying medical conditions. It should be noted that epilepsy participants were consecutively selected for the best possible match of age to the HC cohort. Additionally, the data for each participant was divided into two states: eyes open (EO) and eyes close (EC). For each subject of each eye state, the epoch lasts 10 seconds. Ethics approval for use of the patients' EEGs for the development of new quantitative EEG methods was obtained both from the University of Sheffield and the NHS ethics committees (SMBRER207 and 11/YH/0414). All HC subjects provided informed consent as part of a project approved by the Yorkshire and the Humber (Leeds West) Research Ethics Committee (reference number 14/YH/1070). A resting-state standard EEG recording was used.

Interictal (i.e. seizure-free) EEG data was selected for this paradigm. The recording was resampled to 500 Hz through an FIR filter.

Data was acquired with a Natus Headbox (Optima Medical, Ltd.) at a sampling

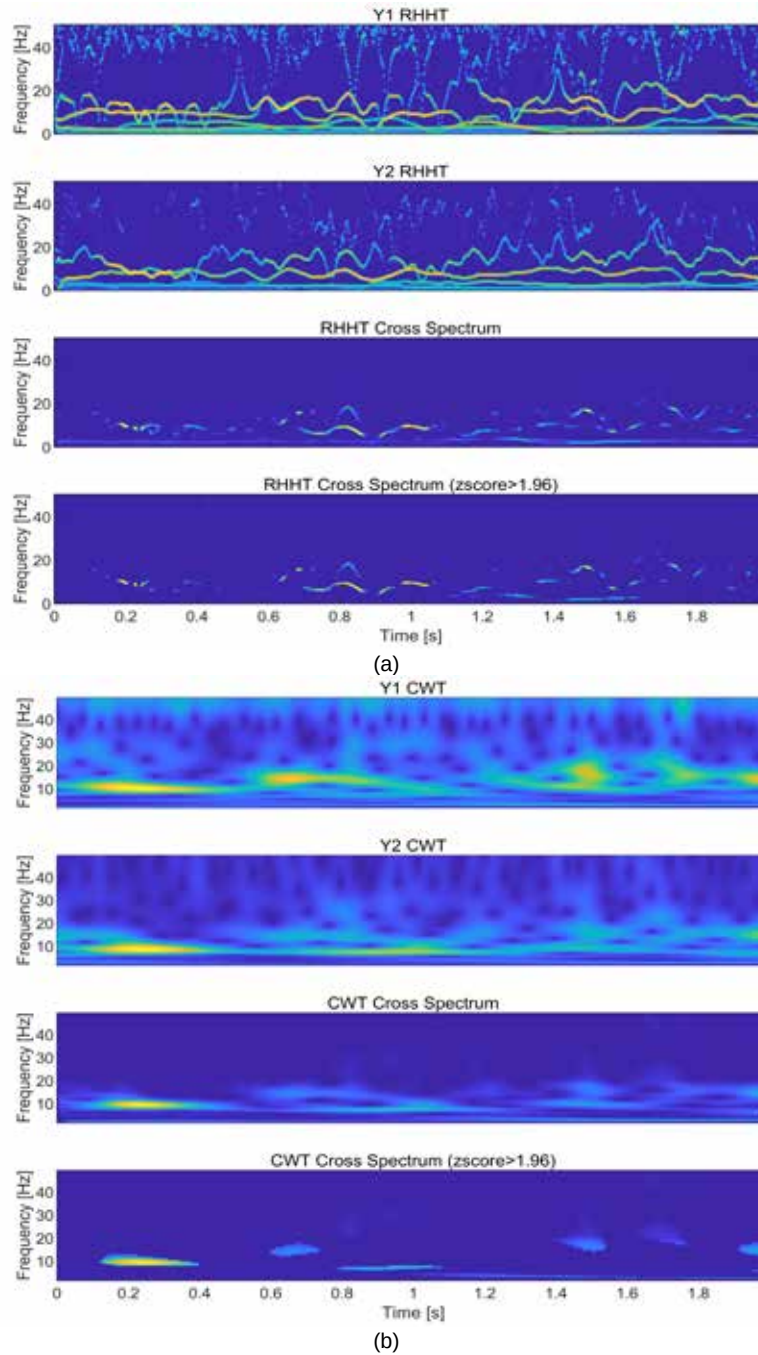


**Figure 3-6 (a) RHHT of Y1 and Y2 and the cross-spectrum significance result with 95% confidence interval (b) CWT of Y1 and Y2 and the cross-spectrum significance result with 95% confidence interval.**



rate of 500 Hz (analogue bandwidth 0.1–200 Hz) with the international 10–20 system of electrode placement positions. The EEG data was recorded from the standard 21 electrodes.

The EEG epochs were collected with Spike2 (version 9) software where data filtering and labeling were also undertaken. The filtering method was performed using the following equation  $y(t) = x(t) - 1/2c \int_{t-c}^{t+c} x(i) di$ , where  $x$  is the input

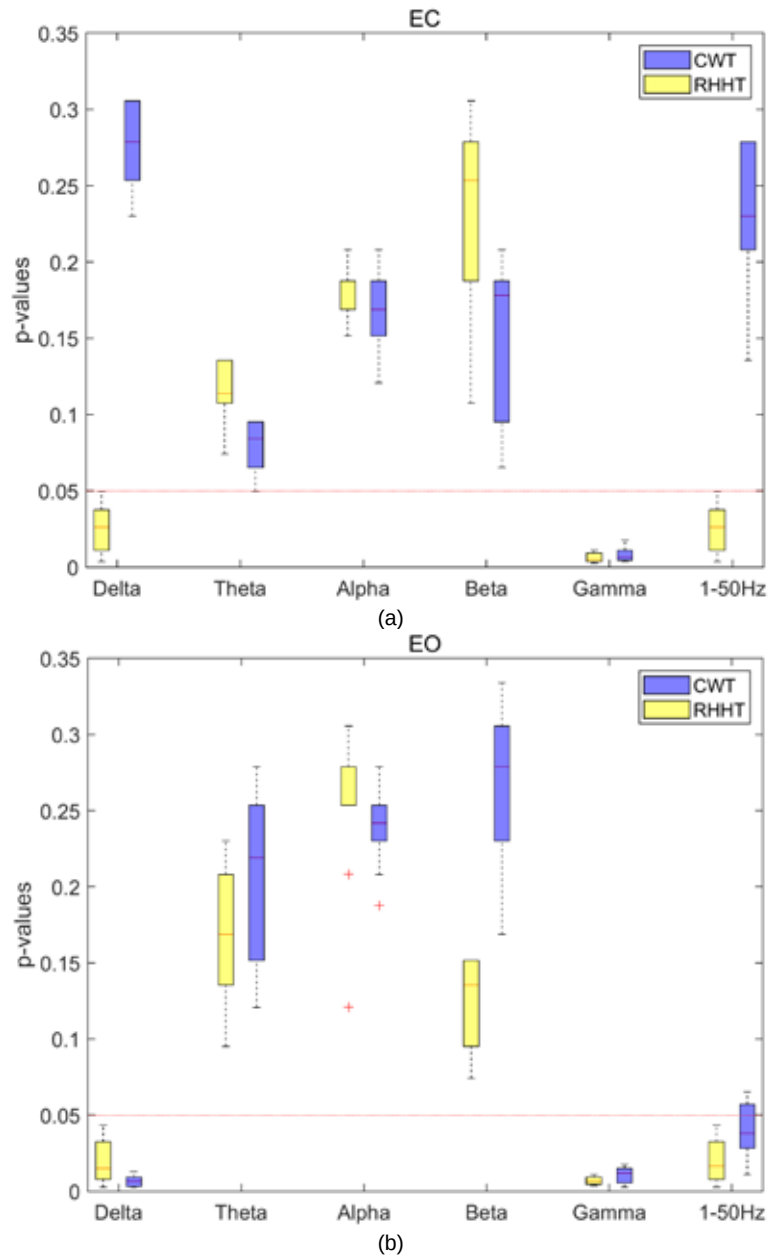


**Figure 3-7 The boxplots of top 10 smallest p-values of all 207 pairs for each frequency band of EC (a) and EO (b) data by RHHT and CWT cross-spectrum.**

EEG signal,  $y$  is the output signal,  $t$  is a discrete-time point and  $c$  is a time constant value. The value of  $c$  is set at 0.2s. The filtering using a time constant result in a high pass filtering of the signal where the cut-off frequency is  $1/2\pi c$ , which is equal to 0.79Hz. All data were selected by a specialised physician to ensure no interictal EEG abnormalities were included. Additionally, care was taken to select relatively artefact-free epochs.

A bipolar montage approach was implemented to avoid volume conduction effects caused by the common reference electrode. The following bipolar channels were available: Fp2-F8, Fp1-F7, 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, O2-O1.

#### **3.1.4.1 RHHT Cross Spectrum**

We used C3-P3 as  $Y_1$  and T3-T5 as  $Y_2$  from one epilepsy patient as an example to compare the performance of RHHT and CWT in tracking the non-stationary association of EEG signals. It has been proven in previous research that the connectivity between this pair can provide an insight to differentiate epilepsy group and healthy control group (Pyrzowski et al., 2015). In the first two rows of Figure 3-6(a), the RHHT spectrums of  $Y_1$  and  $Y_2$  show irregularly distributed energy in the high-frequency part above 20Hz, strongly unevenness of the intrawave frequency variation around 5-20Hz frequency band, and low-frequency oscillations below 5Hz with considerable energy. In the third row, RHHT cross-spectrum between  $Y_1$  and  $Y_2$  shows strong oscillations with sharp frequency and time localization. In the last row, RHHT cross-spectrum within a 95% confidence interval shows the connectivity energy of some intrawave frequency oscillations. In the first two rows of Figure 3-6 (b), the CWT spectrums of  $Y_1$  and  $Y_2$  show smoothed distributed energy in the high-frequency part, some transient energy clusters around intermediate frequency, and almost uniform low-frequency response. In the third row, CWT cross-spectrum between  $Y_1$  and  $Y_2$  is smoother in both time and frequency with spurious harmonics. In the last row, CWT cross-spectrum within a 95% confidence interval shows low connectivity energy in the high-frequency part and strong energy of some

connectivity clusters. Comparing the third and last row, some spurious connectivity energy exists in the CWT cross-spectrum without the significant test but disappears in the result with the significant test.

Comparing Figure 3-6 (a) and Figure 3-6 (b), it is suggested that both cross-spectrums based on RHHT and CWT show similar locations of energy associations in time and frequency axes, but RHHT gives a sharper and more refined definition of the energy than CWT. The intrawave frequency modulations by the RHHT spectrum not only present a clear physical picture of the motion but also eradicate the need for eliminating the spurious harmonics to represent nonlinear and non-stationary signals. All the interesting details pertinent to the similar frequency modulations between the two EEG channels of this paradigm can be seen clearly in the RHHT cross-spectrum, whereas the CWT cross-spectrum only gives a profuse display of harmonics with limited detail.

#### **3.1.4.2 Statistical Inference for Potential Classification Benchmark**

The Mann-Whitney U test (Ruxton, 2006), also known as the Wilcoxon rank-sum test, is a nonparametric test for differences between two groups on a single, ordinal variable with no specific distribution. By contrast, the independent-samples  $t$ -test, also a test of two groups, requires the single variable to be measured at the interval or ratio level, rather than the ordinal level, and to be normally distributed. When data do not meet the parametric assumptions of a normal distribution, the Mann-Whitney U test tends to be more appropriate (McKnight and Najab, 2010). Our data sample size is small and it fails the normal distribution test. Therefore, the Mann-Whitney U test is used in this example to accept or reject the null hypothesis that the epileptic patients and HC samples are samples from continuous distributions with equal medians. If rejected at a critical probability level, typically  $<0.05$ , it infers that the two samples do not come from the same population. The attained  $p$ -values can give an idea of the significance level of the results (Hlinka et al., 2010).

Considering the estimated measures of cross-spectrum between two signals, with 23 bipolar channels used in this Chapter, there are 253 possible combinations ( $C_2^{23}$ ) when any 2 bipolar derivations are paired together. These

were organised in a pairwise manner by taking the first bipolar derivation in the list (F8-F4) and pairing it with every other bipolar derivation according to their order on the list (F8-F4:F7-F3, F8-F4:F4-C4, F8-F4:F3-C3, etc.). The process was repeated for all other channels until the end of the list. However, since each channel is bipolar in nature, any pair with common electrode locations (such as F8-F4 and F4-C4) is neglected as this could lead to misleading high false correlation between the pair. A total of 46 channel pairs has this characteristic, and 207 bipolar pairs are therefore analysed in this Chapter.

In order to compare the effectiveness of cross-spectrum based on RHHT and CWT in feature extraction for the potential classification between epileptic patients and HC, we normalized each 10s epoch of EO and EC for all epileptic patients and HC. Then, the cross-spectrum values of all 207 pairs for EO and EC were calculated. Next, RHHT and CWT cross-spectra were averaged along both the time and frequency axis to obtain a singular value for each pair. Finally, 207 averaged cross-spectrum values were assessed by the Mann-Whitney U test for two groups. Figure 3-9 in the Supplementary Material illustrates the  $p$ -values of each band of EC and EO data obtained through the RHHT and CWT cross-spectrum.

To quantitatively compare the performance of RHHT and CWT cross-spectrum to differentiate these two groups, we selected the top 10 smallest  $p$ -values of the 207 pairs in each frequency band of EO and EC data. Figure 3-7(a) and (b) present the boxplots of EC and EO data, respectively. The red dashed line indicates the critical probability level ( $\alpha=0.05$ ) that can effectively distinguish epileptic patients and HC. For both Figure 3-7(a) and (b), in Theta, Alpha, and Beta bands, the distributions of the 10 smallest  $p$ -values of RHHT and CWT cross-spectrum are all higher than 0.05, which indicates invalid discrimination and will not be discussed further.

Figure 3-7(a), the distributions of the smallest 10  $p$ -values of RHHT cross-spectrum in Delta, Gamma, and 1-50Hz bands are all lower than 0.05, while the distribution of CWT cross-spectrum in Delta and 1-50Hz bands are above the threshold. The smallest 10  $p$ -values of CWT cross-spectrum only distributes

below 0.05 in Gamma band, but the mean value of its distribution is still higher than that of RHHT cross-spectrum.

Figure 3-7(b), distributions of the smallest 10 p-values in Delta, Gamma, and 1-50Hz bands of RHHT cross-spectrum are lower than 0.05 overall, whereas distribution of CWT cross-spectrum is only lower than 0.05 in Delta and Gamma bands. Although the distribution of CWT cross-spectrum in Delta band is lower than that of RHHT cross-spectrum, the distribution of CWT cross-spectrum in Gamma band is higher than that of RHHT cross-spectrum, and the distribution of CWT cross-spectrum in 1-50Hz band spans  $\approx 0.05$ .

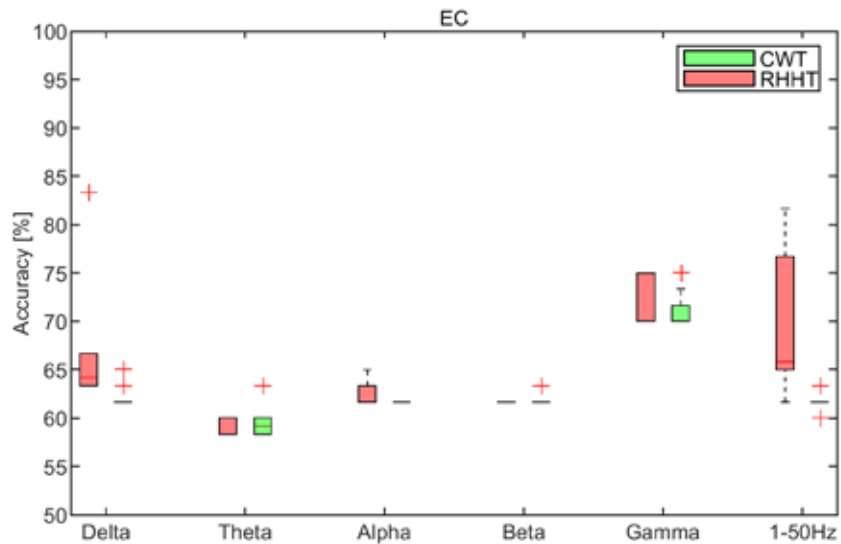
In order to test the effect of epoch length on the  $p$ -values, we split the 10s epoch into different lengths and calculated the corresponding  $p$ -values. By counting the effective bipolar pairs with  $p$ -values less than 0.05, we calculated the mean numbers for each segment length. The statistical results are shown in Table 3-2 in the Supplementary Material. It suggests that for both RHHT and CWT cross-spectrum, the length of data samples has a negligible effect on the number of effective pairs. Similar to the results of 10s data, other data lengths (5s, 2.5s, 1.25s, and 0.625s) confirm the conclusion that RHHT cross-spectrum can reveal more distinguishable bipolar pairs than CWT in most frequency bands.

#### **3.1.4.3 Classification Results using Support Vector Machine (SVM)**

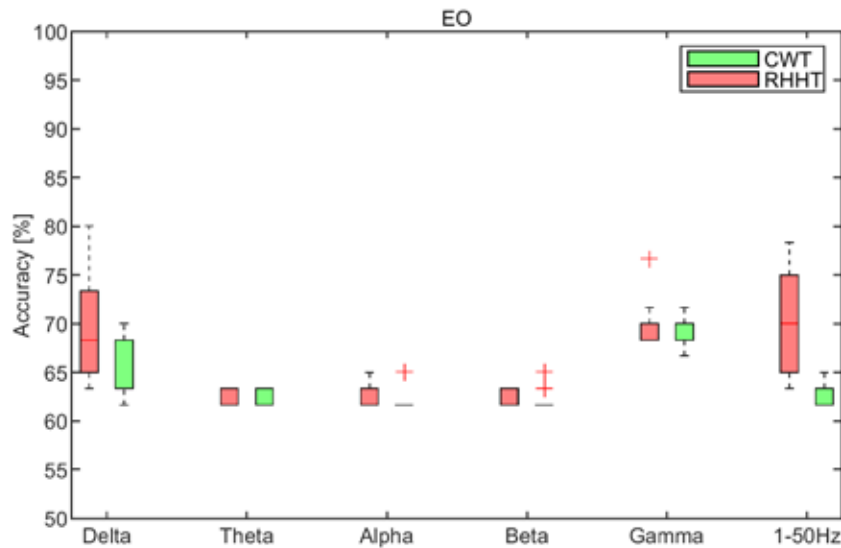
We conducted a direct classification of HC and epilepsy groups using the extracted 207 cross-spectrum features from RHHT and CWT. During the initial development of the classification solution, a few popular machine learning classifiers were implemented including SVM, decision tree and k-nearest neighbors (KNN). SVM was selected as the classifier in this work because it presents slightly better performance than other methods. The classification accuracy was calculated by 10 folder cross-validation.

The boxplots of 10-highest accuracy within all bipolar pairs of EC and EO data are shown in Figure 3-8(a) and (b) respectively. The classification results are in agreement with the  $p$ -values calculated by Mann-Whitney U-test shown in

Figure 3-7. In the three frequency bands with  $p$ -values less than 0.05 (Delta, Gamma, 1-50Hz), the classification results of SVM are better than the other three bands for both eye conditions. The average classification results of the cross-spectrum based on RHHT are higher than CWT although it has a slightly larger variation.



(a)



(b)

**Figure 3-8 The boxplots of top 10 highest classification accuracies of all 207 pairs for each frequency band of EC (a) and EO (b) data by RHHT and CWT cross-spectrum.**

### 3.1.5 Conclusion

This Chapter proposes a novel data analysis framework to track the dynamic association between EEG channels in the time-frequency domain. This framework aims to characterize frequency fluctuations with a high time-frequency resolution and capture the dynamic association in the frequency domain using a new indicator, called Revised Hilbert-Huang Transformation cross-spectrum. It has been demonstrated by the two simulation examples that, compared to CWT, RHHT can estimate the frequency modulations more accurately, as well as detect the abrupt change of frequency, amplitude, and corresponding associations. This characterisation will potentially provide more insight when applied to real EEG signals which are dominated by high non-linearity and non-stationarity. Implementation of the significant test on the proposed RHHT based cross-spectrum allows identification of the significant interactions by taking into account the effects of noise interference. The application of the proposed method on EEG data from 14 epileptic patients and 10 healthy controls further demonstrates the validity of this framework and its potential application in building benchmarks for classification. In the real EEG data paradigm, we have shown with the implementation of a statistical significance approach, the estimated cross-spectrum using RHHT can better distinguish these two groups than that using CWT. However, it should be noted that a further study is required to deeply understand what additional information of RHHT contributes to the improved classification by exploring the ground truth of brain cross-spectrum.

### 3.1.6 Supplementary

**Table 3-1 THE DETAILS OF PARTICIPANTS AND AVAILABLE DATA**

| GROUP    | GENDER | AGE | MEDICATIONS                               |
|----------|--------|-----|-------------------------------------------|
| EPILEPSY | F      | 40  | FLUOXETINE                                |
|          | F      | 45  | LAMOTRIGINE, LACOSAMIDE                   |
|          | F      | 58  | SODIUM VALPROATE,<br>LEVETIRACETAM        |
|          | F      | 29  | NIL                                       |
|          | F      | 29  | CARBAMAZEPINE                             |
|          | F      | 29  | NIL                                       |
|          | F      | 23  | LAMOTRIGINE                               |
|          | F      | 20  | NIL                                       |
|          | F      | 19  | LAMOTRIGINE                               |
|          | M      | 46  | LAMOTRIGINE,<br>LEVETIRACETAM             |
|          | M      | 45  | LAMOTRIGINE, SODIUM<br>VALPROATE          |
|          | M      | 36  | PHENOBARBITAL,<br>LAMOTRIGINE, PREGABALIN |
|          | M      | 28  | NIL                                       |
|          | M      | 19  | NIL                                       |
| HEALTHY  | F      | 41  | NIL                                       |
|          | F      | 52  | NIL                                       |
|          | F      | 61  | NIL                                       |
|          | F      | 23  | NIL                                       |
|          | F      | 23  | NIL                                       |
|          | F      | 22  | NIL                                       |
|          | M      | 49  | NIL                                       |
|          | M      | 48  | NIL                                       |
|          | M      | 30  | NIL                                       |
|          | M      | 22  | NIL                                       |



**Table 3-2 THE NUMBERS OF EFFECTIVE CROSS-CHANNELS FOR CLASSIFICATION WITH RHHT- AND CWT-BASED CROSS-SPECTRUM**

|        |        | EC   |      | EO   |      |
|--------|--------|------|------|------|------|
|        |        | RHHT | CWT  | RHHT | CWT  |
| 1-50Hz | 10s    | 10   | 0    | 12   | 7    |
|        | 5s     | 11   | 0    | 2.5  | 5    |
|        | 2.5s   | 9.8  | 0    | 6.5  | 3    |
|        | 1.25s  | 8    | 0.3  | 4.9  | 3    |
|        | 0.625s | 8.7  | 0.5  | 6    | 3    |
| Gamma  | 10s    | 31   | 28   | 52   | 38   |
|        | 5s     | 34.5 | 27.5 | 54.5 | 37.5 |
|        | 2.5s   | 34.3 | 27.5 | 52.3 | 37.3 |
|        | 1.25s  | 32.4 | 27.6 | 51   | 36.3 |
|        | 0.625s | 32.4 | 26.9 | 50.4 | 36.4 |
| Beta   | 10s    | 0    | 0    | 0    | 0    |
|        | 5s     | 0.5  | 0.5  | 0.5  | 0    |
|        | 2.5s   | 0    | 0.5  | 1.8  | 0    |
|        | 1.25s  | 0.5  | 1    | 3.1  | 0    |
|        | 0.625s | 0.6  | 1.1  | 4.5  | 0.1  |
| Alpha  | 10s    | 0    | 0    | 0    | 0    |
|        | 5s     | 0    | 0    | 0    | 0    |
|        | 2.5s   | 2.5  | 1    | 0    | 0    |
|        | 1.25s  | 2.4  | 2.4  | 0    | 0    |
|        | 0.625s | 3.6  | 2.1  | 0.7  | 0.1  |

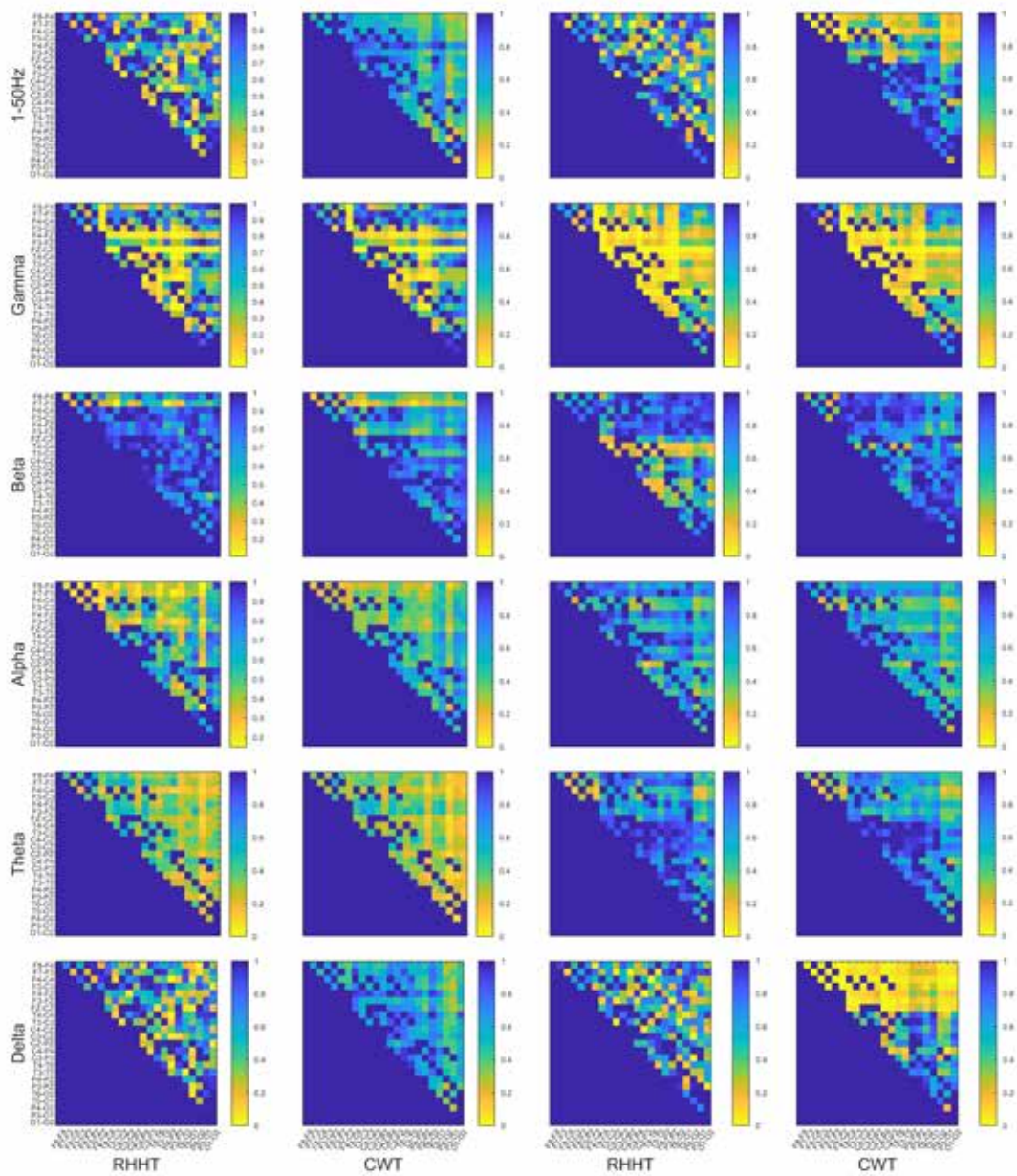
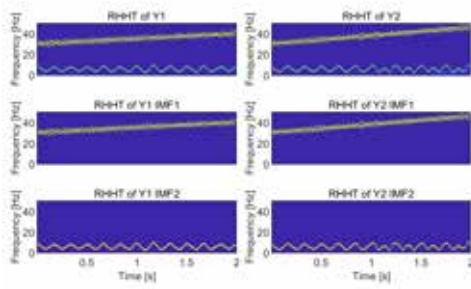
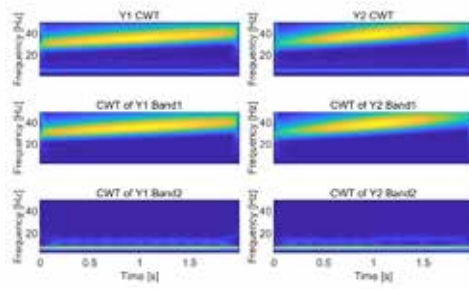


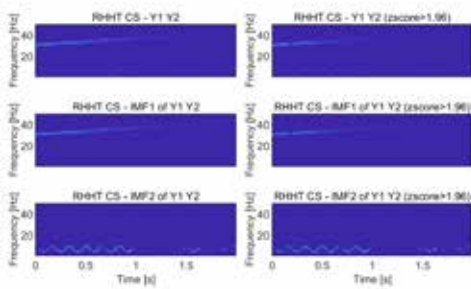
Figure 3-9 Rows from top to bottom show the results for EO data calculated by the frequency bands of 1-50Hz, Gamma, Beta, Alpha, Theta, and Delta respectively. The left two columns (EC data) and the right two columns (EO data) show the p-values of each cross-bipolar channel based on the cross-spectrums by RHHT and CWT respectively.



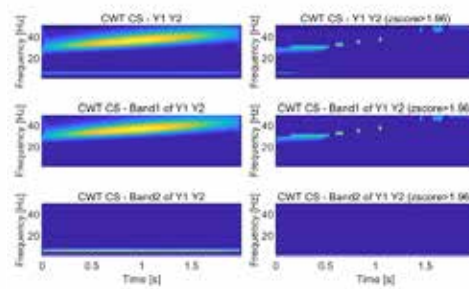
(a) RHHT of Y1, Y2 and IMFs



(b) CWT of Y1, Y2 and bands

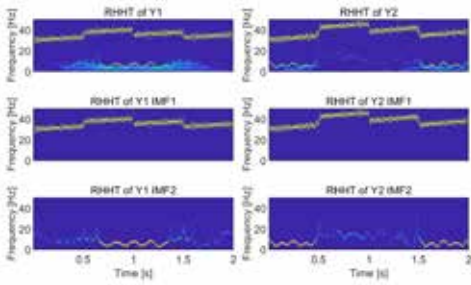


(c) RHHT cross-spectrum of Y1, Y2 and IMFs

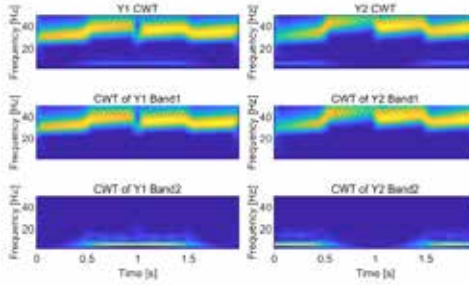


(d) CWT cross-spectrum of Y1, Y2 and bands

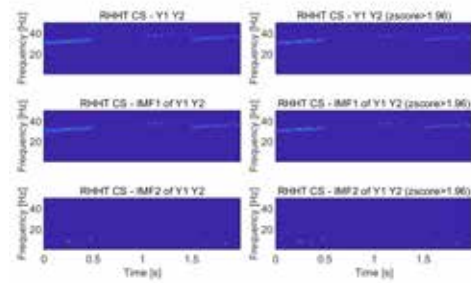
**Figure 3-10 The results of synthetic Examples 1 with SNR being 30dB.**



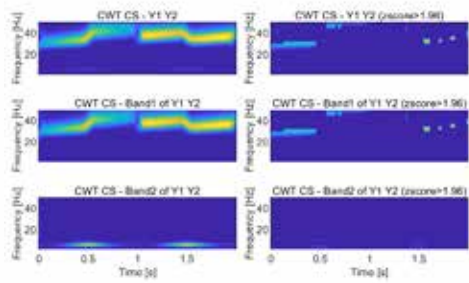
(a) RHHT of Y1, Y2 and IMFs



(b) CWT of Y1, Y2 and bands

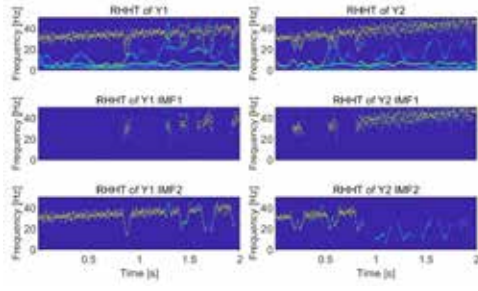


(c) RHHT cross-spectrum of Y1, Y2 and IMFs

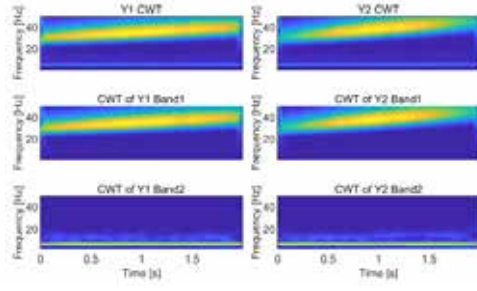


(d) CWT cross-spectrum of Y1, Y2 and bands

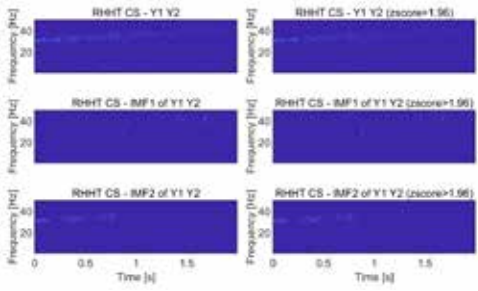
**Figure 3-11 The results of synthetic Examples 2 with SNR being 30dB.**



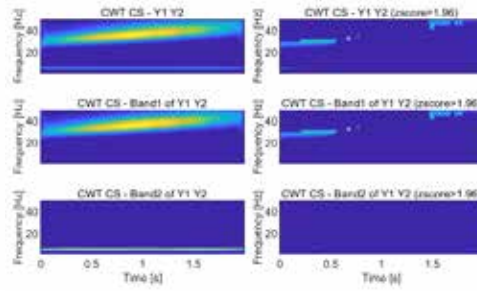
(a) RHHT of Y1, Y2 and IMFs



(b) CWT of Y1, Y2 and bands

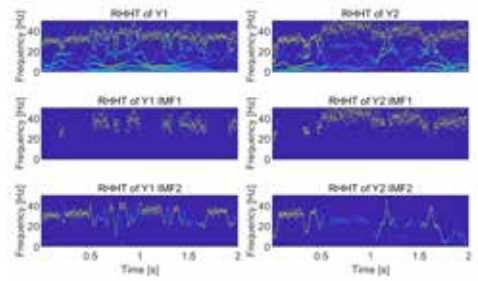


(c) RHHT cross-spectrum of Y1, Y2 and IMFs

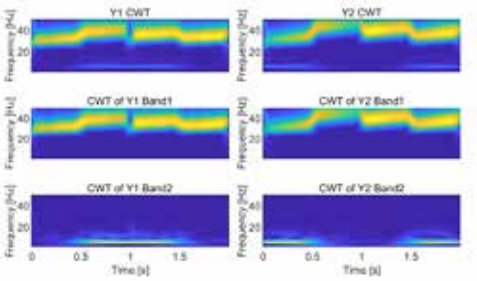


(d) CWT cross-spectrum of Y1, Y2 and bands

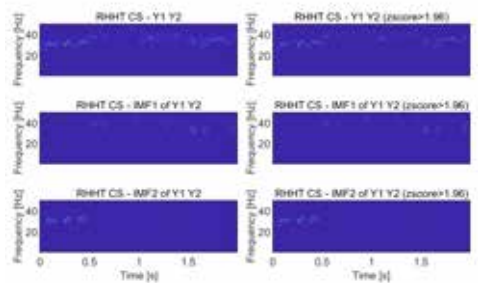
**Figure 3-12 The results of synthetic Examples 1 with SNR being 20dB.**



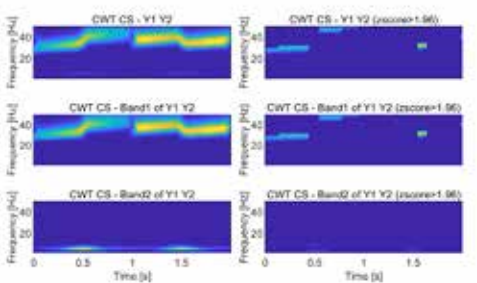
(a) RHHT of Y1, Y2 and IMFs



(b) CWT of Y1, Y2 and bands



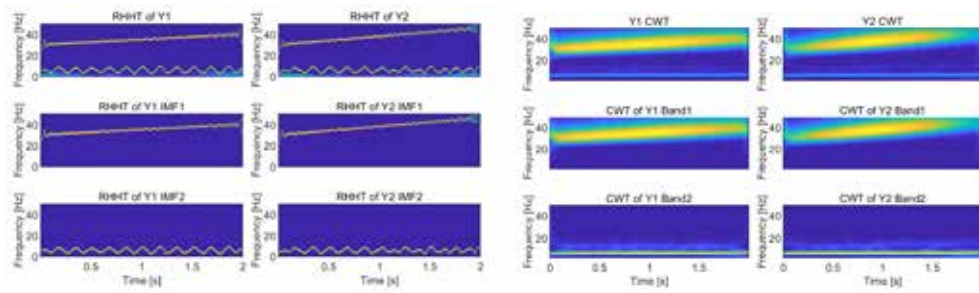
(c) RHHT cross-spectrum of Y1, Y2 and IMFs



(d) CWT cross-spectrum of Y1, Y2 and bands

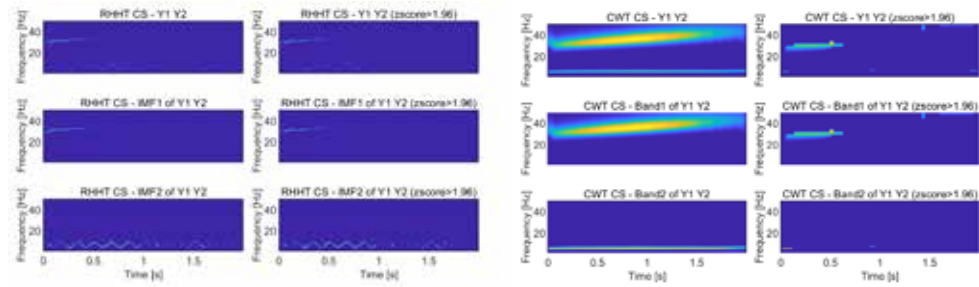
**Figure 3-13 The results of synthetic Examples 2 with SNR being 20dB.**





(a) RHHT of Y1, Y2 and IMFs

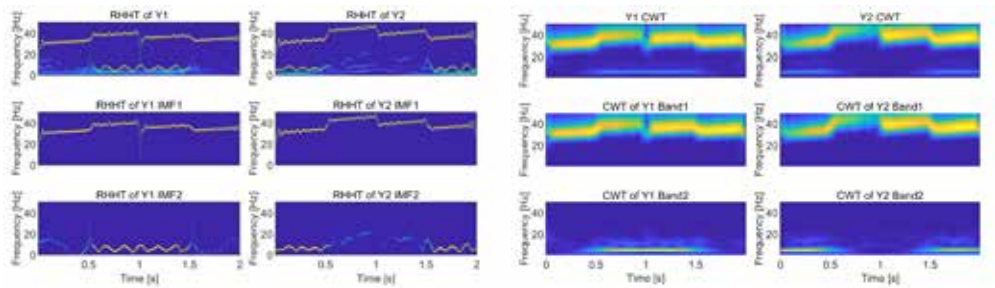
(b) CWT of Y1, Y2 and bands



(c) RHHT cross-spectrum of Y1, Y2 and IMFs

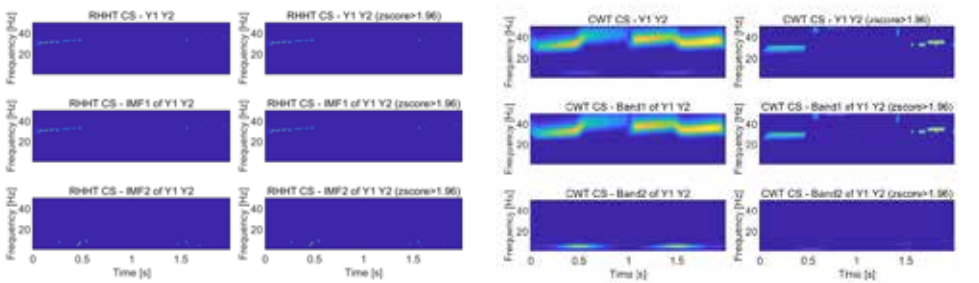
(d) CWT cross-spectrum of Y1, Y2 and bands

**Figure 3-14 The results of synthetic Examples 1 with SNR being 20dB with filtering.**



(a) RHHT of Y1, Y2 and IMFs

(b) CWT of Y1, Y2 and bands



(c) RHHT cross-spectrum of Y1, Y2 and IMFs

(d) CWT cross-spectrum of Y1, Y2 and bands

**Figure 3-15 The results of synthetic Examples 2 with SNR being 20dB with filtering.**

### 3.1.7 References

Akbarian, B., Erfanian, A., 2020. Biomedical Signal Processing and Control A framework for seizure detection using effective connectivity , graph theory , and multi-level modular network. *Biomed. Signal Process. Control* 59, 101878. <https://doi.org/10.1016/j.bspc.2020.101878>

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

Babiloni, C., Lizio, R., Marzano, N., Capotosto, P., Soricelli, A., Triggiani, A.I., Cordone, S., Gesualdo, L., Del Percio, C., 2016. Brain neural synchronization and functional coupling in Alzheimer's disease as revealed by resting state EEG rhythms. *Int. J. Psychophysiol.* 103, 88–102. <https://doi.org/10.1016/j.ijpsycho.2015.02.008>

Boersma, M., Smit, D.J. a., Boomsma, D.I., de Geus, E.J. c., delemarre-Van de Waal, H.A., Stam, C.J., 2013. Growing Trees in Child Brains: Graph Theoretical Analysis of Electroencephalography-Derived Minimum Spanning Tree in 5- and 7-Year-Old Children Reflects Brain Maturation. *Brain Connect.* 3, 50–60. <https://doi.org/10.1089/brain.2012.0106>

Chen, J., Ros, T., Gruzelier, J.H., 2013. Dynamic Changes of ICA-Derived EEG Functional Connectivity in the Resting State 868, 852–868. <https://doi.org/10.1002/hbm.21475>

Chiang, J., Wang, Z.J., McKeown, M.J., 2011. A generalized multivariate autoregressive (GmAR)-based approach for EEG source connectivity analysis. *IEEE Trans. Signal Process.* 60, 453–465.

Cohen, M.X., 2019. A better way to define and describe Morlet wavelets for time-frequency analysis. *Neuroimage* 199, 81–86.

Cohen, M.X., 2017a. Where does EEG come from and what does it mean? *Trends Neurosci.* 40, 208–218.

Cohen, M.X., 2017b. Multivariate cross-frequency coupling via generalized eigendecomposition. *Elife* 6, e21792.

Cohen, M.X., 2014. Analyzing neural time series data: theory and practice. MIT press.

de Cheveigné, A., 2010. Time-shift denoising source separation. *J. Neurosci. Methods* 189, 113–120.

Diessen, E. Van, Numan, T., Dellen, E. Van, Kooi, A.W. Van Der, Boersma, M., Hofman, D., Lutterveld, R. Van, Dijk, B.W. Van, Straaten, E.C.W. Van, Hillebrand, A., Stam, C.J., 2015. Clinical Neurophysiology Opportunities and methodological challenges in EEG and MEG resting state functional brain network research. *Clin. Neurophysiol.* 126, 1468–1481. <https://doi.org/10.1016/j.clinph.2014.11.018>

Erra, R.G., Velazquez, J.L.P., Rosenblum, M., 2017. Neural synchronization from the perspective of non-linear dynamics. *Front. Comput. Neurosci.* 11, 98. <https://doi.org/10.3389/fncom.2017.00098>

Faust, O., Acharya, U.R., Adeli, H., Adeli, A., 2015. Wavelet-based EEG processing for computer-aided seizure detection and epilepsy diagnosis. *Seizure* 26, 56–64.

Friston, K.J., 2011. Functional and Effective Connectivity: A Review. *Brain Connect.* 1, 13–36. <https://doi.org/10.1089/brain.2011.0008>

Guggisberg, A.G., Honma, S.M., Findlay, A.M., Dalal, S.S., Kirsch, H.E., Berger, M.S., Nagarajan, S.S., 2008. Mapping functional connectivity in patients with brain lesions. *Ann. Neurol.* 63, 193–203.

Hamed, M., Salleh, S.-H., Noor, A.M., 2016. Electroencephalographic motor imagery brain connectivity analysis for BCI: a review. *Neural Comput.* 28, 999–1041.

Hlinka, J., Alexakis, C., Diukova, A., Liddle, P.F., Auer, D.P., 2010. NeuroImage Slow EEG pattern predicts reduced intrinsic functional connectivity in the default -subject analysis. *Neuroimage* 53, 239–246. <https://doi.org/10.1016/j.neuroimage.2010.06.002>

- Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.-C., Tung, C.C., Liu, H.H., 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. London. Ser. A Math. Phys. Eng. Sci.* 454, 903–995.
- Jones, S.R., 2016. When brain rhythms aren't 'rhythmic': implication for their mechanisms and meaning. *Curr. Opin. Neurobiol.* 40, 72–80.
- Korats, G., Le Cam, S., Ranta, R., Louis-Dorr, V., 2015. A space-time-frequency dictionary for sparse cortical source localization. *IEEE Trans. Biomed. Eng.* 63, 1966–1973.
- Marrelec, G., Daunizeau, J., Pélégrini-Issac, M., Doyon, J., Benali, H., 2005. Conditional correlation as a measure of mediated interactivity in fMRI and MEG/EEG. *IEEE Trans. Signal Process.* 53, 3503–3516.
- Mazaheri, A., Jensen, O., 2008. Asymmetric amplitude modulations of brain oscillations generate slow evoked responses. *J. Neurosci.* 28, 7781–7787.
- McKnight, P.E., Najab, J., 2010. Mann-Whitney U Test. *Corsini Encycl. Psychol.* 1.
- Mheich, A., Hassan, M., Khalil, M., Berrou, C., Wendling, F., 2015. A new algorithm for spatiotemporal analysis of brain functional connectivity. *J. Neurosci. Methods* 242, 77–81. <https://doi.org/10.1016/j.jneumeth.2015.01.002>
- Mouraux, A., Iannetti, G.D., 2008. Across-trial averaging of event-related EEG responses and beyond. *Magn. Reson. Imaging* 26, 1041–1054.
- Navarro, X., Porée, F., Beuchée, A., Carrault, G., 2015. Denoising preterm EEG by signal decomposition and adaptive filtering: A comparative study. *Med. Eng. Phys.* 37, 315–320. <https://doi.org/10.1016/j.medengphy.2015.01.006>
- Petrantonakis, P.C., Hadjileontiadis, L.J., 2012. Adaptive emotional information retrieval from EEG signals in the time-frequency domain. *IEEE Trans. Signal Process.* 60, 2604–2616.



- Pyrzowski, J., Sieminski, M., Sarnowska, A., Jedrzejczak, J., Nyka, W.M., 2015. Interval analysis of interictal EEG: Pathology of the alpha rhythm in focal epilepsy. *Sci. Rep.* 5, 1–10. <https://doi.org/10.1038/srep16230>
- Qassim, Y.T., Cutmore, T., James, D., Rowlands, D., 2012. FPGA implementation of Morlet continuous wavelet transform for EEG analysis, in: 2012 International Conference on Computer and Communication Engineering (ICCCE). pp. 59–64.
- Rehman, N., Mandic, D.P., 2010. Multivariate empirical mode decomposition. *Proc. R. Soc. A Math. Phys. Eng. Sci.* 466, 1291–1302.
- Rilling, G., Flandrin, P., Gonçalves, P., Lilly, J.M., 2007. Bivariate empirical mode decomposition. *IEEE Signal Process. Lett.* 14, 936–939.
- Ruxton, G.D., 2006. The unequal variance t-test is an underused alternative to Student's t-test and the Mann--Whitney U test. *Behav. Ecol.* 17, 688–690.
- Sakkalis, V., 2011. Review of advanced techniques for the estimation of brain connectivity measured with EEG/MEG. *Comput. Biol. Med.* 41, 1110–1117.
- Schreiber, T., Schmitz, A., 2000. Surrogate time series. *Phys. D Nonlinear Phenom.* 142, 346–382.
- Schumacher, E.M., Stiris, T.A., Larsson, P.G., 2015. Clinical Neurophysiology Effective connectivity in long-term EEG monitoring in preterm infants. *Clin. Neurophysiol.* 126, 2261–2268. <https://doi.org/10.1016/j.clinph.2015.01.017>
- Schutte, N.M., Hansell, N.K., de Geus, E.J.C., Martin, N.G., Wright, M.J., Smit, D.J.A., 2013. Heritability of resting state EEG functional connectivity patterns. *Twin Res. Hum. Genet.* 16, 962–969. <https://doi.org/10.1017/thg.2013.55>
- Tafreshi, T.F., Daliri, M.R., Ghodousi, M., 2019. Functional and effective connectivity based features of EEG signals for object recognition. *Cogn. Neurodyn.* 13, 555–566.

Theiler, J., Eubank, S., Longtin, A., Galdrikian, B., Farmer, J.D., 1992. Testing for nonlinearity in time series: the method of surrogate data. *Phys. D Nonlinear Phenom.* 58, 77–94.

Torres, M.E., Colominas, M.A., Schlotthauer, G., Flandrin, P., 2011. A complete ensemble empirical mode decomposition with adaptive noise, in: 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). pp. 4144–4147.

Uhlhaas, P.J., Singer, W., 2006. Neural Synchrony in Brain Disorders: Relevance for Cognitive Dysfunctions and Pathophysiology. *Neuron* 52, 155–168. <https://doi.org/10.1016/j.neuron.2006.09.020>

van Mierlo, P., Papadopoulou, M., Carrette, E., Boon, P., Vandenberghe, S., Vonck, K., Marinazzo, D., 2014. Functional brain connectivity from EEG in epilepsy: Seizure prediction and epileptogenic focus localization. *Prog. Neurobiol.* 121, 19–35. <https://doi.org/10.1016/j.pneurobio.2014.06.004>

Vecchio, F., Miraglia, F., Quaranta, D., Granata, G., Romanello, R., Marra, C., Bramanti, P., Rossini, P.M., 2016. Cortical connectivity and memory performance in cognitive decline: a study via graph theory from EEG data. *Neuroscience* 316, 143–150.

Vicente, R., Wibral, M., Lindner, M., Pipa, G., 2011. Transfer entropy—a model-free measure of effective connectivity for the neurosciences. *J. Comput. Neurosci.* 30, 45–67.

Wendling, F., Ansari-asl, K., Bartolomei, F., Senhadji, L., 2009. From EEG -based evaluation of interdependence measures 183, 9–18. <https://doi.org/10.1016/j.jneumeth.2009.04.021>

Wendling, F., Chauvel, P., Biraben, A., Bartolomei, F., 2010. From intracerebral EEG signals to brain connectivity: Identification of epileptogenic networks in partial epilepsy. *Front. Syst. Neurosci.* 4, 1–13. <https://doi.org/10.3389/fnsys.2010.00154>

Wu, Z., Huang, N.E., 2009. Ensemble empirical mode decomposition: a noise-assisted data analysis method. *Adv. Adapt. Data Anal.* 1, 1–41.

Zhang, Y., Liu, B., Ji, X., Huang, D., 2017. Classification of EEG signals based on autoregressive model and wavelet packet decomposition. *Neural Process. Lett.* 45, 365–378.

Zhao, Yifan, Laguna, R.C., Zhao, Yitian, Liu, J.J., He, X., Yianni, J., Sarrigiannis, P.G., 2018. A wavelet-based correlation analysis framework to study cerebromuscular activity in essential tremor. *Complexity* 2018. <https://doi.org/10.1155/2018/7269494>

## **3.2 Ultra-high-resolution time-frequency analysis of EEG to characterise brain functional connectivity with the application in Alzheimer's disease.**

### **3.2.1 Introduction**

Alzheimer's disease (AD) is one of the most common neurodegenerative diseases, resulting in the loss of memory and other cognitive impairments (Blinowska et al., 2017; Ferreri et al., 2016). The number of patients affected by AD and the difficulties in treating this disorder provoke massive demands for the early diagnosis of the condition and effective approaches for monitoring disease progression. In the past decades, electroencephalogram (EEG) has attracted significant interest since it is economical, non-invasive and with an ultra-high time resolution. A variety of biomarkers were extracted from EEGs in AD-related research, such as amplitude (Poil et al., 2013), power spectral densities (PSD) (Sankari et al., 2011), phase-related features (Engels et al., 2015; Kent et al., 2021), alpha rhythm power (Angelakis et al., 2007; Babiloni et al., 2013; Benwell et al., 2020; Sadaghiani and Kleinschmidt, 2016; Schmidt et al., 2013) wavelet energy (Jeong et al., 2016), features from graph theory (delEtoile and Adeli, 2017; Miraglia et al., 2016) and brain connectivity estimations (Blinowska et al., 2017; delEtoile and Adeli, 2017; Durongbhan et al., 2019; Vecchio et al., 2017; Yu et al., 2018; Zhao et al., 2020). Peak frequency is also a promising

biomarker in the field of dementia and other neurodegenerative disorders (Grandy ., 2013; Ratti ., 2017). To support the classification of AD based on those features extracted from EEG recordings, many machine learning algorithms have been employed, such as linear discriminant analysis (LDA), logistic regression (LR), random forest, support vector machines (SVM), K-nearest neighbour (KNN) and deep learning (Durongbhan ., 2019; Kulkarni, 2019; Qiao ., 2018; Rajintha S Gunawardena ., 2021; Vecchio ., 2020)

Alpha and Beta peak frequency of a single channel was used as a quantitative EEG (qEEG) measure for the classification between AD and mild cognitive impairment (MCI). MCI represents a transitional period of neurological degeneration from normal aging to AD (Poil ., 2013). That is to say, the peak frequency may be a potential biomarker to evaluate the degree of brain degeneration. Dementia with Lewy bodies (DLB) can also be discriminated from AD since it has a significantly lower peak frequency (van der Zande ., 2018). Besides, different from young adults, older people show significant slowing of individual alpha peak frequency. In a word, there is increasing evidence from the literature suggesting the EEG peak frequency has great potential to advance dementia research and possibly one day to be translated into a clinically useful AD diagnostic tool.

However, most of the previous related techniques are univariate-based methods. That is to say, those techniques are trying to obtain independent features from each EEG channel. However, there is evidence suggesting brain disorders affect information exchange between multiple brain areas, namely brain connectivity (Cao ., 2021b; Hassan ., 2017; McBride ., 2015; Varotto ., 2014). To be more specific, brain connectivity is divided into three well-accepted categories: neuroanatomical brain connectivity, functional brain connectivity, and effective brain connectivity (Abbasvandi and Nasrabadi, 2019; Cao ., 2022). Neuroanatomical connectivity refers to structural links such as synapses or fibre pathways at the microscopic scale of neurons (Cao ., 2022). In terms of effective connectivity, it indicates the directed causal

influence of one neural region over others. On the other hand, functional brain connectivity is defined as statistical strength of covariance and/or correlation between pairs of brain regions, typically estimated with correlation, coherence, and information theory (Seth et al., 2015; Ursino et al., 2007). Many researchers demonstrated that brain connectivity is able to reflect complex cortical interconnections among brain networks and state of independent brain regions (Durongbhan et al., 2019; V. Sakalis, 2011; Tafreshi et al., 2019a; van Mierlo et al., 2014). Furthermore, it can reveal distinct aspects characterising various neurological conditions, such as dementia and other neurodegenerative diseases. There is evidence that both anatomical and functional connections among neural areas are affected in various forms of neurodegeneration (Cao et al., 2022; Pijnenburg et al., 2004). In this regard, measures of functional connectivity from scalp EEG recordings are of keen interest to elucidate the effect of neurodegeneration on intercommunications within widely distributed brain networks.

Analysing time-varying interactions and dynamic brain networks are increasingly attractive and challenging for researchers in the field of neurosciences (Li et al., 2019). Recently, many techniques to estimate dynamic functional connectivity have been developed to extract biomarkers from EEG signals, including short-time Fourier transform (STFT) (Chen et al., 2020; Keijzer et al., 2021; Wendling et al., 2009; Yildirim et al., 2019), wavelet analysis (Handojoseno et al., 2013; Ieracitano et al., 2017; Jeong et al., 2016; Sankari and Adeli, 2011), Error Reduction Ratio (ERR) (Zhao et al., 2020) and Hilbert–Huang transforms (HHT) (Shan et al., 2021). Although those techniques are satisfactory to an extent and can be extended to estimate time-frequency coherence between different EEG channels, there are some limitations because of their principles. STFT computes Fourier spectra on successive sliding windows and the STFT mainly suffers from the trade-off between temporal and spectral resolution (Fu et al., 2014; Moca et al., 2021). Generally, STFT is employed to analyse the linear and non-stationary signals (Mousavi et al., 2020). To avoid the window problem of STFT, Wavelet analysis employs a longer window for lower frequencies and a shorter one for higher frequencies. In

this case, it is more suitable for extracting time-varying information in different frequency bands (V. Sakkalis, 2011). However, wavelet-based methods suffer from the Heisenberg uncertainty principle, the wavelet transform cannot achieve fine resolutions in both time domain and frequency domain simultaneously, and the non-adaptivity once the basis wavelet is set (Furber et al., 2014). Wavelet transform is an advanced technique developed from Fourier analysis, using harmonic waves as its templates. Therefore, facing some similar problems of Fourier spectrum analysis, wavelet is capable to solve inter-wave frequency modulation and cannot solve intra-wave frequency modulation (Shan et al., 2021). However, neural oscillations have not been proved to be sinusoidal (Cohen, 2017b; Jones, 2016; Mazaheri and Jensen, 2008). Wavelet-based approaches have the weakness of dealing with non-sinusoidal oscillations and discriminating them from sinusoidal ones (Cohen, 2017a). Therefore, there is a need to develop novel methods to fully explore the hidden association in the typical nonlinear and non-stationary EEG recordings. Unlike wavelet transform with sinusoidal templates, Empirical Mode Decomposition (EMD) does not need any template assumption of the target signal (Huang et al., 1998), which may improve the ability to extract non-stationary and non-linear EEG information to an extent. EMD analyses the behaviour of non-stationary and nonlinear signals by decomposing them into several intrinsic mode functions (IMFs) that could be further analysed by Hilbert Huang transforms (HHT). Since the decomposition is based on the characteristics of the local time scale, with the HHT, the IMFs generate instantaneous frequencies as functions of time that separately estimates dynamic structures of different transient information. In this case, Shan et al. (2021) developed a new brain connectivity method relying on the revised HHT (RHHT) based on complete ensemble EMD with adaptive noise (CEEMDAN). It is able to capture dynamic interconnection between EEG signals and shows higher time-frequency resolution, compared with wavelet analysis. The "mode mixing" is a nonnegligible issue of the traditional EMD. To be more specific, one mode may represent different-amplitude oscillations or there are oscillations with high similarity found in different modes. To overcome this, Wu and Huang (2009) developed Ensemble EMD (EEMD), which performs

the EMD over an ensemble of the signal plus white Gaussian noise. However, there are still a variety of problems remaining and showing up. For instance, residual noise exists in IMF, and adding different white Gaussian noise to the signal increases the difficulty in controlling the number of IMFs. Torres [et al.](#) (2011) improved the algorithm by proposing the CEEMDAN. The main improvement occurs in adding noise. CEEMDAN tries to add a distinct noise at each step of the decomposition process, while EEMD adds the white Gaussian noise after the extraction of each IMF.

This project proposes a new time-frequency brain functional connectivity analysis framework, to track the non-stationary association of two EEG signals, based on a Revised Hilbert-Huang Transform (RHHT). One of the highlights of this framework is a novel method of estimating the cross-spectrum, based on the decomposed components. Compared with other time-frequency methods, such as CWT, the proposed method provides a more accurate dynamic time-frequency feature recognition and can be used to dissect the details of dynamic connectivity between brain regions. The effectiveness of this proposed method is validated through simulated and clinical EEG data.

Considering the aforementioned findings, it is hypothesised that functional brain network connectivity of patients with AD differs from networks of age-matched healthy controls (HC). The peak frequency extracted from the functional connectivity estimates is also hypothesised to act differentially for these two groups. The present Chapter proposes a framework to accomplish the discrimination between AD and HCs based on wavelet cross-spectrum (WC) and RHHT cross-spectrum, the latter of which is used for the first time in the field of AD. The second innovation of this Chapter is to evaluate a promising biomarker based on the peak frequency of the cross-spectrums (PFoCS). The proposed method is advanced in capturing dynamic interconnection between EEG signals and pointing out frequency-related biomarkers more precisely, which may help us to better understand this brain dysfunction. Then, a novel topographic visualisation method is designed to map the estimated brain

connectivity. Finally, the proposed RHHT technique is critically compared with other approaches in terms of principles, results and limitations.

### **3.2.2 Materials and Methodology**

#### **3.2.2.1 Experimental data and EEG pre-processing**

All patients and healthy controls included in this work were recruited in the Sheffield Teaching Hospital memory clinic and all provided written informed consent. The EEG study underwent ethics approval by the Yorkshire and The Humber (Leeds West) Research Ethics Committee (reference number 14/YH/1070). AD patients had their diagnosis confirmed between one month and up to 2 years prior to their EEG recording while they had mild to moderate cognitive deficits, according to their Mini-mental state examination. All patients and controls had brain magnetic resonance imaging (MRI) scans to eliminate other alternative causes of dementia. For the aged and gender-matched HC cohort, normal MRI brain scans and cognitive assessments were required before their EEG recordings. Further details about recruitment, diagnostic criteria and study design can be found in previously published work (Blackburn *et al.*, 2018).

All participants were younger than 70 years old, including 19 AD patients and 20 HC participants. EEG recordings were undertaken with an XLTEK 128-channel headbox (Optima Medical LTD) and Ag/AgCL electrodes at a sampling frequency of 2 kHz by implementing a modified 10-10 overlapping a 10-20 international system of electrode placement, with a referential montage (linked earlobe reference). Thirty-minute resting state (task-free – participants were instructed to rest and refrain from thinking anything specific) EEG recordings were obtained from each participant including sustained periods of keeping their eyes closed (EC) alternating with periods during which they kept their eyes open (EO). The recordings obtained were subsequently reviewed by a neurophysiologist - on an XLTEK review station - and for each participant, three 12-second artefact-free mini-epochs of EC and EO were selected. To avoid volume conduction effects related to the common reference electrode, 23 bipolar derivations were created from different brain regions: the left frontal (F7–



F3, F3–FZ), right frontal (F8–F4, F4–FZ), left frontocentral, temporocentral, temporal and centroparietal (F3–C3, T3–C3, T3–T5, C3–P3), right frontocentral,

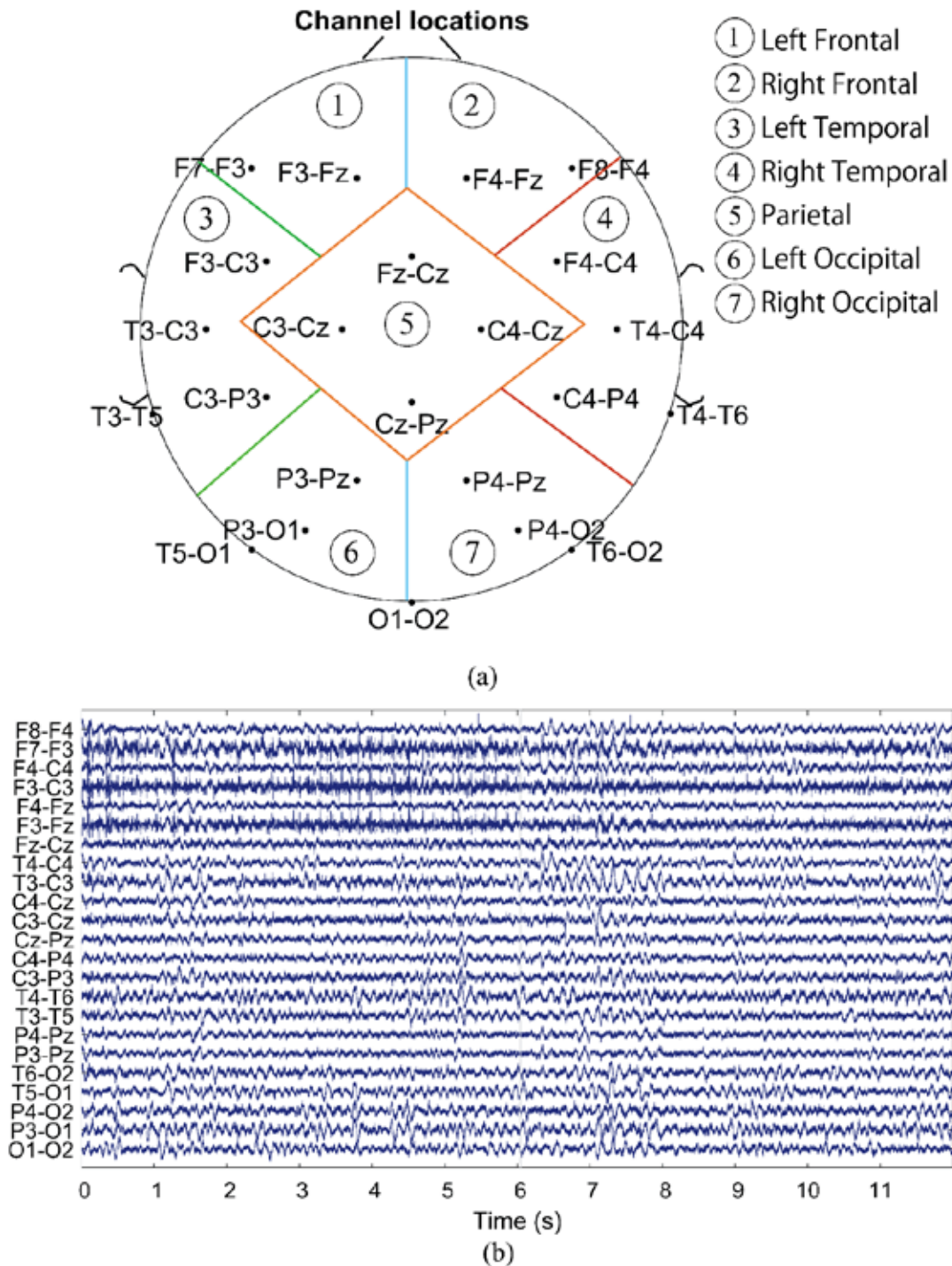


**Figure 3-16 (a) Bipolar channel locations.** The numbers 1, 2, 3, 4, 5, 6 and 7 represent the left frontal, right frontal, left temporal, right temporal, parietal, left occipital and right occipital regions, respectively **(b)** An example of 23-bipolar-channel EEG recordings.

temporocentral, temporal and centroparietal (F4-C4, T4-C4, T4-T6, C4-P4) midline frontocentral, centroparietal and left and right parasagittal central regions (Fz-Cz, Cz-Pz, C3-Cz, C4-Cz), left parietal, parieto-occipital and temporo-occipital areas (P3-Pz, P3-O1, T5-O1), right parietal, parieto-occipital and temporo-occipital areas (P4-Pz, P4-O2, T6-O2) and midline occipital region (O1-O2)).

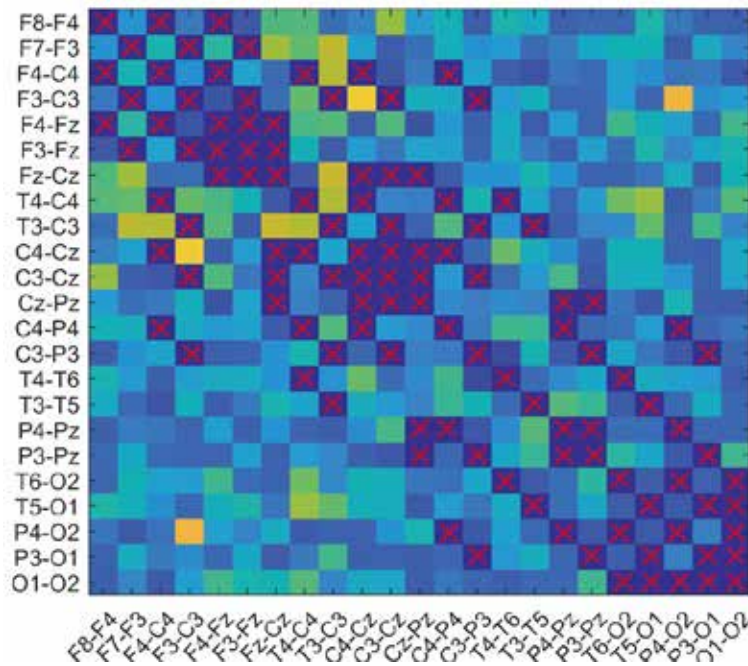


**Figure 3-17 207 possible channel pair combinations and 46 discarded pairs marked with X (symmetry)**

An example of the 23 bipolar channels used in this work and their locations are shown in Figure 3-16. The functional connectivity is calculated by using each pair of EEG signals, and Figure 3-17 shows the 253 possible channel pair combinations ( $C_2^{23}$ ). It cannot be ignored that some combinations have a common EEG channel because a bipolar approach is employed (such as F4-C4 and C4-P4), which may result in a misleading high false connectivity between some of the bipolar channels. To avoid this issue, the authors neglected any pair with common EEG electrodes. Therefore, the 46 channel pairs that have this characteristic include (F8-F4,F4-FZ), (F8-F4,F4-C4), (F7-F3,F3-FZ), (C4-

CZ,CZ-PZ), (T4-C4,C4-P4), (F3-C3,T3-C3), (C4-P4,P4-PZ), (FZ-CZ,C4-CZ), (T3-T5,T5-O1), (C3-CZ,CZ-PZ), (P4-O2,O1-O2), (T6-O2,P4-O2), (T6-O2,O1-O2), (F7-F3,F3-C3), (F4-FZ,FZ-CZ), (T3-C3,C3-CZ), (F3-FZ,FZ-CZ), (T4-C4,T4-T6), (FZ-CZ,CZ-PZ), (C4-CZ,C3-CZ), (P3-PZ,P3-O1), (F3-C3,C3-P3), (F4-C4,F4-FZ), (F4-C4,T4-C4), (F4-C4,C4-CZ), (C3-P3,P3-O1), (C4-P4,P4-O2), (F4-C4,C4-P4), (T5-O1,O1-O2), (CZ-PZ,P4-PZ), (F3-C3,F3-FZ), (P4-PZ,P3-PZ), (F3-C3,C3-CZ), (F4-FZ,F3-FZ), (FZ-CZ,C3-CZ), (T4-C4,C4-CZ), (T3-C3,C3-P3), (T3-C3,T3-T5), (C4-CZ,C4-P4), (C3-CZ,C3-P3), (CZ-PZ,P3-PZ), (C3-P3,P3-PZ), (T4-T6,T6-O2), (P4-PZ,P4-O2), (T5-O1,P3-O1), (P3-O1,O1-O2) (Cao et al., 2021a). In the following processes, only those pair without common channels are considered, while the values of the 46 neglected

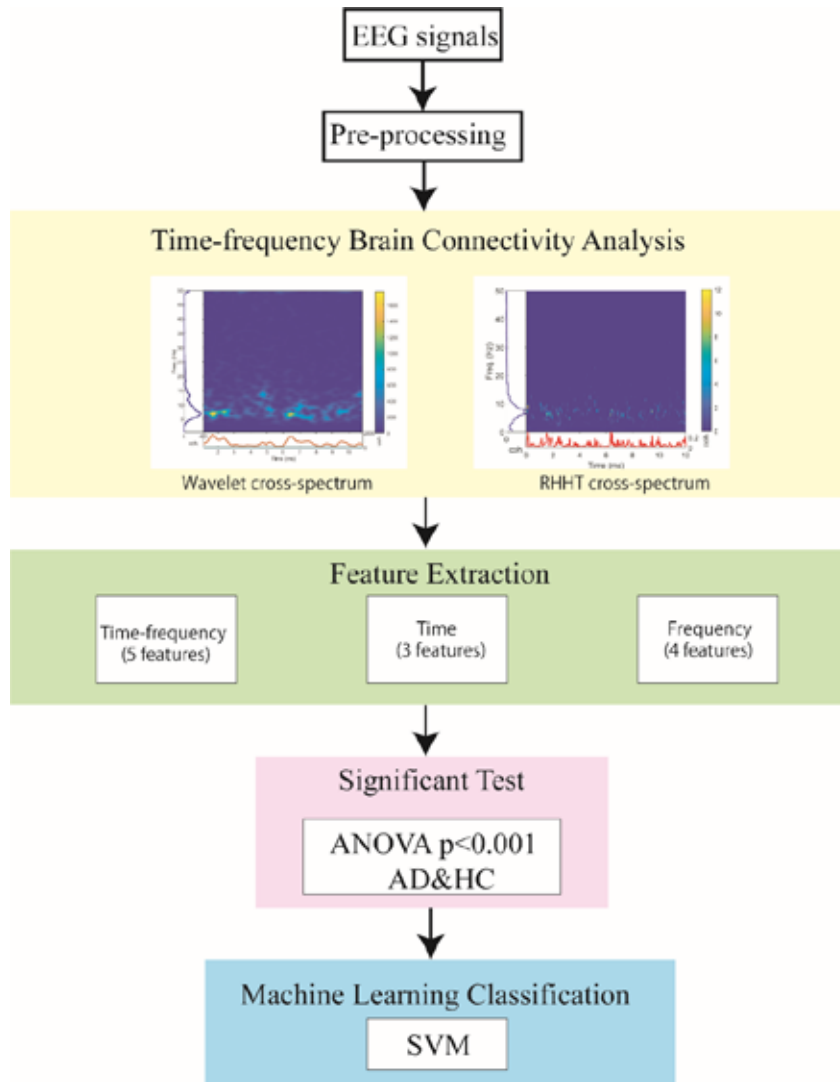


**Figure 3-18 Flow Diagram of the proposed framework**

channels are set as null (Figure 2). Hence, there are 207 channel pairs remaining for the following processing and discussion.

Figure 3-18 presents a flowchart of the proposed framework, including pre-processing, time-frequency brain connectivity analysis, feature extraction, significance test and machine learning classification. Before estimating brain connectivity, data was pre-processed by the following steps: (1) Each signal was filtered to 0-50 Hz; (2) the recordings were down sampled to 100 Hz to decrease computation cost.

### 3.2.2.2 Time-frequency Brain connectivity Methods

The Continuous Wavelet Transform (CWT) of a time series  $x$  is defined as:

$$CWT_x(a, b) = \int_{-\infty}^{+\infty} x(t) \Psi_{a,b}(t) dt \quad (3-30)$$

where  $\Psi$  is the mother wavelet,  $a$  is the scaling parameter and  $b$  is the shifting parameter. In this Chapter, the Morlet wavelet was chosen as the mother wavelet because it is reasonably localized in both time and frequency (Ieracitano et al., 2017). Each scale corresponds to a specific frequency value, so CWT is a function of time and frequency. The wavelet formulation of cross-spectrum between two signals,  $x$  and  $y$ , can be formulated as:

$$WC_{xy}(a, b) = S(W_x(a, b) W_y^*(a, b)) \quad (3-31)$$

where  $W_x(a, b)$ ,  $W_y(a, b)$  are the wavelet transforms of  $x$  and  $y$  at scales  $a$  and position  $b$ ;  $S$  denotes a function of smoothing, and  $*$  means complex conjugate.

Smoothing takes place across scale and time axes; it increases the degree of freedom for each point in the CWT (Sankari et al., 2012). A proper smoothing function for WC application across the time axis  $S_{time}$  is defined for the Morlet:

$$S_{time}(CWT(t, f)) = CWT(t, f) * \frac{-\lambda^2}{c_1^2} \quad (3-32)$$

where  $\lambda = t/a$ ,  $c_1$  is a normalization constant, and  $*$  refers to the convolution operator. The smoothing function across the scale  $S_{scale}$  (frequency) axis is defined as

$$S_{scale}(CWT(t, f)) = CWT(t, f) * c_2 \cdot (0.6a) \quad (3-33)$$

where  $c_2$  is a normalization constant, and  $\cdot$  is the rectangular function. In practice, the 2 convolutions in equations are computed discretely and the normalization coefficients are determined numerically. The width of the rectangular function used in  $S_{scale}$  is determined by the scale-decorrelation length that is empirically determined to be 0.6 for the Morlet wavelet (Sankari et al., 2012).

Firstly, an advanced method of empirical mode decomposition (EMD) is used to decompose each EEG signal, called complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN). In this case, we can obtain a series of complete and oscillatory components, named intrinsic mode functions (IMFs). Secondly, the Hilbert–Huang Transform (HHT) is performed on the IMFs of the signal to capture instantaneous frequency and amplitude features. Finally, the cross-spectrum of each pair of channels is calculated based on the HHT spectrum.

#### 3.2.2.2.1 CEEMDAN

Given a signal  $x(t)$ , the defined operator  $E_j(\cdot)$  produces the  $j$ -th mode of  $x(t)$  by EMD. Let  $n^i \in [0, 1]$ ,  $i = 1, \dots, I$  be white noise where  $I$  is the realisations times of adding noise. Coefficient  $\varepsilon_i$  allows selecting the signal-to-noise ratio (SNR) at each stage. The implementation of the CEEMDAN algorithm can be summarised as follows:

Decompose I realisations  $x(t) + \varepsilon_0 n^i(t)$  by EMD to obtain the first mode  $\widetilde{IMF}_1$  that is defined as:

$$\widetilde{IMF}_1 = \frac{1}{I} \sum_{i=1}^I IMF_1^i \quad (3-34)$$

At the first stage (j=1) calculate the first residue  $r_1(t)$ :

$$r_1(t) = x(t) - \widetilde{IMF}_1 \quad (3-35)$$

Decompose I realisations  $r_1(t) + \varepsilon_1 E_1(n^i)$  by EMD to obtain their first modes and the second CEEMDAN mode  $\widetilde{IMF}_2$  is defined as:

$$\widetilde{IMF}_2 = \frac{1}{I} \sum_{i=1}^I E_1 \left( r_1(t) + \varepsilon_1 E_1(n^i(t)) \right) \quad (3-36)$$

For  $j = 2, 3, \dots, I$ , calculate the j-th residue:

$$r_j(t) = r_{j-1}(t) - \widetilde{IMF}_j \quad (3-37)$$

Decompose realisations  $r_j(t) + \varepsilon_j E_j(n^i(t))$  by EMD to obtain their first modes and the (j+1)<sup>th</sup> CEEMDAN mode  $\widetilde{IMF}_{j+1}$  is defined as:

$$\widetilde{IMF}_{j+1} = \frac{1}{I} \sum_{i=1}^I E_j \left( r_j(t) + \varepsilon_j E_j(n^i(t)) \right) \quad (3-38)$$

Go to step 4 for the next  $j$  and repeat step 4 and 5 until the residue is no longer feasible to be decomposed (the residue does not have at least two extrema).

The final residual  $R(t)$  is written as

$$R(t) = x(t) - \sum_{j=1}^J \widetilde{IMF}_j \quad (3-39)$$

with  $J$  being the total number of modes. The given signal  $x(t)$  can be expressed as:

$$x(t) = \sum_{j=1}^J \widetilde{IMF}_j + R(t) \quad (3-40)$$

CEEMDAN needs to adjust its parameters to obtain a better decomposition of the data (Mousavi et al., 2020; Torres et al., 2011). Noise standard deviation (Nstd), the number of realisations (NR), and the maximum number of sifting iterations (MaxIter) are important parameters for optimising the results of decomposition (Mousavi et al., 2020). Figure 3-19 shows the flowchart of the

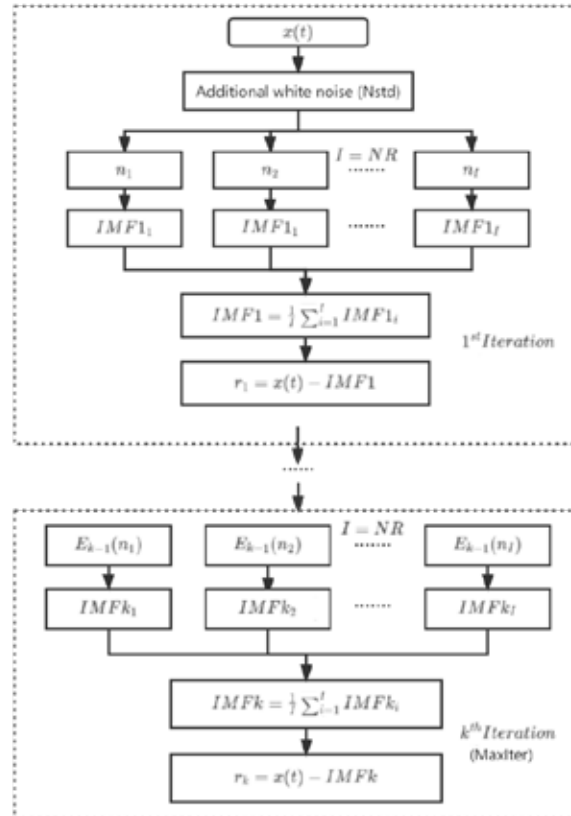


**Figure 3-19 flowchart of CEEMDAN algorithm**

CEEMDAN algorithm. Three parameters control the process of CEEMDAN. To be more specific, noise standard deviation (Nstd) represents the strength of the added white noise. The number of realisations (NR) controls the number of adding noise. The maximum number of sifting iterations indicates the number of decompositions. In the present Chapter, considering both computation efficiency and separation accuracy, these parameters were set as: Nstd = 0.1, NR = 100, and MaxIter = 1000.

### 3.2.2.2.2 Hilbert–Huang Transform

Then, the Hilbert Transform (HT) is applied to each IMF to obtain instantaneous frequencies (IF) and instantaneous amplitude (IA) features. This eventually yields a time-frequency representation (Hilbert spectrum) for each IMF. For an original signal  $x(t)$ , its HT  $\hat{x}(t)$  is obtained as (Huang et al., 1998):

$$\hat{x}(t) = HT(x(t)) = \frac{1}{\pi} P.V. \int_{-\infty}^{\infty} \frac{x(\tau)}{t - \tau} d\tau \quad (3-41)$$

where P.V. denotes Cauchy principal value,  $t$  is the time variable and  $\tau$  is the time interval. In this case, the HHT of the signal is given as follows:

$$z(t) = x(t) + i \hat{x}(t) = a(t)e^{i\theta(t)} \quad (3-42)$$

spectively.

$a(t)$  is the trace envelope and defined as:

$$a(t) = \sqrt{x^2(t) + \hat{x}^2(t)} \quad (3-43)$$

$\theta(t)$  is defined as:

$$\theta(t) = \arctan\left(\frac{\hat{x}(t)}{x(t)}\right) \quad (3-44)$$

Instantaneous frequency  $f(t)$  is defined as the first derivative of  $\theta(t)$ . Thus,

$$f(t) = \frac{1}{2\pi} \frac{d\theta(t)}{dt} \quad (3-45)$$

To prevent ambiguities because of phase unwrapping in Equation (14),  $f(t)$  can be calculated instead from:

$$f(t) = \frac{1}{2\pi} \frac{x(t) \hat{x}'(t) - \hat{x}(t) x'(t)}{x^2(t) + \hat{x}^2(t)} \quad (3-46)$$

where the prime denotes derivative with respect to time. The time-frequency spectrum of signal  $x(t)$  is defined as:

$$RHHT_x(t, f(t)) = a(t)e^{i\theta(t)} \quad (3-47)$$



The time-frequency spectrum of  $\widetilde{IMF}_j(t)$ ,  $t = 1, \dots, N$ , and  $j = 1, \dots, J$ , and  $RHHT(t)$  are defined as:

$$RHHT_{\widetilde{IMF}_j}(t, f_j(t)) = a_j(t)e^{i\theta_j(t)} \quad (3-48)$$

$$RHHT_R(t, f_R(t)) = a_R(t)e^{i\theta_R(t)} \quad (3-49)$$

The time-frequency spectrum of  $x(t)$  is:

$$RHHT_x(t, f(t)) = \sum_{k=1}^K a_j(t)e^{i\theta_k(t)} + a_R(t)e^{i\theta_R(t)} \quad (3-50)$$

A series of instantaneous frequencies are generated from the process, empowering the capability of in-depth signal analysis. At each time point, the number of instantaneous frequencies is corresponding to the number of IMFs. The RHHT method can generate a dozen of IMFs, which results in sparse time-frequency representations.

### 3.2.2.2.3 RHHT Cross-Spectrum

The RHHT cross-spectrum of  $x(t)$  and  $y(t)$  is defined as follows:

$$RC_{xy}(t, f) = \int_{t-\delta}^{t+\delta} RHHT_x(\tau, f) RHHT_y^*(\tau, f) d\tau \quad (3-51)$$

where,  $RHHT_{x_1}(\tau, f)$  and  $RHHT_y(\tau, f)$  are the RHHT coefficients of  $x(t)$  and  $y(t)$  and;  $*$  means complex conjugate;  $\delta$  is the length of the integrating range.

For RHHT cross-spectrum, the selection of  $\delta$  is entirely empirical, independent of the signal performance varied in terms of time-frequency resolution and computation cost. In this Chapter, the length of the integral window of RHHT coherence was set to 4. Given two signals  $x(t)$  and  $y(t)$ , for all IMFs of  $x(t)$  and  $y(t)$ , the RHHT cross-spectrums between  $\widetilde{IMF}_{xj}(t)$  and  $\widetilde{IMF}_{yj}(t)$  can be calculated and analysed separately. An example of RHHT and cross-spectrum is illustrated in Figures S1 and S2 in Supplementary.

As shown in Figure 3-18, both WC and RHHT methods yield a two-dimensional time-frequency cross-spectrum. To reveal the distribution of cross-spectrum in time (horizontal) and frequency (vertical) direction, the average was taken for two techniques, respectively (Eq. (3-52)(3-53)(3-54)(3-55)).

$$WC_{fre}(f) = \frac{1}{t_{max}} \sum_{t=0}^{t_{max}} WC(t, f) \quad (3-52)$$

$$WC_{time}(t) = \frac{1}{f_{max}} \sum_{f=0}^{f_{max}} WC(t, f) \quad (3-53)$$

$$RC_{fre}(f) = \frac{1}{t_{max}} \sum_{t=0}^{t_{max}} RC(t, f) \quad (3-54)$$

$$RC_{time}(t) = \frac{1}{f_{max}} \sum_{f=0}^{f_{max}} RC(t, f) \quad (3-55)$$

In this case, 12 features were captured from the cross-spectrum map and two average curves (Table 3-3).

**Table 3-3 Extracted features from the cross-curves.um and two average curves**

| $WC$ and $RC$               | Max | Peak Frequency | Mean     | Max/Mean | Std |
|-----------------------------|-----|----------------|----------|----------|-----|
| $WC_{fre}$ and $RC_{fre}$   | Max | Peak Frequency | Std      | Max/Mean |     |
| $WC_{time}$ and $RC_{time}$ | Max | Std            | Max/Mean |          |     |

### 3.2.2.3 Single-channel Methods

To compare the efficiency and ability of the developed RHHT technique and PFoCS features to a non-connectivity method, power spectral density (PSD)

was calculated and then the same SVM classifier was applied. The PSD tends to be used to extract features from each band of the EEG recordings. Specifically, PSD in each trial was calculated for each of the 23 channels using Welch's method with 0.5 s windows, 50% overlap and 256 points. Then, power was averaged for Delta (0–4 Hz), Theta (4–8 Hz), Alpha (8–12 Hz) and Beta (12–32 Hz) and Gamma (32–45 Hz) frequency bands.

#### **3.2.2.4 Statistical analysis**

One-way analysis of variance (ANOVA) was employed to evaluate the significance of differences in the investigated features of AD vs HC. In this case, it guides to select channels for each frequency band based on the p-value, preparing for optimising the input features of machine learning. ANOVA is widely used to test the significant difference between the AD and HC groups (Benwell et al., 2020; Engels et al., 2015; Sankari et al., 2011). ANOVA tests the null hypothesis, i.e. means of the tested groups are equal and the p-value indicates the statistical significance. Rejection of the null hypothesis leads to the conclusion that the two groups are statistically different (Cao et al., 2021a). To select promising features that offer significant differentiation between AD and HC and to reduce the computational burden of machine learning, the estimation was undertaken using a threshold  $p < 0.001$ .

#### **3.2.2.5 Machine Learning**

Features were computed between all possible pairs of channels for each of the six independent EEG frequency bands for each subject: Delta (0–4 Hz), Theta (4–8 Hz), Alpha (8–12 Hz), Beta (12–32 Hz), Gamma (32–45 Hz) and Full (0–45 Hz). There were in total  $207 * 6 * 12 = 14904$  (channel\*band\*features) features extracted from WC and RHHT cross-spectrum respectively. If a feature is successfully passed through the significance test ( $p < 0.001$ , ANOVA), it would become an input to the machine learning classifier, as it indicates AD and HCs have a significant difference in terms of this feature. Each feature was normalised to [-1, 1] for the AD and HC group in preparation for machine learning classification.

The Support Vector Machine (SVM) was applied to achieve the classification of AD and HC. SVM constructs an optimal separating hyper-plane in the feature space based on the structure risk minimization principle (Fu et al., 2014). The optimised features extracted from RHHT and WC in six bands are fed into the SVM with radial basis function (RBF) kernel. Different machine learning algorithms were tested and compared, such as K-Nearest Neighbour (KNN), decision tree and SVM, etc. SVM outperforms other methods in terms of classification. Hence, this Chapter mainly utilises the SVM with RBF to represent classification results.

The data from all subjects were mixed up and grouped by the eye state: EC, EO and EC&EO state. To avoid the problem of overfitting in the case of limited data as well as decrease the influence of selecting a training set, a five-fold cross-validation technique is employed. To be more specific, the dataset of each condition was divided into five subsets. Then five iterations are undertaken to ensure each subset is used for training and testing (Berrar, 2018; Cao et al., 2021a). That is to say, for each iteration, 80% of the dataset is used for training and the remaining 20% of the dataset is used for testing. Consequently, the classification result is calculated by averaging the accuracies obtained from 5 iterations.

### **3.2.3 Results**

#### **3.2.3.1 Machine learning classification**

The ten pairs with the highest classification accuracy were chosen to evaluate the efficiency of WC and RHHT cross-spectrum methods in EO, EC and EO&EC state separately (Figure 3-20). It can be observed that RHHT performs better in the Full band for all three eye states. It can achieve about 86% and 89% classification accuracy in EO and EC respectively, while features from the wavelet analysis can only provide classification accuracy of less than 85%. Besides, for the EC condition, the mean and median accuracies of RHHT lay closely to about 85%, which is relatively higher than the results obtained from

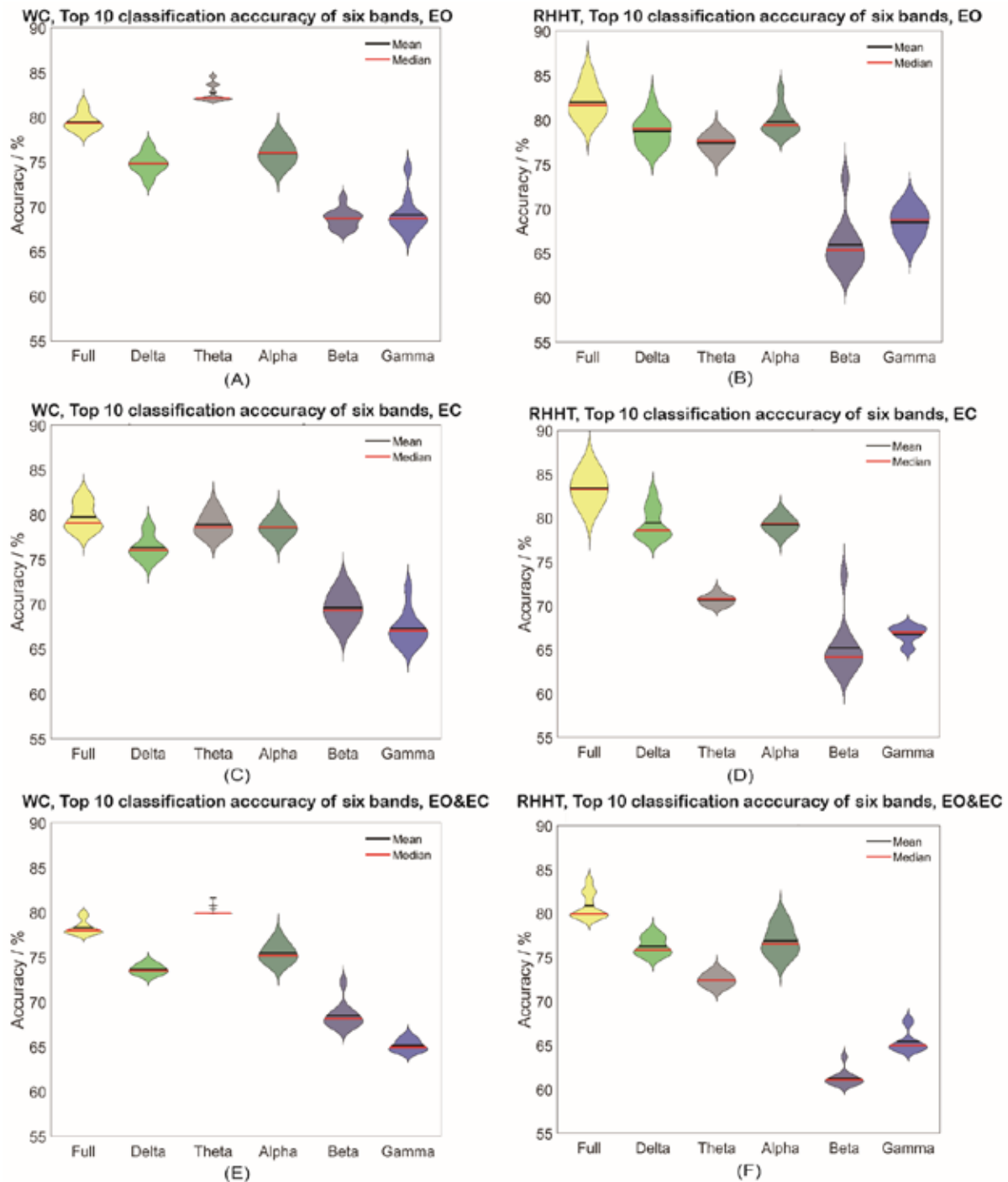


**Figure 3-20 Top ten classification accuracies in EO (A, B), EC (C, D) and EO&EC (E, F) using features from Wavelet (left column) and RHHT cross-spectrum (right column)**

the wavelet-based method (78%). Notably, the peak frequency of frequency-average curve in the Full band most accurately discriminated the AD from the HC group in EO, EC and EO&EC conditions, and RHHT enables the Full band to consistently outperform other bands with respect to both highest and mean

accuracy. However, for the WC method in EO and EO&EC states, the Theta band was superior to other bands, and the best feature was in the Theta band. In the meantime, WC only surpassed RHHT in the Theta band, while RHHT increased the accuracy in the Delta and Alpha band, especially for the EO Group with an approximate 5% increase.

Furthermore, the two techniques did not obtain satisfying outcomes in high-frequency bands (Beta and Gamma), which suggests the effect of AD on the high-frequency bands cannot be tracked with the current methodologies. Those results agree with several previous works that focus on the peak frequency. For instance, Scally et al., (2018) provided evidence that peak frequency brain connectivity in the range of 6-14 Hz is associated with aging. Barry and Blasio (2017) revealed a significant difference between young and older adults in the Delta, Theta, and Alpha bands. Therefore, cognitive performance and neurodegeneration are more related to the change in the low-frequency band rather than the Beta and Gamma band (Grandy et al., 2013; Nir et al., 2010; Posthuma et al., 2001; Smit et al., 2006).

It is noteworthy that the functional connectivity estimates, between the frontocentral midline and occipital derivations, demonstrate the highest accuracy in the Full band in all three eye conditions. For example, the RHHT PFoCS of the pair Fz-Cz:O1-O2 achieved the best classification performance with about 89% accuracy using the proposed machine learning method, which surpasses all features extracted from the wavelet-based method. Therefore, more details and analysis with respect to the Fz-Cz:O1-O2 connections are further discussed below.

### **3.2.3.2 Time-frequency functional connectivity**

To evaluate the efficiency of the wavelet and RHHT-based methods, the time-frequency analysis is implemented on each channel pair. Figure 3-21 represents the cross-spectrum of Fz-Cz:O1-O2 obtained from wavelet (A and B) and RHHT (C and D), as well as the average value in time and frequency axes. The process of RHHT is shown in Figure 3-25 and Figure 3-26 in Supplementary. It is obvious that the frequency resolution of RHHT is higher

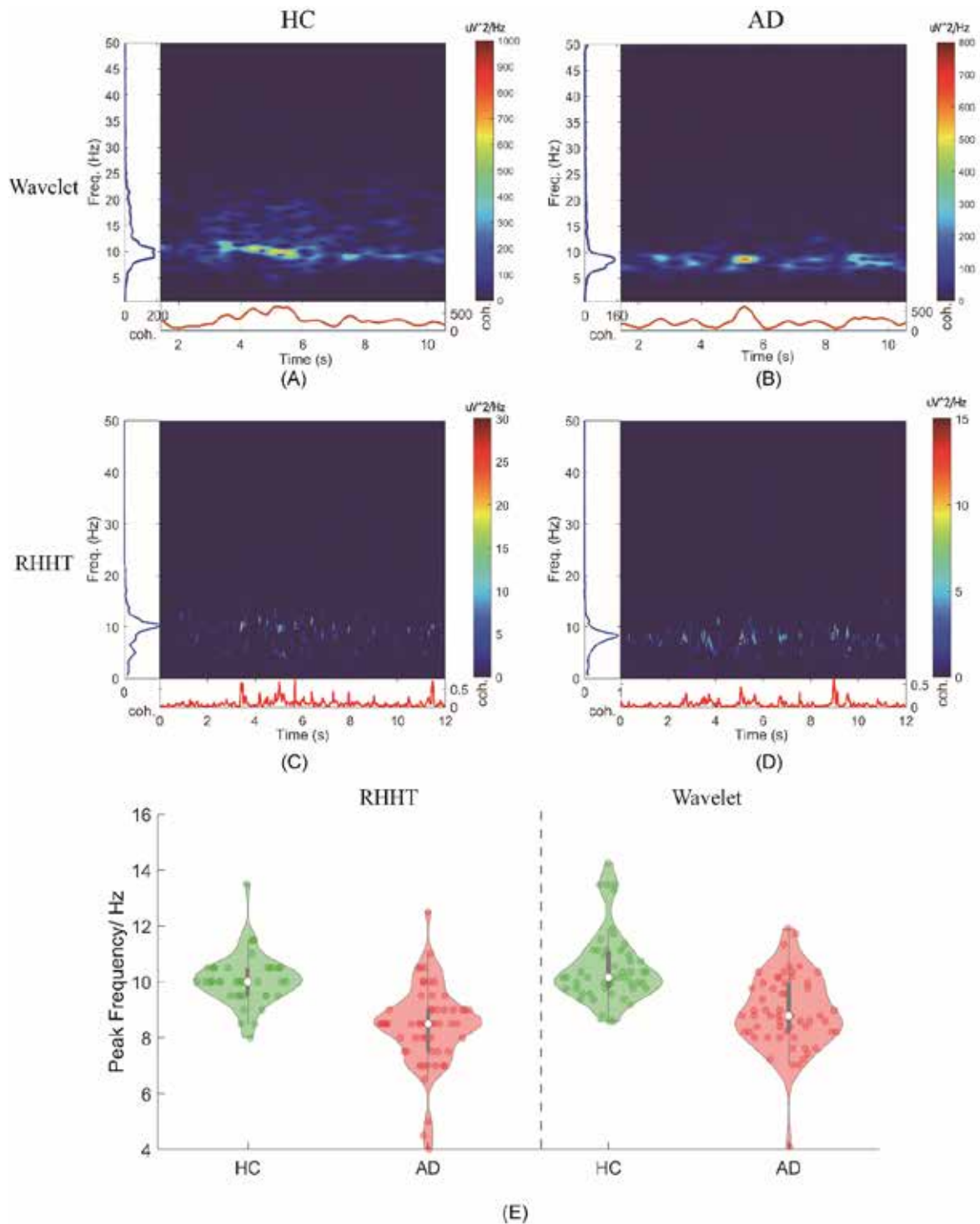


**Figure 3-21 Time-frequency brain connectivity analysis of channel pair Fz-Cz:O1-O2. (a) Wavelet cross-spectrum for a HC participant. (b) Wavelet cross-spectrum for an AD participant. (c) RHHT cross-spectrum for a HC participant. (d) RHHT cross-spectrum for an AD participant. (e) A violin plot to present peak frequency values for channel pair Fz-Cz:O1-O2 ( $p < 0.0001$ , ANOVA)**

than the WC for both AD and HC participants. For instance, the WC method

suggests that the PFoCS is located in the Alpha band (Figure 3-21 (a)), while RHHT is capable to indicate a specific frequency of 10 Hz (Figure 3-21 (c)). Similarly, Figure 3-21 (b) and (d) also exhibit the same pattern in the AD case. Furthermore, it can be observed that two methods have dominating power in the Alpha band. Based on the resulting cross-spectrum of the channel pair Fz-Cz:O1-O2, the PFoCS of the AD subject is lower than the value of the HC subject, observed both from the wavelet and the RHHT cross-spectrum.

Figure 3-21(e) plots the distribution of the PFoCS value of each sample for the Fz-Cz:O1-O2 channel pair. For both RHHT and WC, there was a significant difference in the PFoCS ( $p < 0.0001$ , ANOVA). For both methods, the PFoCS for HCs ranges between 8-14 Hz with an average of 10 Hz, while the AD PFoCS varies from 4 to 12 Hz with an average of 8.5 Hz. Notably, compared with HCs, a majority of AD participants have an obvious decrease in terms of PFoCS in the Alpha band, and RHHT seems to outperform WT in detecting the difference to an extent. Consequently, the classification accuracy can be improved by using the RHHT technique. Moreover, the value of PFoCS varies in different bands, although it was mainly concentrated in the Alpha band. This may explain that the feature extracted from the Full band outperformed the Alpha band to some extent.

### **3.2.3.3 Topographic visualisation of PFoCS**

To further compare the difference between AD and HCs across different brain areas, a novel visualisation method based on a topographic map is proposed to illustrate an extensive brain connectivity map. Figure 3-22 represents the occipital-related connectivity using the PFoCS as an estimation. The topography map represents the distribution of Full-band functional connectivity between channel O1-O2, and every other derivation included in this work. The channel O1-O2 and other contiguous channels that share a common electrode are set to NaN in the resulting map (i.e., P3-O1, T6-O2, T5-O1, and P4-O2). On one hand, from the averaged topographic map, it can be observed that not only Fz-Cz:O1-O2 can significantly differentiate AD from HCs, but also many other areas show increased levels of PFoCS functional connectivity, offering a visual



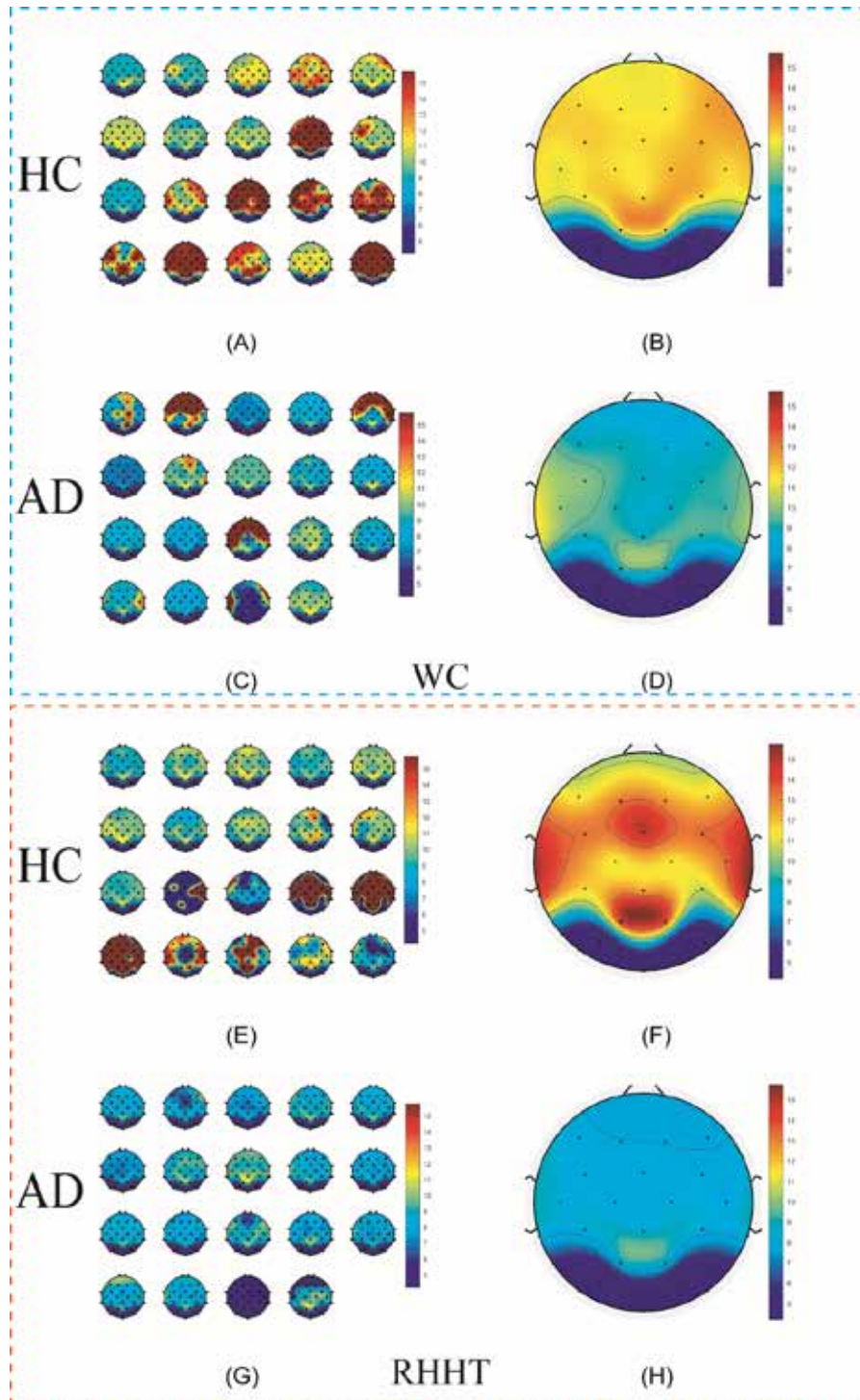


**Figure 3-22 Topographic plots representing Full-band connectivity PFOCS for each participant (Left) and average (Right) extracted by Wavelet and RHHT techniques , by using each available derivation against the occipital O1-O2 channel.**

representation of the striking differences between AD and HC, clearly more prominent for the averaged RHHT-based maps. The latter is in keeping with the

previous significance test results and SVM-based classification.

On the other hand, the variations of estimated distribution for the participants are quite large. This is a general issue in EEG research since EEG has been associated with cognitive, emotional, and motor processes (Bell and Cuevas, 2012). In this respect, participants from the AD and HC groups may have similar characteristics. As a result, the machine learning classifier cannot achieve 100% discrimination between the two cohorts. However, a similar trend can be observed as AD patients have relatively lower PFoCS across widely distributed brain areas. To be more specific, RHHT performed better in revealing the difference between AD and HCs in terms of the PFoCS, resulting from its superior time-frequency resolution and ability to estimate a frequency value.

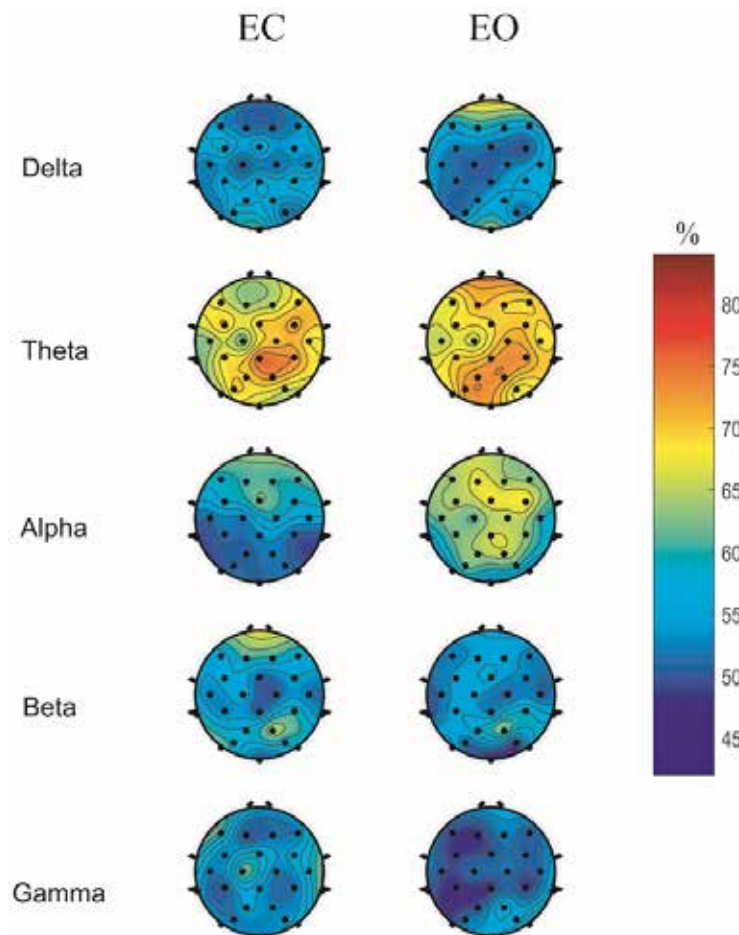


**Figure 3-23 Topographic map of the classification accuracy using PSD**

### **3.2.3.4 PSD results**

The spatial distribution of the classification accuracy is displayed in Figure 3-23. Obviously, the most significant difference occurs in the Theta band. Noticeably, the highest accuracy is approximately 72%. The results obtained by PSD are about 15% lower than that of the RHHT method. That is to say, AD, the most common neurodegenerative disorder, affects not only the function of individual cortical regions but also the intercommunications between them. Furthermore, the abnormalities of AD brain connectivity are more prominent for the present dataset, which is satisfactorily evaluated by the proposed RHHT method.

## **3.2.4 Discussion**

### **3.2.4.1 RHHT vs. WC**

EEG signals are often non-stationary in nature, contain non-linear processes, and their frequency components dynamically change over time (Lotte et al., 2018; Science and Engineering, 2014; Yao and Wang, 2017). Hence, the desired time-frequency analysis methods for quantitative EEG studies should have satisfying resolutions in both time and frequency domain. RHHT can clearly present the cross-spectrum distribution with time and frequency and localise any event on its occurring time as well as its instantaneous frequency (Law et al., 2012; Peng et al., 2005). Specifically, the PFoCS variation of RHHT shows a much higher resolution in comparison to those achieved by the Morlet wavelet spectrum. In this case, even though WC can achieve about 85% classification accuracy in EO condition using the mean of cross-spectrum as a feature, it cannot perform better when it comes to the PFoCS. Figure 3-27 in Supplementary represents a brain connectivity map using the mean cross-spectrum of WC as an estimation. Consequently, RHHT is more suited for EEG brain connectivity analysis in this paradigm and performs better in localising in time PFoCS connectivity strength that translates in better performance in distinguishing between AD and HCs.

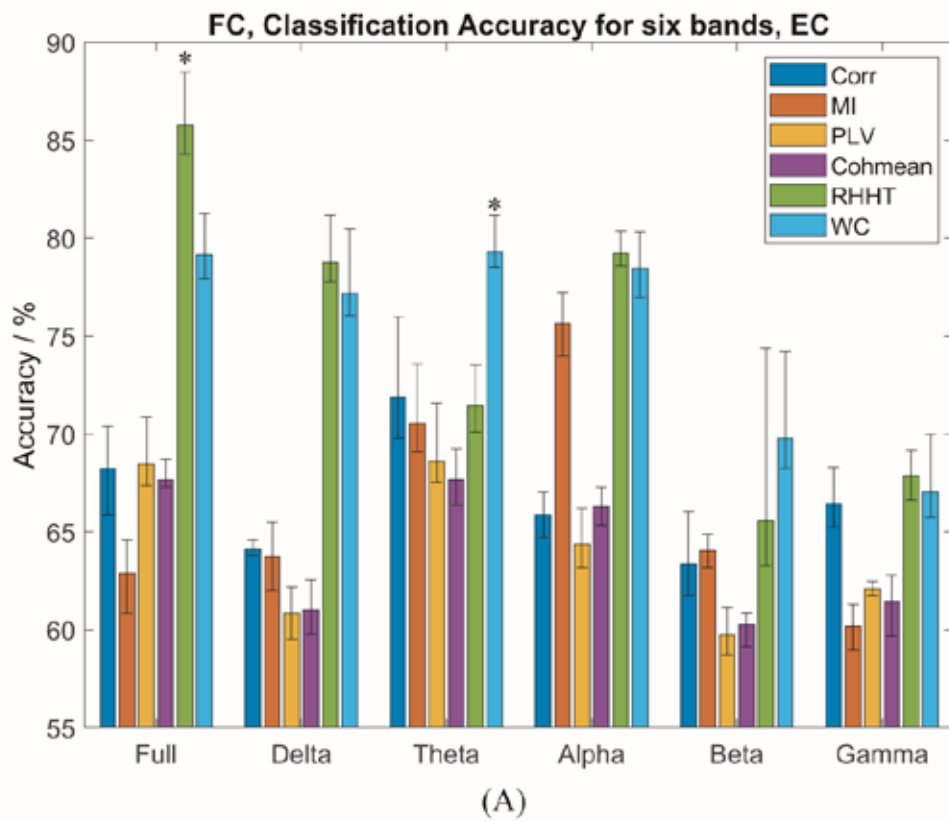
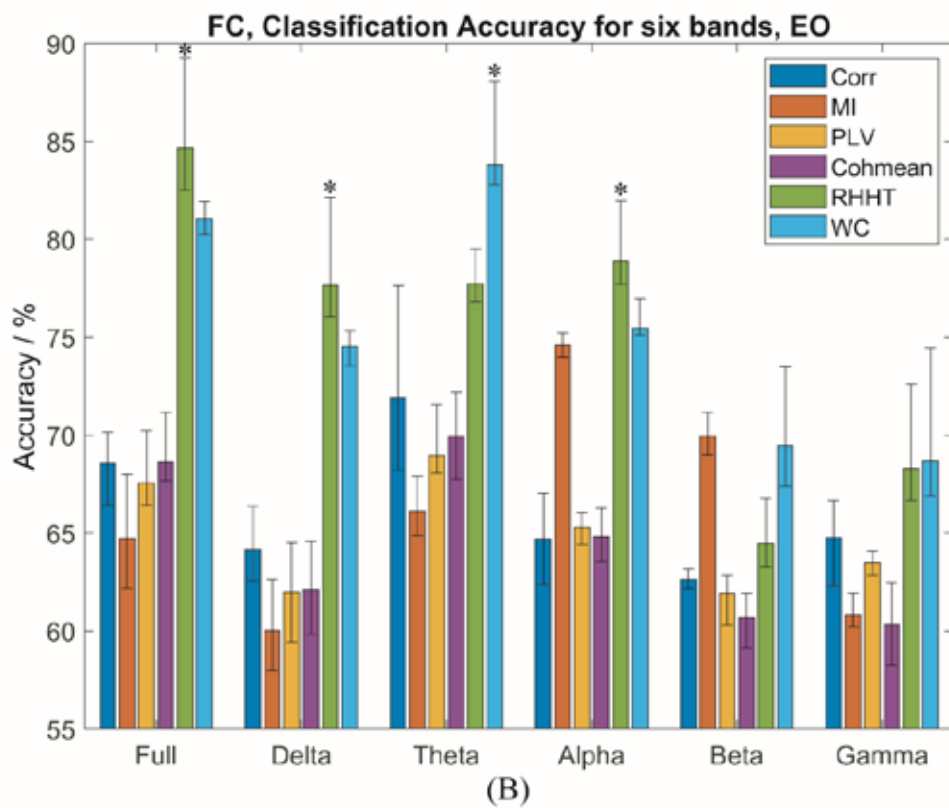


**Figure 3-24 Comparison of the RHHT and previous brain connectivity methods. \* Mann-Whitney Test,  $p < 0.001$ , it means significant outperformance**

### 3.2.4.2 RHHT vs. other functional connectivity methods

In addition to the aforementioned advantages, the RHHT method can track non-linear associations between channels. To demonstrate the superior performance of cross-spectrum RHHT, in distinguishing between AD and HCs, we compared our findings against those obtained by implementing other popular linear and non-linear brain connectivity methods. Phase Locking Value (PLV) (Hassan et al., 2017) and Mutual Information (MI) (Melia et al., 2015; Nimmy et al., 2019) were used to estimate non-linear connections, while Pearson Correlation Coefficient (PCC) (Chen et al., 2018) and Magnitude Squared Coherence (MSC) (Cao et al., 2021a) were calculated with respect to linear brain connectivity. Mutual Information (MI) measure indicates the mutual dependence of two signals, i.e. how much information is shared between two signals (Cover and Thomas, 2005). It is based on a probability function and entropy. Three features are taken from the cross-correlation, namely the maximum value of correlation (CorrMax), the mean value of correlation (CorrMean) and the correlation lag at the maximum value of correlation (CorrLag). Two features are taken from the MSC: maximum (CohMax) and mean values (CohMean). PLV measures the significance of the phase covariance between two signals. PLV values are in the range from 0 to 1. If the phase difference of two signals remains the same, PLV is close to one. A PLV close to zero indicates that there is no phase synchrony between two signals (Cao et al., 2021a). Figure 3-24 illustrates the comparison between the proposed methods and conventional techniques in terms of the top ten channel classification accuracies. The average, highest and lowest accuracy are represented by the bar chart. To statistically compare different techniques, the Mann-Whitney U tends to be more appropriate as the top ten accuracies do not meet the parametric assumptions of a normal distribution (Shan et al., 2021). The Mann-Whitney test was applied to check for significant differences between the accuracies of techniques, using a threshold of  $p < 0.001$ . Overall, the EO condition appears more suitable to distinguish AD patients from HCs, since the classification accuracy is slightly higher in comparison to the EC recordings. For the EC condition (Figure 3-24 (A)), as expected, RHHT and WC were superior

in comparison to others, especially in the Full band, as well as Delta and Alpha band. RHHT performs significantly better than others in the Full band (the accuracy ranging between 84%-88%), while WC provides significantly higher accuracy in the Theta band (78%-82%). For the EC condition (Figure 3-24 (B)), RHHT significantly outperforms other methods in the Full (83%-89%), Delta (76%-82%) and Alpha band (77%-82%). WC keeps its significant superiority in the Theta band, which ranges from 84% to 87%. Therefore, our findings suggest that the non-linear time-frequency method (RHHT) has superior performance in the Full, Delta and Alpha band that contributes in better differentiating AD participants from HCs, as there is a 10%-20% improvement in classification accuracy. The advantages of the proposed RHHT method rely on its distinct signal decomposition and higher resolution in the time and frequency domain.

#### **3.2.4.3 RHHT vs. single-channel method**

This Chapter demonstrates that RHHT cross-spectrum, a functional brain connectivity method implemented in this work to produce strength of associations between multiple EEG channels, could produce a higher classification accuracy, on distinguishing AD participants from a group of age-matched HCs, compared to single-channel methods like PSD. PSD is a widely used single-channel method dealing with EEG recordings, which calculates the distribution of power against the frequencies. Furthermore, the spectral information in different bands can be estimated by PSD. It normally considers Delta, Theta, Alpha, Beta and Gamma bands for EEG, since those bands represent distinct neural activities (Liu et al., 2016; Wang et al., 2017, 2015a, 2015b). Many previous works have shown that PSD can be used to evaluate the changes between HC and AD. For example, it was found that AD groups have relatively higher PSD than HC groups in the Theta band (4-8 Hz), while AD patients may experience the decrease of Alpha power (8-12 Hz) and Beta power (12-32 Hz) (Benwell et al., 2020; Liu et al., 2016).

**Table 3-4 A comparison of performances of the various functional connectivity methods and univariate methods (maximum accuracy achieved by each method)**

| Features            | Reference                                                    | SVM Accuracy |
|---------------------|--------------------------------------------------------------|--------------|
| PSD                 | (Wang et al., 2017)(Benwell et al., 2020)                    | 72.2%        |
| Correlation         | (Chen et al., 2018)                                          | 77.7%        |
| Coherence           | (Sankari et al., 2011; Wang et al., 2015b)                   | 72.11%       |
| Mutual information  | (Babiloni et al., 2016)(Nimmy et al., 2019)                  | 77.23%       |
| Phase locking value | (Engels et al., 2015)(Kabbara et al., 2017)(Su et al., 2021) | 72.11%       |
| WC                  | (Sankari and Adeli, 2011)                                    | 85.2%        |
| RHHT-- PFOCS        | This Chapter                                                 | 88.9%        |

Table 3-4 compares the highest classification results of the popular features extracted from EEG in the literature and the approach proposes in this Chapter, using the same dataset of AD and HCs. Evidently, the developed method obtains the highest accuracy, which is clearly higher than other existing approaches. Notably, the time-frequency methods (i.e., WC and RHHT) perform better than other methods that focus on stationary information with the superiority of approximately 10%. That is to say, dynamic and non-stationary methods may have more potential to reveal hidden changes during the progress of neurodegeneration.

#### 3.2.4.4 Limitations and future direction

It should be noted that the proposed framework has several limitations to be addressed in future research. Firstly, the computation efficiency of the RHHT technique is not satisfying. It could be improved to meet the requirement of real-time feedback, aiming to translate it into a clinically useful diagnostic tool.

Secondly, the validation of the proposed solution is not cross-subject. Cross-subject validation means that the data for training and testing are extracted from different participants. Only this type of approach can test the universality of the proposed method, but it usually requires a significant number of participants to reveal the degree of variability between subjects. In the present study, the size of the dataset is limited. The cross-subject approach is not appropriate to evaluate the model performance and therefore has not been applied. In this case, the data from all subjects were mixed up and then divided into training and testing. Thirdly, to reduce volume conduction effects from a common reference, bipolar derivations were used to assess the degree of differences between various pairs of electrodes for 2 different cohorts of subjects. With this approach – the use of bipolar pairs of electrodes - the effects of volume conduction are reduced but not eliminated. We recognise that this work is based on a sensor level scalp EEG analysis, and we do not claim to be able to precisely localise the spatial characteristics underpinning the EEG sensor findings. Fourthly, the noise tolerating capability of the RHHT technique may need to be further studied. According to our previous study, the performance of RHHT seems to be more sensitive to noise (Shan et al., 2021). Finally, only a two-group classification (HCs vs. AD) has been carried out in the present Chapter, without considering preclinical or prodromal populations or staging of disease severity of AD patients. The proposed method performs effectively in the case of HCs vs. AD, and it may have potential in dealing with the classification of HCs vs MCI vs AD. We are collecting data from MCI and this three-class classification will be tested in further studies.

### **3.2.5 Conclusion**

In this Chapter, we demonstrate that the peak frequency estimated with the RHHT-based brain functional connectivity is reduced in a group of AD patients in comparison to an age-matched HC cohort. A comparison undertaken with various other linear and nonlinear functional connectivity EEG methods shows that RHHT is better suited in detecting this difference, because of its superior resolution and ability to track non-stationary EEG dynamics and possibly non-



linear interconnections. Notably, this Chapter introduces an approach to reveal the areas of interest at an EEG sensor level of AD, which helps to better understand each brain area's contribution to the classification, but this can also be applied in numerous other brain-related disorders.

### 3.2.6 Supplementary

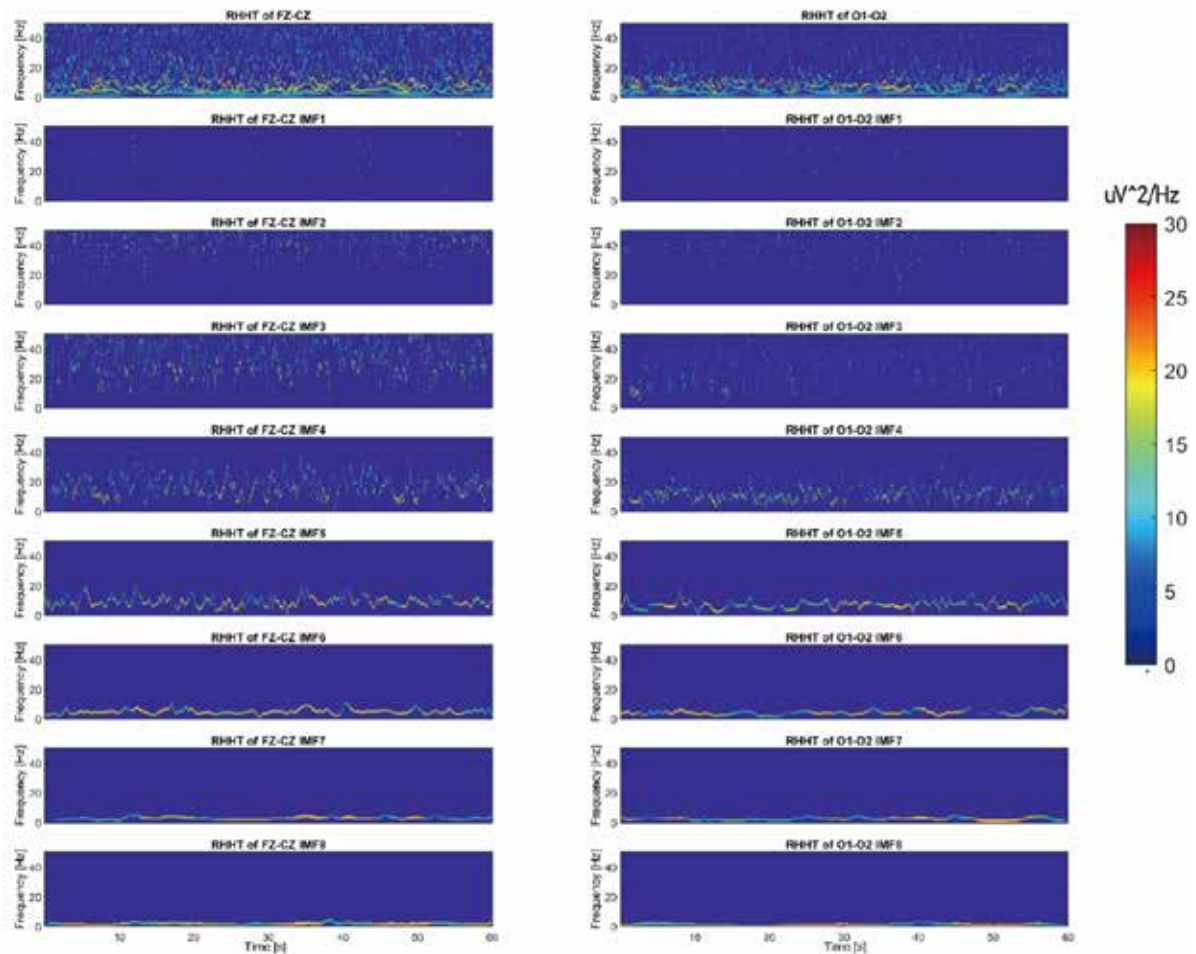


**Figure 3-25** An example of the resulting Revised Hilbert-Huang Transformation (RHHT) based on complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) from the EEG recordings of channel Fz-Cz:O1-O2

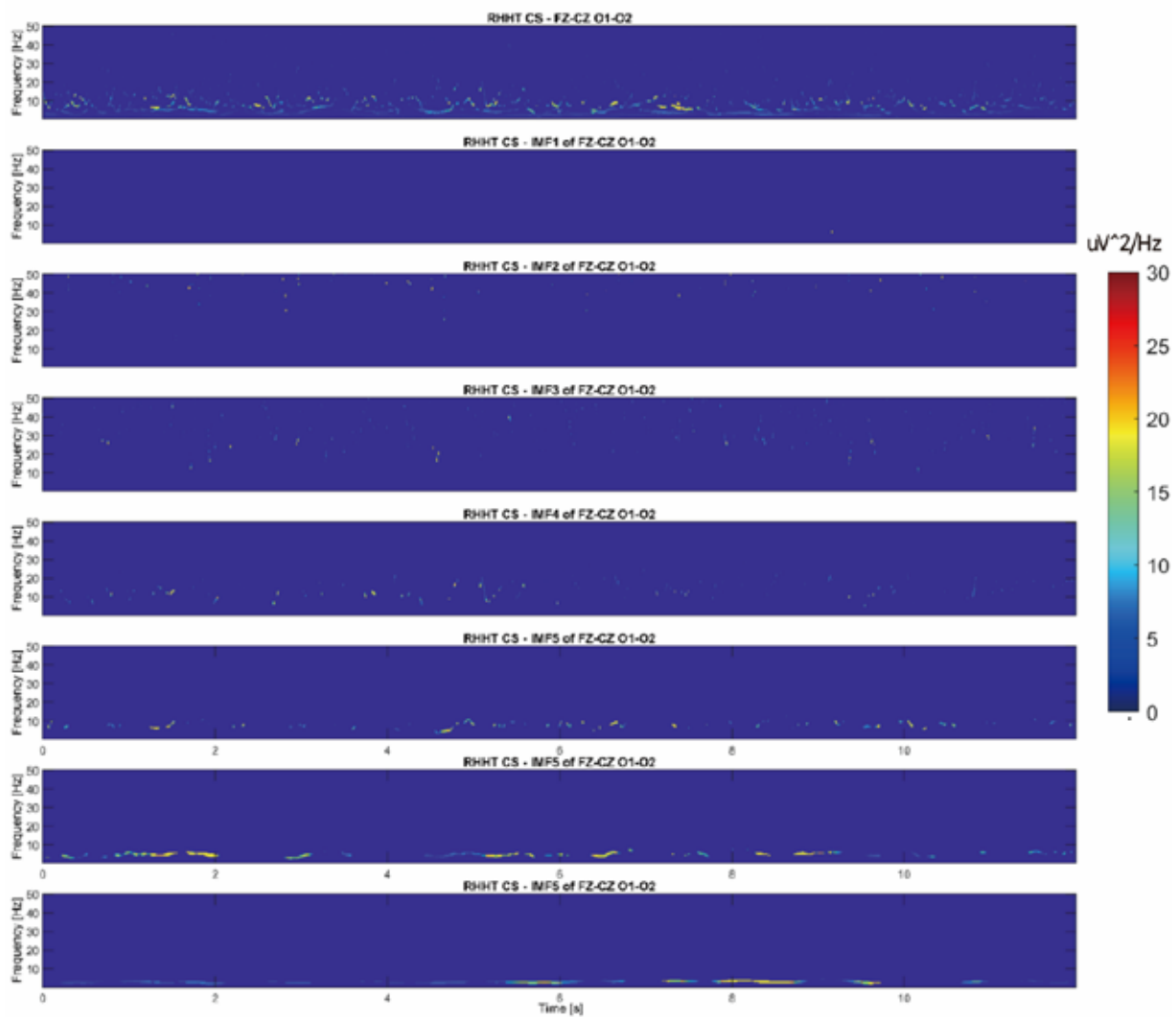


**Figure 3-26 An example of the resulting Revised Hilbert-Huang Transformation (RHHT) cross-spectrum the EEG recordings of channel Fz-Cz:O1-O2**

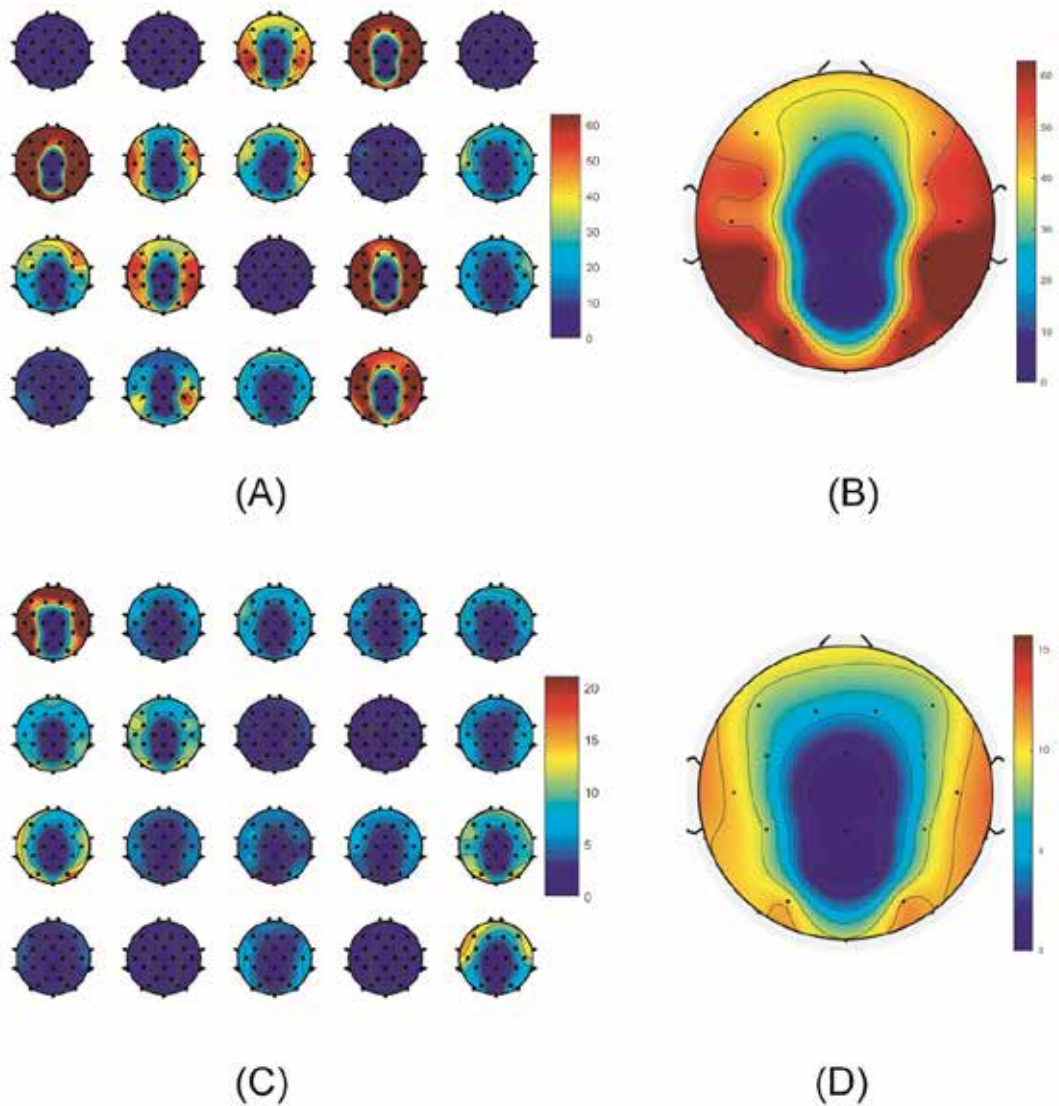


**Figure 3-27** Topographic plots representing mean coherence connectivity for each participant (Left) and average (Right) extracted by Wavelet techniques, by using each available derivation against the parietal Cz-Pz channel.

### 3.2.7 References

Abbasvandi, Z., Nasrabadi, A.M., 2019.

Comput. Biol. Med. 110, 93–107.

<https://doi.org/10.1016/j.combiomed.2019.05.012>

Angelakis, E., Stathopoulou, S., Frymiare, J.L., Green, D.L., Lubar, J.F., Kounios, J., 2007.

. Clin. Neuropsychol. 21, 110–129. <https://doi.org/10.1080/13854040600744839>

Babiloni, C., Del Percio, C., Bordet, R., Bourriez, J.L., Bentivoglio, M., Payoux, P., Derambure, P., Dix, S., Infarinato, F., Lizio, R., Triggiani, A.I., Richardson, J.C., Rossini, P.M., 2013.

. Clin. Neurophysiol. 124, 837–850. <https://doi.org/10.1016/j.clinph.2012.09.017>

Babiloni, C., Lizio, R., Marzano, N., Capotosto, P., Soricelli, A., Triggiani, A.I., Cordone, S., Gesualdo, L., Del Percio, C., 2016.

. Int. J. Psychophysiol. 103, 88–102. <https://doi.org/10.1016/j.ijpsycho.2015.02.008>

Barry, R.J., Blasio, F.M. De, 2017.

. Fac. Soc. Sci. - Pap.

Bell, M.A., Cuevas, K., 2012.

. J. Cogn. Dev. 13, 281–294. <https://doi.org/10.1080/15248372.2012.691143>

Benwell, C.S.Y., Davila-Pérez, P., Fried, P.J., Jones, R.N., Trivison, T.G., Santarnecchi, E., Pascual-Leone, A., Shafi, M.M., 2020.

. Neurobiol. Aging 85, 83–95. <https://doi.org/10.1016/j.neurobiolaging.2019.10.004>

Berrar, D., 2018. . Encycl. Bioinforma. Comput. Biol. ABC Bioinforma. 1–3, 542–545. <https://doi.org/10.1016/B978-0-12-809633-8.20349-X>

Blackburn, D.J., Sarrigiannis, P.G., De Marco, M., Zhao, Y., Venneri, A., Lawrence, S., Unwin, Z.C., Blyth, M., Angel, J., Baster, K., Wilkinson, I.D., Bell, S.M., He, F., Wei, H.L., Billings, S.A., Farrow, T.F.D., 2018.

. Brain  
Sci. 8, 1–19. <https://doi.org/10.3390/brainsci8070134>

Blinowska, K.J., Rakowski, F., Kaminski, M., De Vico Fallani, F., Del Percio, C., Lizio, R., Babiloni, C., 2017.

. Clin. Neurophysiol. 128, 667–680.  
<https://doi.org/10.1016/j.clinph.2016.10.002>

Cao, J., Grajcar, K., Shan, X., Zhao, Y., Zou, J., Chen, L., Li, Z., Grunewald, R., Zis, P., Marco, M. De, Unwin, Z., Blackburn, D., Sarrigiannis, P.G., 2021a.

. Biomed. Signal Process. Control 67, 102554. <https://doi.org/10.1016/j.bspc.2021.102554>

Cao, J., Zhao, Y., Shan, X., Wei, H., Guo, Y., Chen, L., Erkoyuncu, J.A., Sarrigiannis, P.G., 2021b.

. Hum. Brain Mapp. n/a, 1–20.  
<https://doi.org/https://doi.org/10.1002/hbm.25683>

Chen, G., Chen, M., Hong, G., Lu, Y., Zhou, B., Gao, Y., 2020.

. Energies 13.  
<https://doi.org/10.3390/en13040888>

Chen, Y., Cai, L., Wang, R., Song, Z., Deng, B., Wang, J., Yu, H., 2018.

. Phys. A Stat. Mech. its Appl. 490, 171–184. <https://doi.org/10.1016/j.physa.2017.08.009>

Cohen, M.X., 2017a.

. Elife 6, e21792.

Cohen, M.X., 2017b.

Trends Neurosci. 40, 208–218.

Cover, T.M., Thomas, J.A., 2005. Elements of Information Theory, Elements of Information Theory. New York: JOHN WILEY and SONS, INC., 1991. ISBN: 0-471-20061-1. <https://doi.org/10.1002/047174882X>

delEtoile, J., Adeli, H., 2017.

. Neuroscientist 23, 616–626.  
<https://doi.org/10.1177/1073858417702621>

Durongbhan, Pholpat, Zhao, Yifan, Chen, L., Zis, Panagiotis, De Marco, Matteo, Unwin, Zoe C, Venneri, Annalena, He, X., Li, S., Zhao, Yitian, Blackburn, Daniel J, Sarrigiannis, Ptolemaios G, Durongbhan, P, Zhao, Y, Zis, P, De Marco, M, Unwin, Z C, Venneri, A, Blackburn, D J, Sarrigiannis, P G, 2019.

. IEEE Trans. NEURAL Syst. Rehabil. Eng. 27.  
<https://doi.org/10.1109/TNSRE.2019.2909100>

Engels, M.M.A., Stam, C.J., Flier, W.M. Van Der, Scheltens, P., Waal, H. De, Straaten, E.C.W. Van, 2015.

. BMC Neurol. 1–8.  
<https://doi.org/10.1186/s12883-015-0400-7>

Ferreri, F., Vecchio, F., Vollero, L., Guerra, A., Petrichella, S., Ponzio, D., Määttä, S., Mervaala, E., Könönen, M., Ursini, F., Pasqualetti, P., Iannello, G., Rossini, P.M., Di Lazzaro, V., 2016.

. Hum. Brain Mapp. 37, 2083–2096. <https://doi.org/10.1002/hbm.23158>

Fu, K., Qu, J., Chai, Y., Dong, Y., 2014.

. Biomed. Signal Process. Control 13, 15–22. <https://doi.org/10.1016/j.bspc.2014.03.007>

Grandy, T.H., Werkle-Bergner, M., Chicherio, C., Schmiedek, F., Lövdén, M., Lindenberger, U., 2013.

Psychophysiology 50, 570–582. <https://doi.org/10.1111/psyp.12043>

Handojoseno, A.M.A., Shine, J.M., Nguyen, T.N., Tran, Y., Lewis, S.J.G., Nguyen, H.T., 2013.

. Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS 4263–4266. <https://doi.org/10.1109/EMBC.2013.6610487>

Hassan, M., Chaton, L., Benquet, P., Delval, A., Leroy, C., Plomhause, L., Moonen, A.J.H., Duits, A.A., Leentjens, A.F.G., van Kranen-Mastenbroek, V., Defebvre, L., Derambure, P., Wendling, F., Dujardin, K., 2017.

. Neurolmage Clin. 14, 591–601. <https://doi.org/10.1016/j.nicl.2017.03.002>

Huang, Norden E, Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.-C., Tung, C.C., Liu, H.H., 1998.

. Proc. R. Soc. London. Ser. A Math. Phys. Eng. Sci. 454, 903–995.

Huang, Norden E., Shen, Z., Long, S.R., Wu, M.C., Snin, H.H., Zheng, Q., Yen, N.C., Tung, C.C., Liu, H.H., 1998.

. Proc. R. Soc. A Math. Phys. Eng. Sci. 454, 903–995. <https://doi.org/10.1098/rspa.1998.0193>

Ieracitano, C., Duun-Henriksen, J., Mammone, N., Foresta, F. La, Morabito, F.C., 2017.

. Proc. Int. Jt. Conf. Neural Networks 2017-May, 1297–1304. <https://doi.org/10.1109/IJCNN.2017.7966002>

Jeong, D.H., Kim, Y. Do, Song, I.U., Chung, Y.A., Jeong, J., 2016.

. Entropy 18. <https://doi.org/10.3390/e18010008>

Jones, S.R., 2016.

. *Curr. Opin. Neurobiol.* 40, 72–80.

Kabbara, A., Falou, W. El, Khalil, M., Eid, H., Hassan, M., 2017.

. *Int. Conf. Adv.*

*Biomed. Eng. ICABME* 2017-Octob, 0–3.

<https://doi.org/10.1109/ICABME.2017.8167549>

Keijzer, H.M., Tjepkema-Cloostermans, M.C., Klijn, C.J.M., Blans, M., van Putten, M.J.A.M., Hofmeijer, J., 2021.

. *Clin. Neurophysiol.* 132, 157–

164. <https://doi.org/10.1016/j.clinph.2020.10.024>

Kent, B.A., Feldman, H.H., Nygaard, H.B., 2021.

. *Prog.*

*Neurobiol.* 197, 101902. <https://doi.org/10.1016/j.pneurobio.2020.101902>

Kulkarni, N., 2019.

, in: Kulkarni, A.J., Satapathy,

S.C., Kang, T., Kashan, A.H. (Eds.), *Proceedings of the 2nd International Conference on Data Engineering and Communication Technology*. Springer Singapore, Singapore, pp. 395–403.

Law, L.S., Kim, J.H., Liew, W.Y.H., Lee, S.K., 2012.

. *Mech. Syst. Signal Process.* 33, 197–

211. <https://doi.org/10.1016/j.ymssp.2012.06.004>

Li, Y., Lei, M., Cui, W., Guo, Y., Wei, H.L., 2019.

. *IEEE*

*Trans. Biomed. Eng.* 66, 3509–3525.



Liu, X., Zhang, C., Ji, Z., Ma, Y., Shang, X., Zhang, Q., Zheng, W., Li, X., Gao, J., Wang, R., Wang, J., Yu, H., 2016.

. Cogn. Neurodyn. 10, 121–133. <https://doi.org/10.1007/s11571-015-9367-8>

Lotte, F., Bougrain, L., Cichocki, A., Clerc, M., Congedo, M., Rakotomamonjy, A., Yger, F., 2018.

. J. Neural Eng. 15. <https://doi.org/10.1088/1741-2552/aab2f2>

Mazaheri, A., Jensen, O., 2008.

. J. Neurosci. 28, 7781–7787.

McBride, J., Zhao, X., Munro, N., Jicha, G., Smith, C., Jiang, Y., 2015.

. J. Healthc. Eng. 6, 55–70. <https://doi.org/10.1260/2040-2295.6.1.55>

Melia, U., Guaita, M., Vallverdú, M., Embid, C., Vilaseca, I., Salamero, M., Santamaria, J., 2015.

. Med. Eng. Phys. 37, 297–308. <https://doi.org/10.1016/j.medengphy.2015.01.002>

Miraglia, F., Vecchio, F., Bramanti, P., Rossini, P.M., 2016.

. Clin. Neurophysiol. 127, 1261–1268. <https://doi.org/10.1016/j.clinph.2015.07.040>

Moca, V. V., Bârzan, H., Nagy-

. Nat. Commun. 12, 1–18. <https://doi.org/10.1038/s41467-020-20539-9>

Mousavi, A.A., Zhang, C., Masri, S.F., Gholipour, G., 2020.

(Switzerland) 20. <https://doi.org/10.3390/s20051271>

Nimmy, J.T., Subha, D.P., Menon, R.N., 2019.

. IEEE Reg. 10 Annu. Int.

Conf. Proceedings/TENCON 2019-October, 2295–2299.  
<https://doi.org/10.1109/TENCON.2019.8929242>

Nir, R.R., Sinai, A., Raz, E., Sprecher, E., Yarnitsky, D., 2010.

. Brain

Res. 1344, 77–86. <https://doi.org/10.1016/j.brainres.2010.05.004>

Peng, Z.K., Tse, P.W., Chu, F.L., 2005.

. Mech. Syst. Signal Process. 19, 974–988.

<https://doi.org/10.1016/j.ymssp.2004.01.006>

Pijnenburg, Y.A.L., Vd Made, Y., Van Cappellen Van Walsum, A.M., Knol, D.L., Scheltens, P., Stam, C.J., 2004.

.

Clin. Neurophysiol. 115, 1332–1339.  
<https://doi.org/10.1016/j.clinph.2003.12.029>

Poil, S.S., de Haan, W., van der Flier, W.M., Mansvelder, H.D., Scheltens, P., Linkenkaer-Hansen, K., 2013.

. Front. Aging Neurosci. 5, 1–12.

<https://doi.org/10.3389/fnagi.2013.00058>

Posthuma, D., Neale, M.C., Boomsma, D.I., De Geus, E.J.C., 2001.

. Behav. Genet. 31, 567–579.

<https://doi.org/10.1023/A:1013345411774>

Qiao, J., Lv, Y., Cao, C., Wang, Z., Li, A., 2018.

. Front. Aging Neurosci. 10, 1–12.  
<https://doi.org/10.3389/fnagi.2018.00417>

Rajintha S Gunawardena, S.A., Sarrigiannis, P.G., Blackburn, D.J., He, F., 2021.  
<https://doi.org/10.1101/2021.10.15.464451>

Ratti, E., Waninger, S., Berka, C., Ruffini, G., Verma, A., 2017.  
 . Front. Hum. Neurosci. 11, 1–7. <https://doi.org/10.3389/fnhum.2017.00398>

Sadaghiani, S., Kleinschmidt, A., 2016.  
 . Trends Cogn. Sci. 20, 805–817. <https://doi.org/10.1016/j.tics.2016.09.004>

Sakkalis, V., 2011.  
 . Comput. Biol. Med. 41, 1110–1117.  
<https://doi.org/10.1016/j.compbiomed.2011.06.020>

Sankari, Z., Adeli, H., 2011.  
 . J. Neurosci. Methods 197, 165–170. <https://doi.org/10.1016/j.jneumeth.2011.01.027>

Sankari, Z., Adeli, H., Adeli, A., 2012.  
 . Clin. EEG Neurosci. 43, 268–278.  
<https://doi.org/10.1177/1550059412444970>

Sankari, Z., Adeli, H., Adeli, A., 2011.  
 . Clin. Neurophysiol. 122, 897–906. <https://doi.org/10.1016/j.clinph.2010.09.008>

Scally, B., Burke, M.R., Bunce, D., Delvenne, J.F., 2018.  
 . Neurobiol. Aging 71, 149–155.  
<https://doi.org/10.1016/j.neurobiolaging.2018.07.004>

Schmidt, M.T., Kanda, P.A.M., Basile, L.F.H., da Silva Lopes, H.F., Baratho, R., Demario, J.L.C., Jorge, M.S., Nardi, A.E., Machado, S., Ianof, J.N., Nitrini, R., Anghinah, R., 2013.

. Front. Aging Neurosci. 5, 1–6.  
<https://doi.org/10.3389/fnagi.2013.00060>

Science, C., Engineering, S., 2014.  
4, 84–91.

Seth, A.K., Barrett, A.B., Barnett, L., 2015.  
. J. Neurosci. 35, 3293–3297.  
<https://doi.org/10.1523/JNEUROSCI.4399-14.2015>

Shan, X., Huo, S., Cao, J., Yang, L., Zou, J., Chen, L., Li, Z., 2021.

. IEEE Trans. Neural Syst. Rehabil. Eng.  
Shan. <https://doi.org/10.1109/TNSRE.2021.3076311>

Smit, C.M., Wright, M.J., Hansell, N.K., Geffen, G.M., Martin, N.G., 2006.  
. Int. J. Psychophysiol. 61, 235–243.  
<https://doi.org/10.1016/j.ijpsycho.2005.10.004>

Su, R., Li, X., Li, Z., Han, Y., Cui, W., Xie, P., Liu, Y., 2021.

. Biomed. Signal Process. Control 70,  
103000. <https://doi.org/10.1016/j.bspc.2021.103000>

Tafreshi, T.F., Daliri, M.R., Ghodousi, M., 2019.

. Cogn. Neurodyn. 13, 555–566. <https://doi.org/10.1007/s11571-019-09556-7>

Torres, M.E., Colominas, M.A., Schlotthauer, G., Flandrin, P., 2011.  
4144–4147.

Ursino, M., Zavaglia, M., Astolfi, L., 2007.

351–365. <https://doi.org/10.1007/s00422-006-0122-4>

van der Zande, J.J., Gouw, A.A., Steenoven, I. van, Scheltens, P., Stam, C.J., Lemstra, A.W., 2018.

. *Front. Aging Neurosci.* 10, 1–10.

<https://doi.org/10.3389/fnagi.2018.00190>

van Mierlo, P., Papadopoulou, M., Carrette, E., Boon, P., Vandenberghe, S., Vonck, K., Marinazzo, D., 2014.

. *Prog.*

*Neurobiol.* 121, 19–35. <https://doi.org/10.1016/j.pneurobio.2014.06.004>

Varotto, G., Fazio, P., Rossi Sebastiano, D., Duran, D., D'Incerti, L., Parati, E., Sattin, D., Leonardi, M., Franceschetti, S., Panzica, F., 2014.

. *Clin. Neurophysiol.* 125, 63–68.  
<https://doi.org/10.1016/j.clinph.2013.06.016>

Vecchio, F., Miraglia, F., Alù, F., Menna, M., Judica, E., Cotelli, M., Rossini, P.M., 2020.

. *J. Alzheimer's Dis.* 75, 1253–1261. <https://doi.org/10.3233/JAD-200171>

Vecchio, F., Miraglia, F., Piludu, F., Granata, G., Romanello, R., Caulo, M., Onofrj, V., Bramanti, P., Colosimo, C., Rossini, P.M., 2017.

. *Brain Imaging Behav.* 11, 473–485. <https://doi.org/10.1007/s11682-016-9528-3>

Wang, R., Wang, J., Li, H., Chen, Y., 2015a.

. *Proc. 2015 27th Chinese Control*

*Decis. Conf. CCDC 2015* 1822–1826.  
<https://doi.org/10.1109/CCDC.2015.7162215>

Wang, R., Wang, J., Yu, H., Wei, X., Yang, C., Deng, B., 2015b.  
. Cogn. Neurodyn. 9, 291–304. <https://doi.org/10.1007/s11571-014-9325-x>

Wang, R., Yang, Z., Wang, J., Shi, L., 2017.  
. Chinese Control Conf. CCC 5549–5554. <https://doi.org/10.23919/ChiCC.2017.8028238>

Wendling, F., Ansari-asl, K., Bartolomei, F., Senhadji, L., 2009.  
183, 9–18. <https://doi.org/10.1016/j.jneumeth.2009.04.021>

Wu, Z., Huang, N.E., 2009.  
. Adv. Adapt. Data Anal. 1, 1–41.

Yao, W., Wang, J., 2017.  
Phys. A Stat. Mech. its Appl. 484, 276–281.  
<https://doi.org/10.1016/j.physa.2017.04.181>

Yildirim, O., Talo, M., Ay, B., Baloglu, U.B., Aydin, G., Acharya, U.R., 2019.  
. Comput. Biol. Med. 113, 103387. <https://doi.org/10.1016/j.compbio.2019.103387>

Yu, H., Lei, X., Song, Z., Wang, J., Wei, X., Yu, B., 2018.  
. Phys. A Stat. Mech. its Appl. 506, 1093–1103.  
<https://doi.org/10.1016/j.physa.2018.05.009>

Zhao, Yifan, Zhao, Yitian, Durongbhan, P., Chen, L., Liu, J., Billings, S.A., Zis, P., Unwin, Z.C., De Marco, M., Venneri, A., Blackburn, D.J., Sarrianni, P.G., 2020.  
. IEEE Trans. Med. Imaging 39, 1571–1581.  
<https://doi.org/10.1109/TMI.2019.2953584>

## 4 The Application of Functional Connectivity on Brain Disease Classification

This chapter is based on two published journal papers. ‘Using interictal seizure-free EEG data to recognise patients with epilepsy based on machine learning of brain functional connectivity’ is published in

. ‘Spatial – temporal graph convolutional network for Alzheimer classification based on brain functional connectivity imaging of electroencephalogram’ is published in .

Cao, J., Grajcar, K., Shan, X., Zhao, Y., Zou, J., Chen, L., Li, Z., Grunewald, R., Zis, P., Marco, M. De, Unwin, Z., Blackburn, D., Sarrigiannis, P.G., 2021a.

. Biomed. Signal Process. Control 67, 102554. <https://doi.org/10.1016/j.bspc.2021.102554>

Shan, X., Cao, J., Huo, S., Sarrigiannis, P.G., Zhao, Y., Chen, L., 2022.

. Hum. Brain Mapp. 1–16. <https://doi.org/10.1002/hbm.25994>

### 4.1 Using interictal seizure-free EEG data to recognise patients with epilepsy based on machine learning of brain functional connectivity.

#### 4.1.1 INTRODUCTION

Epilepsy is one of the most common neurological disorders and it can affect people of various age, gender and ethnic origin. According to the World Health Organization in 2019, there were worldwide around 50 million people with the condition. Epilepsy is characterised by recurrent seizures and carries variable degrees of morbidity and mortality depending on the electroclinical characteristics of the seizures. Some types of epileptic seizures carry the potential to put patients and others at risk of harm, for example, if they occur during cooking, driving or swimming. Seizure detection and epilepsy diagnosis have therefore attracted many studies using various measurements. Electroencephalography (EEG) has become one of the most important

approaches to assist the diagnosis of epilepsy. EEG is a non-invasive and painless technology for measuring brain activity that is also economical, easy to administer and widely available in most hospitals. When compared with other methods that provide information about the anatomical structure of the brain, such as MRI and fMRI, EEG offers ultra-high time resolution (Pievani, 2011), which is critical to understand brain function. Synchronous networks form and dissipate in the range of 100-300ms which is thought to be the meaningful operational brain temporal scale (Varela, 2001).

Empirical interpretation of the EEG is largely based on recognising abnormal frequencies in specific biological states (e.g. wakefulness versus sleep (Brodbeck, 2012; Lioi, 2017)), the spatial-temporal and morphological (e.g. sharp waves, spikes etc.) characteristics of paroxysmal (Dash, 2018) or persistent discharges (Renzel, 2017), reactivity to external stimuli and activation procedures (like a period of hyperventilation (Watanabe, 2018) or intermittent photic stimulation (Visani, 2010)). Despite being useful in many instances, these practical approaches for interpreting EEGs leave important linear or nonlinear interactions between various brain network anatomical constituents, buried undetected within the recordings as such interactions are far beyond the observational capabilities of any specialised physician in this field (Sarrigiannis, 2014, 2018; Blackburn, 2018). Another limitation is that previous studies are based on univariate methods, focusing on single EEG recording channels. Studying brain functional connectivity among multiple channels (i.e., multivariate methods), by examining the magnitudes of temporal correlations or coherence with frequency, is increasingly being recognised as an important approach with many advantages. Various abnormalities in brain functional connectivity have been reported for numerous brain disorders (Uhlhaas and Singer, 2006; Guevara Erra, Perez Velazquez and Rosenblum, 2017) but to the best of our knowledge, this has not yet translated in a diagnostic method for clinical use. More efforts are required to further explore the brain connectivity measures as new biomarkers for the diagnosis of epilepsy. Different methods were previously introduced to measure brain connectivity, based on linear or nonlinear association in the time,



frequency, and time-frequency domains. As the most widely studied type of brain connectivity, functional connectivity can be presented by magnitude squared coherence (Sakkalis, 2011; Battaglia and Brovelli, 2019; Tafreshi, Daliri and Ghodousi, 2019), Minimum Description Length (Sakkalis, 2011; Salman, Grover and Shankar, 2018), phase synchronisation (Sakkalis, 2006), phase-locking value (Lotte, 2018; Bedo, Ribary and Ward, 2020), robust synchronisation (Sakkalis, 2006; Delgado-Restituto, Romaine and Rodríguez-Vázquez, 2019), non-linear correlation coefficient (Wendling, 2010), correlation (Horstmann, 2010) and power distributions (Rosch, 2017) etc. The prospect to use those techniques to aid the diagnosis of epilepsy remains unknown.

Fortunately, the large majority of patients with epilepsy experience seizures only infrequently, although there are exceptions to this rule, for example, the type of seizures characterizing children with various forms of absence seizures; absences can frequently occur daily and can, in addition, be easily precipitated by a brief period of hyperventilation. The EEG recording of a typical absence seizure is shown in Figure 4-1 and can be separated into three states: seizure (ictal), pre-seizure (pre-ictal) and seizure-free (interictal). The difference between a seizure and interictal EEG segment is clearly visible in terms of the amplitude and morphology of the generalised epileptiform discharges dominating the ictal phase. Thus, epilepsy diagnosis when one of the patient's typical seizures is captured can be relatively straightforward. Nonetheless, as most seizures in adults with epilepsy occur rather infrequently, the interictal EEG (i.e., recorded while the patient is not experiencing any epileptic seizures)

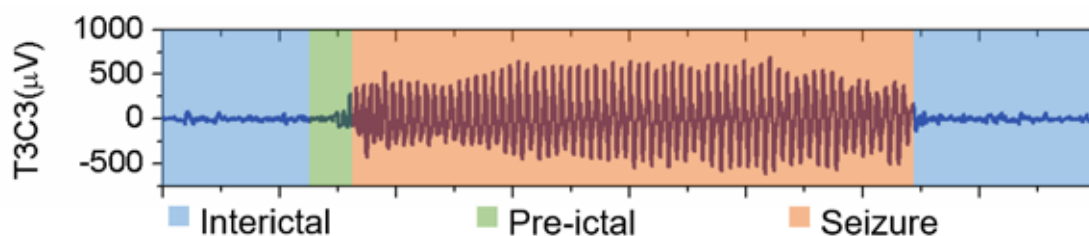


**Figure 4-1 A typical EEG recording before during and after a generalised epileptic seizure (typical absence) recorded from the left centrotemporal area. The ictal segment (i.e., seizure) contains large magnitude epileptic discharges, which can be easily distinguished from the seizure-free recording (interictal).**

is the most commonly available data for clinicians attempting to diagnose epilepsy. The interictal epileptiform discharge (i.e., sharp wave and/or spike) recognised empirically by the reporting EEG physician, is an expression of the abnormal neuronal and brain network behaviour, a demonstration of cortical hypersynchrony and hyperexcitability. However, a first EEG in adult patients, subsequently proven to have epilepsy, has very low sensitivity, ranging between 29-55%, that can go up to 80 to 90% on repeats of the examination (Pillai and Sperling, 2006).

EEG is one of the most useful diagnostic procedures for epilepsy. It provides evidence for the diagnosis, classification and management of different types of epilepsy (Smith, 2005; Noachtar and Rémi, 2009). It should be noted that most of the current research in this topic focus on seizure detection, in subjects that have epilepsy while relatively limited work is centred on normal in appearance, interictal EEG epochs in patients where there is independent strong evidence for the electroclinical diagnosis of epilepsy. Amin . (2020) proposed a novel method based on wavelet analysis and arithmetic coding to achieve efficient classification between epileptic seizure signals and seizure-free signals, while this Chapter tended to classify epilepsy patients from healthy people merely using seizure-free EEGs. Similarly, to identify epileptic seizures from EEG signals, Dhiman & Priyanka (2014) proposed a novel scheme based on discrete wavelet packet transform and Zhu (2014) proposed a fast weighted horizontal visibility graph constructing algorithm. In studies aiming to diagnose epilepsy, most papers use data that include ictal EEG recordings, such as (Fani, Azemi and Boashash, 2011; Xie and Krishnan, 2013; Kaya and

Erfanian, 2020), where seizure detection is also involved. However, analysis of interictal seizure-free EEG has gradually attracted more attention recently. For instance, Lopes . (2019) suggested that the interictal EEG supported by many publications can also contribute to epilepsy diagnosis and they developed a framework to classify focal and generalized epilepsy by extracting Ictogenic Spread as biomarker from normal EEG. Horstmann . (2010) found differences in the functional networks of seizure-free intervals of patients with

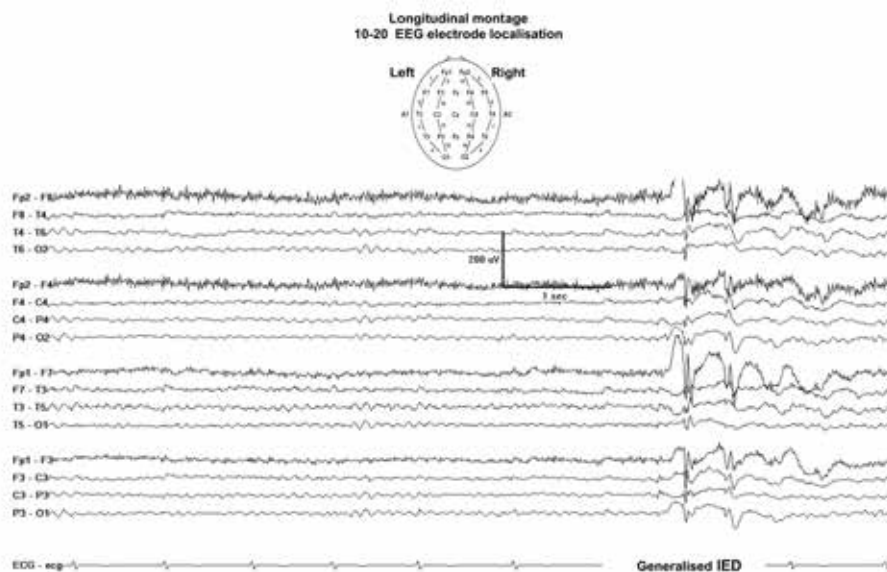


**Figure 4-2 EEG recording from one of the patients of the epileptic cohort showing an interictal epileptiform discharge (IED). This demonstrates a clear generalised distribution in keeping with the patient’s diagnosis of generalised epilepsy. Of note, the EEG is “normal-looking”, i.e. Empirically classified as normal, prior to the generalised paroxysmal discharge. “Normal-looking” EEG epochs were the dominant feature on this interictal EEG recording, a common occurrence in this group of patients.**

focal epilepsy (treated with anti-epileptic drugs) and healthy controls as measured from EEG/MEG. Besides, the existence of alpha rhythm abnormalities in patients with epilepsy was explored based on the information obtained from interictal scalp EEG recordings (Pyrzowski , 2015). In the appropriate clinical context, the presence of a generalised interictal epileptiform discharge (IED) further strongly supports the clinical diagnosis of generalised epilepsy which will be the focus of this work.

The research question of this Chapter is to examine if appropriate analysis of “normal-looking”, i.e. empirically classified as normal, interictal EEG data can identify adult patients with generalised epilepsy. This task is unlikely to be solved by direct inspection of EEG recordings due to the lack of obvious patterns to distinguish patients with epilepsy from age and gender-matched healthy controls. Achieving this aim could greatly improve the diagnostic sensitivity and specificity of interictal EEG, reducing diagnostic times and cost

while ensuring that epileptic patients can promptly initiate appropriate treatment. Furthermore, the EEG recording time and repeat EEG recordings can be kept to a minimum leading to much more efficient use of resources and increased patient satisfaction. Another challenge of this Chapter was to determine whether the “normal-looking” interictal EEG recordings of patients with non-epileptic attacks (the ictal EEG is normal during those episodes) occurring within the context of a psychogenic disorder, called non-epileptic attack disorder (NEAD) could be differentiated from equivalent recordings of patients with epilepsy. NEAD is a brain-related disorder which involves psychogenic non-epileptic seizures (Sheldon and Agrawal, 2019). According to Milán-Tomás (2018), about 20% to 40% of patients diagnosed with epilepsy also have NEAD rendering differentiation between the two conditions a challenging task for physicians. Furthermore, frequently NEAD patients receive several anti-epileptic drugs (AED) due to their seizures being pharmacoresistant as they are psychogenic in nature. As a result misdiagnosis of NEAD may cause patients serious iatrogenic arrhythmia (Brown, 2011). Therefore, a computer-aided classification between NEAD and epilepsy becomes important but very challenging as the two conditions can co-exist in the same subject and appropriate data labelling is problematic.

Different from most of the existing studies either focus on single EEG channel or EEG data with seizures, this Chapter proposes a novel framework to diagnose epilepsy based on interictal seizure-free EEG data only using a set of estimates of linear and nonlinear brain functional connectivity. A machine learning approach is then employed to classify (a) the epilepsy group (EG) vs healthy controls (HC), and (b) EG vs NEAD, followed by a visualisation of classification results. Although the methods to calculate the connectivity are not new, the attempt to systematically evaluate their potential on interictal seizure-free EEG data in time, frequency and time-frequency domains has not been done before, particularly for the challenging classification of EG vs NEAD.

## 4.1.2 METHODOLOGY

### 4.1.2.1 Dataset

In this Chapter we have retrospectively selected video EEGs (vEEGs) with occasional generalised IEDs, providing electrophysiological evidence for the diagnosis of epilepsy. All selected patients were isolated from the Royal Hallamshire Hospital (Sheffield, UK) Department of Neurophysiology database with the following inclusion criteria: standard interictal EEG available containing at least one well defined generalised IED (Figure 4-2); age between late teens to 61 years (based on a cohort of HCs available from previous work to ensure no significant age differences between groups occurred); their EEGs included periods of wakefulness with eyes open (EO) and eyes closed (EC) epochs; previous history of at least one witnessed generalised tonic-clonic seizure without any other known type of seizures. We have also selected a cohort of patients with NEAD where we have captured at least one typical psychogenic attack on their vEEG recording for which there was no evidence of ictal or IED, no dynamically evolving ictal EEG patterns and no clear history of other types of seizure. The following exclusion criteria were also applied: learning difficulties; sleep deprivation the night before the EEG was recorded; known history of drug addiction; refractory epilepsy; any other known neurological disorder other than epilepsy or NEAD. Some of the NEAD patients received various AEDs but there was no convincing evidence on their past medical history (reviewed during their EEG recording) to suggest epileptic seizures. However admittedly this cannot be entirely excluded for this group of patients.

A Natus Headbox (Optima Medical, Ltd.) at a sampling rate of 500Hz (analogue bandwidth 0.1–200Hz) and a standard international 10–20 system of electrode placement positions were used for the recordings for all subjects. The EEG data was recorded from the standard 21 electrodes of the 10-20 system of electrode placement. The investigated EEG data comes from three groups: HCs, EG and NEAD. 10 HCs (6 females, mean age  $37\pm15y$ ), 15 EG (10 females, mean age  $33\pm12y$ ) and 14 NEAD cases (10 females, mean age  $33\pm13y$ ) were collected, details of which can be seen in Table 4-3 of Supplementary material. It should

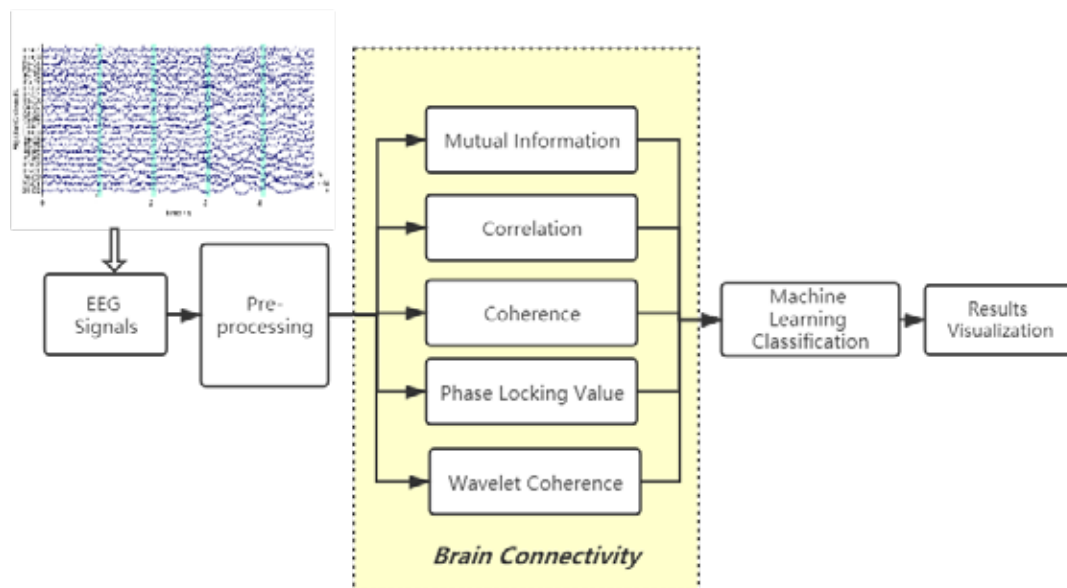


**Figure 4-3 Flow Diagram for EG vs. HC and EG vs. NEAD classification**

be noted that EG and NEAD participants were consecutively selected for the best possible match of age to our healthy control cohort. Additionally, the data for each participant was divided into two states: eyes open (EO) and eyes closed (EC). For each HC participant in each eye state, 3 trials were collected and each epoch lasts 12 seconds. For each NEAD and EG participant in each eye state, 2 trials were collected and each epoch lasts 12 seconds. The total number of available data for each group is therefore similar to ensure the fairness of training and validation, as shown in Table 4-3. All HC participants provided informed consent as part of a project approved by the Yorkshire and the Humber (Leeds West) Research Ethics Committee (reference number 14/YH/1070). For the retrospectively selected EEG data for the NEAD and EG cohorts, ethics approval for use of patients' EEGs for the development of novel qEEG methodologies was granted both from the University of Sheffield and the NHS ethics committees (SMBRER207 and 11/YH/0414).

Bipolar and unipolar (i.e., referential) derivations are the two main types of EEG recordings used in everyday clinical practice. The unipolar montage uses one channel as a source and the other as a reference that is usually fixed; this is the

default mode of recording EEG data in routine clinical practice. The bipolar montage is obtained by subtraction of two unipolar derivations. This Chapter used bipolar derivations to minimise the effects of volume conduction introduced by a common reference (Fein , 1988; Nunez , 1997). We have produced 23 bipolar channels calculated by the 21-channel unipolar recordings after excluding the frontopolar electrodes (Fp1 and Fp2) due to their vicinity to the eyes that results in high levels of eye blink artefacts.

The EEG epochs were collected with Spike2 (version 9) software where data filtering and labelling was also undertaken. The filtering method was performed using the below equation :

$$y(k) = x(k) \frac{1}{2c \sum_{i=k-c/f_s}^{k+c/f_s} x(i)} \quad (4-1)$$

where  $x$  is the input EEG signal,  $y$  is the output signal,  $k$  is a discrete-time point,  $f_s$  is the sample rate and  $c$  is a time constant value. The value of  $c$  is set at 0.2s. The filtering using a time constant result in a high pass filtering of the signal where the cut-off frequency  $f_c$  is:

$$f_c = \frac{1}{2\pi c} \quad (4-2)$$

The value of  $f_c$  is equal to 0.79Hz.

All data were selected by a specialised physician in clinical neurophysiology to ensure no interictal EEG abnormalities were included. Additionally, care was taken to select relatively artefact free epochs. Furthermore, Figure 4-3 shows the proposed framework, including pre-processing, features extraction, classification and result visualization, details of which are explained below. Sensor level brain functional connectivity measures, estimated by five methods, were extracted from EEG data as features for further classification.

#### 4.1.2.2 Data pre-processing

Before estimating functional connectivity, data was pre-processed by the following steps:

1. For the groups of EG (15 subjects) and NEAD (14 subjects), there are 2 epochs for EO and EC for each subject. For the HC group (10 subjects), there are 3 epochs for EO and EC for each participant. To increase the number of samples for classification, each dataset (12 seconds) was segmented into 3 mini-epochs, 4 seconds each. In total, there were 528 labelled samples for 3 groups (EG:180, NEAD:168 and HC:180), prepared for feature extraction and classification. As brain interactions are usually highly dynamic, prolonged EEG segments can obscure this important characteristic (Durongbhan , 2019).
2. Each segment of data was normalised in the range from -1 to 1 for all bipolar channels. This step was performed mainly for visualisation purposes at the beginning of this work to search for any possible artefacts and does not affect any of the connectivity measures used in this Chapter.
3. In this framework, it is tended to extract features in each frequency band. Therefore, each data segment was filtered to produce six frequency bands: full band (no filtering), Alpha (8-15Hz), Beta (15-32Hz), Gamma (>32Hz), Delta (<4Hz), Theta (4-8Hz). This operation was performed using an FIR filter with an order of 600. The frequency ranges of the Delta and full bands were modified to eliminate the effect of the time constant filtering.

**Table 4-1 The properties of the estimated functional connectivity**

| Measure             | Linear | Nonlinear | Time | Frequency | Time-Frequency |
|---------------------|--------|-----------|------|-----------|----------------|
| Mutual information  |        |           |      |           |                |
| Correlation         |        |           |      |           |                |
| Coherence           |        |           |      |           |                |
| Phase locking value |        |           |      |           |                |
| Wavelet coherence   |        |           |      |           |                |

#### 4.1.2.3 Functional connectivity estimation

This Chapter implements 5 different measures in time, frequency and time-frequency domains to represent the brain functional connectivity between two EEG signals  $x_i$  and  $y_i$ . These measures with their extension are then



constructed as classification features. The properties of these measures are presented in Table 4-1.

Mutual Information (MI) measure indicates the mutual dependence of two signals, i.e. how much information is shared between two signals (Cover and Thomas, 2005). It is based on a probability function and entropy. The entropy of a signal  $X$  with a length of  $n$  is expressed as

$$H(X) = \sum_{i=1}^n p(x_i) \log p(x_i) \quad (4-3)$$

where  $p(x_i)$  is the probability function and values  $x_i$  ( $i = 1, 2, 3, \dots, n$ ) represent all possible values of the signal  $X$ . If the entropy of the signal is high, then the signal contains a lot of different values. For low entropy, the signal is more organised, for example, more values are repeated in the signal. A joint probability  $p(x_i, y_j)$  is the probability that  $X = x_i$  and  $Y = y_j$ . The mutual information  $I(X; Y)$  between two signals  $X$  and  $Y$  is expressed as

$$I(X; Y) = \sum_{i=1}^n \sum_{j=1}^m p(x_i, y_j) \log \frac{p(x_i, y_j)}{p(x_i)p(y_j)} \quad (4-4)$$

where  $n$  and  $m$  are the length of the signal  $X$  and  $Y$  respectively. A total of 256 bins were used to calculate the value of mutual information.

This well-known measure shows how two signals are correlated with each other corresponding to a time shift between these two signals. In this Chapter, the cross-correlation used is calculated using the following formula

$$R_{xy}[n] = \frac{\sum_{k=-\infty}^{\infty} x[k]y[k+n]}{\sqrt{\sum_{k=-\infty}^{\infty} x[k]^2} \sqrt{\sum_{k=-\infty}^{\infty} y[k]^2}} \quad (4-5)$$

where  $x[k]$  and  $y[k]$  are discrete signals and  $n$  is the lag index. Cross-correlation is a measure of the linear correlation (dependence) between two signals, giving a value between -1 and 1 inclusive, where 1 indicates a total

positive correlation, 0 indicates no correlation, and -1 indicates a total negative correlation. Three features are taken from the cross-correlation, namely the maximum value of correlation (CorrMax), the mean value of correlation (CorrMean) and the correlation lag at the maximum value of correlation (CorrLag). It should be noted that the absolute value of the correlation is taken into account in terms of the maximum value of the correlation. This is due to the fact that two EEG signals can also be negatively correlated.

The coherence in the frequency domain is estimated using magnitude squared coherence (MSC), expressed as

$$C_{xy}(f) = \frac{|G_{xy}(f)|^2}{G_{xx}(f)G_{yy}(f)} \quad (4-6)$$

where  $G_{XY}(f)$  is the cross-spectral density of signals  $X$  and  $Y$ .  $G_{XX}(f)$  and  $G_{YY}(f)$  are the auto-spectral density of these signals. The step size of the frequency is  $f_s/nFFT = 500/2048 = 0.2441$  Hz. Two features are taken from the MSC: maximum (CohMax) and mean values (CohMean). Although these features are extracted for six frequency bands, MSC is computed only for signal filtered to the full band range (2Hz-60Hz). To obtain MSC for different bands, MSC is segmented using the appropriate frequency ranges for each band and the mean and maximum values of the MSC are taken from these frequency segments.

The Phase Locking Value (PLV) was first introduced by Lachaux . (Lachaux , 1999) in 1999. PLV measures the significance of the phase covariance between two signals. It is defined as

$$PLV(t) = \frac{1}{N} \left| \sum_{n=1}^N e^{i\theta(t,n)} \right| \quad (4-7)$$

where  $\theta(t, n)$  is the difference in the phase of two signals. PLV values are in the range from 0 to 1. If the phase difference of two signals remains the same, PLV

is close to one. A PLV close to zero indicates that there is no phase synchrony between two signals.

To compute the instantaneous phase of each EEG signal, a Hilbert transform is used. For the signal  $x(t)$ , this transform allows creating an analytic signal  $a_1(t)$  as:

$$a_1(t) = x(t) + ih_1(t) \quad (4-8)$$

where  $i$  is an imaginary number and  $h_1(t)$  is the Hilbert transform of  $x(t)$ . The instantaneous phase is defined as an angle between the real and imaginary part of the analytic signal and is written as

$$\theta_1(t) = \arctan\left(\frac{h_1(t)}{x(t)}\right) \quad (4-9)$$

The same method can be applied to the signal  $y(t)$ . Lachaux (Lachaux, 1999) used  $N$  trials to obtain statistically significant values of PLV. This stage is omitted in this project, thus  $N = 1$  during the calculation of the PLV. Additionally, the differences between the instantaneous phases are computed for all points in time and the differences are summed to obtain the PLV. To normalise the PLV, the obtained sum is divided by the signal length. The final equation to calculate PLV is then rewritten as

$$PLV = \frac{1}{T} \left| \sum_{i=1}^T e^{i\theta(t_i)} \right| \quad (4-10)$$

where  $T$  represents the size of a time window and  $t_i$  is a discrete point in time.

The wavelet formulation of coherence between two signals,  $x$  and  $y$ , and in the frequency  $w$  and time  $t$  domain, can be formulated as (Zhao, 2018):

$$co_{xy}(w, t) = \frac{|S_{xy}(w, t)|^2}{S_x(w, t) S_y(w, t)} \quad (4-11)$$

$$S_{xy}(w, t) = \mathbb{E}(W_x(w, t) \overline{W_y(w, t)})$$

where  $S_{xy}(w, t)$  is the wavelet cross-spectrum between  $x$  and  $y$  and  $S_x(w, t)$ ,  $S_y(w, t)$  are the corresponding auto-spectrums. Working with two single signals (single realisation) usually requires using a smoothing operator (see  $f(\cdot)$  operator in Eq. (12)), and ergodicity properties should be assumed (Sairamya, 2018).

$$\widehat{coh}_{xy}^2(w, t) = \frac{|f(S_{xy}(w, t))|^2}{f(S_x(w, t)) f(S_y(w, t))} \quad (4-12)$$

Two features are extracted from the wavelet coherence: mean value (WCohMean) and maximum value (WCohMax). As wavelet coherence is in the time-frequency domain, mean and maximum values are taken from all time points and the relevant frequency ranges. The mother wavelet was used for this Chapter.

#### 4.1.2.4 Classification using machine learning.

Considering the estimated measures of functional connectivity between two signals, with the 23 channels (shown in the supplementary Table 4-4), used in this Chapter, there are 253 possible combinations ( $C_2^{23}$ ) when any 2 bipolar derivations are paired together. These were organised in a pairwise manner by taking the first bipolar derivation in the list (F8-F4) and pairing it with every other bipolar derivation according to their order on the list (F8-F4:F7-F3, F8-F4:F4-C4, F8-F4:F3-C3, ...). The process was repeated for all other channels until the end of the list. However, since each channel is bipolar in nature, any pair with common electrode locations (such as F8-F4 and F4-C4) is neglected as this could lead to misleading high false correlation between the pair. The 46 channel pairs that have this characteristic include F8-F4:F4-C4, F8-F4:F4-FZ, F7-F3:F3-C3, F7-F3:F3-FZ, F4-C4:F4-FZ, F4-C4:T4-C4, F4-C4:C4-CZ, F4-C4:C4-P4, F3-C3:F3-FZ, F3-C3:T3-C3, F3-C3:C3-CZ, F3-C3:C3-P3, F4-FZ:F3-FZ, F4-FZ:FZ-CZ, F3-FZ:FZ-CZ, FZ-CZ:C4-CZ, FZ-CZ:C3-CZ, FZ-CZ:CZ-PZ, T4-C4:C4-CZ, T4-C4:C4-P4, T4-C4:T4-T6, T3-C3:C3-CZ, T3-C3:C3-P3, T3-C3:T3-T5, C4-CZ:C3-CZ, C4-CZ:CZ-PZ, C4-CZ:C4-P4, C3-CZ:CZ-PZ, C3-CZ:C3-P3, CZ-PZ:P4-PZ, CZ-PZ:P3-PZ, C4-P4:P4-PZ, C4-P4:P4-O2, C3-P3:P3-PZ, C3-

P3:P3-O1, T4-T6:T6-O2, T3-T5:T5-O1, P4-PZ:P3-PZ, P4-PZ:P4-O2, P3-PZ:P3-O1, T6-O2:P4-O2, T6-O2:O1-O2, T5-O1:P3-O1, T5-O1:O1-O2, P4-O2:O1-O2, P3-O1:O1-O2. A total of 207 channel pairs is therefore analysed in this Chapter.

The total number of feature values for this work is based on 9 estimations (PLV, MI, Corrmax, corrMean, CorrLag, MaxCoh, MeanCoh, MaxWCoh and MeanWCoh)  $\times$  6 bands (Full, Delta, Theta, Alpha, Beta and Gamma)  $\times$  207 pairs  $\times$  528 samples = 5,901,984. To explore the statistically significant differences between EG vs HC, and EG vs NEAD, one-way analysis of variance (ANOVA) was employed to select channels for each band and the estimation was undertaken using  $\alpha < 0.00001$  and  $\alpha < 0.05$  for two classification tasks respectively. ANOVA tests the null hypothesis, i.e. means of the tested groups are equal and the  $p$ -value indicates the statistical significance. Rejection of the null hypothesis leads to the conclusion that the two groups are statistically different. The selection of the threshold was based on previous studies (Orekhova et al., 2014; Vecchio et al., 2016).

The K-Nearest Neighbour (KNN) algorithm was applied to perform the classification of the selected features. During the initial development of classification solutions, different machine learning algorithms were analysed, such as Support Vector Machine (SVM), decision tree and KNN etc. KNN tended to present performance superiority in this Chapter compared with other methods and was selected as the main classifier in this Chapter. During the development of KNN, different values of  $k$  were tested. Initially,  $k$  was equal to 1 and was gradually increased to find the optimal one. Finally,  $k$  was set to 15 as it showed the best classification accuracy. For higher values of  $k$  (17 to 40) the accuracy was not improving and for a big  $k$  (above 40), the accuracy was dropping. Euclidean distance was applied in the KNN classification because of its better interpretability and performance (Prasath et al., 2017).

The dataset was divided into 10 subsets and then cross-validation was undertaken. For each iteration of the 10-fold cross-validation, different subsets are used for training and testing. In the first iteration, the first subset is used for testing and the remaining subsets are used for training. The second iteration

uses the second subset for testing and so on. To obtain the final result, an average of 10 classification accuracies is computed. Each accuracy comes from a single iteration of k-fold cross-validation. The 10-fold cross-validation is performed for 5 times by reshuffling data to gather 5 accuracy results statistically for each classification task. In the 10-fold cross-validation, 475 samples were used for training and 53 samples were used for testing.

To further evaluate the machine learning algorithms performance, accuracy (Accu), sensitivity (Sens) and specificity (Spec) were calculated, which are defined as:

$$Accu = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (4-13)$$

$$Sens = \frac{TP}{TP + FN} \times 100\% \quad (4-14)$$

$$Spec = \frac{TN}{TN + FP} \times 100\% \quad (4-15)$$

where TP = True Positive, FN = False Negative, TN = True Negative, FP = False Positive. Moreover, the receiver operating characteristic (ROC) curve, and the area under the ROC curve (AUC) (Pyrzowski, 2015; Lotte, 2018) were used to assess the goodness of classification to select appropriate machine learning methods, k value for KNN and features. Specifically, ROC is constructed from true positive rate (TPR = sensitivity) in the vertical axis and false positive rate (FPR = 1-specificity) in the horizontal axis (Blinowska, 2017). Besides, both ANOVA and multiple comparisons were performed to statistically compare the results of different brain connectivity estimations from different bands.

### 4.1.3 RESULTS

The section aims to report the classification results based on a single feature selected from different connectivity features, bipolar pairs and frequency bands.

#### 4.1.3.1 Healthy Controls vs Epilepsy Group

Considering the considerable number of possible channel combinations (207), the one-way ANOVA for a threshold of  $< 0.00001$  was used to determine if the difference of each feature between HC and EG is statistically significant. Only the pairs with metrics that are significantly different between these two groups were selected for further classification. For example, for in the Beta band, 67 out of 207 pairs showed a significant difference between these two groups. It is expected that statistically, the selected pairs will provide relatively high classification accuracy using the machine learning classifiers. As shown in Figure 4-13 in Supplementary material, the pairs with a small  $p$ -value present a more distinguishable distribution of features than the ones with a relatively large  $p$ -value. However, it should be noted that the  $p$ -value cannot fully represent or replace the classification performance due to two reasons: (1) it focuses more on the mean of each group's features while machine learning classifiers pay more attention to the distribution of features, and (2) the cross-

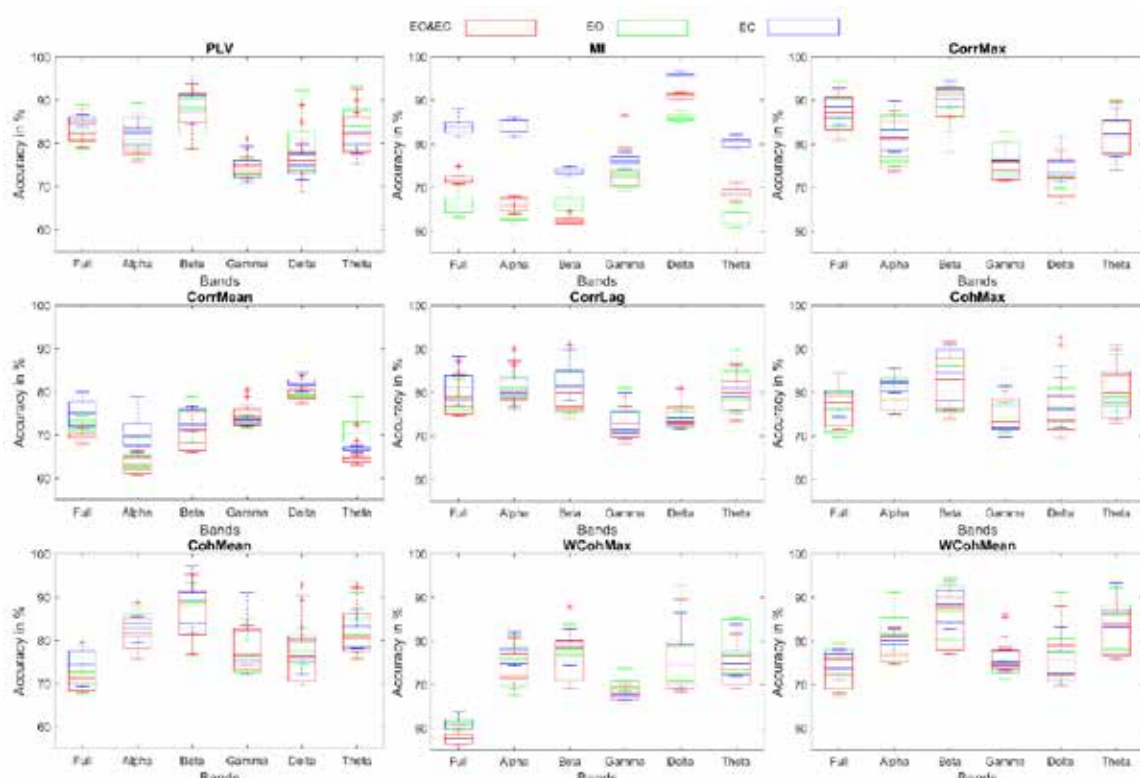
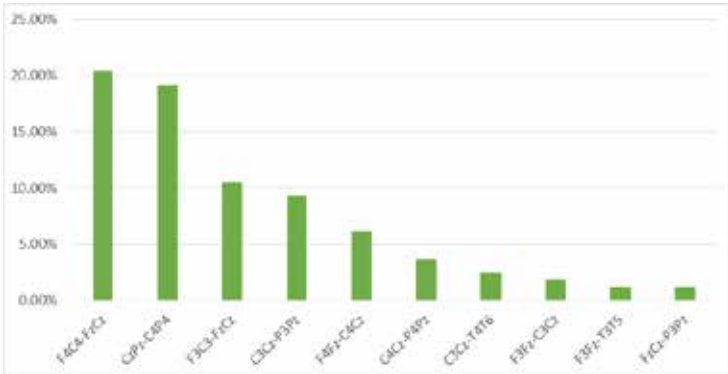


**Figure 4-4** The box charts of the classification accuracy of HC vs EG against various frequency bands for the 9 selected measures, where only the top 10 pairs in terms of classification accuracy were considered. Each box shows highest, lowest, and median accuracy.

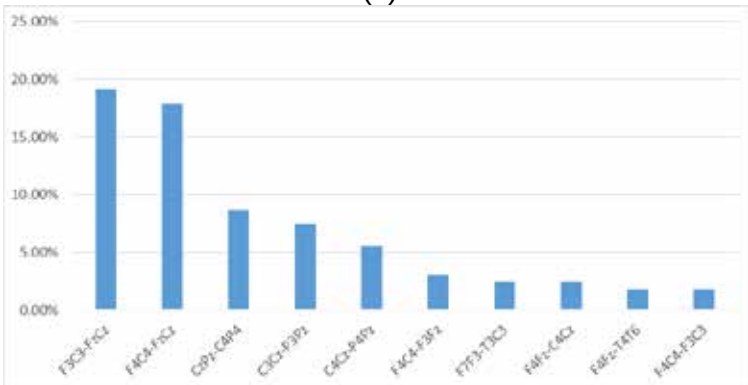
validation usually is used to determine the classification accuracy where testing data are not sampled, while to calculate the  $\chi^2$ -value all samples are used.

Figure 4-4 plots the box charts of the classification accuracy of EG vs HC, against various frequency bands for the 9 selected measures, where only the top 10 pairs in terms of classification accuracy were considered. For the Theta band, all measures have similar performance with around 80% accuracy except

( $<70\%$ ) For the Delta band,  $\chi^2$  has exceptionally high performance with over 95% accuracy in EC while other bands have much lower accuracy ( $<80\%$ ). The rest 8 measures have only 70-80% accuracy. The Gamma band has relatively low accuracy across all measures ( $<80\%$ ). For the Beta band,  $\chi^2$ ,  $\chi^2$ ,  $\chi^2$  and  $\chi^2$  have relatively good performance ( $>85\%$ ) while others are less than 80%. Overall, the Beta band has the best



(a)



(b)

**Figure 4-5 The top 10 bipolar pairs based on the percentage of appearance of top 3 classification performance between HC and EG for different connectivity measures and bands; (a) EO; (b) EC**



performance across all measures. The Alpha band has decent accuracy across all measures (around 80%). In terms of eye state, has the most distinguishable performance between EC and EO. Specifically, for all bands except the Beta band, the pattern of EC>(EC&EO)>EO can be observed. For other features' performance seems not to be significantly affected by the eye condition to an extent while combining EO and EC tends to slightly decrease the ability to discriminate EG from HC. This observation suggests that MI is more appropriate to classify these two groups for EC than EO and that the samples with different eye states should not be mixed.

To investigate which bipolar pairs consistently produced high classification accuracy, Figure 4-5 plots the top 10 bipolar pairs based on the percentage of appearance of top 3 classification performance for different connectivity measures and bands. It can be observed that the top 3 pairs (F4C4-FZCZ, CZPZ-C4P4, F3C3-FZCZ) are identical for EO and EC. More detailed results using in the Beta band during the EC state are shown in Figure 4-6. Besides, Figure 4-8 represented the areas in a head map corresponding to the highest classification accuracy and the discrimination ability of identified features were evaluated by ROC and AUC.

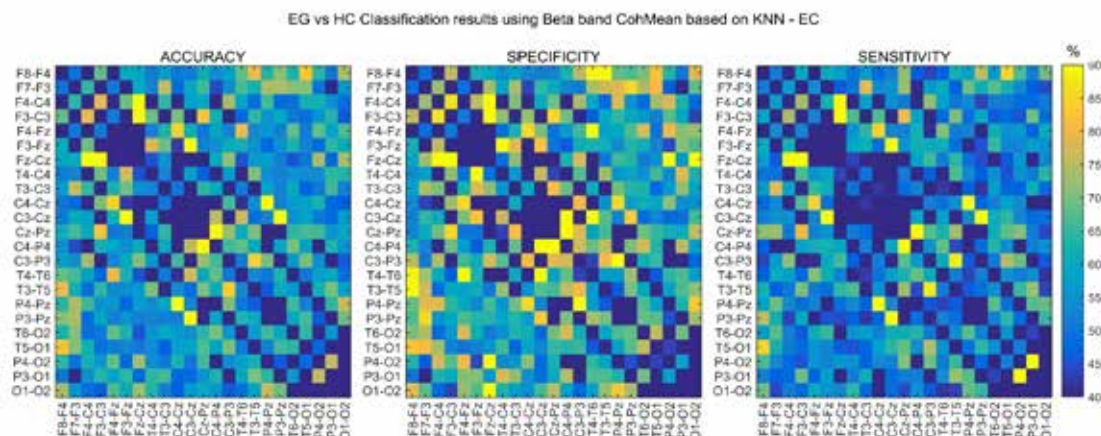


**Figure 4-6 The heatmap for the HC vs EG classification performance based on CohMean in Beta band for the EC state.**

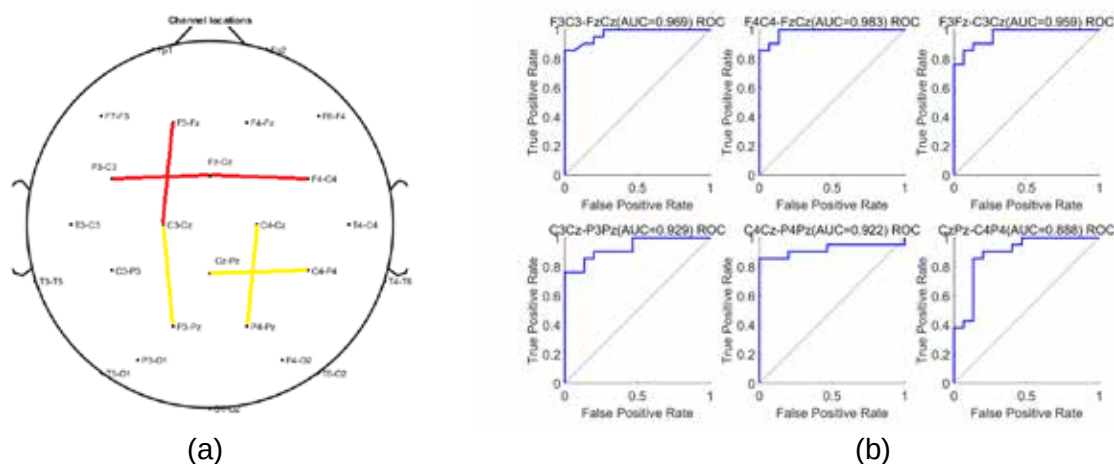


**Figure 4-8 (a)** The locations of 6 channel pairs corresponding to the highest classification accuracy (>90%) for the HC vs EG based on CohMean in the Beta band in the EC state (Red area is better than Yellow area); **(b)** The ROC and AUC of the 6 pairs.

To reveal the overall performance against different frequency bands, Figure 4-14 in Supplementary material shows the top accuracy using the bipolar pair F4C4-FZCZ for different features and frequency bands. The result was achieved by averaging 5 trials. The best classification result (97.22%) was

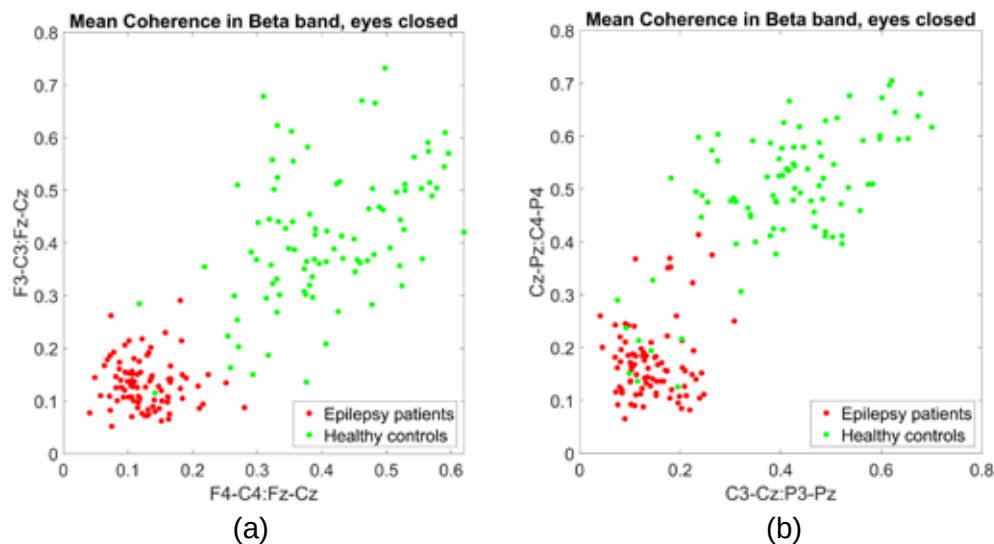


**Figure 4-7** The distribution of mean coherence values in the Beta band of 4 channel pairs corresponding to the highest classification accuracy (>90%) for HC vs EG in the EC state.

achieved using in the Beta band during the EC state for the bipolar pair F4C4-FZCZ. It can be observed that overall, the Beta band has the best performance for EO, EC, and EO & EC, while the Gamma band has the worst performance consistently, which confirms the observations in Figure 4-4. Besides, the evaluation of other machine learning methods and K selection refer to KNN supported by ROC and AUC, shown in Figure 4-17 and Figure 4-18 in Supplementary material. To further understand how the selected features contribute to the classification, Figure 4-7 plots the clustering of

in the Beta band of F4C4-FZCZ, F3C3-FzCz, CzPz-C4P4 and C3Cz-P3Pz in the EC state, all of which have high classification accuracy (>90%). It can be observed that, for the HC subjects, the value of coherence is much higher than that of the EG subjects. Besides, it is also observed that the distribution of these features in EG is more concentrated, which may explain the classification sensitivity is always high, almost 100%, but with relatively low specificity. The machine learning has also been implemented using two features. The accuracy increases slightly, but there are still some HC samples misclassified as EG. The results are as shown in Figure 4-18 in Supplementary material.

#### 4.1.3.2 Epilepsy Group vs NEAD

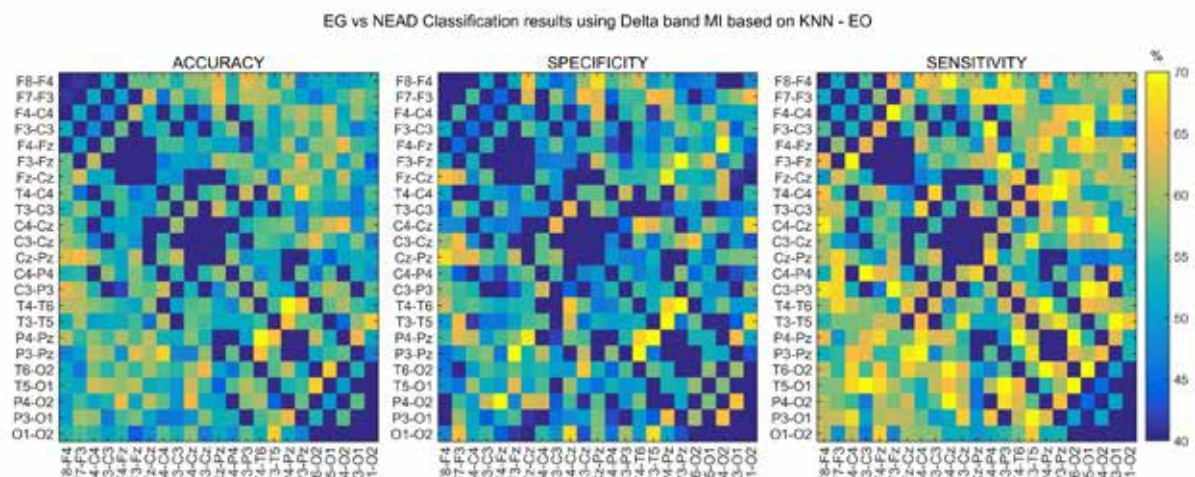


**Figure 4-9** The heatmap for the NEAD vs EG classification performance based on MI in Delta band for the EO state.

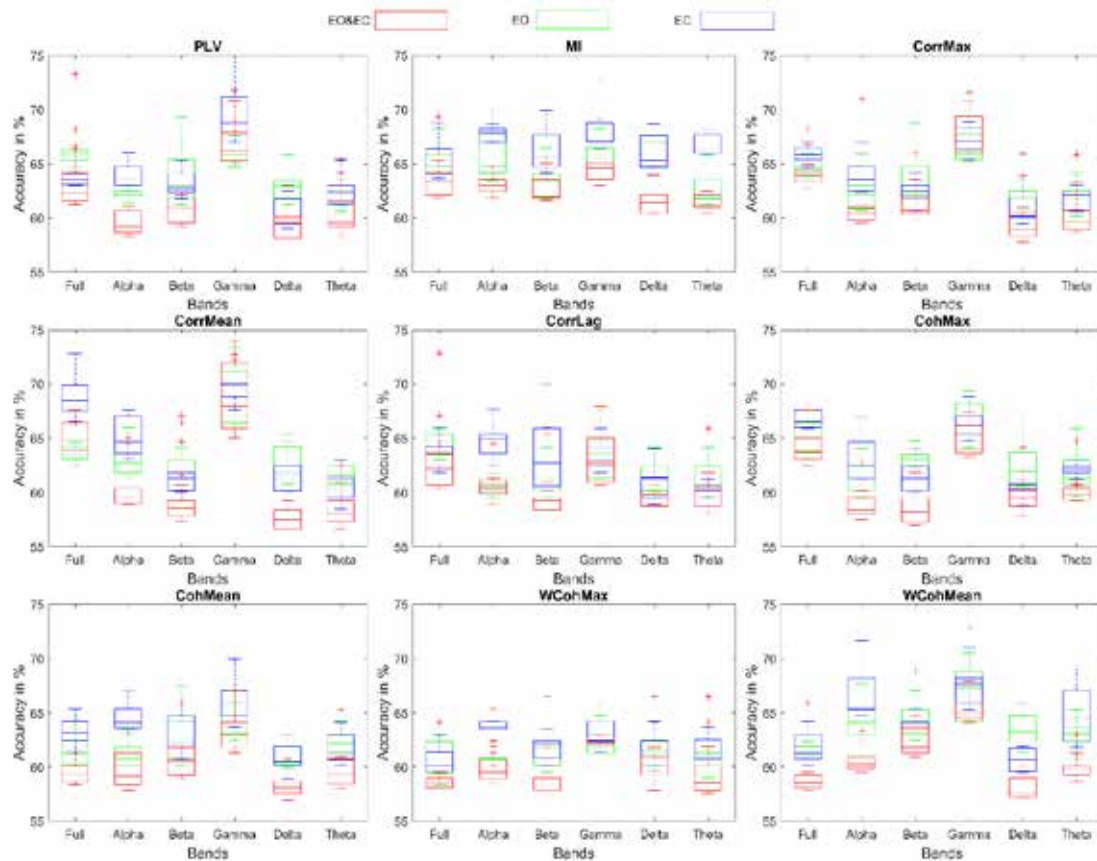


**Figure 4-10 The box charts of the classification accuracy of NEAD and EG against various frequency bands for the 9 selected measures, where only the top 10 pairs in terms of classification accuracy were considered. Each box shows highest, lowest, and median accuracy.**

Figure 4-10 plots the box charts of the classification accuracy of NEAD vs EG against different frequency bands using 9 distinct estimations. It can be observed that the overall accuracy is much lower than that of HC vs EG (about 20% less). Overall, the Gamma band produced a better performance (65-70%) than other 5 bands for all features. Especially for , the Gamma band produces significantly higher accuracy which was witnessed by ANOVA and multiple comparisons ( $p < 0.001$ ).

Results (see Figure 4-15) show that the best classification result (74.44%) was achieved using MI in the Delta band during the EO state for the bipolar pair T4T6-P4PZ. The second-best accuracy (74.24%) was achieved using PLV in the Gamma band during the EO state for the bipolar pair T3C3-CZPZ. Overall,

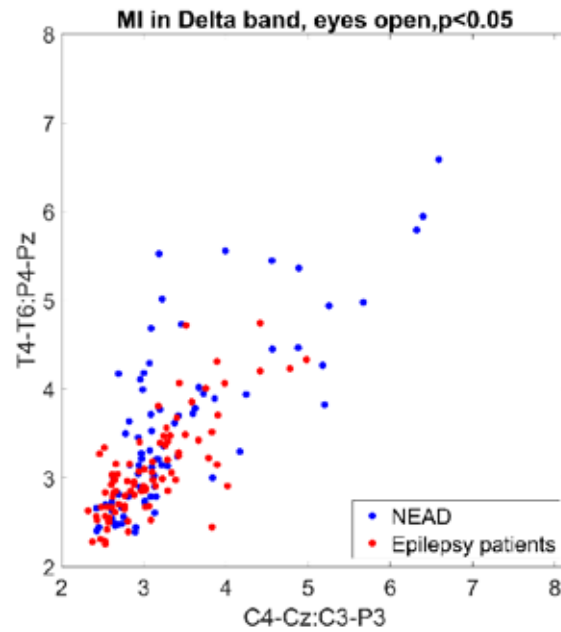


**Figure 4-11** The distribution of MI values in Delta band of two channel pairs corresponding highest classification accuracy (>70%) for the NEAD vs EG for the EO state.

the Gamma band has the best performance across all features and eye states. To further explore how the highest classification was produced, Figure 4-11 presents the scatter plot of MI in the Delta band for T4T6-P4PZ and C4Cz-C3P3. The means of those two bipolar pairs for the EG is slightly lower than those of NEAD. ANOVA suggests that the means of both features for the two groups are significantly different (  $<0.05$ ).

#### 4.1.4 DISCUSSIONS

This Chapter proposes a multivariate approach to investigate the potential of leveraging the association between two channels (i.e., EEG sensor level functional connectivity) to classify different groups. It should be noted that most of the state-of-the-art for this topic use the features extracted from single channels. Durongbhan . (2019) proposed a univariate method that uses the response of five frequency bands of each channel for classification. Similarly, a typical univariate method, power spectral density (PSD), was also applied to EEG recordings and it was found that theta (4-9 Hz) PSD ratio can contribute to evaluating the influence of neurofeedback training for epilepsy patients (Zhao , 2009). Wan (2019) suggested that alpha rhythm (8-12 Hz) PSD



observed in EEG over human posterior cortex differs in distinct groups. Therefore, the univariate approach was applied to our database to critically compare the proposed and the existing method.

The spatial distributions of the frequency response of each channel are shown in Figure 4-19, Figure 4-20, Figure 4-21, Figure 4-22 and Figure 4-23 of Supplementary material, where the subjects are divided into 5 groups: EG with medication, EG without medication, NEAD with medication, NEAD without medication, HC. Figure 4-12 shows the detailed spatial distribution of classification accuracy using the univariate method, where four classification methods were tested. It is

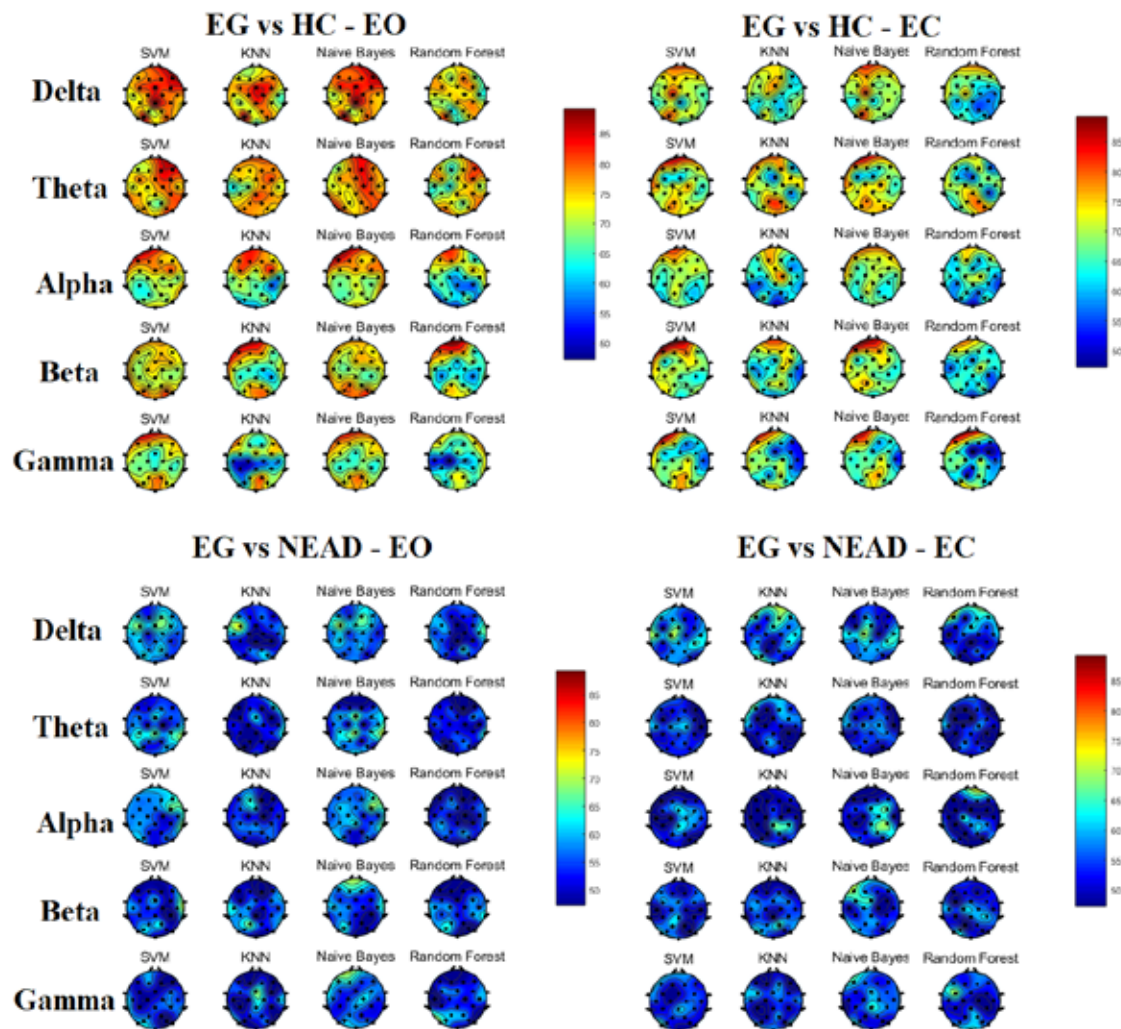


Figure 4-12 The spatial distribution of classification accuracy using the univariate method, where the same colormap scale is used.

**Table 4-2 THE COMPARISON OF THE HIGHEST CLASSIFICATION USING A SINGLE FEATURE BETWEEN THE UNIVARIATE APPROACH AND THE PROPOSED APPROACH**

|            | UNIVARIATE APPROACH |                   | The proposed approach |                   |
|------------|---------------------|-------------------|-----------------------|-------------------|
|            | EO                  | EC                | EO                    | EC                |
| HC vs EG   | 89.60%<br>(Delta)   | 82.80%<br>(Delta) | 94.75%<br>(Beta)      | 97.22%<br>(Beta)  |
| NEAD vs EG | 72.20%<br>(Theta)   | 70.20%<br>(Delta) | 74.44%<br>(Delta)     | 74.24%<br>(Gamma) |

observed that Naïve Bayes has a relatively good performance. Table 4-2 shows the comparison of the best classification result between the univariate approach (Naïve Bayes was used for classification) and the proposed multivariate approach for the same data. For HC vs EG, the proposed method based on multivariate analysis produced higher classification accuracy in comparison to the classic univariate method, particularly for EC, there is 15% improvement. It should be noted that the band with the best performance is different. For NEAD vs EG, the proposed method performs slightly better than the univariate approach for both EO and EC states.

#### **4.1.5 CONCLUSIONS**

In this work, we implement a framework comprising various sensor level functional brain connectivity features, on EEG epochs classified as normal on empirical inspection of interictal traces from a cohort of patients with generalised epilepsy. Despite no interictal abnormalities like IEDs - the empirical diagnostic hallmark of epilepsy - were included in this work, our framework achieved a very high classification accuracy between HC and EG cohorts, which outperforms the state-of-the-art single channel findings. Based on results from scalp EEG sensors, an accuracy of 97% was found for the functional connectivity estimates between the parasagittal frontal and midline regions for EC state and involved the beta frequency band. Unintuitively, as epilepsy in the ictal phase is typically characterised by the transient occurrence

of abnormal, excessive or synchronous neuronal brain activity (Fisher, 2014), our interictal findings show lower levels of synchronisation in this frontal region, in the beta band, in the Epilepsy patients in comparison to the findings from healthy controls (Figure 4-7). This interictal deficit in the Beta band synchronisation in patients with generalised epilepsy was previously shown in a cohort of juvenile myoclonic epilepsy. This in part involved the frontal and parietal brain areas (Clemens, 2013) that also consistently produced high classification accuracy in our work. However, our results have to be interpreted with caution as previous univariate EEG data analysis showed some spectral density differences in various phenotypes of generalised epilepsy syndromes (Clemens, 2012), namely patients with absence seizures versus juvenile myoclonic epilepsy and epilepsy with generalised tonic-clonic seizures in isolation. The importance of this type of brain network connectivity approach was also shown from another angle in a study where decoupling of functional and structural connectivity, based on fMRI and tractography estimates was demonstrated for a cohort of patients with idiopathic generalised epilepsy in comparison to age-matched healthy controls (Zhang, 2011). More work is required in well-characterised large cohorts of patients with different forms of generalised and focal epilepsies, ideally in drug naïve studies although clinical decisions render this latter requirement difficult to achieve.

Our findings demonstrate that seizure-free EEG recordings contain invisible information that can be revealed with appropriate methodology to identify patients with generalised epilepsy. The low sensitivity of a single EEG, at most at 55% (Pillai and Sperling, 2006), has been a long-standing diagnostic problem in clinical neurophysiology and translates to a high financial cost (several EEGs are frequently required to record IEDs) can result in time delay to treatment initiation and can also produce high levels of anxiety to patients, carers and relatives. Our work demonstrates that significant coherence data are confounded by power and phase effects, outside the observational capabilities of reporting physicians, can be successfully used to reveal spatiotemporal deficits in brain network organisation and behaviours that could be translated in clinical useful diagnostic tools. EEG epochs during periods of EO and EC differ



significantly, reflecting the dynamic brain network changes between the two conditions. The discrepancy in the classification accuracy between EO and EC for HC vs the EG, based on MI, a technique able to capture nonlinear connectivity, demonstrates that advantages of different methods might have to be implemented in conjunction in future work for the development of a robust diagnostic framework. The results of the classification accuracy between EG and NEAD also demonstrates that there cannot be a one size fits all approach as the findings differ significantly when compared to the EG vs HC results. Specifically, the highest classification accuracies were achieved for synchronisation estimates within different frequency ranges (delta for EO and gamma for EC), involving different brain areas -the parietal and temporal regions - while additionally they were based on the two nonlinear methods (MI and PLV) implemented in this work. The roughly 20% lower classification accuracy (about 73%) between EG and NEAD is very reassuring, as not infrequently NEAD patients can also be experiencing epileptic seizures (Milán-Tomás, 2018). This potential coexistence of the two conditions blurs the boundaries of “gold standard” labelling between the two, in keeping with the lower classification accuracy found in this work. Several EG and NEAD patients included in this work were receiving various AEDs which could influence the EEG recordings, extensively described in previous work (Höller, Helmstaedter and Lehnertz, 2018). However, 9 of our patients (5 from EG and 4 from NEAD) were on no medication. Therefore, it is highly unlikely that the remarkable classification accuracy of this Chapter, particularly between HC and EG, was due to the effect of medication on EEG.

#### **4.1.6 Supplementary**

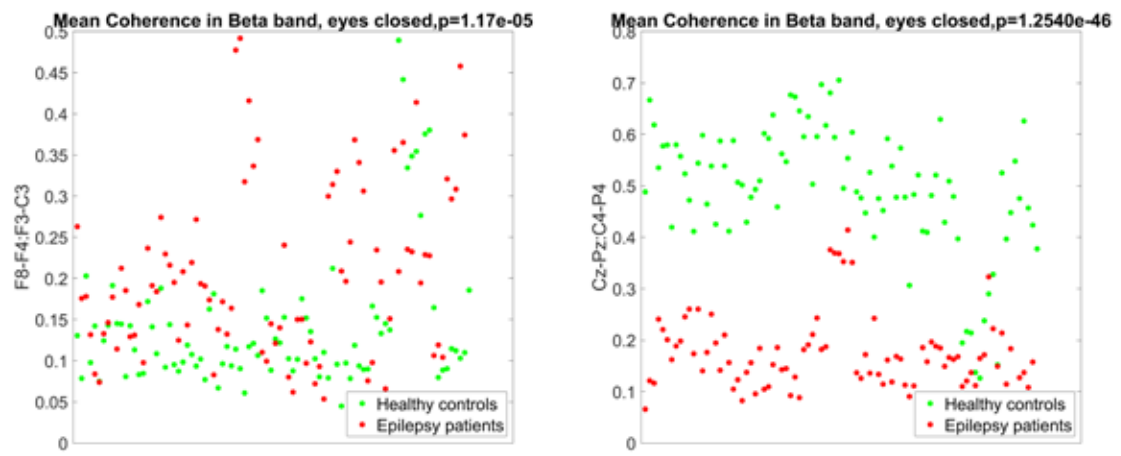


**Figure 4-13 The distribution of Mean coherence in the Beta band of two channel pairs with different p-value**

**Table 4-3 THE DETAILS OF PARTICIPANTS AND AVAILABLE DATA**

| Group    | Gender | Age | Medications                                  | Segment |    | Total |    |
|----------|--------|-----|----------------------------------------------|---------|----|-------|----|
|          |        |     |                                              | EO      | EC | EO    | EC |
| Epilepsy | F      | 39  | Diazepam, Propranolol                        | 6       | 6  | 90    | 90 |
|          | F      | 40  | Fluoxetine                                   | 6       | 6  |       |    |
|          | F      | 45  | Lamotrigine, Lacosamide                      | 6       | 6  |       |    |
|          | F      | 58  | Sodium valproate, Levetiracetam              | 6       | 6  |       |    |
|          | F      | 29  | Nil                                          | 6       | 6  |       |    |
|          | F      | 29  | Carbamazepine                                | 6       | 6  |       |    |
|          | F      | 29  | Nil                                          | 6       | 6  |       |    |
|          | F      | 23  | Lamotrigine                                  | 6       | 6  |       |    |
|          | F      | 20  | Nil                                          | 6       | 6  |       |    |
|          | F      | 19  | Lamotrigine                                  | 6       | 6  |       |    |
|          | M      | 46  | Lamotrigine, Levetiracetam                   | 6       | 6  |       |    |
|          | M      | 45  | Lamotrigine, Sodium valproate                | 6       | 6  |       |    |
|          | M      | 36  | Phenobarbital, Lamotrigine, Pregabalin       | 6       | 6  |       |    |
|          | M      | 28  | Nil                                          | 6       | 6  |       |    |
|          | M      | 19  | Nil                                          | 6       | 6  |       |    |
| Healthy  | F      | 41  | Nil                                          | 9       | 9  | 90    | 90 |
|          | F      | 52  | Nil                                          | 9       | 9  |       |    |
|          | F      | 61  | Nil                                          | 9       | 9  |       |    |
|          | F      | 23  | Nil                                          | 9       | 9  |       |    |
|          | F      | 23  | Nil                                          | 9       | 9  |       |    |
|          | F      | 22  | Nil                                          | 9       | 9  |       |    |
|          | M      | 49  | Nil                                          | 9       | 9  |       |    |
|          | M      | 48  | Nil                                          | 9       | 9  |       |    |
|          | M      | 30  | Nil                                          | 9       | 9  |       |    |
|          | M      | 22  | Nil                                          | 9       | 9  |       |    |
|          | F      | 39  | Nil                                          | 6       | 6  | 84    | 84 |
| NEAD     | F      | 42  | Phenytoin, Levetiracetam                     | 6       | 6  |       |    |
|          | F      | 49  | Perampanel, Clobazam                         | 6       | 6  |       |    |
|          | F      | 49  | Antidepressants, SSRI                        | 6       | 6  |       |    |
|          | F      | 55  | Nil                                          | 6       | 6  |       |    |
|          | F      | 29  | Nil                                          | 6       | 6  |       |    |
|          | F      | 28  | Sertraline, Amitriptyline                    | 6       | 6  |       |    |
|          | F      | 29  | Nil                                          | 6       | 6  |       |    |
|          | F      | 23  | Sertraline, Co-codamol, Antihistamine        | 6       | 6  |       |    |
|          | F      | 24  | Propranolol, Levothyroxine                   | 6       | 6  |       |    |
|          | F      | 19  | Levetiracetam, Pregabalin                    | 6       | 6  |       |    |
|          | M      | 55  | Epilim, Setraline, Gabapentin, Levetiracetam | 6       | 6  |       |    |
|          | M      | 34  | Pregabalin, Lamotrigine, SSRI                | 6       | 6  |       |    |
|          | M      | 23  | Nil                                          | 6       | 6  |       |    |
|          | M      | 19  | Lamotrigine, Amitriptyline                   | 6       | 6  |       |    |

**Table 4-4 THE LIST OF 23 BIPOLAR CHANNELS**

|         |         |         |         |         |
|---------|---------|---------|---------|---------|
| 'F8-F4' | 'F7-F3' | 'F4-C4' | 'F3-C3' | 'F4-Fz' |
| 'F3-Fz' | 'Fz-Cz' | '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' | 'O1-O2' |         |         |

| Eyes state                | Feature  | Band   |        |        |        |        |        |
|---------------------------|----------|--------|--------|--------|--------|--------|--------|
|                           |          | Full   | Alpha  | Beta   | Gamma  | Delta  | Theta  |
| Eyes open                 | PLV      | 88.97% | 88.22% | 93.84% | 78.14% | 92.07% | 92.61% |
|                           | MI       | 70.75% | 65.30% | 69.03% | 76.99% | 88.04% | 66.04% |
|                           | CorrMax  | 94.41% | 87.82% | 92.86% | 84.01% | 82.01% | 90.15% |
|                           | CorrMean | 74.45% | 67.79% | 77.30% | 79.49% | 83.00% | 78.20% |
|                           | CorrLag  | 87.13% | 87.46% | 89.61% | 81.02% | 80.67% | 88.29% |
|                           | MaxCoh   | 83.90% | 86.26% | 91.62% | 80.76% | 92.71% | 88.68% |
|                           | MeanCoh  | 79.10% | 88.27% | 93.74% | 83.78% | 92.28% | 90.84% |
|                           | MaxWCoh  | 63.67% | 81.69% | 82.23% | 73.62% | 91.39% | 86.25% |
|                           | MeanWCoh | 78.80% | 91.51% | 94.75% | 80.70% | 90.60% | 92.52% |
| Eyes closed               | PLV      | 86.82% | 85.14% | 94.42% | 79.31% | 84.83% | 89.95% |
|                           | MI       | 88.08% | 86.36% | 75.32% | 86.24% | 97.11% | 81.70% |
|                           | CorrMax  | 92.72% | 88.48% | 94.40% | 80.47% | 80.79% | 88.73% |
|                           | CorrMean | 80.01% | 76.88% | 76.99% | 79.54% | 84.13% | 68.37% |
|                           | CorrLag  | 88.17% | 90.50% | 90.61% | 80.53% | 77.01% | 87.15% |
|                           | MaxCoh   | 83.66% | 83.92% | 91.62% | 81.27% | 87.35% | 90.73% |
|                           | MeanCoh  | 80.37% | 88.60% | 97.22% | 91.26% | 89.13% | 92.64% |
|                           | MaxWCoh  | 62.52% | 81.34% | 88.97% | 70.64% | 86.79% | 83.69% |
|                           | MeanWCoh | 78.98% | 83.46% | 93.97% | 85.93% | 85.16% | 93.07% |
| Eyes open and eyes closed | PLV      | 89.61% | 87.21% | 93.73% | 81.50% | 89.56% | 91.79% |
|                           | MI       | 75.14% | 68.27% | 63.54% | 78.33% | 91.79% | 71.68% |
|                           | CorrMax  | 93.07% | 86.92% | 93.30% | 81.35% | 78.83% | 89.76% |
|                           | CorrMean | 71.89% | 65.24% | 77.26% | 76.87% | 81.63% | 73.77% |
|                           | CorrLag  | 84.97% | 86.70% | 91.45% | 79.52% | 80.73% | 86.70% |
|                           | MaxCoh   | 84.36% | 85.48% | 91.50% | 85.08% | 90.11% | 90.06% |
|                           | MeanCoh  | 76.03% | 87.15% | 95.02% | 83.62% | 90.33% | 91.86% |
|                           | MaxWCoh  | 58.56% | 79.55% | 82.19% | 71.07% | 89.50% | 82.23% |
|                           | MeanWCoh | 78.82% | 84.31% | 92.91% | 85.64% | 87.36% | 93.29% |

**Figure 4-14 The best classification accuracy of HC vs EG using F4C4-FZCZ for each connectivity feature, frequency band and eye state.**

| Eyes state                | Feature  | Band   |        |        |        |        |        |
|---------------------------|----------|--------|--------|--------|--------|--------|--------|
|                           |          | Full   | Alpha  | Beta   | Gamma  | Delta  | Theta  |
| Eyes open                 | PLV      | 73.08% | 65.12% | 67.78% | 74.24% | 65.25% | 65.42% |
|                           | MI       | 68.03% | 68.93% | 66.86% | 68.14% | 74.44% | 64.78% |
|                           | CorrMax  | 66.75% | 65.88% | 68.55% | 71.72% | 64.62% | 65.07% |
|                           | CorrMean | 66.43% | 65.78% | 65.18% | 71.04% | 65.25% | 64.75% |
|                           | CorrLag  | 71.95% | 63.85% | 64.67% | 67.02% | 68.56% | 64.39% |
|                           | MaxCoh   | 67.65% | 64.01% | 66.29% | 69.27% | 66.72% | 67.23% |
|                           | MeanCoh  | 63.86% | 65.35% | 67.78% | 67.90% | 64.48% | 64.44% |
|                           | MaxWCoh  | 62.44% | 65.43% | 64.07% | 64.00% | 65.35% | 63.73% |
| Eyes closed               | MeanWCoh | 64.44% | 66.97% | 66.03% | 70.90% | 68.09% | 64.84% |
|                           | PLV      | 66.92% | 66.24% | 65.97% | 74.24% | 63.45% | 67.54% |
|                           | MI       | 67.61% | 70.54% | 69.48% | 73.61% | 68.33% | 68.36% |
|                           | CorrMax  | 68.72% | 67.54% | 67.92% | 71.05% | 65.14% | 64.21% |
|                           | CorrMean | 72.93% | 67.21% | 65.95% | 72.90% | 62.54% | 62.49% |
|                           | CorrLag  | 67.65% | 67.92% | 68.82% | 69.31% | 65.23% | 64.31% |
|                           | MaxCoh   | 70.50% | 68.18% | 64.16% | 71.26% | 64.14% | 62.81% |
|                           | MeanCoh  | 66.21% | 65.96% | 66.90% | 70.15% | 63.91% | 65.08% |
| Eyes open and eyes closed | MaxWCoh  | 63.45% | 67.05% | 70.33% | 66.51% | 64.31% | 64.90% |
|                           | MeanWCoh | 63.96% | 72.22% | 69.13% | 73.37% | 63.33% | 67.90% |
|                           | PLV      | 66.81% | 63.16% | 64.68% | 71.29% | 61.63% | 61.61% |
|                           | MI       | 64.75% | 64.06% | 64.49% | 66.31% | 65.76% | 63.50% |
|                           | CorrMax  | 66.70% | 62.03% | 63.53% | 70.87% | 62.83% | 63.66% |
|                           | CorrMean | 68.56% | 64.27% | 59.88% | 73.59% | 59.74% | 60.64% |
|                           | CorrLag  | 65.05% | 62.25% | 64.97% | 67.28% | 61.51% | 61.88% |
|                           | MaxCoh   | 67.96% | 63.70% | 61.38% | 67.62% | 60.16% | 60.72% |
|                           | MeanCoh  | 61.85% | 62.68% | 65.83% | 66.55% | 61.68% | 65.63% |
|                           | MaxWCoh  | 60.61% | 62.82% | 61.33% | 64.27% | 61.60% | 60.75% |
|                           | MeanWCoh | 61.27% | 64.51% | 64.94% | 71.96% | 61.15% | 63.23% |

Figure 4-15 The best classification accuracy of NEAD vs EG for each connectivity feature, frequency band and eye state.

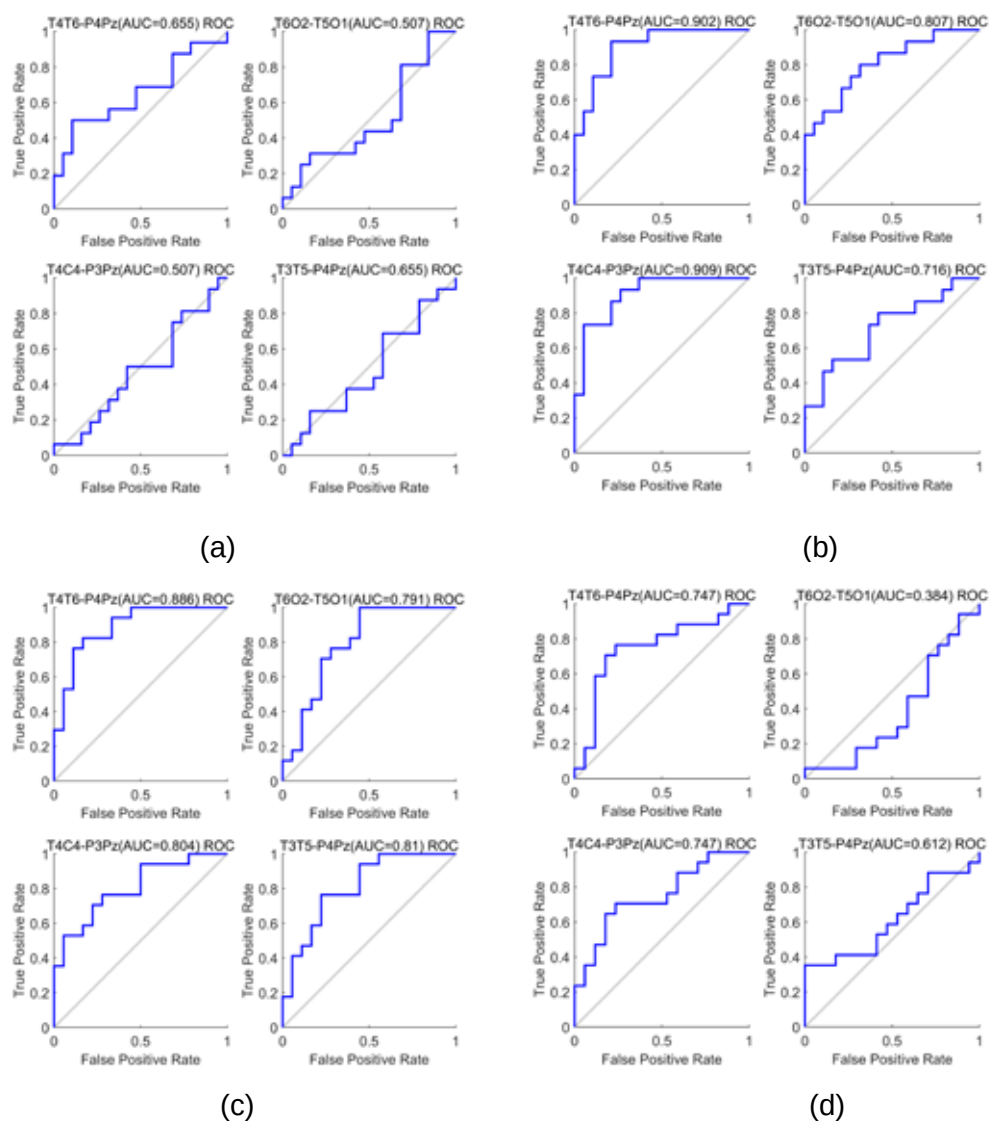


**Figure 4-16 The evaluation of best classifier using features providing highest accuracy classification between NEAD and EG (a) SVM (b) optimized SVM (c) KNN, k=15 (d) decision tree**

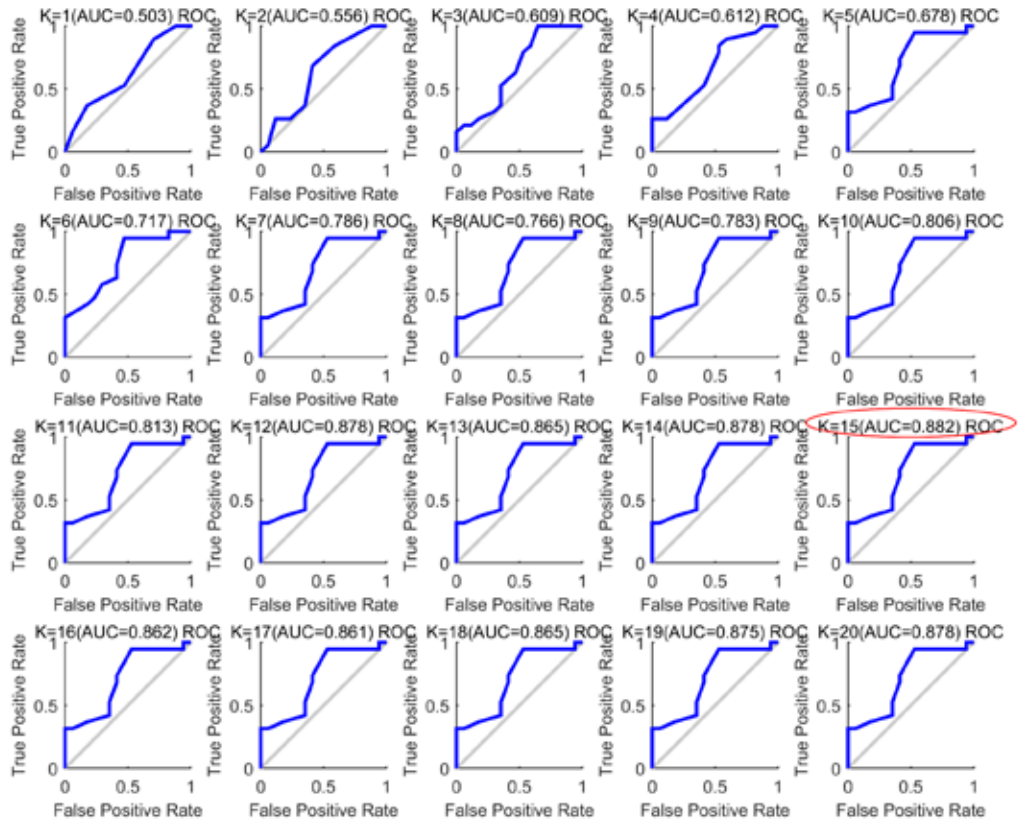


Figure 4-17 The evaluation of the best K value of KNN based on ROC and AUC

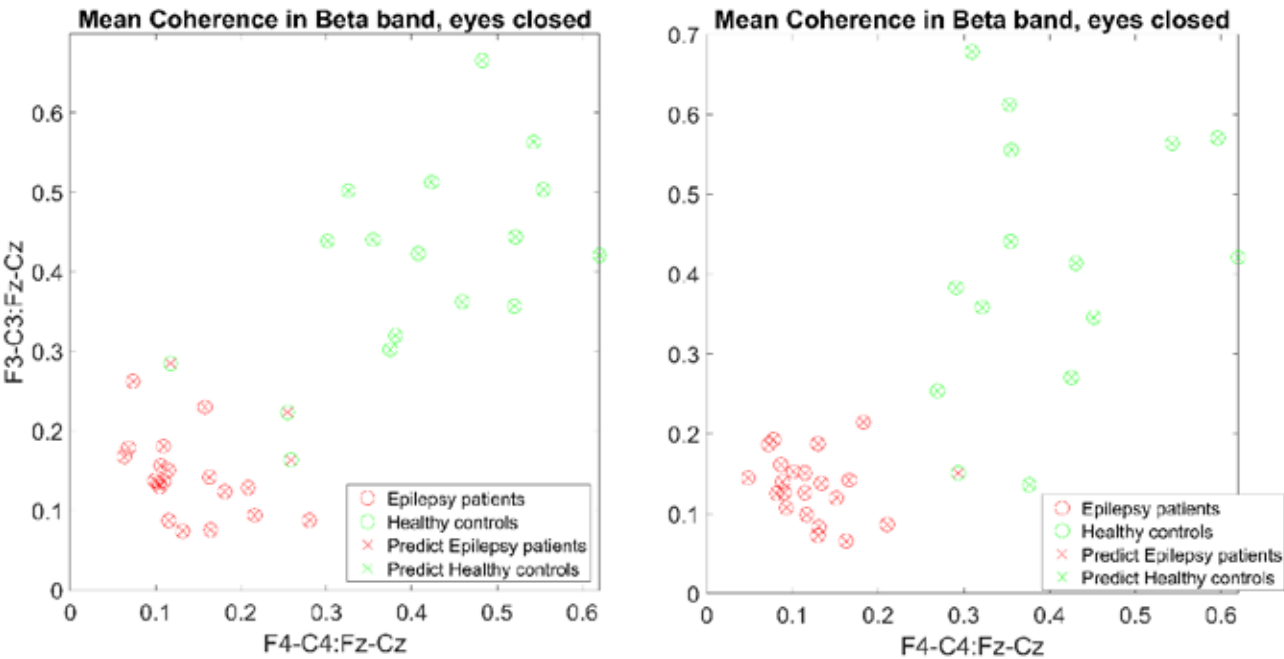


Figure 4-18 The classification results using two features for EG vs HC.



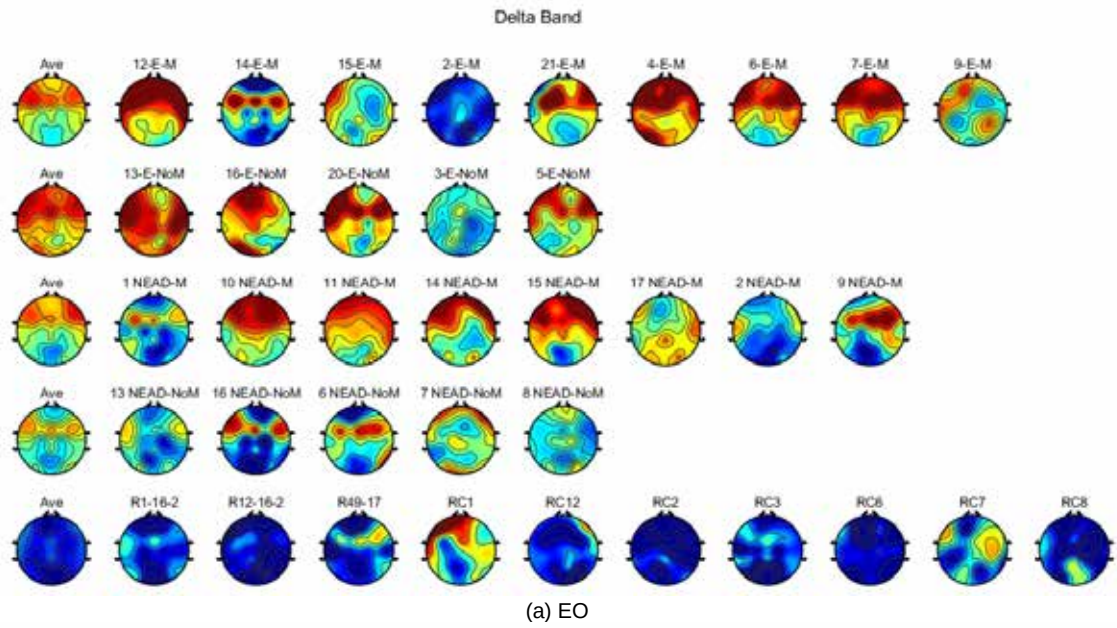
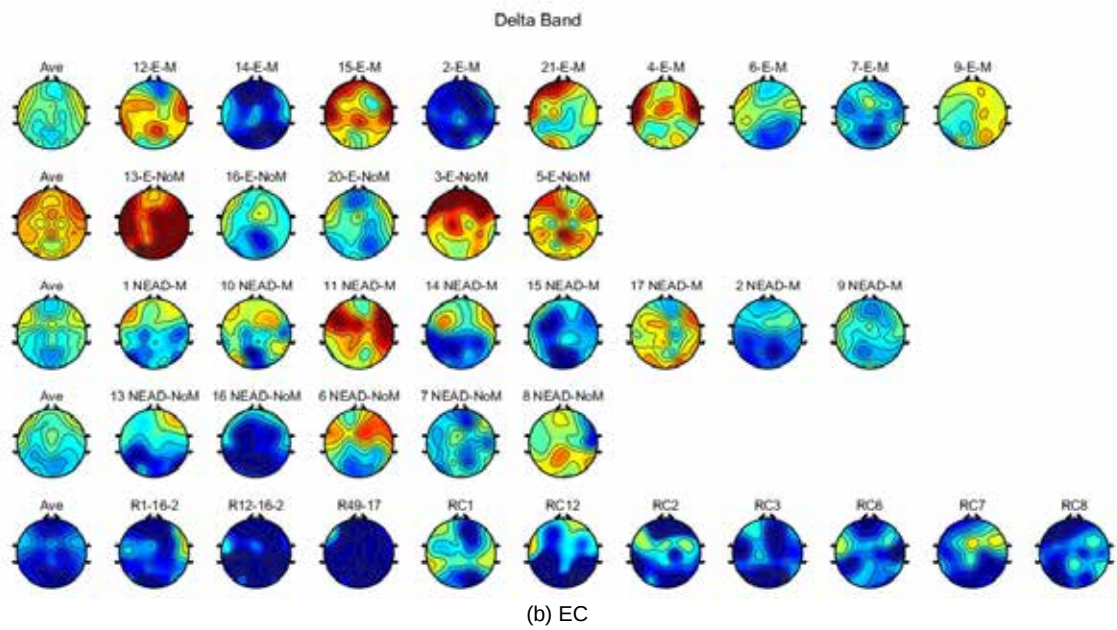
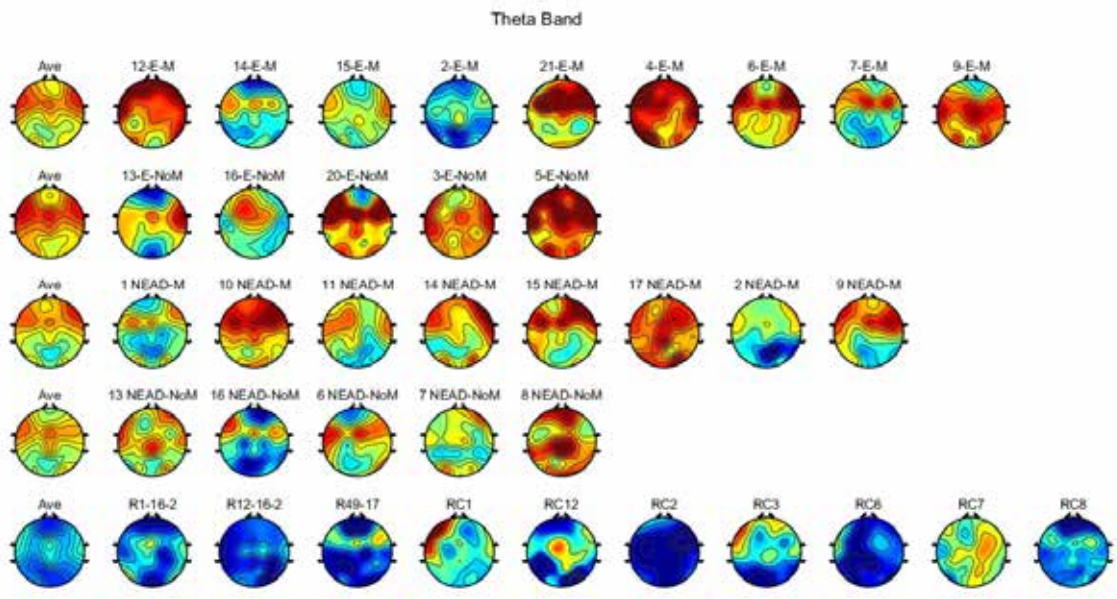
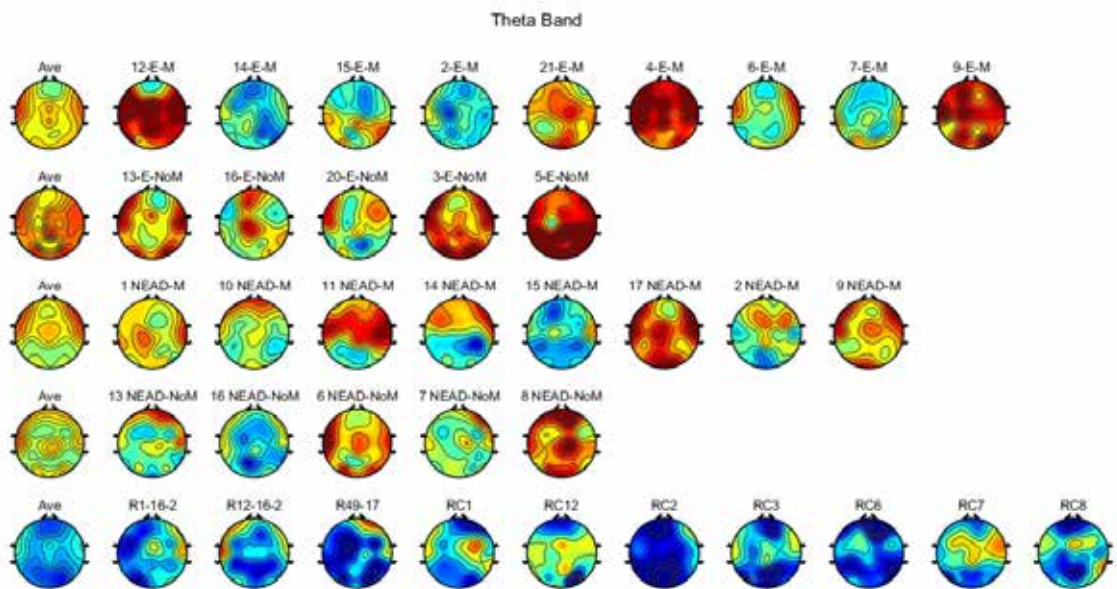


**Figure 4-19** Spatial distribution of frequency response of each channel for Delta band, where the subjects are divided into 5 groups: EG with medication (row 1), EG without medication (row 2), NEAD with medication (row 3), NEAD without medication (row 4), HC (row 5). The same colormap scale is used across all plots.





(a) EO



(b) EC

**Figure 4-20 Spatial distribution of frequency response of each channel for Theta band, where the subjects are divided into 5 groups: EG with medication (row 1), EG without medication (row 2), NEAD with medication (row 3), NEAD without medication (row 4), HC (row 5). The same colormap scale is used across all plots**

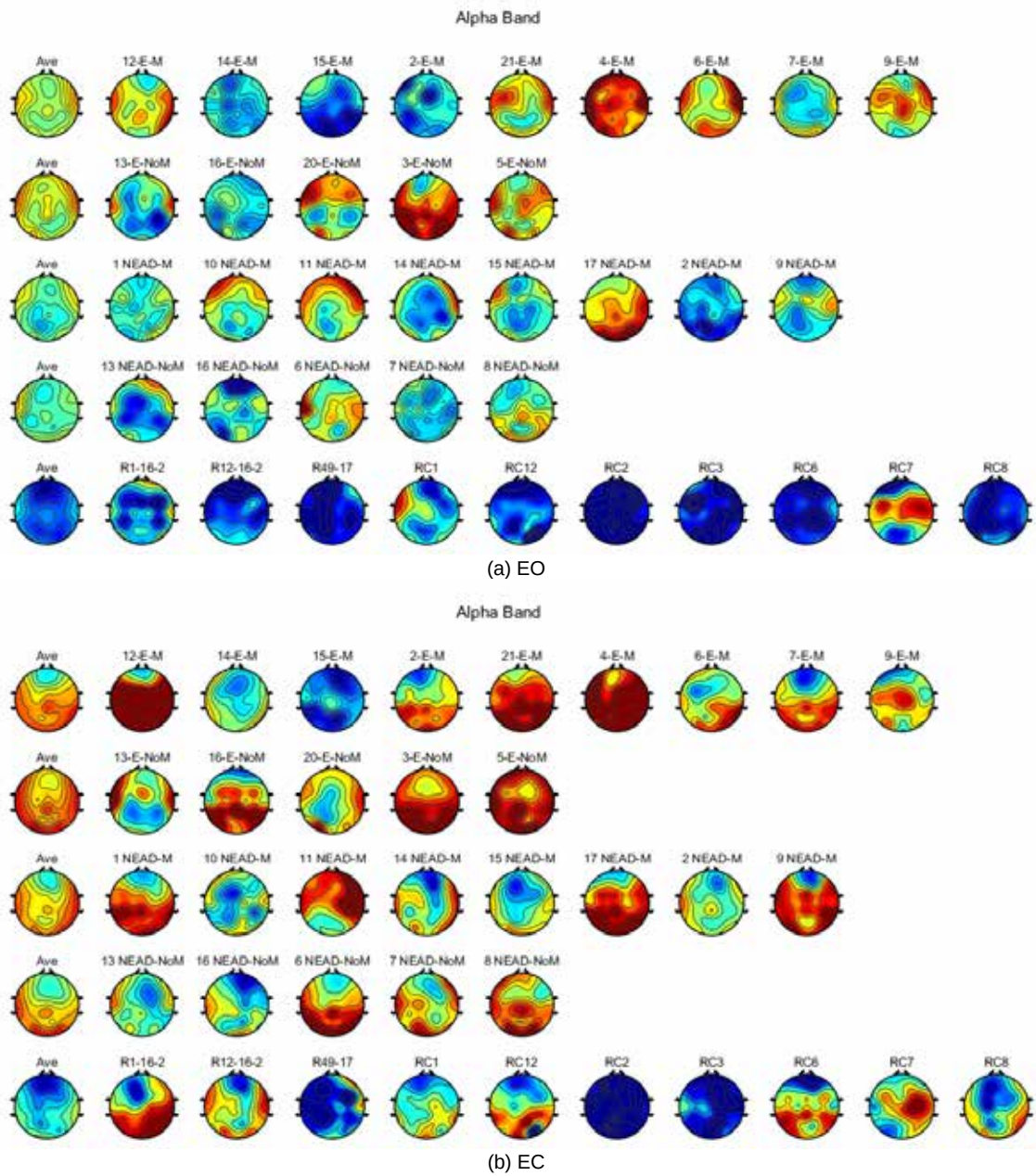


**Figure 4-21** Spatial distribution of frequency response of each channel for Alpha band, where the subjects are divided into 5 groups: EG with medication (row 1), EG without medication (row 2), NEAD with medication (row 3), NEAD without medication (row 4), HC (row 5). The same colormap scale is used across all plots.



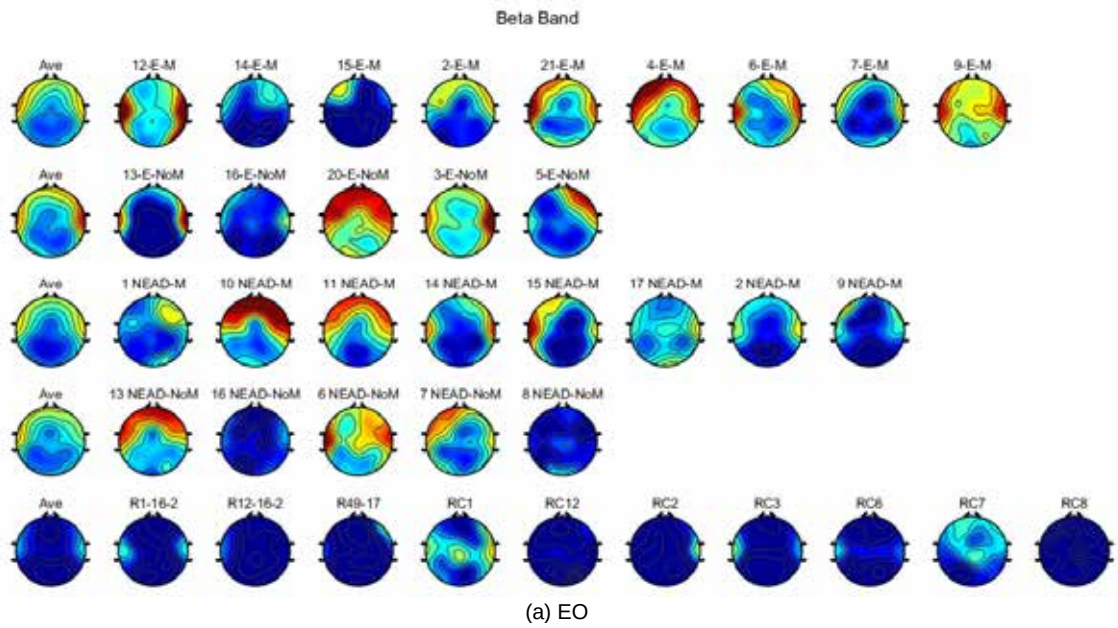
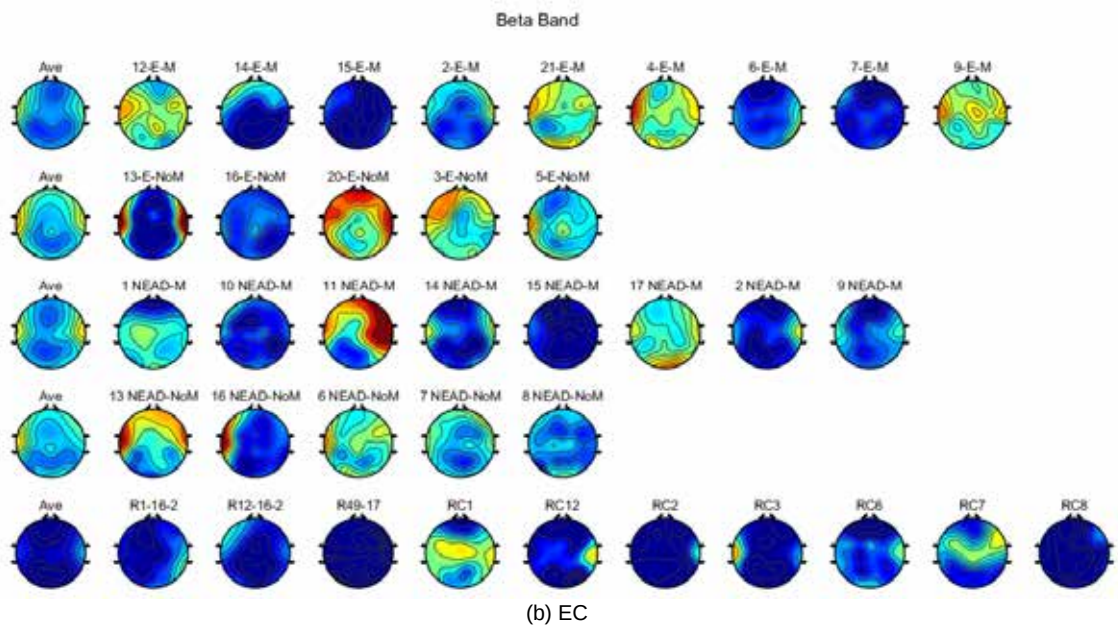


**Figure 4-22 Spatial distribution of frequency response of each channel for Beta band, where the subjects are divided into 5 groups: EG with medication (row 1), EG without medication (row 2), NEAD with medication (row 3), NEAD without medication (row 4), HC (row 5). The same colormap scale is used across all plots.**

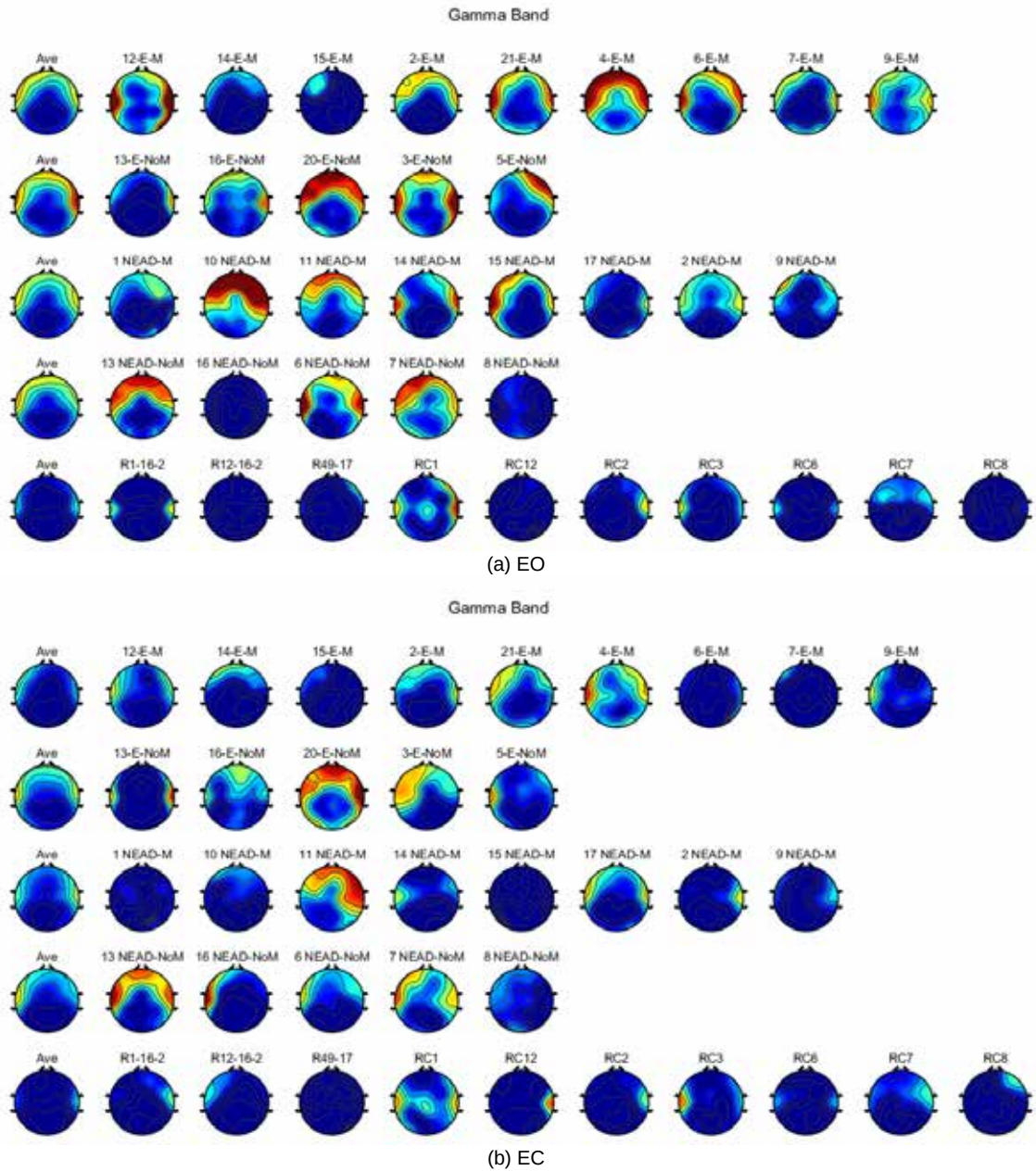


**Figure 4-23** Spatial distribution of frequency response of each channel for Gamma band, where the subjects are divided into 5 groups: EG with medication (row 1), EG without medication (row 2), NEAD with medication (row 3), NEAD without medication (row 4), HC (row 5). The same colormap scale is used across all plots.

#### 4.1.7 REFERENCES

Akbarian, B., Erfanian, A., 2020. Biomedical Signal Processing and Control A framework for seizure detection using effective connectivity , graph theory , and multi-level modular network. *Biomed. Signal Process. Control* 59, 101878. <https://doi.org/10.1016/j.bspc.2020.101878>

Amin, H.U., Yusoff, M.Z., Ahmad, R.F., 2020. A novel approach based on wavelet analysis and arithmetic coding for automated detection and diagnosis of epileptic seizure in EEG signals using machine learning techniques. *Biomed. Signal Process. Control* 56. <https://doi.org/10.1016/j.bspc.2019.101707>

insights from computational methods. *Cogn. Neurosci.*

Bedo, N., Ribary, U., Ward, L.M., 2020. Fast Dynamics of Cortical Functional and Effective Connectivity during Word Reading 9, 1–24. <https://doi.org/10.1371/journal.pone.0088940>

Blackburn, D.J., Sarrigiannis, P.G., De Marco, M., Zhao, Y., Venneri, A., Lawrence, S., Unwin, Z.C., Blyth, M., Angel, J., Baster, K., Wilkinson, I.D., Bell, S.M., He, F., Wei, H.L., Billings, S.A., Farrow, T.F.D., 2018. 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 Sci.* 8, 1–19. <https://doi.org/10.3390/brainsci8070134>

Blinowska, K.J., Rakowski, F., Kaminski, M., De Vico Fallani, F., Del Percio, C., Lizio, R., Babiloni, C., 2017. Functional and effective brain connectivity for discrimination between Alzheimer’s patients and healthy individuals: A study on resting state EEG rhythms. *Clin. Neurophysiol.* 128, 667–680. <https://doi.org/10.1016/j.clinph.2016.10.002>

Brodbeck, V., Kuhn, A., von Wegner, F., Morzelewski, A., Tagliazucchi, E., Borisov, S., Michel, C.M., Laufs, H., 2012. EEG microstates of wakefulness and NREM sleep. *Neuroimage* 62, 2129–2139. <https://doi.org/10.1016/j.neuroimage.2012.05.060>

Brown, R.J., Syed, T.U., Benbadis, S., LaFrance, W.C., Reuber, M., 2011. Psychogenic nonepileptic seizures. *Epilepsy Behav.* 22, 85–93. <https://doi.org/10.1016/j.yebeh.2011.02.016>

Clemens, B., Puskás, S., Besenyei, M., Emri, M., Opposits, G., Kis, S.A., Hollódy, K., Fogarasi, A., Kondákor, I., Füle, K., Bense, K., Fekete, I., 2012. EEG-LORETA endophenotypes of the common idiopathic generalized epilepsy syndromes. *Epilepsy Res.* 99, 281–292. <https://doi.org/10.1016/j.eplepsyres.2011.12.008>

Clemens, B., Puskás, S., Besenyei, M., Spisák, T., Opposits, G., Hollódy, K., Fogarasi, A., Fekete, I., Emri, M., 2013. Neurophysiology of juvenile myoclonic epilepsy: EEG-based network and graph analysis of the interictal and immediate preictal states. *Epilepsy Res.* 106, 357–369. <https://doi.org/10.1016/j.eplepsyres.2013.06.017>

Cover, T.M., Thomas, J.A., 2005. *Elements of Information Theory*, Elements of Information Theory. New York: JOHN WILEY and SONS, INC., 1991. ISBN: 0-471-20061-1. <https://doi.org/10.1002/047174882X>

Dash, G.K., Rathore, C., Jeyaraj, M.K., Wattamwar, P., Sarma, S.P., Radhakrishnan, K., 2018. Interictal regional paroxysmal fast activity on scalp EEG is common in patients with underlying gliosis. *Clin. Neurophysiol.* 129, 946–951. <https://doi.org/10.1016/j.clinph.2018.02.007>

Delgado-Restituto, M., Romaine, J.B., Rodríguez-Vázquez, Á., 2019. Phase Synchronization Operator for On-Chip Brain Functional Connectivity Computation. *IEEE Trans. Biomed. Circuits Syst.* 13, 957–970. <https://doi.org/10.1109/TBCAS.2019.2931799>

Dhiman, R., Saini, J.S., Priyanka, 2014. Genetic algorithms tuned expert model for detection of epileptic seizures from EEG signatures. *Appl. Soft Comput.* 19, 8–17. <https://doi.org/10.1016/j.asoc.2014.01.029>

Durongbhan, P., Zhao, Yifan, Chen, L., Zis, P., De Marco, M., Unwin, Z.C., Venneri, A., He, X., Li, S., Zhao, Yitian, Blackburn, D.J., Sarrigiannis, P.G.,

2019. A Dementia Classification Framework Using Frequency and Time-Frequency Features Based on EEG Signals. *IEEE Trans. Neural Syst. Rehabil. Eng.* 27, 826–835. <https://doi.org/10.1109/TNSRE.2019.2909100>

Erra, R.G., Velazquez, J.L.P., Rosenblum, M., 2017. Neural synchronization from the perspective of non-linear dynamics. *Front. Comput. Neurosci.* 11, 98. <https://doi.org/10.3389/fncom.2017.00098>

Fani, M., Azemi, G., Boashash, B., 2011. EEG-based automatic epilepsy diagnosis using the instantaneous frequency with sub-band energies. 7th Int. Work. Syst. Signal Process. their Appl. WoSSPA 2011 187–190. <https://doi.org/10.1109/WOSSPA.2011.5931447>

Fein, G., Raz, J., Brown, F.F., Merrin, E.L., 1988. Common reference coherence data are confounded by power and phase effects. *Electroencephalogr. Clin. Neurophysiol.* 69, 581–584. [https://doi.org/https://doi.org/10.1016/0013-4694\(88\)90171-X](https://doi.org/https://doi.org/10.1016/0013-4694(88)90171-X)

Fisher, R.S., Acevedo, C., Arzimanoglou, A., Bogacz, A., Cross, J.H., Elger, C.E., Engel, J., Forsgren, L., French, J.A., Glynn, M., Hesdorffer, D.C., Lee, B.I., Mathern, G.W., Moshé, S.L., Perucca, E., Scheffer, I.E., Tomson, T., Watanabe, M., Wiebe, S., 2014. ILAE Official Report: A practical clinical definition of epilepsy. *Epilepsia* 55, 475–482. <https://doi.org/10.1111/epi.12550>

Höller, Y., Helmstaedter, C., Lehnertz, K., 2018. Quantitative Pharmacoelectroencephalography in Antiepileptic Drug Research. *CNS Drugs* 32, 839–848. <https://doi.org/10.1007/s40263-018-0557-x>

Horstmann, M.T., Bialonski, S., Noennig, N., Mai, H., Prusseit, J., Wellmer, J., Hinrichs, H., Lehnertz, K., 2010. State dependent properties of epileptic brain networks: Comparative graph-theoretical analyses of simultaneously recorded EEG and MEG. *Clin. Neurophysiol.* 121, 172–185. <https://doi.org/10.1016/j.clinph.2009.10.013>

classification epileptic EEG signals. *Australas. Phys. Eng. Sci. Med.* 41, 721–730. <https://doi.org/10.1007/s13246-018-0669-0>

Lachaux, J.-P., Rodriguez, E., Martinerie, J., Varela, F., 1999. Measuring Phase Synchrony in Brain Signals. *Hum. Brain Mapp.* <https://doi.org/10.1017/S0007680500048066>

Lioi, G., Bell, S.L., Smith, D.C., Simpson, D.M., 2017. Directional connectivity in the EEG is able to discriminate wakefulness from NREM sleep. *Physiol. Meas.* 38, 1802–1820. <https://doi.org/10.1088/1361-6579/aa81b5>

Lopes, M.A., Perani, S., Yaakub, S.N., Richardson, M.P., Goodfellow, M., Terry, J.R., 2019. Revealing epilepsy type using a computational analysis of interictal EEG. *Sci. Rep.* 9, 1–10. <https://doi.org/10.1038/s41598-019-46633-7>

Lotte, F., Bougrain, L., Cichocki, A., Clerc, M., Congedo, M., Rakotomamonjy, A., Yger, F., 2018. A review of classification algorithms for EEG-based brain-computer interfaces: A 10 year update. *J. Neural Eng.* 15. <https://doi.org/10.1088/1741-2552/aab2f2>

Milán-Tomás, Á., Persyko, M., Del Campo, M., Shapiro, C.M., Farcnik, K., 2018. An Overview of Psychogenic Non-Epileptic Seizures: Etiology, Diagnosis and Management. *Can. J. Neurol. Sci.* <https://doi.org/10.1017/cjn.2017.283>

Noachtar, S., Rémi, J., 2009. The role of EEG in epilepsy: A critical review. *Epilepsy Behav.* 15, 22–33. <https://doi.org/10.1016/j.yebeh.2009.02.035>

Nunez, P.L., Srinivasan, R., Westdorp, A.F., Wijesinghe, R.S., Tucker, D.M., Silberstein, R.B., Cadusch, P.J., 1997. EEG coherency I: Statistics, reference electrode, volume conduction, Laplacians, cortical imaging, and interpretation at multiple scales. *Electroencephalogr. Clin. Neurophysiol.* 103, 499–515. [https://doi.org/10.1016/S0013-4694\(97\)00066-7](https://doi.org/10.1016/S0013-4694(97)00066-7)

Orekhova, E. V., Elsabbagh, M., Jones, E.J., Dawson, G., Charman, T., Baron-Cohen, S., Bedford, R., Bolton, P., Fernandes, J., Ganea, N., Garwood, H., Gliga, T., Hudry, K., Murias, M., Ribeiro, H., Tucker, L., Volein, A., Webb, S.J.,



2014. EEG hyper-connectivity in high-risk infants is associated with later autism. *J. Neurodev. Disord.* 6, 1–11. <https://doi.org/10.1186/1866-1955-6-40>
- Pievani, M., de Haan, W., Wu, T., Seeley, W.W., Frisoni, G.B., 2011. Functional network disruption in the degenerative dementias. *Lancet Neurol.* 10, 829–843. [https://doi.org/10.1016/S1474-4422\(11\)70158-2](https://doi.org/10.1016/S1474-4422(11)70158-2)
- Pillai, J., Sperling, M.R., 2006. Interictal EEG and the Diagnosis of Epilepsy. *Epilepsia* 47, 14–22. <https://doi.org/10.1111/j.1528-1167.2006.00654.x>
- Prasath, V.B.S., Alfeilat, H.A.A., Hassanat, A.B.A., Lasassmeh, O., Tarawneh, A.S., Alhasanat, M.B., Salman, H.S.E., 2017. Distance and Similarity Measures Effect on the Performance of K-Nearest Neighbor Classifier -- A Review 1–39. <https://doi.org/10.1089/big.2018.0175>
- Pyrzowski, J., Sieminski, M., Sarnowska, A., Jedrzejczak, J., Nyka, W.M., 2015. Interval analysis of interictal EEG: Pathology of the alpha rhythm in focal epilepsy. *Sci. Rep.* 5, 1–10. <https://doi.org/10.1038/srep16230>
- Renzel, R., Baumann, C.R., Mothersill, I., Poryazova, R., 2017. Persistent generalized periodic discharges: A specific marker of fatal outcome in cerebral hypoxia. *Clin. Neurophysiol.* 128, 147–152. <https://doi.org/10.1016/j.clinph.2016.10.091>
- Rosch, R., Baldeweg, T., Moeller, F., Baier, G., 2017. Network dynamics in the healthy and epileptic developing brain. *Netw. Neurosci.* 2, 41–59. <https://doi.org/10.1162/netn>
- Sairamya, N.J., Thomas George, S., Balakrishnan, R., Subathra, M.S.P., 2018. An effective approach to classify epileptic EEG signal using local neighbor gradient pattern transformation methods. *Australas. Phys. Eng. Sci. Med.* 41, 1029–1046. <https://doi.org/10.1007/s13246-018-0697-9>
- Sakkalis, V., 2011. Review of advanced techniques for the estimation of brain connectivity measured with EEG/MEG. *Comput. Biol. Med.* 41, 1110–1117. <https://doi.org/10.1016/j.compbiomed.2011.06.020>

- Sakkalis, V., Xanthopoulos, P., Zervakis, M., Yang, Y., Micheloyannis, S., 2006. Assessment of linear and non-linear EEG synchronization measures for evaluating mild epileptic signal patterns. *Test* 13, 26–28.
- Salman, H., Grover, J., Shankar, T., 2018. Electroencephalographic Motor Imagery Brain Connectivity Analysis for BCI: A Review 2733, 2709–2733. <https://doi.org/10.1162/NECO>
- Sarrigiannis, P.G., Zhao, Y., He, F., Billings, S.A., Baster, K., Rittey, C., Yianni, J., Zis, P., Wei, H., Hadjivassiliou, M., Grünwald, R., 2018. The cortical focus in childhood absence epilepsy; evidence from nonlinear analysis of scalp EEG recordings. *Clin. Neurophysiol.* 129, 602–617. <https://doi.org/10.1016/j.clinph.2017.11.029>
- Sarrigiannis, P.G., Zhao, Y., Wei, H.-L., Billings, S.A., Fotheringham, J., Hadjivassiliou, M., 2014. Quantitative EEG analysis using error reduction ratio-causality test; validation on simulated and real EEG data. *Clin. Neurophysiol.* 125, 32–46. <https://doi.org/10.1016/j.clinph.2013.06.012>
- Sheldon, R.J.G., Agrawal, N., 2019. Functional non-epileptic attacks: essential information for psychiatrists. *BJPsych Bull.* 43, 182–187. <https://doi.org/10.1192/bjb.2019.34>
- Smith, S.J.M., 2005. EEG in the diagnosis, classification, and management of patients with epilepsy. *J. Neurol. Neurosurg. Psychiatry* 76, ii2–ii7. <https://doi.org/10.1136/jnnp.2005.069245>
- Tafreshi, T.F., Daliri, M.R., Ghodousi, M., 2019. Functional and effective connectivity based features of EEG signals for object recognition. *Cogn. Neurodyn.* 13, 555–566. <https://doi.org/10.1007/s11571-019-09556-7>
- Uhlhaas, P.J., Singer, W., 2006. Neural Synchrony in Brain Disorders: Relevance for Cognitive Dysfunctions and Pathophysiology. *Neuron* 52, 155–168. <https://doi.org/10.1016/j.neuron.2006.09.020>

- Varela, F., Lachaux, J.-P., Rodriguez, E., Martinerie, J., 2001. The brainweb: Phase synchronization and large-scale integration. *Nat. Rev. Neurosci.* 2, 229–239. <https://doi.org/10.1038/35067550>
- Vecchio, F., Miraglia, F., Quaranta, D., Granata, G., Romanello, R., Marra, C., Bramanti, P., Rossini, P.M., 2016. Cortical connectivity and memory performance in cognitive decline: A study via graph theory from EEG data. *Neuroscience* 316, 143–150. <https://doi.org/10.1016/j.neuroscience.2015.12.036>
- Vijay Anand, S., Shantha Selvakumari, R., 2019. Noninvasive method of epileptic detection using DWT and generalized regression neural network. *Soft Comput.* 23, 2645–2653. <https://doi.org/10.1007/s00500-018-3630-y>
- Visani, E., Varotto, G., Binelli, S., Fratello, L., Franceschetti, S., Avanzini, G., Panzica, F., 2010. Photosensitive epilepsy: Spectral and coherence analyses of EEG using 14 Hz intermittent photic stimulation. *Clin. Neurophysiol.* 121, 318–324. <https://doi.org/10.1016/j.clinph.2009.12.003>
- Wan, L., Huang, H., Schwab, N., Tanner, J., Rajan, A., Lam, N.B., Zaborszky, L., Li, C. shan R., Price, C.C., Ding, M., 2019. From eyes-closed to eyes-open: Role of cholinergic projections in EC-to-EO alpha reactivity revealed by combining EEG and MRI. *Hum. Brain Mapp.* 40, 566–577. <https://doi.org/10.1002/hbm.24395>
- Watanabe, H., Terada, K., Suzuki, N., Ishisaka, M., Naitoh, Y., Ishihara, R., Shimoeda, H., Konagaya, T., Inoue, Y., 2018. Effect of hyperventilation on seizures and EEG findings during routine EEG. *Clin. Neurophysiol.* 129, e38. <https://doi.org/10.1016/j.clinph.2018.02.097>
- Wendling, F., Chauvel, P., Biraben, A., Bartolomei, F., 2010. From intracerebral EEG signals to brain connectivity: Identification of epileptogenic networks in partial epilepsy. *Front. Syst. Neurosci.* 4, 1–13. <https://doi.org/10.3389/fnsys.2010.00154>

Xie, S., Krishnan, S., 2013. Wavelet-based sparse functional linear model with applications to EEGs seizure detection and epilepsy diagnosis. *Med. Biol. Eng. Comput.* 51, 49–60. <https://doi.org/10.1007/s11517-012-0967-8>

Zhang, Z., Liao, W., Chen, H., Mantini, D., Ding, J.-R., Xu, Q., Wang, Z., Yuan, C., Chen, G., Jiao, Q., Lu, G., 2011. Altered functional–structural coupling of large-scale brain networks in idiopathic generalized epilepsy. *Brain* 134, 2912–2928. <https://doi.org/10.1093/brain/awr223>

Zhao, L., Wu, W., Liang, Z., Hu, G., 2009. Changes in EEG measurements in intractable epilepsy patients with neurofeedback training. *Prog. Nat. Sci.* 19, 1509–1514. <https://doi.org/10.1016/j.pnsc.2009.03.010>

Zhao, Yifan, Laguna, R.C., Zhao, Yitian, Liu, J.J., He, X., Yianni, J., Sarrigiannis, P.G., 2018. A Wavelet-Based Correlation Analysis Framework to Study Cerebromuscular Activity in Essential Tremor. *Complexity* 2018, 1–15. <https://doi.org/10.1155/2018/7269494>

Zhu, G., Li, Y., Wen, P.P., 2014. Epileptic seizure detection in EEGs signals using a fast weighted horizontal visibility algorithm. *Comput. Methods Programs Biomed.* 115, 64–75. <https://doi.org/10.1016/j.cmpb.2014.04.001>

## **4.2 Spatial – temporal graph convolutional network for Alzheimer classification based on brain functional connectivity imaging of electroencephalogram.**

### **4.2.1 Introduction**

Alzheimer's disease (AD) is the most common form of dementia, resulting in the loss of memory and cognitive impairments. Most commonly it occurs among elderly people over the age of 65, but it may also occur earlier (Brookmeyer et al., 1998, 2007). According to the World Health Organization, more than 55 million people are currently diagnosed with some type of dementia (Organization). Due to increasing global ageing, the technologies to diagnose AD more effectively and accurately are highly demanded. Numerous studies have shown that synaptic dysfunction is an early feature of AD and that

decreased synaptic density in the neocortex and limbic regions could account for AD-associated disturbances in brain function (Terry et al., 1991; Masliah et al., 2001; Scheff et al., 2006, 2007). Being more extensive than the corresponding neuronal loss when analysed in the same brain regions, synaptic dysfunction has been proved to be the best neuropathological correlate of cognitive impairment in AD patients (Davies et al., 1987; DeKosky and Scheff, 1990; Terry et al., 1991; Scheff et al., 2007). The pathological progression of AD leads to cortical disconnections and manifests as functional connectivity alterations (Nobukawa et al., 2020). Electroencephalography (EEG) is a non-invasive diagnostic method for studying the bioelectrical function alterations and degeneration of the brain. Consisting of scalp electric potential differences, EEG is one of the first measurements that directly reflect the functioning of synapses in real-time (Jelic, 2005; Michel et al., 2009). In contrast to functional MRI or PET which detect indirect metabolic signals, EEG offers several additional attractions: noninvasiveness, high time-resolution, wide availability, low cost, and direct access to neuronal signalling (Michel et al., 2009; Smailovic et al., 2018).

For the diagnosis of AD based on EEG, there are two main development areas in recent years. The first area is statistically topological brain function characteristics. Finding statistical biomarkers of AD is based on analysing graph topographic or dynamic patterns. For graph topography, Jalili (Jalili, 2017) constructed functional brain networks and relevant graph theory metrics based on EEG for discriminating AD from Healthy Controls (HC). Similarly, Tylová (Tylová et al., 2018) proposed the permutation entropy for measuring the chaotic behaviour of EEG and observed a statistically significant decrease in permutation entropy at all channels of AD. Fan (Fan et al., 2018) employed multiscale entropy as the biomarker to characterise the nonlinear complexity at multiple temporal scales to capture the topographic pattern of AD. For dynamic patterns, Zhao (Zhao et al., 2019) proposed a method to measure nonlinear dynamics of functional connectivity for distinguishing between AD and HC, and Tait (Tait et al., 2020) developed a biomarker by combining microstate transitioning complexity and the spectral measure. The

second area is data-driven machine learning or deep learning methods with various input features for the AD classification. Researchers have proven that SVM (Trambaiolli et al., 2017; Song et al., 2018; Tavares et al., 2019; Nobukawa et al., 2020), fuzzy model (Yu et al., 2019), KNN (Safi and Safi, 2021), bagged trees (Oltu et al., 2021), and artificial neural network (ANN) (Triggiani et al., 2017; Ieracitano et al., 2020; Rodrigues et al., 2021) can help test the validity of the input topographic or dynamic patterns or features, such as multiscale entropy vector (Song et al., 2018), functional connectivity (Song et al., 2018; Yu et al., 2019; Nobukawa et al., 2020), cepstral distances (Rodrigues et al., 2021), Hjorth parameters (Safi and Safi, 2021), coherence (Oltu et al., 2021), bispectrum (Ieracitano et al., 2020), power spectral density (Tavares et al., 2019; Oltu et al., 2021), frequency bands (Trambaiolli et al., 2017; Triggiani et al., 2017), and wavelet transform (Ieracitano et al., 2020; Oltu et al., 2021).

Among these machine learning methods, it has been proven that deep learning has exceptional performance in terms of the accuracy of classification. Convolutional neural networks (CNN), a leading deep learning structure for data on Euclidean space, outperform the above machine learning algorithms in classification accuracy (Craik et al., 2019). Based on the Fast Fourier Transform for extracting spectral features of EEG for AD diagnosis, Bi and Wang (Bi and Wang, 2019) developed a discriminative convolutional Boltzmann Machine; Ieracitano (Ieracitano et al., 2019) and Deepthi (Deepthi et al., 2020) proposed a CNN model respectively. Based on combining latent factors output by the encoder part of variational auto-encoder of EEG, Li et al. extracted characteristics of AD (Li et al., 2021a). Based on the time-frequency analysis using CWT, Huggins et al. (Huggins et al., 2021) proposed an AlexNet model for the classification of AD, mild cognitive impairment subjects, and HC. With a connections matrix from EEG, Alves et al. (Alves et al., 2021) presented a CNN for classifying AD and schizophrenia. The above CNN applications on AD classification (Bi and Wang, 2019; Ieracitano et al., 2019; Deepthi et al., 2020; Huggins et al., 2021; Li et al., 2021a) focus more on learning the locally and continuously changed multi-scaled features on the Euclidean space from the EEG signals, neglecting the functional connectivity features. Although Alves

(Alves et al., 2021) used the connection of EEG channels as the CNN input, it neglected the temporal EEG channels features and the input connections topology feature cannot be modelled effectively due to the arranged order of EEG channels. The key convolutional filters on CNN structures in the above applications cannot fully mine the multiscale topological interactive information of EEG channels.

Considering the complexity of EEG signals in the spatial and temporal domain, how to extract more abstract geometric features for better generalisation using the deep learning methods remains tremendously troubling. The structure-function connectivity network of EEG is non-Euclidean data because the channels are discrete and discontinuous in the spatial domain. Each EEG channel can be considered as a node and there is a cross-channel interaction between nodes. Instead, geometric graph-based deep learning methods would provide a more suitable way to learn the cross-channel topologically associated features of EEG. Building neural networks under the graph theory, graph convolutional neural networks (GCNs) have been developed specifically to handle highly multi-relational graph data by jointly leveraging node-specific sequential features and cross-nodes topologically associative features in the graph domain (Sperduti and Starita, 1997; Gori et al., 2005; Scarselli et al., 2008; Gallicchio and Micheli, 2010). In recent two years, GCNs have been applied in the diagnoses of various brain disorders, such as children's ASD evaluation (Zhang et al., 2021), detection of epileptic (Zeng et al., 2020), (Zhao et al., 2021), seizure prediction (Li et al., 2021b), and epilepsy classification (Chen et al., 2020). As far as we are concerned, there are no AD diagnostic approaches based on GCN-related models.

To enable this application, an adjacency matrix, representing the topological association between different EEG channels, must be constructed as the key input of GCN. EEG functional networks are widely used in cognitive neuroscience, for example, decision making (Si et al., 2019), emotion recognition (Li et al., 2019b), and Schizophrenia research (Li et al., 2019a). Traditional statistical functional correlation measures include Pearson

correlation (PC) (Chen et al., 2020; Zhao et al., 2021), Tanh nonlinearity (Li et al., 2021b), average correlation coefficients (Zeng et al., 2020), and covariance (Zhang et al., 2021), which cannot fully represent the complex brain connectivity. Many advanced methods to measure functional connectivity (FC) have been developed, such as Phase locking Values (PLV) and Phase Lag Index (PLI) in the time domain (Mormann et al., 2000; Van Mierlo et al., 2014; Franciotti et al., 2019), Magnitude Squared Coherence (MSC) and Imaginary Part of Coherence (IPC) in the frequency domain (Babiloni et al., 2005; Wendling et al., 2009; Van Diessen et al., 2015; Al-Ezzi et al., 2021), and Wavelet Coherence (WC) in the time-frequency domain (Franciotti et al., 2019). These values can measure the degree of synchronisation between different brain regions and alterations in complex behaviours produced by the interaction among widespread brain regions (Babiloni et al., 2005; Sakkalis, 2011; Van Mierlo et al., 2014; Tafreshi et al., 2019), which have been proved important for AD classification using the statistical (Jalili, 2017; Zhao et al., 2019), and machine learning methods (Song et al., 2018; Yu et al., 2019; Nobukawa et al., 2020). However, the research on the combination of FC and GCN is limited, especially for AD-related research. Using these efficient FCs to construct the input adjacency matrix of GCN may promisingly provide more insightful information for the brain function interaction and lead to a higher classification accuracy of brain-related disorders.

In this Chapter, a novel Spatial-Temporal GCN (ST-GCN) is proposed to classify AD from HC, benefiting from the adjacency matrix constructed by a variety of FC measures and the raw EEG recordings. We tested six adjacency matrices based on PC, MSC, IPC, WC, PLV, and PLI using EEG recordings from patients with AD and HCs. ST-GCN can jointly leverage the cross-channel topological connectivity features and channel-specific temporal features. To the best of the authors' knowledge, this is the first attempt for GCN to distinguish between AD and HC based on EEG recordings.



## 4.2.2 Methods

### 4.2.2.1 Spatial-Temporal Graph Convolutional Network

In 1997, Sperduti and Starita firstly adopted neural networks to direct acyclic graphs (Sperduti and Starita, 1997), which motivated the early studies on GCNs (Gori et al., 2005; Scarselli et al., 2008; Gallicchio and Micheli, 2010). Currently, there are two basic approaches to generalising convolutions to structure graph

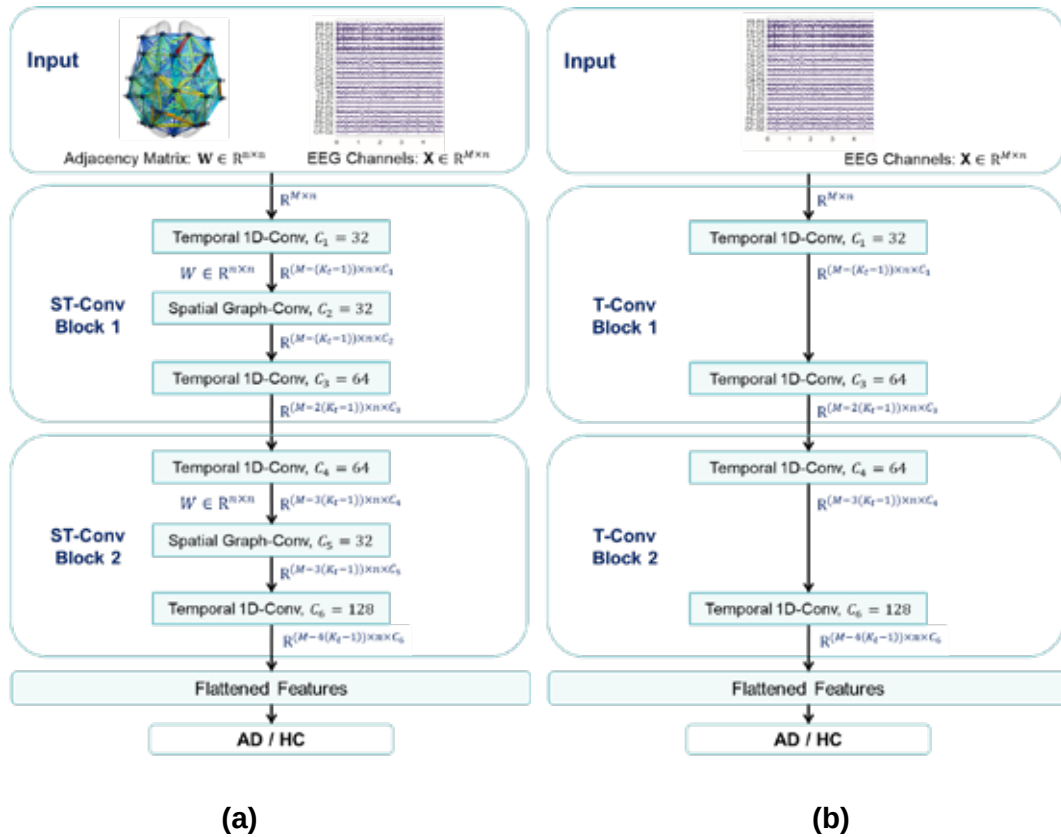


Figure 4-24 The flowchart of (a) the proposed ST-GCN framework and (b) the T-CNN framework for comparison.  $X$  is sized  $M \times N$  ( $M = 25$ , representing the length of the mini-epoch channel;  $N = 23$ , representing the 23 channels) and the input  $W$  is sized  $N \times N$ . The  $(i, j)$ th entry of the adjacency matrix  $W$  denotes spatial coupling correlation strength between the  $i$ th and  $j$ th of all the 23 different channels, the detailed calculation of which is presented in Section 2.2.  $K_t$  is the size of the temporal 1D-Conv filter, set as 3 here.  $C_l$  ( $l = 1, 2, \dots, 6$ ) is the number of filters in each layer. The estimated computational complexity of ST-GCN and T-CNN are 54M and 52M respectively.

data forms: spatial-based and spectral-based GNNs. Spatial-based GNNs define graph convolutions by rearranging vertices into certain grid forms which can be processed by normal convolutional operations (Niepert et al., 2016; Yu et al., 2017). Bruna et al. (2013) presented the first prominent spectral-based GCNs by applying convolutions in spectral domains with graph Fourier transforms. Since then, there have been increasing improvements, approximations, and extensions on spectral-based GNNs (Henaff et al., 2015; Defferrard et al., 2016; Kipf and Welling, 2016; Levie et al., 2018) to reduce the computational complexity from  $\mathcal{O}(n^2)$  to  $\mathcal{O}(n)$  (Defferrard et al., 2016; Kipf and Welling, 2016). Visually, a graph convolution can handle the complexity of graph data by generalising a 2D convolution, motivated by the successful applications of CNNs in Euclidean space (Wu et al., 2020). Being considered as special graph data, each pixel of an image can be taken as a node whose neighbours are determined by a filter and a 2D convolution takes the weighted average of adjacent pixel values of each node. Similarly, graph convolutions can be performed by taking the weighted average of a node's neighbourhood information, which is unordered and variable in size, and different from images.

As shown in Figure 4-24(a), the proposed architecture of ST-GCN is composed of two Spatial-Temporal Convolutional Blocks (ST-Conv Blocks), each of which is formed with one spatial graph convolution layer (Spatial Graph-Conv) and two sequential convolution layers (Temporal 1D-Conv). ST-Conv block can be stacked based on the complexity of specific cases. Layer normalisation is utilised within every ST-Conv block to prevent overfitting. The EEG channels with the adjacency matrix are uniformly processed by ST-Conv Blocks to explore spatial and temporal dependence coherently. A flattened layer integrates comprehensive features to generate the final AD/HC classification.

For comparison, we designed a structure of a classical temporal convolutional neural network (T-CNN) shown in Figure 4-24 (b), inspired by some related works (Deepthi et al., 2020; Huggins et al., 2021; Li et al., 2021a). The difference between these two model structures is that T-CNN only has EEG channels at the input without the adjacency matrix, and no spatial graph

convolution unit is used in the feature layer in the middle of the T-CNN. Other aspects of T-CNN are the same as ST-GCN. For brevity, we illustrate the structural details of each part of ST-GCN in the following.

#### 4.2.2.2 Spatial Graph-Conv

Based on the concept of a spectral graph convolution, we introduce the notion of a graph convolution operator “ $\ast_{\mathcal{G}}$ ”, multiplying a signal  $x \in \mathbb{R}^n$  in the spatial space with a kernel  $\theta$ ,

$$\theta \ast_{\mathcal{G}} x = \theta(L)x = \theta(U\Lambda U^T)x = U\theta(\Lambda)U^T x, \quad (4-16)$$

where the graph Fourier basis  $U \in \mathbb{R}^{n \times n}$  is a matrix of eigenvectors of the normalised graph Laplacian  $L = I_n - D^{-\frac{1}{2}}WD^{-\frac{1}{2}} = U\Lambda U^T \in \mathbb{R}^{n \times n}$  ( $I_n$  is an identity matrix,  $D \in \mathbb{R}^{n \times n}$  is the diagonal degree matrix with  $D_{ii} = \sum_j W_{ij}$ ,  $\Lambda \in \mathbb{R}^{n \times n}$  is the diagonal matrix of eigenvalues of  $L$ , and filter  $\theta(\Lambda)$  is a diagonal matrix). A graph signal  $x$  is filtered by a kernel  $\theta$  with multiplication between  $\theta$  and the graph Fourier transform  $U^T x$  (Shuman et al., 2013).

We utilise Chebyshev polynomials and 1<sup>st</sup>-order approximations (Kipf and Welling, 2016) here to reduce expensive computations of kernel  $\theta$  in Eq. (1) due to its  $\mathcal{O}(n^2)$  complex multiplications. Kernel  $\theta$  can be restricted to a polynomial of  $\Lambda$  as  $\theta(\Lambda) = \sum_{k=0}^{K-1} \theta_k \Lambda^k$ , where  $\theta \in \mathbb{R}^K$  is a vector of polynomial coefficients and  $K$  is the kernel size determining the maximum radius of the convolution from central nodes. Chebyshev polynomial  $T_k(x)$  is traditionally used to approximate kernels as a truncated expansion of order  $K-1$  as  $\theta(\Lambda) \approx \sum_{k=0}^{K-1} \theta_k T_k(\tilde{\Lambda})$  by rescaled  $\tilde{\Lambda} = 2\Lambda/\lambda_{max} - I_n$ , where  $\lambda_{max}$  denotes the largest eigenvalue of  $L$  (Hammond et al., 2011). Then the graph convolution in Eq. (1) can be rewritten as:

$$\theta \ast_{\mathcal{G}} x = \theta(L)x = \sum_{k=0}^{K-1} \theta_k T_k(\tilde{L})x, \quad (4-17)$$

where  $T_k(\tilde{L}) \in \mathbb{R}^{n \times n}$  is the Chebyshev polynomial of order  $k$  evaluated at the scaled Laplacian  $\tilde{L} = 2L/\lambda_{max} - I_n$ . By recursively computing  $\tilde{L}$ -localised

convolutions through a polynomial approximation, the computational cost of Eq. (1) can be reduced to  $\mathcal{O}(K|\mathcal{E}|)$  as Eq. (2). By stacking multiple localised graph convolutional layers with a first-order approximation of graph Laplacian, a layer-wise linear formulation can be defined (Kipf and Welling, 2016). Further assumption  $\lambda_{max} \gg 2$  can be made, due to the scaling and normalisation in neural networks. Thus, Eq. (2) can be simplified to

$$\theta \mathcal{G} x = \theta_0 x + \theta_1 \left( \frac{2}{\lambda_{max}} L - I_n \right) x = \theta_0 x + \theta_1 \left( D^{-\frac{1}{2}} W D^{-\frac{1}{2}} \right) x, \quad (4-18)$$

where  $\theta_0$  and  $\theta_1$  are two shared parameters of the kernel.  $\theta_0$  and  $\theta_1$  are replaced by a single parameter  $\theta$  by letting  $\theta_0 = \theta_1 = \theta$  to constrain parameters and stabilise numerical performances. By renormalising and with  $\tilde{W} = W + I_n$  and  $\tilde{D}_{ii} = \sum_j \tilde{W}_{ij}$  separately, graph convolution can be expressed as

$$\theta \mathcal{G} x = \theta (I_n + D^{-\frac{1}{2}} W D^{-\frac{1}{2}}) x = \theta (\tilde{D}^{-\frac{1}{2}} \tilde{W} \tilde{D}^{-\frac{1}{2}}) x \quad (4-19)$$

The graph convolution operator “ $\mathcal{G}$ ” defined on  $x \in \mathbb{R}^n$  can be extended to multi-dimensional tensors. For a signal with  $C_i$  channels  $X \in \mathbb{R}^{n \times C_i}$ , the graph convolution can be generalised to

$$y_j = \sum_{i=1}^{C_i} \theta_{i,j} (L) x_i \in \mathbb{R}^n, 1 \leq j \leq C_0 \quad (4-20)$$

with the  $C_i \times C_0$  vectors of Chebyshev coefficients  $\theta_{i,j} \in \mathbb{R}^K$  ( $y_j$  is the output after graph convolution,  $C_i$  and  $C_0$  are the sizes of input and output of the feature maps, respectively). A graph convolution for 2D variables is denoted as “ $\theta \mathcal{G} X$ ” with  $\theta \in \mathbb{R}^{K \times C_i \times C_0}$ . The input of ST-GCN is composed of frames of EEG channels graph as shown in Figure 1(a). Each frame  $X_t$  can be regarded as a matrix whose column  $i$  is a  $C_i$ -dimensional value of  $X_t$  at the  $i^{\text{th}}$  node in graph  $\mathcal{G}_t$ , as  $X \in \mathbb{R}^{n \times C_i}$  (in this case,  $C_i=1$ ). For each time step of , the equal graph convolution operation with the same kernel  $\theta$  is imposed on  $X_t \in \mathbb{R}^{n \times C_i}$  in parallel. Thus, the graph convolution can be generalised to 3D variables, noted as “ $\theta \mathcal{G}$ ” with  $\chi \in \mathbb{R}^{M \times n \times C_i}$ .

#### 4.2.2.3 Temporal 1D-Conv

Inspired by (Gehring et al., 2017) that CNNs have the superiority of fast training in sequential-series analysis, we employ an entire convolutional structure on a temporal axis to capture sequential dynamic behaviours of EEG recordings. As shown in Figure 1(a), a sequential convolutional layer contains a 1D convolution with a width  $K_t$  kernel followed by ReLu (a rectified linear unit function) as a nonlinearity. For each node in graph  $\mathcal{G}$ , its corresponding sequential convolution explores  $K_t$  neighbours of input elements, leading to shortening the length of sequences by  $K_t - 1$  each time. Thus, an input of a sequential convolution for each node can be regarded as a length- sequence with  $C_i$  channels as  $Y \in \mathbb{R}^{M \times C_i}$ . The convolution kernel  $\Gamma \in \mathbb{R}^{K_t \times C_i \times C_0}$  is designed to map the input  $Y$  to a single output  $P \in \mathbb{R}^{(M-K_t+1) \times (C_0)}$ . Similarly, the temporal convolution can be generalised to 3D variables by employing the same convolution kernel  $\Gamma$  to every channel node in  $\mathcal{G}$  equally, noted as “ $\Gamma \circ \mathcal{T} \mathcal{Y}$ ” with  $\mathcal{Y} \in \mathbb{R}^{M \times n \times C_i}$ .

The input and output of ST-Conv blocks are all 3D tensors. For input  $x^l \in \mathbb{R}^{M \times n \times C_l}$  of block , the output  $x^{l+1} \in \mathbb{R}^{(M-2(K_t-1)) \times n \times C^{l+1}}$  is computed by

$$x^{l+1} = \Gamma_1^l \circ \mathcal{T} \text{ReLU}(\Theta^l \circ \mathcal{G}(\Gamma_0^l \circ \mathcal{T} x^l)) \quad (4-21)$$

where  $\Gamma_0^l$  and  $\Gamma_1^l$  are the upper and lower temporal kernels within block , respectively;  $\Theta^l$  is the spectral kernel of a graph convolution;  $\text{ReLU}(\cdot)$  denotes a rectified linear unit function. After stacking three ST-Conv blocks, the output features are fused as a flattened layer (Figure 1(a)). We can obtain a final output  $Z \in \mathbb{R}^{n \times c}$  from the fully connected layer and calculate the classification result by applying a sigmoid transformation as  $\hat{cl} = \text{sig}(Z\omega + b)$ , where  $\omega \in \mathbb{R}^c$  is a weight vector and  $b$  is a bias. We use the cross entropy loss to measure the classification performance.

All models were trained on the CPU of DELL DESKTOP-D3UM3P9 with the Tensorflow platform in Microsoft Windows 10, and the optimiser used here is the Adam optimisation.

#### 4.2.2.4 Adjacency Matrix

The spatial information carried by EEG signals plays an important role in AD/HC classification (Babiloni et al., 2005; Sakkalis, 2011; Van Mierlo et al., 2014; Tafreshi et al., 2019). The adjacency matrix  $R^{n \times n}$  of each mini-epoch, representing the spatial correlation along with the channel signals  $X \in R^{n \times M}$  is one of the inputs of the ST-GCN model shown in Figure 1(a). In some studies on brain disorders based on GCN (Chen et al., 2020; Zeng et al., 2020; Li et al., 2021b; Zhang et al., 2021; Zhao et al., 2021), the relationship of different channels of EEG is short of effective prior guidance and the adjacency matrix cannot ensure the utilisation of the coupling information between each channel. To address these issues, we firstly apply functional connectivity, which has been proven to be useful in AD classification, to construct the adaptive adjacency matrix to extract spatial coupling features.

The raw EEG signals of each channel and the association of channels are modelled by a graph. The nodes of the graph denote the feature vector of EEG signals, which are the raw mini-epoch EEG data. The adjacency matrix is constructed by Pearson correlation analysis and five functional connectivities of the mini-epoch EEG channels without any pre-set threshold to avoid the potential risk caused by manual selection, such as in (Chen et al., 2020; Zhang et al., 2021; Zhao et al., 2021). Specifically, the  $a_{ij}$  entry of the adjacency matrix denotes spatial coupling strength between the  $i$  and  $j$  channels. Thus,  $a_{ij}$  indicates that all channels are interconnected with different weights. Pair-wise correlation analysis among these 23 different channels is conducted using five functional connectivities and Pearson correlation for comparison, the calculation of which is presented below.

The most well-known functional connectivity measure is the correlation, also called the Pearson correlation coefficient. It calculates the instantaneous linear interdependency between two signals based on the amplitudes of the signals in

the time domain and it ranges from -1 to 1. The Pearson correlation coefficient between signal  $X$  and  $Y$  can be defined as follows:

$$\rho_{xy} = \frac{E[(x-\mu_x)(y-\mu_y)]}{\sigma_x \sigma_y} \quad (4-22)$$

where  $E$  is the expected value,  $\mu_x$  and  $\mu_y$  are the mean values and  $\sigma_x$  and  $\sigma_y$  are the standard deviations of  $X$  and  $Y$  time series.

MSC is a linear method to estimate the interconnections between the PSD (Power Spectral Density) of two signals in the frequency domain. The MSC of signals  $X$  and  $Y$  can be written as:

$$MSC_{xy}(f) = C_{xy}^2 = \frac{S_{xy}(f)^2}{|S_{xx}(f)| \times |S_{yy}(f)|} \quad (4-23)$$

where  $S_{xx}(f)$  and  $S_{yy}(f)$  are the PSD of signal  $X$  and  $Y$  respectively, and  $S_{xy}(f)$  is the cross PSD at frequency  $f$ .

To avoid the volume conduction effects, instead of looking at the magnitude squared coherency, the imaginary part of the coherency is calculated by

$$IC_{xy} = RC_{xy} + IC_{xy} \quad (4-24)$$

WC is generally acknowledged as a qualitative estimator that can represent the dynamic relations in the time-frequency domain between signals (Tafreshi et al., 2019). The wavelet transform is defined as the convolution of the input  $x$  with a Wavelet family  $\theta(u)$ :

$$W_x(t, f) = \int_{-\infty}^{+\infty} x(u) \theta_{t,f}(u) du$$

Given input signals  $x$  and  $y$ , wavelet cross-spectrum around time  $t$  and frequency  $f$  can be derived by the Wavelet transforms of  $x$  and  $y$ :

$$CW_{xy}(t, f) = \int_{t-\delta/2}^{t+\delta/2} W_x(\tau, f) W_y^*(\tau, f) d\tau$$

where  $*$  defines the complex conjugate and  $\delta$  is assumed as a frequency-depending time scalar. WC at the time  $t$  and frequency  $f$  is derived as:

$$WC_{xy}(t, f) = \frac{|CW_{xy}(t, f)|}{|CW_{xx}(t, f) \times CW_{yy}(t, f)|^{1/2}} \quad (4-27)$$

Phase synchronization assumes that two oscillation signals without amplitude synchronization can have phase synchronization. The PLV is high-frequently utilised to obtain the strength of phase synchronisation (Van Mierlo et al., 2014). The instantaneous phase of a signal  $x$  is given by:

$$\phi_x(t) = \arctan \frac{\tilde{x}(t)}{x(t)} \quad (4-28)$$

where  $\tilde{x}(t)$  is the Hilbert Transform of  $x(t)$ , defined as:

$$\tilde{x}(t) = \frac{1}{\pi} PV \int_{-\infty}^{+\infty} \frac{x(\tau)}{t - \tau} d\tau \quad (4-29)$$

where  $PV$  refers to the Cauchy principal value. The  $PLV$  for two signals is then defined as:

$$PLV = \left| \frac{1}{N} \sum_{j=0}^{N-1} e^{j(\phi_x(jT) - \phi_y(jT))} \right| \quad (4-30)$$

where  $T$  defines the sampling period and  $N$  indicates the sample number of each signal. The range of  $PLV$  is between 0 to 1, where 0 shows a lack of synchronization and 1 indicates strict phase synchronization.



Similarly to the calculation of PLV, PLI captures the asymmetry of the distribution of phase differences between two signals and is calculated based on the relative phase difference between the two signals:

$$PLI = \left| E \left[ \text{sign} \left( \phi_x(j, t) - \phi_y(j, t) \right) \right] \right| \quad (4-31)$$

where  $\phi_x(j, t) - \phi_y(j, t)$  is the phase difference between two signals, *sign* stands for signum function, *E* is the expected value, and || indicates the absolute value. *PLI* values range between 0 and 1, where 0 can indicate possibly no coupling and 1 refers to perfect phase locking.

#### 4.2.2.5 Dataset and preprocessing

All patients and healthy controls included in this work were recruited in the Sheffield Teaching Hospital memory clinic and all provided written informed consent. The EEG study underwent ethics approval by the Yorkshire and The Humber (Leeds West) Research Ethics Committee (reference number 14/YH/1070). AD patients had their diagnosis confirmed between one month and up to two years before recording their EEG while they had mild to moderate cognitive deficits, according to their Mini-mental state examination. All AD subjects had brain MRI scans to eliminate other alternative causes of dementia. For the aged and gender-matched HC cohort, normal MRI brain scans and cognitive assessments were required before their EEG recordings. Further details about recruitment, diagnostic criteria, and study design can be found in the previously published work (Blackburn et al., 2018).

All participants were younger than 70 years old, including 19 AD patients and 20 HC participants. EEG recordings were undertaken with an XLTEK 128-channel headbox (Optima Medical LTD) and Ag/AgCL electrodes at a sampling frequency of 2 kHz by implementing a modified 10-10 overlapping a 10-20 international system of electrode placement, with a referential montage (linked earlobe reference). Thirty-minute resting state (task-free – participants were instructed to rest and refrain from thinking anything specific) EEG recordings were obtained from each participant including sustained periods of keeping their eyes closed (EC) alternating with periods during which they kept their eyes

open (EO). The recordings obtained were subsequently reviewed by a neurophysiologist - on an XLTEK review station. For each participant, three 12-second artefact-free epochs of EC and EO were selected. To avoid volume conduction effects related to the common reference electrode, 23 bipolar derivations were created: 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 O1–O2.

For both EO and EC of 19 ADs and 20 HCs, there are 3 artefact-free epochs, each of which lasts 12 seconds. We first applied the Butterworth filter for every epoch to subsample the EEG signals at 100Hz and obtained the corresponding six bands (Delta, Theta, Alpha, Beta, Gamma, and Full band of 0-48Hz). Then, data segmentation is utilised to obtain 5472 mini-epochs for ADs and 5760 mini-epochs for HCs with a window size of 25 data points without overlapping. For each mini-epoch signal, six functional connectivities (PC, MSC, IPC, WC, PLV, and PLI) were calculated to obtain the corresponding adjacency matrices. Finally, for each measure of EC or EO, 3744 samples (1824 for ADs and 1920 for HCs) were selected randomly covering two-thirds of each 12s epoch and split for 10-folder cross-validation to generate the training dataset and validation dataset. The remaining 1872 samples (912 for ADs and 960 for HCs) were considered as the testing dataset. The purpose of the validation dataset during training is to obtain the model loss on the validation dataset after the optimization of each epoch. If the loss of the validation dataset is decreasing, we save the optimised model until the end of all training epochs, which can prevent the model from over-fitting.

## **4.2.3 Results**

### **4.2.3.1 Adjacency Matrix from different FCs**

To compare the adjacency matrices calculated by six functional connectivity estimations, we take the Full band of one mini-epoch of AD (Figure 4-25) and HC (Figure 4-26) as an example to see their difference. For the mini-epoch signal of both AD and HC, PC generates higher coupling between channels than the five functional connectivity methods. The adjacency matrix calculated

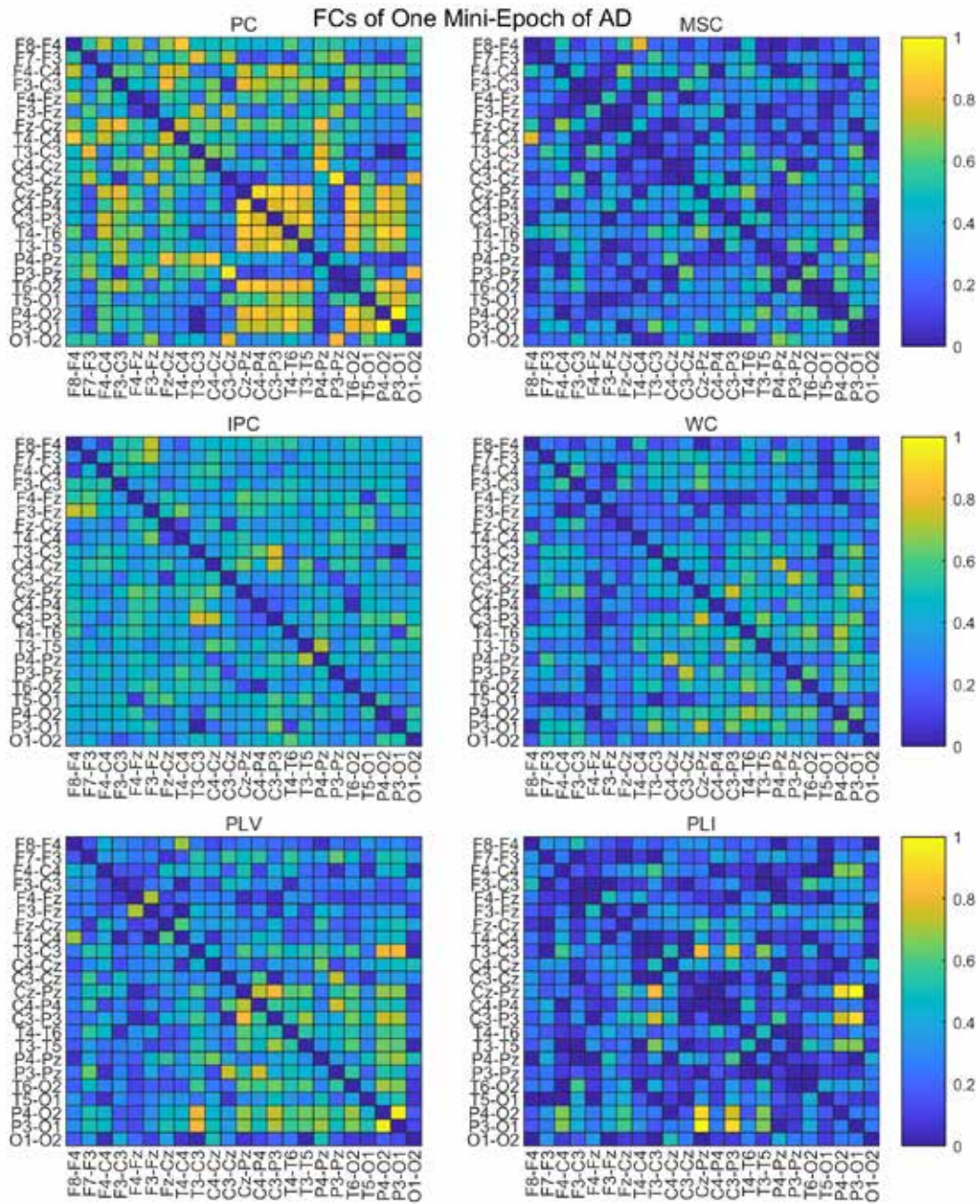


**Figure 4-25 The adjacency matrices of one mini-epoch of AD from six FC methods**

by MSC is relatively lower than the others and WC and PLV show similar coupling distribution with PC but with overall lower strength. IPC produces the most uniform connection distribution. PLI produces the sparsest connection distribution, and there are some strong connection points close to 1. Therefore,



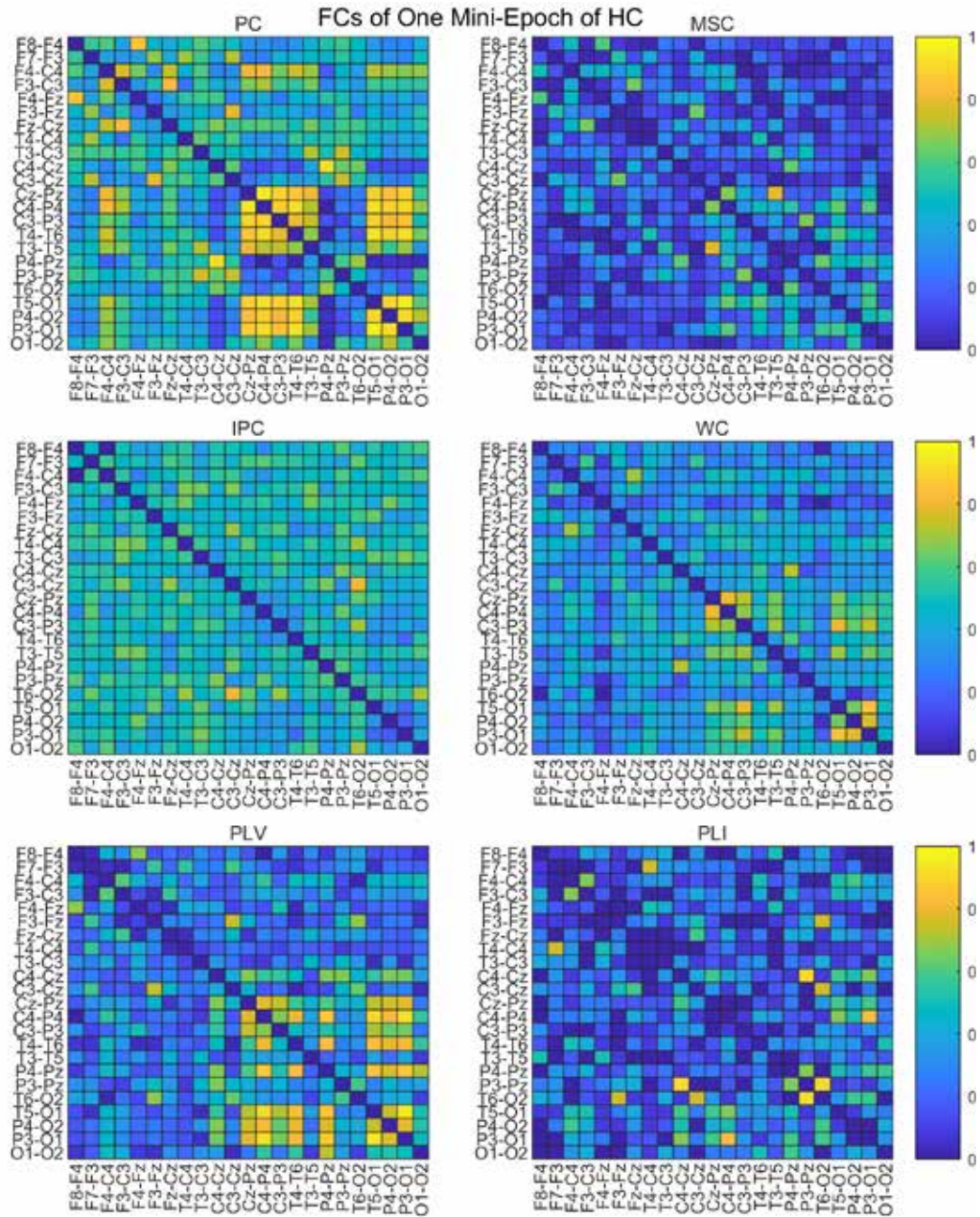


**Figure 4-26 The adjacency matrices of one mini-epoch of HC from six FC methods**

adjacency matrices from the six calculation methods have differences in the coupling analysis between EEG channels, which may have an impact on the classification performance of the ST-GCN method.

#### 4.2.3.2 Overall classification performance

We conducted a direct comparison of classifying AD and HC participants using the extracted 6 adjacency matrices by training the ST-GCN and T-CNN methods on the EO and EC states and six frequency bands independently. Table 4-5 and Table 4-6 show the corresponding testing accuracies from ST-

**Table 4-5 Classification accuracy for HC and AD in EC and EO states within 6 bands by ST-GCN with different adjacency matrices.**

| Eye States | FC  | Delta | Theta | Alpha | Beta | Gamma | Full |
|------------|-----|-------|-------|-------|------|-------|------|
| EC         | PC  | 72.1  | 78.2  | 78.3  | 82.5 | 81.2  | 90.3 |
|            | MSC | 72.2  | 78.1  | 78.9  | 81.9 | 80.8  | 90.2 |
|            | IPC | 73.8  | 78.4  | 80.2  | 82.3 | 82.1  | 90.7 |
|            | WC  | 73.9  | 78.7  | 81.0  | 83.2 | 82.5  | 92.3 |
|            | PLV | 73.0  | 78.6  | 79.8  | 82   | 82.2  | 91.1 |
|            | PLI | 73.9  | 78.2  | 79.6  | 82.3 | 81.1  | 92.1 |
| EO         | PC  | 70.9  | 75.6  | 76.8  | 77.4 | 75.5  | 88.1 |
|            | MSC | 71.3  | 75.7  | 75.4  | 77.8 | 75.5  | 88.3 |
|            | IPC | 71.0  | 76.0  | 77.4  | 78.1 | 75.6  | 88.7 |
|            | WC  | 71.2  | 76.8  | 78.2  | 78.5 | 76    | 89.4 |
|            | PLV | 71.0  | 75.8  | 76.0  | 78.1 | 75.9  | 87.2 |
|            | PLI | 70.3  | 75.0  | 75.6  | 76.5 | 75.8  | 87.9 |

Abbreviations: PC, Pearson Correlation; MSC, Magnitude Squared Coherence; IPC, Imaginary Part of Coherence; WC, Wavelet Coherence; PLV, Phase Locking Value; PLI Phase Lag Index.

**Table 4-6 Classification accuracy for HC and AD in EC and EO states within 6 bands by T-CNN with different adjacency matrices.**

| Eye States | FC  | Delta | Theta | Alpha | Beta | Gamma | Full |
|------------|-----|-------|-------|-------|------|-------|------|
| EC         | PC  | 69.0  | 76.7  | 77.3  | 80.3 | 75.9  | 88.1 |
|            | MSC | 70.4  | 76.1  | 74.0  | 80.7 | 74.8  | 86.8 |
|            | IPC | 69.4  | 77.8  | 77.1  | 78.2 | 77.5  | 88.8 |
|            | WC  | 71.0  | 78.0  | 76.7  | 80.3 | 80.5  | 89.0 |
|            | PLV | 71.4  | 78.2  | 76.9  | 77.9 | 77.4  | 88.1 |
|            | PLI | 68.5  | 76.6  | 78.7  | 80.6 | 75.5  | 87.4 |
| EO         | PC  | 67.9  | 75.3  | 77.2  | 71.3 | 74.3  | 85.6 |
|            | MSC | 67.0  | 74.8  | 76.9  | 75.0 | 69.8  | 85.9 |
|            | IPC | 65.0  | 76.1  | 76.5  | 72.9 | 74.1  | 86.4 |
|            | WC  | 70.1  | 75.9  | 77.9  | 76.2 | 73.7  | 86.3 |
|            | PLV | 65.8  | 74.5  | 75.9  | 71.5 | 73.5  | 85.0 |
|            | PLI | 67.3  | 74.9  | 76.9  | 71.6 | 71.9  | 85.7 |

Abbreviations: PC, Pearson Correlation; MSC, Magnitude Squared Coherence;

GCN and T-CNN models, respectively. For both methods, the classification accuracy of all frequency bands and adjacency matrix calculation methods in

the EC state is higher than the corresponding EO state, which is consistent with previous research results (Barry et al., 2007; Wan et al., 2019). For different frequency bands, when the 1-48Hz full-band signal and the corresponding adjacency matrix are used as the input, the mean classification accuracies are about 91.1% by ST-GCN and 88.0% by T-CNN respectively, which are higher than all sub-bands data. For ST-GCN, the mean classification accuracy of Beta band data is the highest (about 82.3% for EC and 77.7% for EO), and the mean classification accuracy of Delta band data is the lowest (about 73.1% for EC and 71.0% for EO). For T-CNN, the mean classification accuracy of Alpha band

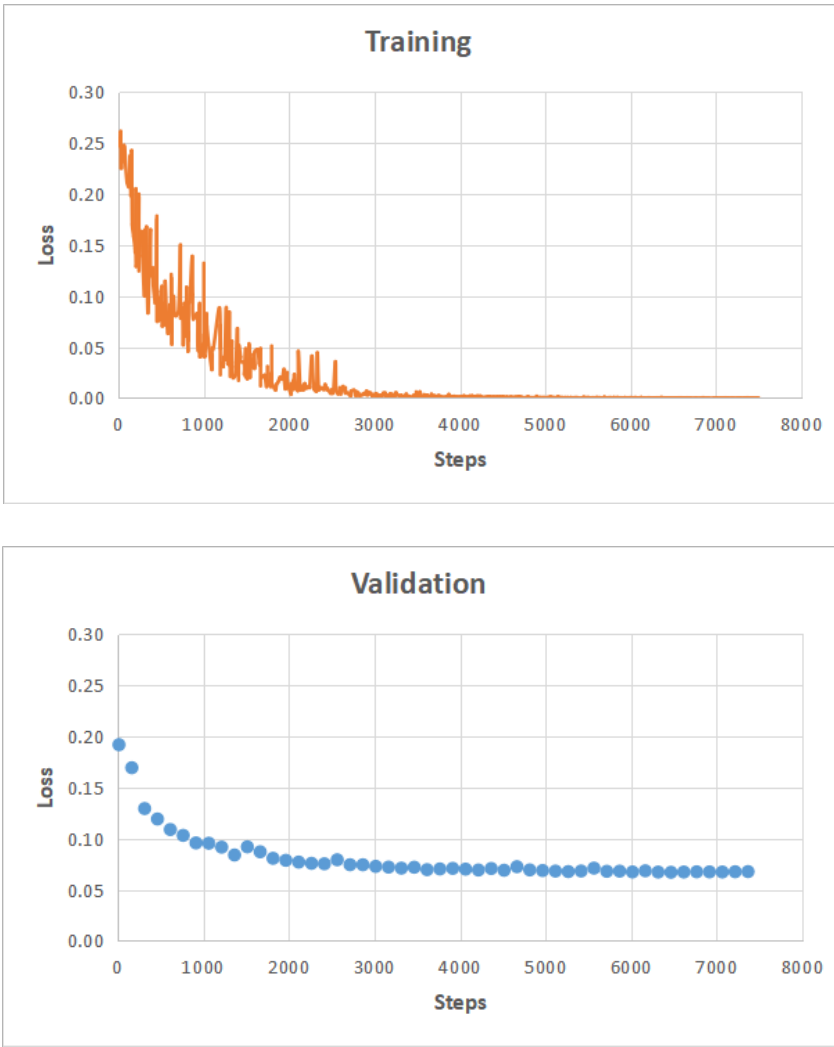


**Figure 4-27 Training and convergence process for Full band data with WC connectivity by the ST-GCN model**

data is the highest (about 76.8% for EC and EO), and the mean classification



accuracy of Delta band data is the lowest (about 68.6% for EC and EO). Overall, ST-GCN outperforms T-CNN on both eye states and almost all sub-band data, which indicates that spatial topology constraints can indeed mine EEG features, resulting in improved classification accuracy.

It is worth noting that for different adjacency matrix calculation methods, the classification accuracy of the WC method is better than all other methods, which proves the effectiveness of time-frequency analysis for extracting coupling features. For ST-GCN, the overall accuracy of the five functional connection methods is higher than that of the PC method, indicating that the usage of functional connectivity as the adjacency matrix can improve the performance of the ST-GCN model to a certain extent.

The training and convergence process for Full band data with WC connectivity by the proposed ST-GCN model is shown in Figure 4-27.

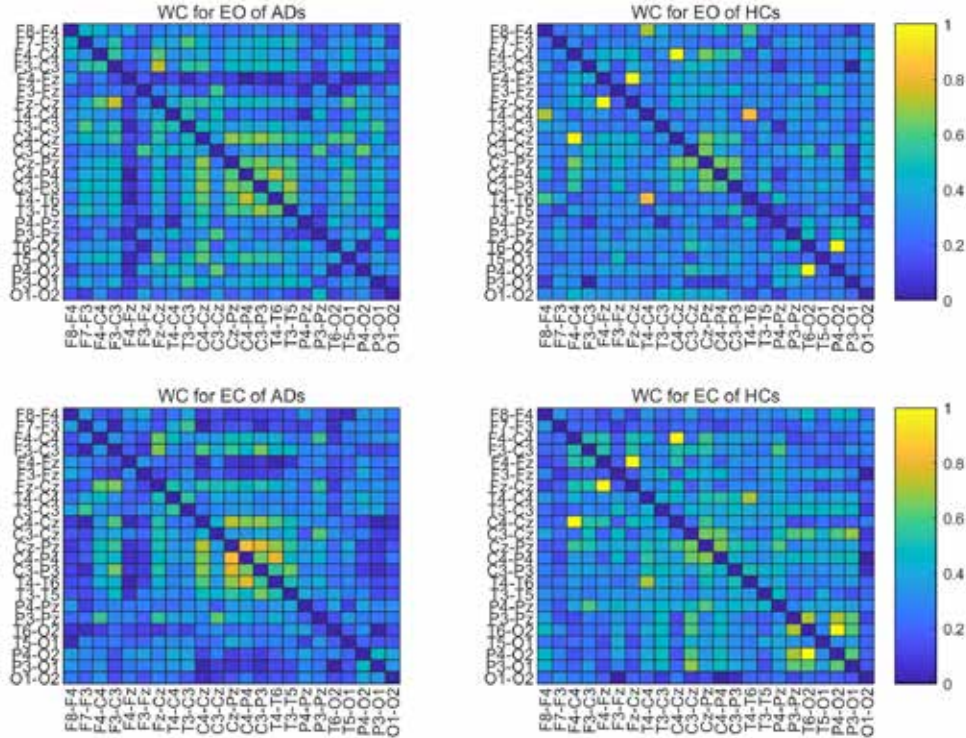


**Figure 4-28 The averaged WC adjacency matrices of ADs and HCs**

#### 4.2.3.3 Wavelet Coherence as Adjacency Matrix

To further explore the effectiveness of WC which has the best classification performance as the adjacency matrix, we conduct a statistical analysis of the WCs of all AD and HC. By averaging the WC adjacency matrices of the Full band of all EO and EC mini epochs for AD and HC respectively, as shown in Figure 4-28, we can analyse their statistical characteristics. For ADs, the EC state has a slightly higher coupling strength than EO in the temporooccipital area (the middle part of the adjacency matrix), while for HCs, the EC state has a bit higher overall coupling strength than EO. For both EO and EC, the strongest inter-channel correlation of AD is less than 0.8 (the middle part of the left two images of Figure 4-28 – temporooccipital area), while HC has some inter-channel correlations close to 1 (the upper-left frontocentral and bottom-right posterior area of the right two images of Figure 4-28). For both EO and EC, the

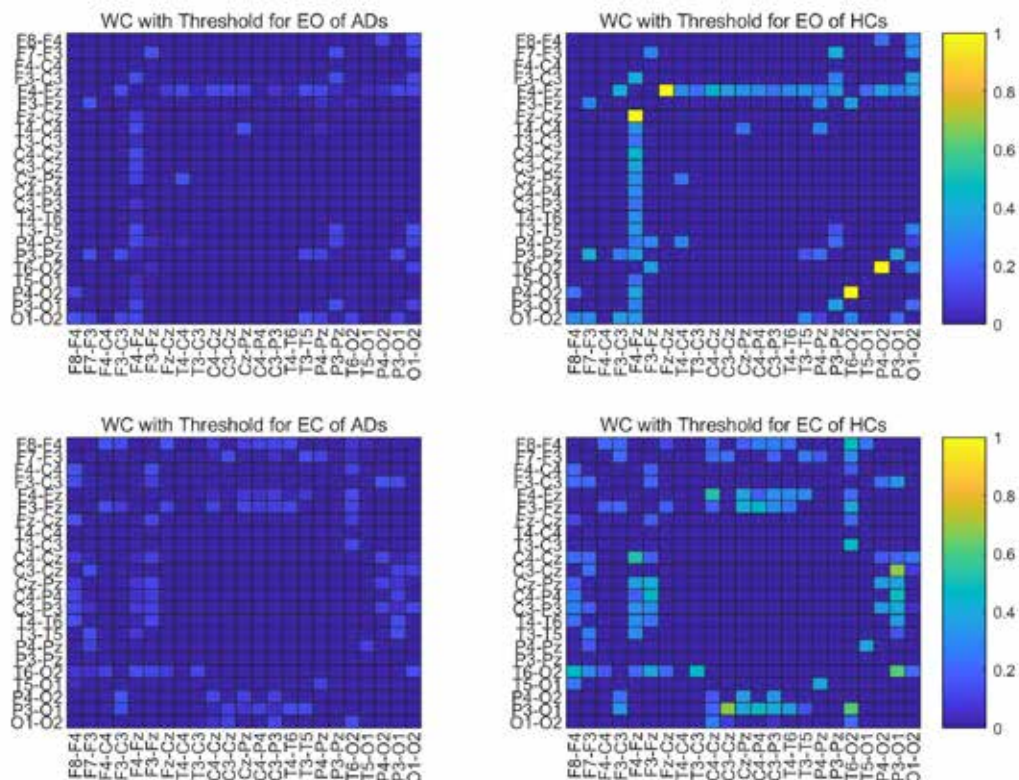


Figure 4-29 Threshold of averaged WC adjacency matrices of ADs and HCs



connectivity between ADs channels is lower than that of HCs in the bottom-right corner (posterior area).

To visually show the connectivity between the channels in the bottom-left corner for both EO and EC states, we thresholded the left two AD adjacency matrices in Figure 4-29 to screen out the channels with connectivity below 0.15. Figure 4-29 reports the corresponding cross-channel indexes for HCs. Through the software BrainNet, the 3D connectivity distributions in the brain of ADs and HCs are visually displayed for EO (Figure 4-31 and Figure 4-) and EC (Figure 4-32 and Figure 4-) respectively. Comparing Figure 4-31 and Figure 4- for EO, the connectivity of HCs between the channels at the front and back parts of the brain are much higher than those of ADs. P4-O2 to T6-O2 and F4-Fz to Fz-Cz have the highest coupling strengths (close to 0.6) for channel pairs of HCs, and the corresponding coupling strengths of ADs are close to 0. Comparing Figure 4-32 and Figure 4- for EC, the connectivity of HCs between the channels at the back of the brain is much higher than that of ADs. C3-Cz to P3-O1 and P3-O1 to T6-O2 have the highest coupling strengths (close to 0.6) for channel pairs of HCs, and the corresponding coupling strengths of ADs are as low as 0.1.

#### **4.2.4 Conclusions**

We proposed a Spatial-temporal Graph Convolutional Network (ST-GCN) for classifying Alzheimer's Disease and Healthy Controls groups by jointly leveraging cross-channel topological association features and channel-specific temporal features of EEG recordings. Different from the currently leading GCN applications for diagnosing brain disorders, this method utilises brain functional connectivity methods for exploring the complex interactive information between EEG channels as well as the single-channel-based dynamic information. The main goal of this work was to determine whether the cross-channel topologically associated features constrained by the functional connectivity can reveal more hidden information in data and extend the applicability of the GCN-based algorithm. For the clinical AD and HC EEG recordings, ST-GCN has exhibited superior performance in achieving the highest classification accuracy with wavelet coherence as the adjacency matrix. For the tested dataset, the overall

classification accuracy of ST-GCN is higher than the classical T-CNN method

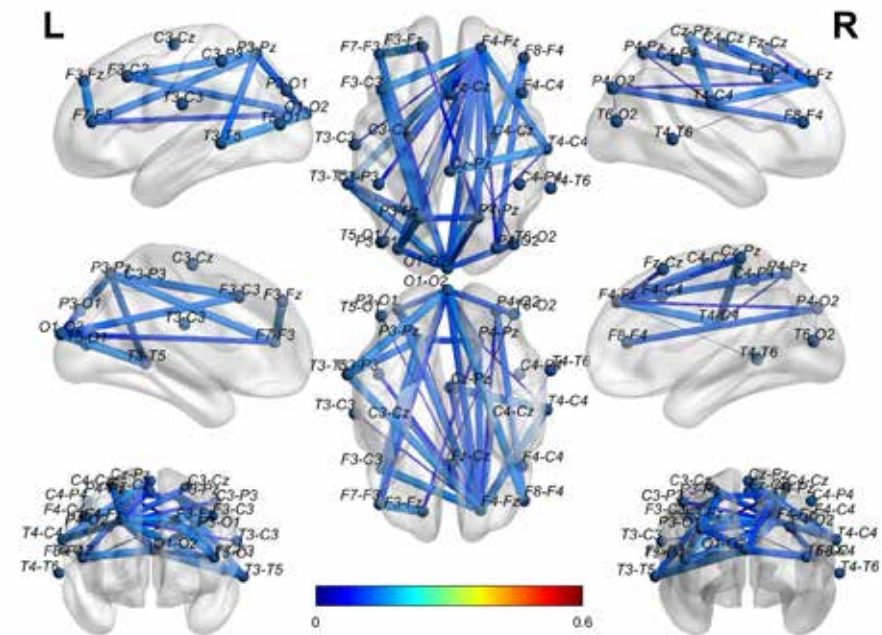


**Figure 4-31**The 3D brain mapping of averaged WC adjacency matrices for EO state of ADs

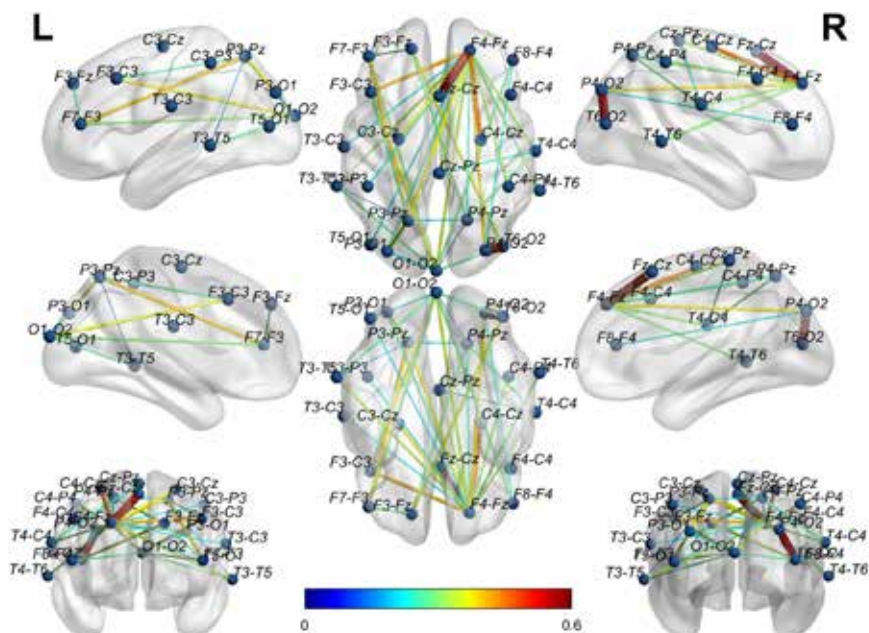


**Figure 4-30 The 3D brain mapping of averaged WC adjacency matrices for EO state of HCs**

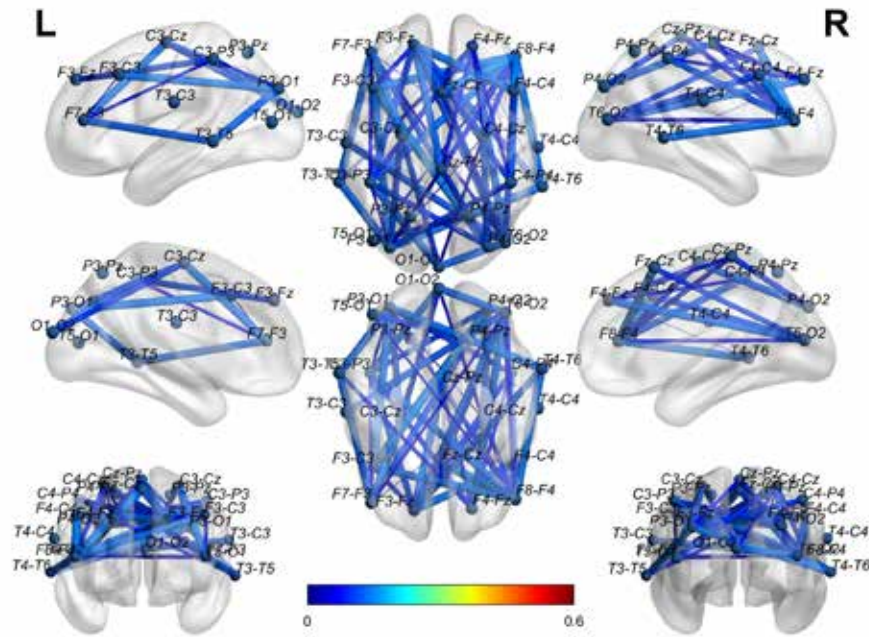


Figure 4-32 The 3D brain mapping of averaged WC adjacency matrices for EC state of ADs

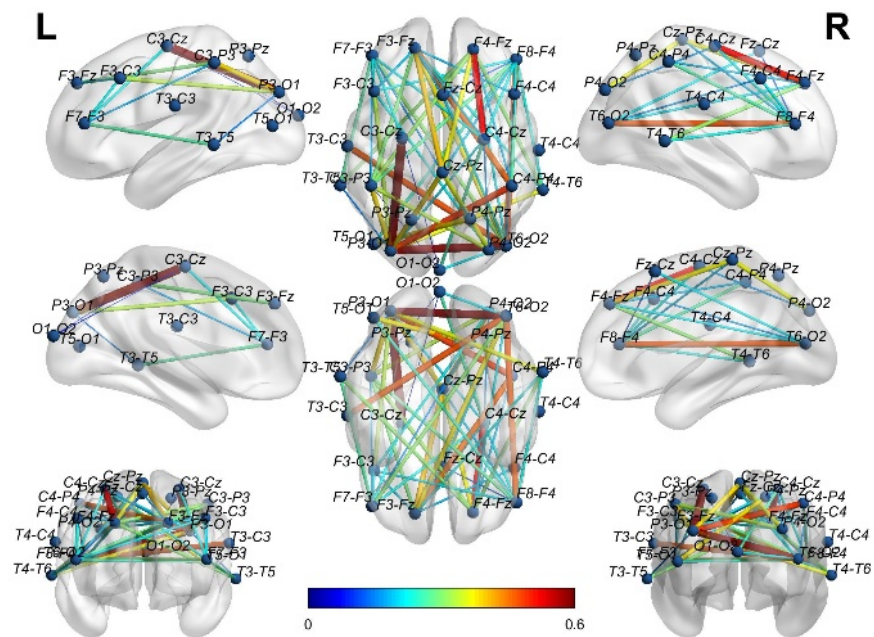


Figure 4-33 The 3D brain mapping of averaged WC adjacency matrices for EC state of HCs

on both eye states and different frequency bands, which suggests that spatial topology constraints can indeed mine brainwave features and thereby improve

the classification accuracy. Different from existing studies for AD diagnosis that are based on either topological brain function characteristics of EEG (Jalili, 2017; Fan et al., 2018; Tylová et al., 2018; Zhao et al., 2019; Tait et al., 2020) or temporal dynamic features (Bi and Wang, 2019; Ieracitano et al., 2019; Deepthi et al., 2020; Huggins et al., 2021; Li et al., 2021a), the proposed ST-GCN considers both the adjacency matrix of functional connectivity from multiple EEG channels and corresponding dynamics of signal EEG channel simultaneously. It has the potential to pick up the anomaly of AD not only in the frequency response of local areas but also in the functional connectivity across different regions. Furthermore, the visualisation of wavelet coherence adjacency matrices increases the transparency of this solution by providing evidence of brain anomaly in terms of functional connectivity. This investigation is important as it will increase the trust in the developed AI-based solution. This algorithm lays a potentially effective strategy for the applications of other brain disorders.

In the present Chapter, due to the limited number of subjects in the dataset, the Leave-One-Subject-Out or cross-subject validation is not discussed to avoid biased conclusions caused by data insufficiency. The accuracy of a hand-out cross-subject validation by ST-GCN and T-CNN can be found in Table 4-7 and Table 4-8 in the Supplementary Material. In this case, 65% of the subjects were used for training and the remaining 35% were used for testing. Although the overall accuracy is dropped significantly for both methods, the proposed ST-GCN still outperforms T-CNN.

To reduce volume conduction effects from a common reference, bipolar derivations were used to assess the degree of differences between various pairs of electrodes for 2 different cohorts of subjects. With this approach – the use of bipolar pairs of electrodes - the effects of volume conduction are reduced but not eliminated. We recognise that this work is based on a sensor level scalp EEG analysis, and we do not claim to be able to precisely localise the spatial characteristics underpinning the EEG sensor findings.

#### 4.2.5 Supplementary

**Table 4-7 ST-GCN Model performance in a hand-out validation, where 12 ADs and 13 HCs were used for training and validation and the remaining 7 ADs, and 7 HCs were used for testing. Overall, the classification accuracy of the full-band data is still the best, with an average of 70.5% for EC and 68.9% for EO.**

|  |      |      |      |      |      |      |
|--|------|------|------|------|------|------|
|  | 56.9 | 67.1 | 65.6 | 58.0 | 45.8 | 70.9 |
|  | 56.6 | 67.7 | 65.1 | 59.1 | 51.7 | 66.7 |
|  | 56.9 | 67.9 | 63.9 | 58.1 | 49.0 | 70.6 |
|  | 58.5 | 69.8 | 67.9 | 59.3 | 49.6 | 72.1 |
|  | 51.0 | 68.1 | 66.3 | 57.9 | 48.6 | 71.1 |
|  | 56.1 | 69.3 | 64.3 | 55.7 | 46.8 | 71.5 |
|  | 53.0 | 68.8 | 63.8 | 61.3 | 54.7 | 68.2 |
|  | 57.7 | 67.2 | 63.1 | 56.0 | 52.3 | 67.3 |
|  | 54.7 | 68.6 | 65.3 | 55.3 | 56.9 | 69.6 |
|  | 57.1 | 69.1 | 65.4 | 61.0 | 57.1 | 70.2 |
|  | 55.7 | 66.6 | 61.9 | 56.1 | 55.4 | 69.6 |
|  | 55.3 | 68.1 | 63.5 | 54.9 | 52.8 | 68.5 |

Abbreviations: PC, Pearson Correlation; MSC, Magnitude Squared Coherence;  
 IPC, Imaginary Part of Coherence; WC, Wavelet Coherence; PLV, Phase  
 Locking Value; PLI, Phase Lag Index.

**Table 4-8 T-CCN Model performance in a hand-out validation, where 12 ADs and 13 HCs were used for training and validation and the remaining 7 ADs and 7 HCs were used for testing. Overall, the classification accuracy of the full-band data is lower than the ST-GCN model.**

|  |      |      |      |      |      |      |
|--|------|------|------|------|------|------|
|  | 52.9 | 70.6 | 62.6 | 54.6 | 50.3 | 66.4 |
|  | 58.2 | 67.2 | 62.2 | 56.5 | 47.6 | 64.4 |
|  | 52.8 | 66.5 | 63.3 | 60.1 | 44.9 | 66.5 |
|  | 55.9 | 64.8 | 62.4 | 59.7 | 50.6 | 65.5 |
|  | 54.1 | 69.5 | 62.8 | 59.6 | 57.8 | 67.6 |
|  | 56.8 | 67.5 | 61.8 | 55.5 | 53.3 | 66.1 |
|  | 50.9 | 69.2 | 64.4 | 60.9 | 55.7 | 66.5 |
|  | 48.7 | 68.3 | 64.8 | 54.9 | 54.7 | 66.1 |
|  | 53.7 | 65.3 | 66.3 | 57.8 | 55.0 | 64.4 |
|  | 50.8 | 69.2 | 66.2 | 58.4 | 56.7 | 69.5 |
|  | 52.5 | 66.8 | 60.5 | 57.7 | 52.7 | 66.7 |
|  | 49.8 | 68.2 | 64.3 | 57.7 | 55.0 | 61.7 |

Abbreviations: PC, Pearson Correlation; MSC, Magnitude Squared Coherence;  
IPC, Imaginary Part of Coherence; WC, Wavelet Coherence; PLV, Phase  
Locking Value; PLI, Phase Lag Index.

#### 4.2.6 References

- Al-Ezzi, A., Kamel, N., Faye, I., and Gunaseli, E. (2021). Analysis of default mode network in social anxiety disorder: Eeg resting-state effective connectivity study. 21. doi:10.3390/s21124098.
- Alves, C. L., Pineda, A. M., Roster, K., Thielemann, C., and Rodrigues, F. A. (2021). EEG functional connectivity and deep learning for automatic diagnosis of brain disorders: Alzheimer's disease and schizophrenia.
- Babiloni, F., Cincotti, F., Babiloni, C., Carducci, F., Mattia, D., Astolfi, L., . (2005). Estimation of the cortical functional connectivity with the multimodal integration of high-resolution EEG and fMRI data by directed transfer function. 24, 118–131. doi:10.1016/j.neuroimage.2004.09.036.
- Barry, R. J., Clarke, A. R., Johnstone, S. J., Magee, C. A., and Rushby, J. A. (2007). EEG differences between eyes-closed and eyes-open resting conditions. 118, 2765–2773.
- Bi, X., and Wang, H. (2019). Early Alzheimer's disease diagnosis based on EEG spectral images using deep learning. 114, 119–135.
- Blackburn, D. J., Sarrigiannis, P. G., De Marco, M., Zhao, Y., Venneri, A., Lawrence, S., . (2018). 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. 8, 1–19. doi:10.3390/brainsci8070134.
- Brookmeyer, R., Gray, S., and Kawas, C. (1998). Projections of Alzheimer's disease in the United States and the public health impact of delaying disease onset. 88, 1337–1342.
- Brookmeyer, R., Johnson, E., Ziegler-Graham, K., and Arrighi, H. M. (2007). Forecasting the global burden of Alzheimer's disease. 3, 186–191.

Bruna, J., Zaremba, W., Szlam, A., and LeCun, Y. (2013). Spectral networks and locally connected networks on graphs. .

Chen, X., Zheng, Y., Niu, Y., and Li, C. (2020). Epilepsy Classification for Mining Deeper Relationships between EEG Channels based on GCN. in , 701–706.

Craik, A., He, Y., and Contreras-Vidal, J. L. (2019). Deep learning for electroencephalogram (EEG) classification tasks: a review. 16, 31001.

Davies, C. A., Mann, D. M. A., Sumpter, P. Q., and Yates, P. O. (1987). A quantitative morphometric analysis of the neuronal and synaptic content of the frontal and temporal cortex in patients with Alzheimer's disease. 78, 151–164.

Deepthi, L. D., Shanthi, D., and Buvana, M. (2020). An intelligent Alzheimer's disease prediction using convolutional neural network (CNN). 11, 12–22.

Defferrard, M., Bresson, X., and Vandergheynst, P. (2016). Convolutional neural networks on graphs with fast localized spectral filtering. 29, 3844–3852.

DeKosky, S. T., and Scheff, S. W. (1990). Synapse loss in frontal cortex biopsies in Alzheimer's disease: correlation with cognitive severity. 27, 457–464.

Fan, M., Yang, A. C., Fuh, J.-L., and Chou, C.-A. (2018). Topological pattern recognition of severe Alzheimer's disease via regularized supervised learning of EEG complexity. 12, 685.

Franciotti, R., Falasca, N. W., Arnaldi, D., Famà, F., Babiloni, C., Onofrj, M., . (2019). Cortical Network Topology in Prodromal and Mild Dementia Due to Alzheimer's Disease: Graph Theory Applied to Resting State EEG. 32, 127–141. doi:10.1007/s10548-018-0674-3.



- Gallicchio, C., and Micheli, A. (2010). Graph echo state networks. in *Neural Networks for Signal Processing*, 1–8.
- Gehring, J., Auli, M., Grangier, D., Yarats, D., and Dauphin, Y. N. (2017). Convolutional sequence to sequence learning. in *Proceedings of the 34th International Conference on Machine Learning*, 1243–1252.
- Gori, M., Monfardini, G., and Scarselli, F. (2005). A new model for learning in graph domains. in *Proceedings of the 22nd International Conference on Machine Learning*, 729–734.
- Hammond, D. K., Vandergheynst, P., and Gribonval, R. (2011). Wavelets on graphs via spectral graph theory. *SIAM Review*, 30, 129–150.
- Henaff, M., Bruna, J., and LeCun, Y. (2015). Deep convolutional networks on graph-structured data. *arXiv preprint arXiv:1506.05161*.
- Huggins, C. J., Escudero, J., Parra, M. A., Scally, B., Anghinah, R., de Araújo, A. V. L., et al. (2021). Deep Learning of resting-state electroencephalogram signals for 3-class classification of Alzheimer’s Disease, Mild Cognitive Impairment and Healthy Ageing. *Frontiers in Aging Neuroscience*, 13, 645821.
- Ieracitano, C., Mammone, N., Bramanti, A., Hussain, A., and Morabito, F. C. (2019). A convolutional neural network approach for classification of dementia stages based on 2D-spectral representation of EEG recordings. *IEEE Transactions on Biomedical Engineering*, 323, 96–107.
- Ieracitano, C., Mammone, N., Hussain, A., and Morabito, F. C. (2020). A novel multi-modal machine learning based approach for automatic classification of EEG recordings in dementia. *IEEE Transactions on Biomedical Engineering*, 123, 176–190.
- Jalili, M. (2017). Graph theoretical analysis of Alzheimer’s disease: Discrimination of AD patients from healthy subjects. *Frontiers in Aging Neuroscience*, 384, 145–156.

Jelic, V. (2005). EEG in dementia, review of the past, view into the future. *Journal of Clinical Neurophysiology*, 24, 245–304.

Kipf, T. N., and Welling, M. (2016). Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.03903*.

Levie, R., Monti, F., Bresson, X., and Bronstein, M. M. (2018). Cayleynets: Graph convolutional neural networks with complex rational spectral filters. *Journal of Machine Learning Research*, 19, 97–109.

Li, F., Wang, J., Liao, Y., Yi, C., Jiang, Y., Si, Y., et al. (2019a). Differentiation of schizophrenia by combining the spatial EEG brain network patterns of rest and task P300. *Journal of Clinical Neurophysiology*, 27, 594–602.

Li, K., Wang, J., Li, S., Yu, H., Zhu, L., Liu, J., et al. (2021a). Feature Extraction and Identification of Alzheimer's Disease based on Latent Factor of Multi-Channel EEG. *Journal of Clinical Neurophysiology*, 29, 1557–1567.

Li, P., Liu, H., Si, Y., Li, C., Li, F., Zhu, X., et al. (2019b). EEG based emotion recognition by combining functional connectivity network and local activations. *Journal of Clinical Neurophysiology*, 66, 2869–2881.

Li, Y., Liu, Y., Guo, Y.-Z., Liao, X.-F., Hu, B., and Yu, T. (2021b). Spatio-Temporal-Spectral Hierarchical Graph Convolutional Network With Semisupervised Active Learning for Patient-Specific Seizure Prediction. *Journal of Clinical Neurophysiology*.

Masliah, E., Mallory, M., Alford, M., DeTeresa, R., Hansen, L. A., McKeel, D. W., et al. (2001). Altered expression of synaptic proteins occurs early during progression of Alzheimer's disease. *Journal of Clinical Neurophysiology*, 56, 127–129.

Michel, C. M., Koenig, T., Brandeis, D., Gianotti, L. R. R., and Wackermann, J. (2009). *Journal of Clinical Neurophysiology*. Cambridge University Press.

Mormann, F., Lehnertz, K., David, P., and Elger, C. E. (2000). Mean phase coherence as a measure for phase synchronization and its application to the EEG of epilepsy patients. *Journal of Neuroscience*, 20(14), 358–369.

Niepert, M., Ahmed, M., and Kutzkov, K. (2016). Learning convolutional neural networks for graphs. in *Proceedings of the 33rd International Conference on Machine Learning*, 2014–2023.

Nobukawa, S., Yamanishi, T., Kasakawa, S., Nishimura, H., Kikuchi, M., and Takahashi, T. (2020). Classification methods based on complexity and synchronization of electroencephalography signals in Alzheimer's disease. *Frontiers in Aging Neuroscience*, 11, 255.

Oltu, B., Ak\csahin, M. F., and Kibaro\uglu, S. (2021). A novel electroencephalography based approach for Alzheimer's disease and mild cognitive impairment detection. *Frontiers in Aging Neuroscience*, 13, 102223.

Organization, W. H. Dementia.

Rodrigues, P. M., Bispo, B. C., Garrett, C., Alves, D., Teixeira, J. P., and Freitas, D. (2021). Lacsogram: a New EEG Tool to Diagnose Alzheimer's Disease. *Frontiers in Aging Neuroscience*, 13, 102223.

Safi, M. S., and Safi, S. M. M. (2021). Early detection of Alzheimer's disease from EEG signals using Hjorth parameters. *Frontiers in Aging Neuroscience*, 13, 102338.

Sakkalis, V. (2011). Review of advanced techniques for the estimation of brain connectivity measured with EEG/MEG. *Frontiers in Aging Neuroscience*, 13, 1110–1117. doi:10.1016/j.compbimed.2011.06.020.

Scarselli, F., Gori, M., Tsoi, A. C., Hagenbuchner, M., and Monfardini, G. (2008). The graph neural network model. *Frontiers in Aging Neuroscience*, 10, 61–80.

Scheff, S. W., Price, D. A., Schmitt, F. A., DeKosky, S. T., and Mufson, E. J. (2007). Synaptic alterations in CA1 in mild Alzheimer disease and mild cognitive impairment. *Annals of the New York Academy of Sciences*, 68, 1501–1508.

Scheff, S. W., Price, D. A., Schmitt, F. A., and Mufson, E. J. (2006). Hippocampal synaptic loss in early Alzheimer's disease and mild cognitive impairment. *Annals of the New York Academy of Sciences*, 27, 1372–1384.

Shuman, D. I., Narang, S. K., Frossard, P., Ortega, A., and Vandergheynst, P. (2013). The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains. *IEEE Signal Processing Magazine*, 30, 83–98.

Si, Y., Wu, X., Li, F., Zhang, L., Duan, K., Li, P., et al. (2019). Different decision-making responses occupy different brain networks for information processing: a study based on EEG and TMS. *NeuroImage*, 29, 4119–4129.

Smailovic, U., Koenig, T., Kåreholt, I., Andersson, T., Kramberger, M. G., Winblad, B., et al. (2018). Quantitative EEG power and synchronization correlate with Alzheimer's disease CSF biomarkers. *NeuroImage*, 63, 88–95.

Song, Z., Deng, B., Wang, J., and Wang, R. (2018). Biomarkers for Alzheimer's disease defined by a novel brain functional network measure. *NeuroImage*, 66, 41–49.

Sperduti, A., and Starita, A. (1997). Supervised neural networks for the classification of structures. *Neural Networks*, 8, 714–735.

Tafreshi, T. F., Daliri, M. R., and Ghodousi, M. (2019). Functional and effective connectivity based features of EEG signals for object recognition. *NeuroImage*, 13, 555–566. doi:10.1007/s11571-019-09556-7.

Tait, L., Tamagnini, F., Stothart, G., Barvas, E., Monaldini, C., Frusciante, R., et al. (2020). EEG microstate complexity for aiding early diagnosis of Alzheimer's disease. *NeuroImage*, 10, 1–10.

- Tavares, G., San-Martin, R., Ianof, J. N., Anghinah, R., and Fraga, F. J. (2019). Improvement in the automatic classification of Alzheimer's disease using EEG after feature selection. in *2019 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 1264–1269.
- Terry, R. D., Masliah, E., Salmon, D. P., Butters, N., DeTeresa, R., Hill, R., et al. (1991). Physical basis of cognitive alterations in Alzheimer's disease: synapse loss is the major correlate of cognitive impairment. *Annals of the New York Academy of Sciences*, 30, 572–580.
- Trambaiolli, L. R., Spolaôr, N., Lorena, A. C., Anghinah, R., and Sato, J. R. (2017). Feature selection before EEG classification supports the diagnosis of Alzheimer's disease. *NeuroImage*, 128, 2058–2067.
- Triggiani, A. I., Bevilacqua, V., Brunetti, A., Lizio, R., Tattoli, G., Cassano, F., et al. (2017). Classification of healthy subjects and Alzheimer's disease patients with Dementia from cortical sources of resting state EEG rhythms: A study using artificial neural networks. *NeuroImage*, 10, 604.
- Tylová, L., Kukal, J., Hubata-Vacek, V., and Vyšata, O. (2018). Unbiased estimation of permutation entropy in eeg analysis for alzheimer's disease classification. *Entropy*, 39, 424–430.
- Van Diessen, E., Numan, T., Van Dellen, E., Van Der Kooi, A. W., Boersma, M., Hofman, D., et al. (2015). Opportunities and methodological challenges in EEG and MEG resting state functional brain network research. *NeuroImage*, 126, 1468–1481.
- Van Mierlo, P., Papadopoulou, M., Carrette, E., Boon, P., Vandenberghe, S., Vonck, K., et al. (2014). Functional brain connectivity from EEG in epilepsy: Seizure prediction and epileptogenic focus localization. *NeuroImage*, 121, 19–35.
- Wan, L., Huang, H., Schwab, N., Tanner, J., Rajan, A., Lam, N. B., et al. (2019). From eyes-closed to eyes-open: Role of cholinergic projections in EC-to-EO

alpha reactivity revealed by combining EEG and MRI. 40, 566–577.

Wendling, F., Ansari-Asl, K., Bartolomei, F., and Senhadji, L. (2009). From EEG signals to brain connectivity: a model-based evaluation of interdependence measures. 183, 9–18.

Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., and Philip, S. Y. (2020). A comprehensive survey on graph neural networks. 32, 4–24.

Yu, B., Yin, H., and Zhu, Z. (2017). Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting.

Yu, H., Lei, X., Song, Z., Liu, C., and Wang, J. (2019). Supervised network-based fuzzy learning of EEG signals for Alzheimer's disease identification. 28, 60–71.

Zeng, D., Huang, K., Xu, C., Shen, H., and Chen, Z. (2020). Hierarchy graph convolution network and tree classification for epileptic detection on electroencephalography signals.

Zhang, S., Chen, D., Tang, Y., and Zhang, L. (2021). Children ASD Evaluation Through Joint Analysis of EEG and Eye-Tracking Recordings With Graph Convolution Network. 15.

Zhao, Y., Dong, C., Zhang, G., Wang, Y., Chen, X., Jia, W., . (2021). EEG-Based Seizure detection using linear graph convolution network with focal loss. 208, 106277.

Zhao, Y., Zhao, Y., Durongbhan, P., Chen, L., Liu, J., Billings, S. A., . (2019). Imaging of nonlinear and dynamic functional brain connectivity based on EEG recordings with the application on the diagnosis of Alzheimer's disease. 39, 1571–1581.

## 5 The Application of Functional Connectivity on Mental workload Classification

This Chapter is based a journal paper, 'EEG/fNIRS Based Workload Classification Using Functional Brain Connectivity and Machine Learning', published in .

Cao, J., Garro, E.M., Zhao, Y., 2022.

. Sensors 22, 7623.

### 5.1 Introduction

Mental workload refers to the amount of working memory required to complete a task in a specified time. Its assessment has attracted many researchers, and workload has been characterised by a variety of physiological sensor data. Investigation of mental workload in neuroscience is significant for a variety of reasons. First, a person's high cognitive workload will affect learning capacity and cause distraction (Anderson ., 2011). Second, since there is a limit to the size of a cognitive workload, there is also a limit to an individual's performance in a given cognitive activity (Bashivan ., 2016). As a result, assessing mental workload is important for preventing accidents in many areas (Zhang and Li, 2017). Table 1 compares various popular neuroimaging modalities for evaluating mental workloads: such as functional near-infrared spectroscopy (fNIRS), electroencephalography (EEG)/Magnetoencephalography (MEG), functional magnetic resonance imaging (fMRI), and position emission tomography (PET).

Because it has the advantages of low cost and high temporal sampling rate, EEG has been well-accepted in the field of disease prediction (Hussain and Park, 2021), sleep stages (Hussain ., 2022), and brain stimulation for different neurological workloads (Hussain ., 2021) as well as mental workload evaluation (Kakkos ., 2021). A substantial number of studies have reported a significant EEG spectral correlation with workload in stereotypical frequency bands: such as delta (1–4 Hz), theta (4–7 Hz), alpha (8–15 Hz), and beta (16–31 Hz) (Aghajani ., 2017; Kuanar ., 2018; Pfeifer ., 2018). Several popular machine learning methods have been applied using EEG

features such as support vector machine (SVM) (So et al., 2017), naive bayes (Zhang and Li, 2017) and linear discriminant analysis (LDA) (Maimon et al., 2020; Mühl et al., 2014). Although those methods achieved satisfactory mental workload classification results, it was notable that most EEG-based features were extracted from a single channel that was univariate-based and neglected association between channels. As a multivariate approach, functional brain connectivity (FBC) is statistically interdependent among spatially distant neurophysiological regions (Cao et al., 2022, 2021; Shan et al., 2021). It has been proven that FBC reveals the underlying function of different brain regions and their complex cortical intercommunication, which helps improve understanding of many neurological conditions including brain-related disorders and emotions (Cao et al., 2021; Shan et al., 2021; Zhao et al., 2020). Kakkos et al. (2021) fed univariate spectrum power features and FBC estimations from EEG into several machine-learning classifiers and achieved promising results in two-level workload discrimination. However, the potential of FBC in multiclass workload classification problems, particularly in combination with other sensing modalities, has not been fully explored.

**Table 5-1 Comparison of four neuroimaging techniques.**

| <b>Specification</b>           | <b>fNIRS</b> | <b>EEG/MEG</b>                               | <b>fMRI</b>     | <b>PET</b> |
|--------------------------------|--------------|----------------------------------------------|-----------------|------------|
| <b>Spatial resolution</b>      | 2–3 cm       | 5–9 cm                                       | 0.3 mm voxels   | 4 mm       |
| <b>Penetration depth</b>       | Brain cortex | Brain cortex for EEG/deep structures for MEG | Whole head      | Whole head |
| <b>Temporal sampling rates</b> |              | >1000 Hz                                     | 1–3 Hz          | <0.1 Hz    |
| <b>Range of possible tasks</b> | Enormous     | Limited                                      | Limited         | Limited    |
| <b>Robustness to motion</b>    | Very good    | Limited                                      | Limited         | Limited    |
| <b>Range of</b>                | Everyone     | Everyone                                     | Limited, can be | Limited    |



| possible participants |                           | challenging for children/patients |            |        |
|-----------------------|---------------------------|-----------------------------------|------------|--------|
| <b>Sounds</b>         | Silent                    | Silent                            | Very noisy | Silent |
| <b>Portability</b>    | Yes, for portable systems | Yes, for portable EEG systems     | None       | None   |
| <b>Cost</b>           | Low                       | Low for EEG; high for MEG         | High       | High   |

In recent decades, fNIRS has grown rapidly as a tool for monitoring functional brain activity in a wide range of applications and populations. fNIRS devices detect two hemodynamic signals, oxygenated (HbO) and deoxygenated (HbR) hemoglobin, from the cortical surface at a spatial resolution of 2–3 cm (Alhudhaif, 2021; Pinti et al., 2020; Zaman and Md. Rabiul Islam, 2021). One of the main reasons for the increased interest in using fNIRS for cognitive activities is that it is resistant to motion artefacts (Herold et al., 2018), which is usually a big problem for EEG data acquisition. Furthermore, fNIRS can be more precise in brain activation areas due to its relatively high spatial resolution. As a result, fNIRS overcomes some shortcomings of EEG. The importance of including both HbO and HbR for analysis has been emphasised by a few studies because their combination provides a more comprehensive assessment of cortical activation (Herff et al., 2014; Lim et al., 2020; Mandrick et al., 2016; Saadati et al., 2019). The majority of related studies has focused on using mean values (Asgher et al., 2019; Putze et al., 2019), standard deviation (Asgher et al., 2019) and slope (Herff et al., 2014; Lim et al., 2020).

Some researchers explored the effectiveness of using both EEG and fNIRS information for n-back workload classification. Liu et al. (2016) employed LDA and obtained 68.1% classification accuracy in the n-back working-memory task using a combined EEG–fNIRS approach, but it used univariate features based on a single channel. Saadati et al. (2020) used deep neural networks and hybrid EEG–fNIRS features. It was claimed that the classification accuracy is considerably higher than that of EEG or fNIRS alone. However, there is very limited research on EEG brain connectivity combined with fNIRS, so the

potential of using both signals to discriminate multi-level workloads requires further exploration.

In this Chapter, we propose a hybrid EEG–fNIRS approach to discriminate among multi-level mental workloads: univariate frequency and bivariate FBC features are extracted from EEG, and biomarkers of HbO and HbR are estimated from fNIRS. Overall, combining EEG and fNIRS tended to provide two distinct sources of information on the brain including electrical activity and hemodynamic responses; this combination has the benefits of non-invasiveness, robustness to motion, availability for all possible participants, silence, portability, and cost-effectiveness. The novelty of this Chapter is summarised in four folds:

To the best of the authors' knowledge, this Chapter is the first to use combined features of EEG-based FBC and fNIRS for workload estimation.

This Chapter explores different linear and nonlinear FBC representations in the time and frequency domains with their associated effect on classification accuracy.

This Chapter reports the contribution of different regions to the classification accuracy of the two sensing modalities.

Topographic and heat maps were used to reveal distinct areas where the greatest change occurred at different workload levels.

## **5.2 Materials and Methods**

As shown in Figure 5-1, the proposed framework contains four main steps: data pre-processing, feature extraction, feature selection, and machine-learning classification. The detail of each step is as follows:

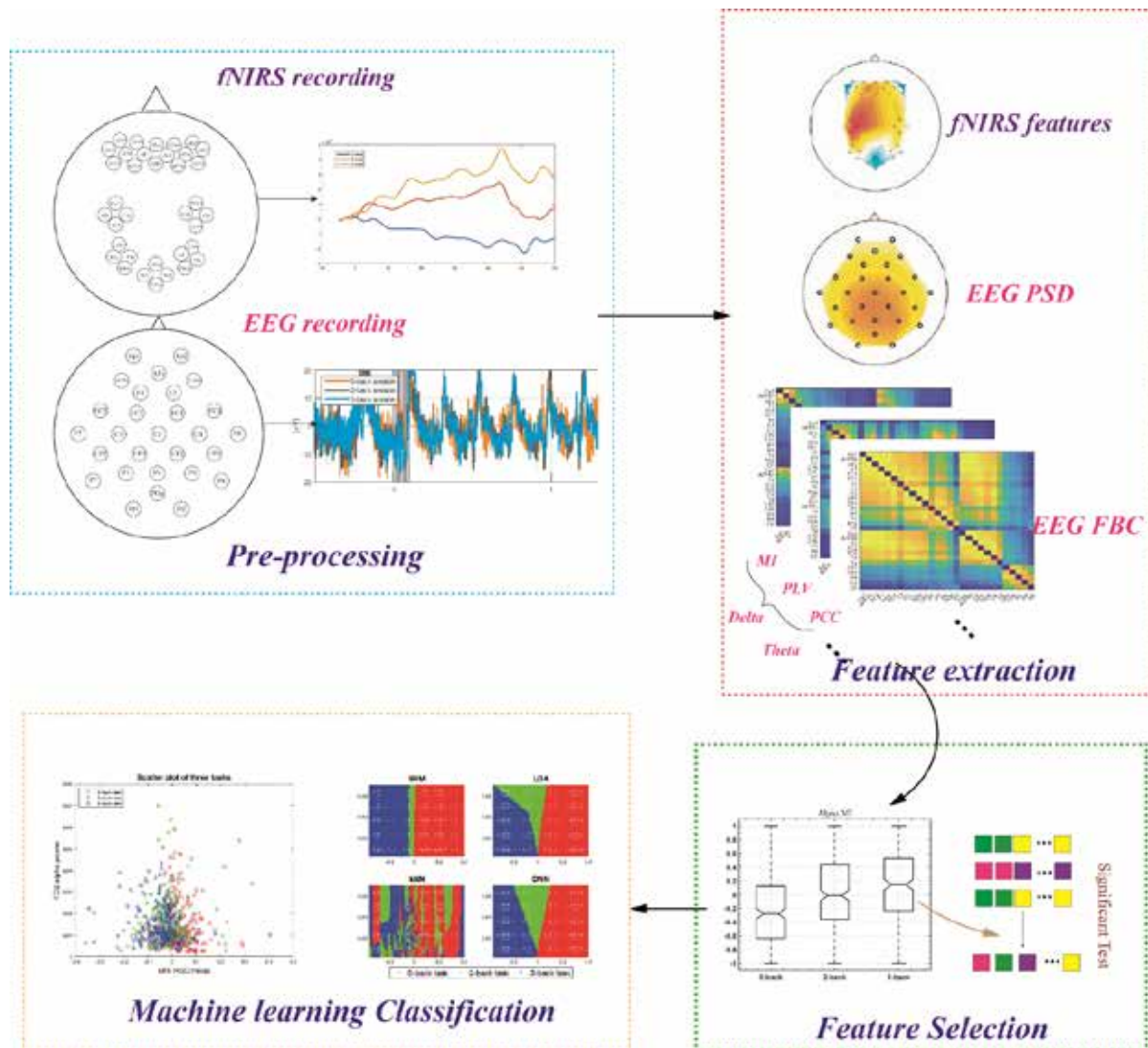


**Figure 5-1** Flowchart of the proposed framework. The pipeline contains four main steps: pre-processing, feature extraction, feature selection and machine-learning classification.

### 5.2.1 Dataset

This Chapter made use of a dataset gathered by Shin . (2018) at the Technische Universität Berlin. The dataset comprised scalp recordings—30 EEG channels and 36 fNIRS channels—for mental workload during n-back tasks. The channels and their locations are shown in Figure 5-9 in Supplementary. These activities were divided into four categories: 0-, 2-, and 3-back tasks, as well as rest between tasks. Twenty-six healthy, right-handed people took part, and the dataset was divided into three sessions, each with three randomly organized sets of 0-, 2-, and 3-back tasks, meaning that each

participant completed nine sets of n-back tasks. A single task consisted of a 2 s instruction indicating the type of task (0-, 2-, or 3-back), a 40 s task period that consists of 20 trials, a 1 s stop period, and a 20 s rest period (see Figure 5-2). Therefore, there were  $26 \times 3 \times 9 = 702$  tasks available for all participants. All EEG and fNIRS signals were captured at the same time.

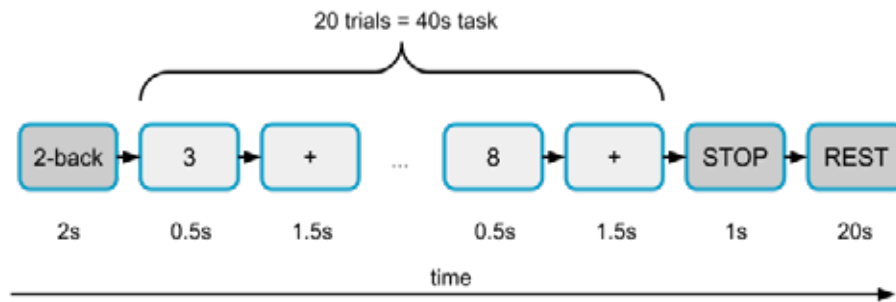


**Figure 5-2** Layout of a set in the experiment. A single task consisted of a 2 s instruction indicating the type of task (0-, 2-, or 3-back), a 40 s task period that consisted of 20 trials, a 1 s stop period, and a 20 s rest period. Each participant completed nine sets of n-back tasks.

## 5.2.2 Signal pre-processing and Feature extraction

### 5.2.2.1 fNIRS

fNIRS data were pre-processed using the BBCI toolbox in MATLAB R2019b (Blankertz et al., 2010). The sampling rate was 10Hz. Initially, HbR and HbO values were calculated using the modified Beer–Lambert equation (mBLL) from the fNIRS optical density (Khan et al., 2014). A sample of HbR and HbO values for each participant is shown in Supplementary Figure 5-10 and Figure 5-11. Data augmentation was performed to create small informative segments. To reduce noise and artefacts, fNIRS signals were passed through a third-order digital Butterworth filter between 0 and 0.04 Hz. Additionally, baseline correction was applied to the fNIRS signals to remove the intra-individual variance of the starting values. In this step, the segments were normalised by subtracting the median value of the pre-stimulus baseline from the signal in each segment (Pfeifer et al., 2018).

It should be noted that there was a general 6 s delay between the stimulus representation and peak cortical hemodynamics. This delay was determined by the task and HbR and HbO concentrations. Normally, the cerebral hemodynamic response does not return to baseline until 10 s after stimulus presentation. However, agreement on an ideal time window for analysis had yet to be reached because the best temporal length depended on the paradigm used and participant characteristics, such as age (Herold et al., 2018). This Chapter conducted a sensitivity analysis to identify the optimal time window to produce the most accurate mental workload estimate and a size of 5s was used. The window slides through the whole 40 s period with a 1 s step. This analysis was performed independently for each participant.

#### **5.2.2.2 EEG**

EEG data were also pre-processed using the BBCI toolbox in MATLAB R2019b, and resampling was done at 200 Hz. The improved weight-adjusted second-order blind identification (iWASOBI) method in the automatic artifact removal (AAR) toolbox in EEGLAB was used to gain ocular artifact rejection. Initially, data augmentation was done by segmenting data samples into smaller but still informative segments. Then, the data were bandpass-filtered between 1 and 45 Hz using a third-order Butterworth digital filter. The EEG epochs were extracted

spectral density (PSD) was calculated for three frequency bands of EEG recordings: delta (0.5–4 Hz), theta (4–7 Hz), and alpha (8–15 Hz) since previous studies indicated that low-frequency information made more contributions for measuring mental workload (Brouwer et al., 2012; Kakkos et al., 2021). The FBC was estimated using four methods: Pearson correlation coefficient (PCC), mutual information (MI) in the time domain, magnitude squared coherence (MSC), and phase-locking value (PLV) in the frequency domain. The principal details are given as follows:

The PCC was able to evaluate the linear interdependency between two signals in the time domain. The PCC between signals  $x$  and  $y$  were

$$\rho_{xy} = \frac{E[(x - \mu_x)(y - \mu_y)]}{\sigma_x \sigma_y} \quad (5-1)$$

where  $E$  is the expected value;  $\mu_x$  and  $\mu_y$  are the mean values; and  $\sigma_x$  and  $\sigma_y$  are the standard deviations of the  $x$  and  $y$  time series.

MSC is a linear method to estimate interconnections between two signals in the frequency domain calculated by PSD. The MSC of signals  $x$  and  $y$  can be

$$\text{MSC}_{xy}(f) = C_{xy}^2 = \frac{S_{xy}(f)^2}{|S_{xx}(f)| \times |S_{yy}(f)|} \quad (5-2)$$

written as

where  $S_{xx}(f)$  and  $S_{yy}(f)$  are the PSDs of signals  $x$  and  $y$ , respectively; and  $S_{xy}(f)$  is the cross PSD at frequency  $f$ .

According to information theory, the MI of two random variables,  $x$  and  $y$ , shows how one is informative for the other one. Let,  $P(x)$  and  $P(y)$  be the probability distributions of random variables  $x$  and  $y$ , respectively. The entropy of  $x$  and  $y$  is defined as

$$H(x) = \sum_{j=1}^N P(x_j) \log_b(P(x_j)) \quad (5-3)$$

$$H(y) = \sum_{j=1}^N P(y_j) \log_b(P(y_j)) \quad (5-4)$$

where  $N$  defines window length.  $H(y|x)$  and  $H(x,y)$  represent conditional entropy and joint entropy between  $x$  and  $y$ , defined respectively as

$$H(x,y) = E_x[E_y[\log_b P(x,y)]] \quad (5-5)$$

$$H(y|x) = E_x[E_y[\log_b P(y|x)]] \quad (5-6)$$

where  $\bar{\cdot}$  is the expected value function. The MI of two random variables  $x$  and  $y$  is computed as follows:

$$MI(x, y) = H(x) + H(y) - H(x, y) = H(y) - H(y|x) \quad (5-7)$$

$MI(x, y) = 0$  if and only if random variables  $X$  and  $Y$  are statistically independent. Notably, the MI is a nonlinear method in the time domain,

Phase synchronisation (PS) assumes that two oscillation systems without amplitude synchronisation can have phase synchronisation. The phase locking value (PLV) is frequently used to obtain the phase synchronisation strength (Cao et al., 2022). The instantaneous phase of a signal  $X$  is given by:

$$\phi_x(t) = \arctan \frac{\tilde{x}(t)}{x(t)} \quad (5-8)$$

where  $\tilde{x}(t)$  is the Hilbert transform of  $x(t)$  which is defined as

$$\tilde{x}(t) = \frac{1}{\pi} PV \int_{-\infty}^{+\infty} \frac{x(\tau)}{t - \tau} d\tau \quad (5-9)$$

where  $PV$  refers to the Cauchy principal value. The PLV for two signals is defined as

$$PLV = \left| \frac{1}{N} \sum_{j=0}^{N-1} e^{j(\phi_x(j \cdot t) - \phi_y(j \cdot t))} \right| \quad (5-10)$$

where  $t$  defines the sampling period, and  $N$  indicates the sample number of each signal (Mormann et al., 2000). The range of PLV was from 0 to 1, where 0 showed a lack of synchronisation and 1 indicated strict phase synchronisation. Notably, the PLV is a nonlinear method in the frequency domain.

### 5.2.3 Feature Selection and Fusion

A large number of features were extracted from EEG and fNIRS. To be more specific, considering three frequency bands (delta, theta and alpha), 28 channels and four FBC methods, there were  $3 \times 28 = 84$  PSD features and  $3 \times 28 = 84$  PLV features from the EEG recording.

According to the time window analysis of the fNIRS signals, the top-10 best time windows were chosen. Considering the number of channels, there were  $10 \times 36 = 360$  features for fNIRS. The next step was to feed the extracted features into machine learning classifiers to classify three workloads/tasks. To avoid the overfitting problem of machine learning and compare the combined methods fairly with the methods using a single type of feature, a statistical significance test is used to reduce the feature number. One-way analysis of variance (ANOVA) was used to evaluate the significance of differences in the 0-back vs. 2-back vs. 3-back features. The  $p$ -value was the criterion for selecting the significant features. As a result, the top-10 features with the smallest  $p$ -values were individually selected from EEG-based and fNIRS-based techniques as classifier input. Furthermore, the top-5 features from EEG (univariate features only) and fNIRS, respectively, were combined, resulting in 10 hybrid features for comparison purposes.

#### **5.2.4 Machine-Learning Classification**

The SVM was applied to achieve workload classification. It constructed an optimal separating hyperplane in the feature space based on the structural risk minimization principle. The selected features extracted from EEG and fNIRS were fed into the SVM with a radial basis function (RBF) kernel. Different machine-learning algorithms were tested and compared, such as the k-nearest neighbour (KNN), decision tree and LDA. The SVM outperformed other methods in classification. Hence, this Chapter mainly used the SVM with RBF to represent classification results. To avoid overfitting in the case of limited data, a five-fold cross-validation technique was employed. To be more specific, the dataset of each condition was divided into five subsets, and then five iterations were undertaken to ensure each subset was used for training and testing (Cao et al., 2021). That is to say, for each iteration, 80% of the dataset was used for training and the remaining 20% for testing. Consequently, the classification result was calculated by averaging the accuracies from 5 iterations. Totally, there were  $3 \text{ workload levels} \times 3 \text{ series} \times 3 \text{ sessions} \times 26 \text{ participants} = 702$



samples. Before being fed into the classifier, the features were normalised from -1 to 1 for each participant to reduce the influence of individual differences.

## 5.3 Results

### 5.3.1 Time Interval Selection

The selection of the time interval relied on the classification performance implemented on each participant. Figure 5-3(A) represents the mean classification accuracy for all participants using fNIRS-based features against the moving time window, and sustained growth was observed during the first 30 s. After a 25 s oscillation, the accuracy reached its peak when the 45–50 s time window was used. Consequently, the 10 time-windows in the range of 45 to 54 s were selected for the next step of feature extraction. Figure 5-3 (B)(C)(D) illustrates the changes in classification accuracy against the length of the time

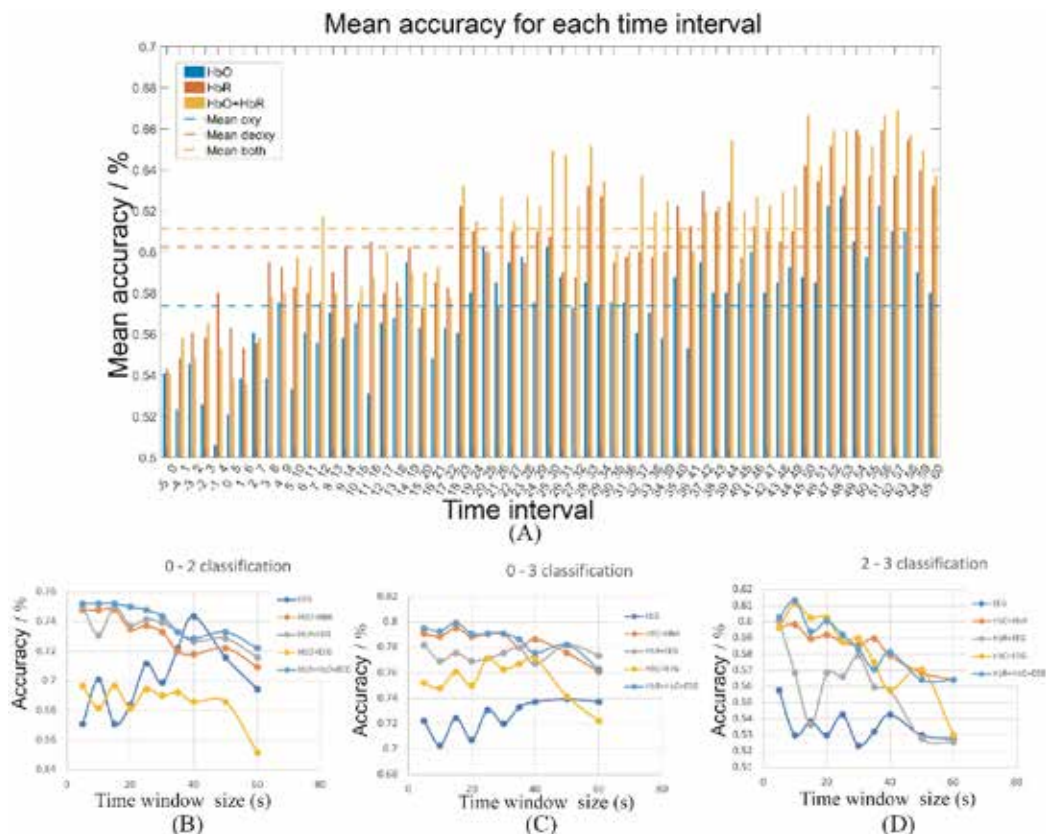


**Figure 5-3 Time window analysis (A) Time interval analysis for fNIRS features; Time window-size evaluation for EEG and fNIRS features for (B) 0-back vs. 2-back, (C) 0-back vs. 3-back, (D) 2-back vs. 3-back.**

window of fNIRS and EEG. Notably, the accuracy of using fNIRS-based features decreased along with the window-size increment for all three classification groups. However, the EEG-based method performed better following the window-size increment and peaks at 40 s, particularly for 0-back vs. 2-back and 0-back vs. 3-back. As a result, the final window-size selections for fNIRS and EEG were 5 and 40 s, respectively. Furthermore, it indicated that the fused features outperformed features from a single modality.

### 5.3.2 Machine-Learning Classification Performance

To select the optimal FBC features, four different methods (MI, PCC, MSC and PLV) were tested individually, and the nonlinear time-domain method, MI, was found to provide the highest classification accuracy. Figure 5-4 shows the comparison of the four estimations in the three bands for top-10 average classification accuracy. The error bar shows the accuracy from each iteration of cross-validation. Therefore, MI was selected as the EEG-based FBC feature for the following analysis.

To classify multi-level mental workload, the classification task was separated into three groups: 0-back vs. 2-back, 0-back vs. 3-back and 2-back vs. 3-back. The performances using EEG-based features only, fNIRS-based features only, and hybrid features were evaluated and shown in Table 5-2, Table 5-3 and

|                   |       | EEG   |     | fNIRS |     | EEG + fNIRS |
|-------------------|-------|-------|-----|-------|-----|-------------|
|                   |       | PSD   | FBC | HBO   | HBR |             |
| <b>0-back vs.</b> | Delta | 65%   | 63% |       |     | 74%         |
|                   | Theta | 69%   | 72% | 62%   | 72% | 75%         |
| <b>3-back</b>     |       | Alpha | 71% | 77%   |     | <b>83%</b>  |

Table 5-4. To ensure classification fairness, each classification task used 10 features as the input. The features were selected according to the significance test and a sample is given in Supplementary Figure 5-12. The EEG alpha band information had the best performance in discriminating the three workload levels

for both univariate (PSD) and bivariate features (FBC). Meanwhile, the results also suggested that the FBC features performed better with an approximately 5% accuracy increment for all three sub-tasks. When it came to fNIRS, HbR outperformed HbO, but the accuracies were both significantly lower than for EEG-based FBC features, particularly for 0-back vs. 2-back and 0-back vs. 3-back. Other references suggested that classifiers, such as LDA, SVM and CNN, achieved higher accuracy using HbR indicators (Alhudhaif, 2021; Zaman and Md. Rabiul Islam, 2021).

Overall, the fused features (EEG–fNIRS) improved classification performance. For the 0-back vs. 2-back and 0-back vs. 3-back tasks, the hybrid method obtains the highest accuracy with 77% (Table 2) and 83% (Table 3), which suggests, as expected, there is more difference between 0-back and 3-back than between 0-back and 2-back. However, the difference between 2-back and 3-back was small, as evident by a much lower accuracy. Notably, the results suggested that the hybrid features did not have superiority in all tasks. As shown in

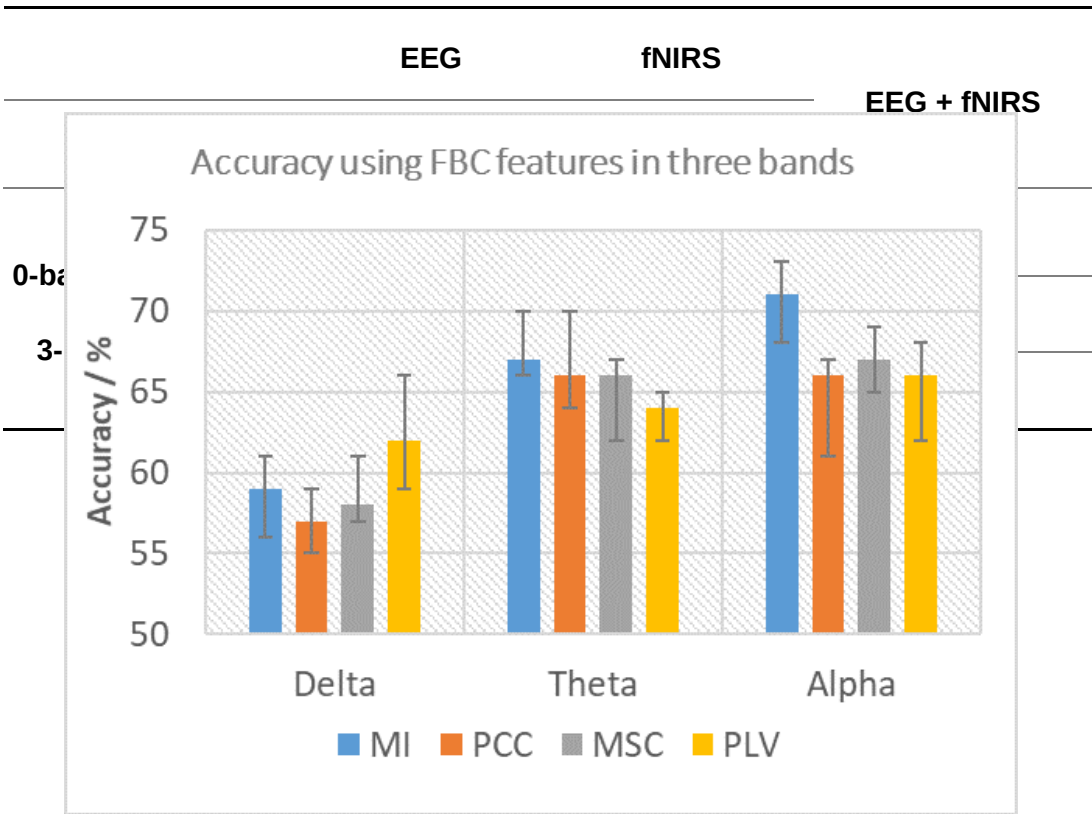


**Figure 5-4 Comparison of four FBC estimations (MI, PCC, MSC and PLV) in terms of the average of the Top 10 classification accuracies along with maximum and minimum value.**

Table 5-4, the FBC features in the alpha band had the best performance (62%) but the fused features had only 59%. Nevertheless, 2-back and 3-back were difficult to distinguish for any features.

**Table 5-2 SVM classification accuracy of 0-back vs. 2-back using different features.**

|                      |       | EEG |     | fNIRS |     | EEG + fNIRS |
|----------------------|-------|-----|-----|-------|-----|-------------|
|                      |       | PSD | FBC | HbO   | HbR |             |
| 0-back vs.<br>2-back | Delta | 66% | 67% |       |     | 72%         |
|                      | Theta | 68% | 73% | 62%   | 68% | 75%         |
|                      | Alpha | 70% | 74% |       |     | 77%         |

**Table 5-3 SVM classification accuracy of 0-back vs. 3-back using different features.**

|                      |       | EEG |     | fNIRS |     | EEG + fNIRS |
|----------------------|-------|-----|-----|-------|-----|-------------|
|                      |       | PSD | FBC | HBO   | HBR |             |
| 0-back vs.<br>3-back | Delta | 65% | 63% |       |     | 74%         |
|                      | Theta | 69% | 72% | 62%   | 72% | 75%         |
|                      | Alpha | 71% | 77% |       |     | 83%         |

**Table 5-4 SVM classification accuracy of 2-back vs. 3-back using different features.**

|  |  | EEG |     | fNIRS |     | EEG + fNIRS |
|--|--|-----|-----|-------|-----|-------------|
|  |  | PSD | FBC | HBO   | HBR |             |

|                   |       |     |            |     |     |     |
|-------------------|-------|-----|------------|-----|-----|-----|
| <b>2-back vs.</b> | Delta | 52% | 60%        |     |     | 57% |
|                   | Theta | 56% | 61%        | 60% | 61% | 58% |
| <b>3-back</b>     | Alpha | 55% | <b>62%</b> |     |     | 59% |

To further evaluate the machine learning algorithms performance, accuracy (     ), sensitivity (     ) and specificity (     ) were calculated:

$$Accu = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (5-11)$$

$$Sens = \frac{TP}{TP + FN} \times 100\% \quad (5-12)$$

$$Spec = \frac{TN}{TN + FP} \times 100\% \quad (5-13)$$

where TP = True Positive; FN = False Negative; TN = True Negative; and FP = False Positive. Moreover, the receiver operating characteristic (ROC) curve, and the area under the ROC curve (AUC) (Lotte     ., 2018; Pyrzowski     ., 2015) were used to assess the goodness of classification. Specifically, the ROC was constructed from the true positive rate (TPR = sensitivity) in the vertical axis and the false positive rate (FPR = 1-specificity) in the horizontal axis (Blinowska     ., 2017). The resulting accuracy, sensitivity , specificity and AUC are shown in Table 5-5. The ROC curves for three binary classification tasks is shown in Figure 5-5

**Table 5-5 Performance of classification for 3 binary classification tasks.**

| <b>Alpha Hybrid Features</b> | <b>Accuracy</b> | <b>Specificity</b> | <b>Sensitivity</b> | <b>AUC</b> |
|------------------------------|-----------------|--------------------|--------------------|------------|
| 0-back vs. 2-back            | 77%             | 79%                | 76%                | 0.8332     |
| 0-back vs. 3-back            | 83%             | 84%                | 80%                | 0.9501     |
| 2-back vs. 3-back            | 59%             | 57%                | 63%                | 0.6721     |

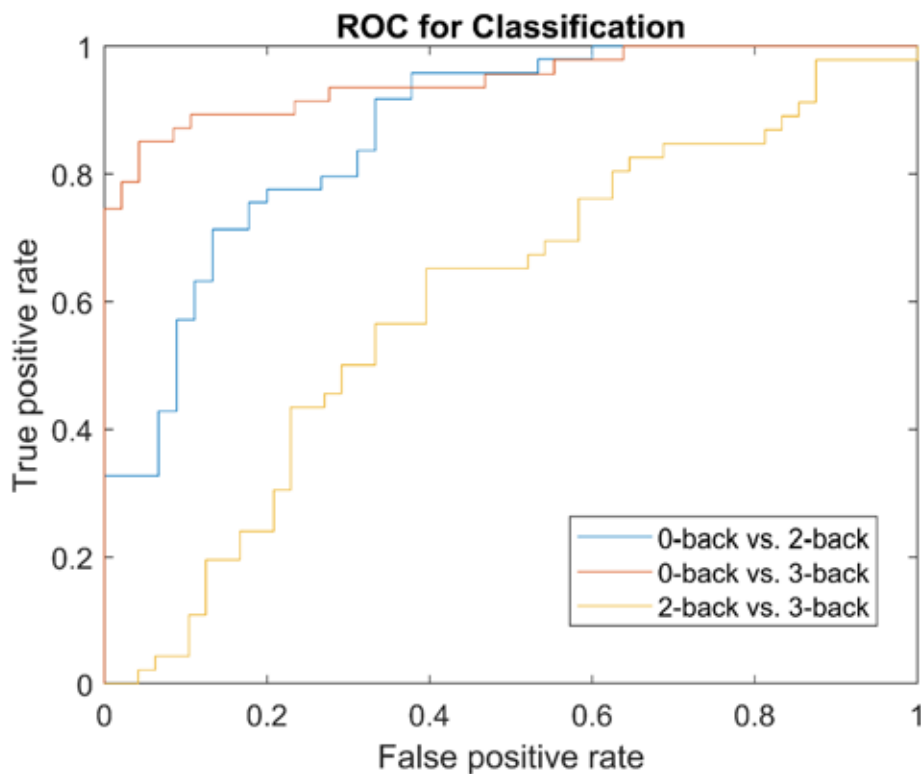


**Figure 5-5** The receiver operating characteristic (ROC) curves for three binary classification tasks: 0-back vs. 2-back, 0-back vs. 3-back, and 2-back vs. 3-back.

### 5.3.3 Visualisation

To further explore the difference among the three workload levels, a distinct visualisation method was employed. A topographic map was used to represent the PSD distribution of the EEG alpha-band (Figure 5-6), which provided about 70% classification accuracy for the 0-back vs. 2-back and 0-back vs. 3-back tasks. The averaged PSD distribution across all participants, illustrated by the left column, suggested that the posterior area of 0-back had much higher PSD than 2-back and 3-back, while other areas had similar PSD distribution. It seemed that, during the low workload level, there was more brain activity in the alpha band in the posterior area than during high workload. It matched the classification result, which revealed that the posterior midline occipital (POz), left occipital (O1) and right occipital (O2) channels contributed more than the others. The patterns of 2-back and 3-back are very similar for the whole brain, which explains the low classification accuracy (55%) in

|                   |       | EEG |     | fNIRS |     | EEG + fNIRS |
|-------------------|-------|-----|-----|-------|-----|-------------|
|                   |       | PSD | FBC | HBO   | HBR |             |
| <b>0-back vs.</b> | Delta | 65% | 63% |       |     | 74%         |
|                   | Theta | 69% | 72% | 62%   | 72% | 75%         |
| <b>3-back</b>     | Alpha | 71% | 77% |       |     | <b>83%</b>  |

Table 5-4. The individual PSD distribution, illustrated in the middle column of Figure 6, indicates the difficulty of the classification to an extent. Furthermore, to validate the observation, the right column of Figure 5-6 shows the accuracy of using each channel's PSD as the input. As expected, the posterior area can provide more than 70% accuracy for 0-back vs. 2-back and 0-back vs. 3-back tasks.

The topographic map of HbR features is shown in Figure 5-7. Similar to EEG, the averaged HbR distribution of 0-back is significantly different from that of 2-back and 3-back, shown in the left column. More specifically, the frontal-right area has increased HbR following the increment of workload level while the

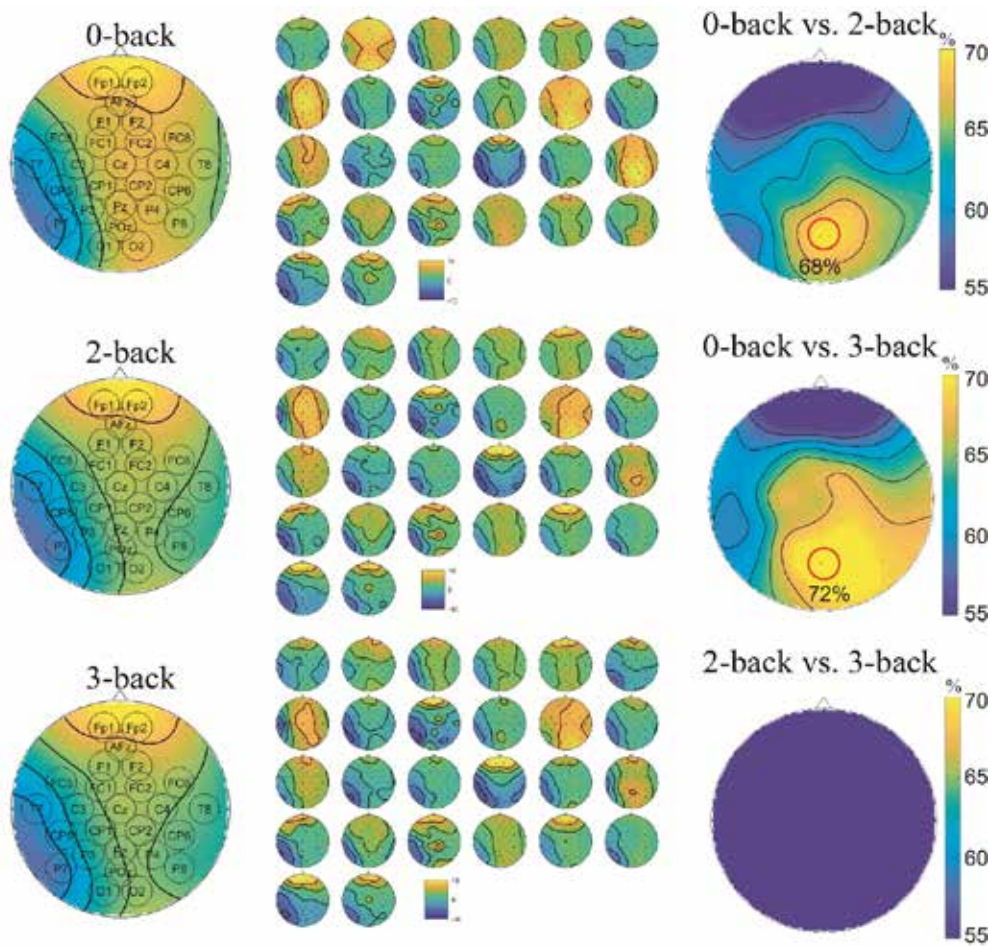


**Figure 5-6 Topographic map of the EEG alpha-band PSD. Left: average; Middle: each participant; Right: Accuracy using each-channel PSD as the input. The area that provides the highest accuracy is highlighted.**



frontal-centre and middle-left areas have decreased HbR following the increment of workload level. All these findings have been supported by the classification result of individual channels (Figure 5-7). There is no significant difference between 2-back and 3-back in terms of the overall pattern. The individual HbR distribution is illustrated in the middle column of Figure 5-7. Furthermore, to validate the observation, the right column of Figure 5-7 shows the accuracy of using each channel's HbR as the input. It is noted that the right frontal area can provide more than 70% accuracy for the 0-back vs. 3-back

task. Interestingly, the accuracy of the posterior area (PPOz) was close to 70% for 0-back vs. 2-back and 0-back vs. 3-back tasks, which was not easy to observe from the feature visualisation. It also matched the findings in the EEG analysis.

To visualise the FBC features, a heat map was used as shown in Figure 5-8. The maps for individual participants are illustrated in Figure 8A–C for 0-back, 2-

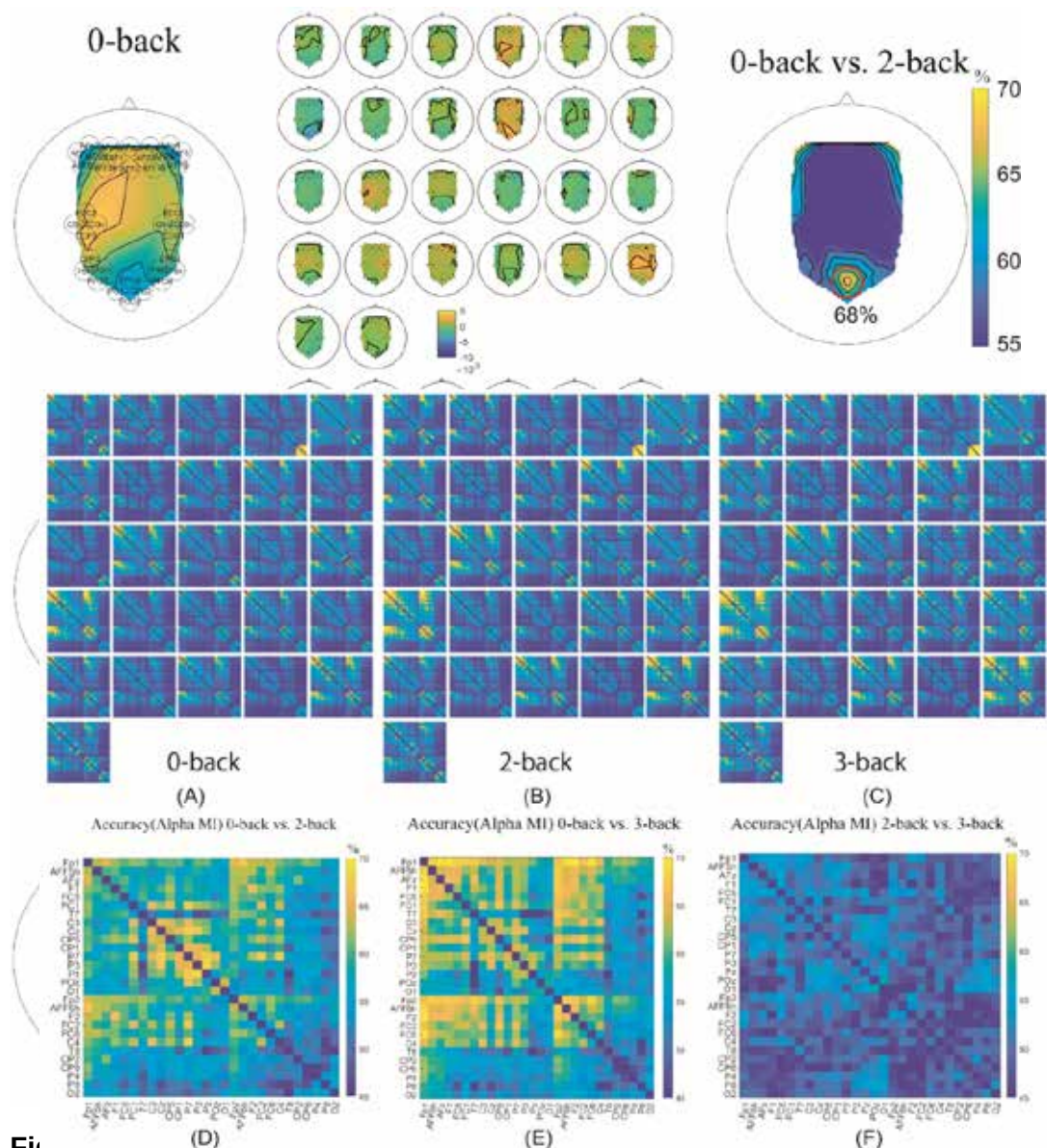


**Figure 5-8** The heat map of MI FBC features and the accuracy results. (A–C) shows the MI value for each participant in 0-back, 2-back, and 3-back task. (D–F) represents the classification accuracy for 0-back vs. 2-back, 0-back vs. 3-back and 2-back vs. 3-back, respectively, using each pair of EEG channels as the input where the FBC value was estimated by MI.

back and 3-back respectively. The accuracy of the three classification tasks is illustrated in Figure 5-8(D)–(F) for 0-back vs. 2-back, 0-back vs. 3-back, and 2-back vs. 3-back, respectively. It shows that each participant had a similar FBC pattern estimated by MI, while the value of different regions varied. Furthermore, it helped us to understand the differential contribution of the various brain regions for mental workload discrimination. The functional brain connectivity between frontal channels and Fp1 estimated by MI had a significant increase when the workload level became higher. It was also proved in Figure 8D–F that the following pairs left frontopolar–anterior midline frontal (Fp1:AFz), left frontopolar–left frontal (Fp1:F1), left frontopolar–right frontopolar (Fp1:Fp2) and left frontopolar–right frontal (Fp1:F2) provided relatively higher classification accuracy when differentiating 0-back from 2-back and 3-back.

## 5.4 Discussion

A comparative analysis of previous research and the proposed work employing EEG and fNIRS in mental workload classification is shown in Table 5-6. This Chapter now discusses the results in detail from three aspects: EEG vs. fNIRS, univariate vs. multivariate features, and independent vs. hybrid feature.

**Table 5-6 A comparative analysis of the previous research and the proposed work.**

| Reference                | Study Setting            | Classifier | Accuracy                                          |
|--------------------------|--------------------------|------------|---------------------------------------------------|
| Liu (Liu 2016)           | 0-, 1-, 2- N-back        | LDA        | 64.4% (EEG)<br>55.6% (fNIRS)<br>68.1% (EEG+fNIRS) |
| Aghajani (Aghajani 2017) | 0-, 1-, 2-, 3- N-back    | SVM        | 85.9% (EEG)<br>74.8% (fNIRS)<br>90.9% (EEG+fNIRS) |
| Nguyen (Nguyen 2017)     | Simulated driving system | FLDA       | 73.7% (EEG)<br>70.5% (fNIRS)<br>79.2% (EEG+fNIRS) |

|                                       |                 |             |                         |
|---------------------------------------|-----------------|-------------|-------------------------|
| <b>Saadati<br/>(Saadati<br/>2020)</b> | N-back          | DNN, SVM    | 67.0% (EEG-DNN)         |
|                                       | DSR             |             | 80.0% (fNIRS-DNN)       |
|                                       | Word generation |             | 87.0% (EEG+fNIRS-DNN)   |
|                                       | LHand vs RHand  |             | 82% (EEG+fNIRS-SVM)     |
| <b>Chu<br/>, n.d.)</b>                | Mental workload | SVM, RF, DT | 55.4% (EEG-RF)          |
|                                       |                 |             | 69.2% (fNIRS-RF)        |
|                                       |                 |             | 78.3% (EEG+fNIRS-RF)    |
|                                       |                 |             |                         |
| <b>Proposed study</b>                 | 0-, 2-, 3-back  | SVM         | 77% (0-back vs. 2-back) |
|                                       |                 |             | 83% (0-back vs. 3-back) |
|                                       |                 |             | 59% (2-back vs. 3-back) |
|                                       |                 |             |                         |

Abbreviations: LDA—Linear discriminant analysis; SVM—Support Vector Machine; FLDA—Fisher Linear Discriminant Analysis; DNN—Deep Neural Network; RF—Random Forest; DT—Decision Tree; DSR—discrimination/selection response task.

#### 5.4.1 EEG vs. fNIRS

On one hand, EEG needed a longer data length to suggest difference between different-level workloads. To be more specific, a 40 s time window was the most suitable, while 5 s was suggested for fNIRS. That is to say, fNIRS required less response time to support a satisfactory classification accuracy, which meant it may be more efficient in actual application. On the other hand, the EEG-based features, especially the FBC, represented obvious advantages over the fNIRS-based features in classification accuracy although the FBC methods were more complicated and entailed a higher computational cost.

The best region for assisting in the discrimination of the mental workload was different for EEG and fNIRS. Specifically, the posterior area performed the best for EEG (POz) in the alpha band and fNIRS had superiority in the right frontal region (AF8). Some studies suggested similar findings. Brouwer . (Brouwer

., 2012) found the alpha power of the midline parietal (Pz) region in EEG recordings significantly decreased with memory load, effectively distinguishing 2-back from 0-back. Chu . (Chu ., n.d.) stated that the alpha-power of O1 indicated differences between multi-level workloads. Regarding fNIRS, the prefrontal areas were well-accepted for measuring variations in mental workload (Han ., 2019; Karran ., 2019; Midha ., 2021). However, there was limited research pointing out a determined channel that contributes the most. Our study narrowed down the region (right frontal) to support the discrimination of workloads, as evidenced by the topographic visualisation of the machine-learning classification results.

#### **5.4.2 Univariate vs. Multivariate Features**

Considering the EEG features, the bivariate FBC approaches obtained more satisfactory accuracy compared to the univariate PSD features. The results of this Chapter provide evidence to support the hypothesis that the FBC not only estimated the informational intercommunication of separate brain regions but also tracked distinct changes for different levels of workload. There are other supporting studies for this hypothesis in the literature on workload classification. Pei . (Pei ., 2021) suggested the fusion of band power and FBC features, which were estimated by PLV and the phase lag index (PLI), enhanced the classification performance of workload identification. The PLI-based FBC was also used by Kakkos . (Kakkos ., 2021), and the study implied that using FBC emphasised its ability to serve as a promising indicator for different workload levels. Our framework employed four FBC estimations that illustrated connections with various properties, and MI outperformed PCC, MSC and PLV for the highest classification accuracy. In this case, the proposed framework deepened the use of the FBC technique in the field of mental workload discrimination. Furthermore, it implied that, among different levels of workload, the greatest changes occurred in nonlinear brain connectivity.

#### **5.4.3 Independent vs. Hybrid Feature**

The hybrid features of EEG and fNIRS outperformed the independent category of features in classification results, achieving the highest accuracy of 77% for 0-

back against 2-back and 83% for 0-back against 3-back. It meant that different methods explored distinct information and became complementary to each other thereby improving classification performance. The results agreed with the conclusion in the literature (Karran et al., 2019; Nguyen et al., 2017; Saadati et al., 2020). The present Chapter is an advance on previous studies because it generated new knowledge about regional information by comparing the foci of independent types of features. To an extent, it paved the way to use an EEG–fNIRS hybrid sensor in real-world workload classifications.

## **5.5 Conclusions**

In this Chapter, a novel solution relying on hybrid EEG–fNIRS features was proposed to deal with multi-level mental workload classification supported by machine-learning classifiers. To be more specific, the univariate PSD and four bivariate FBC features were extracted from an EEG recording in three frequency bands. With the assistance of HbO and HbR indicators from fNIRS, the fused features improved classification performance. Moreover, topographic and heat-map visualisation indicated distinct regions for EEG and fNIRS that represented difference among 0-back, 2-back and 3-back. Overall, the FBC technique based on an EEG recording proved its value in mental workload classification, and accuracy improvement emphasised the effectiveness of the hybrid EEG–fNIRS. The one limitation of this Chapter was that there was a volume conduction effect in the EEG dataset, but the high classification accuracy suggested that the functional connectivity was effective for classifying different workloads. One potential future work would be to use bipolar channels rather than unipolar channels or to pre-process the data to mitigate volume conduction.

## 264



264

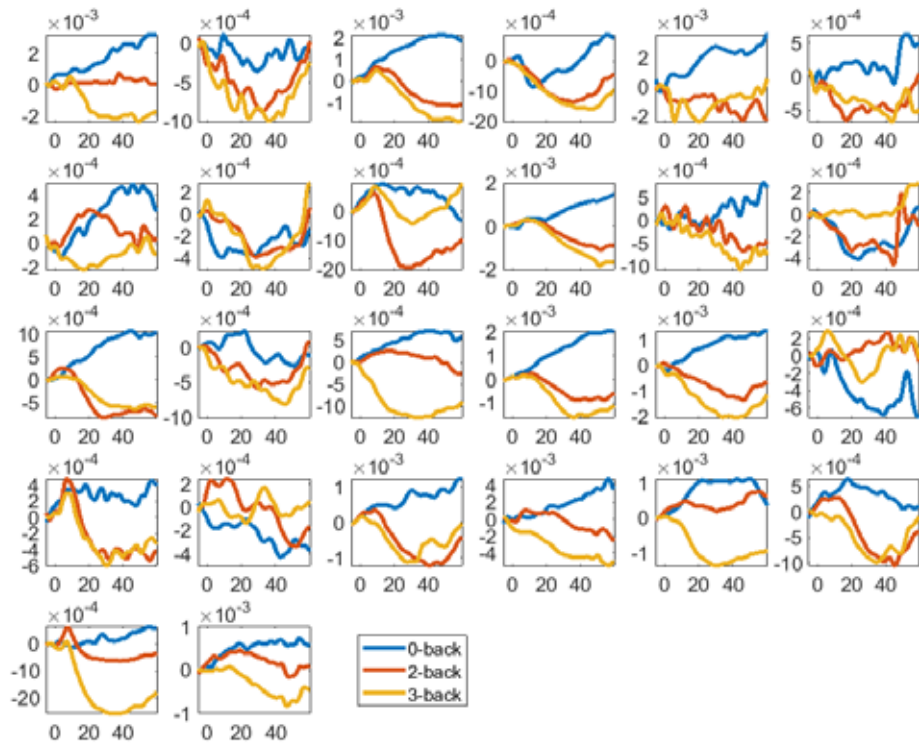


**Figure 5-11 fNIRS average HbR value of each of 26 participants in three levels of workload**

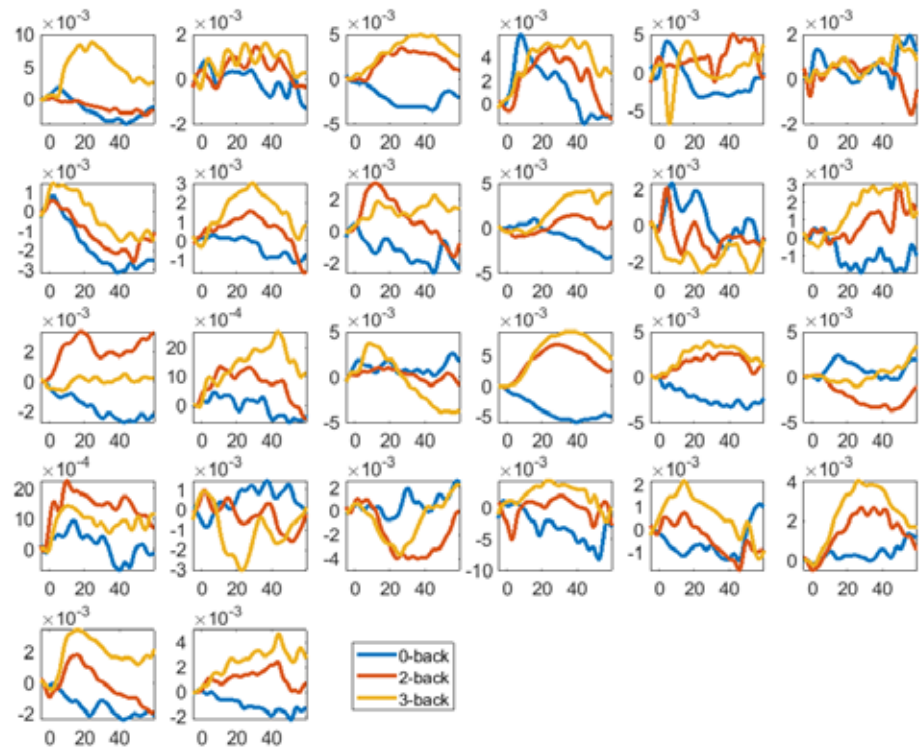


**Figure 5-10 fNIRS average HbO value of each of 26 participants in three levels of workload**



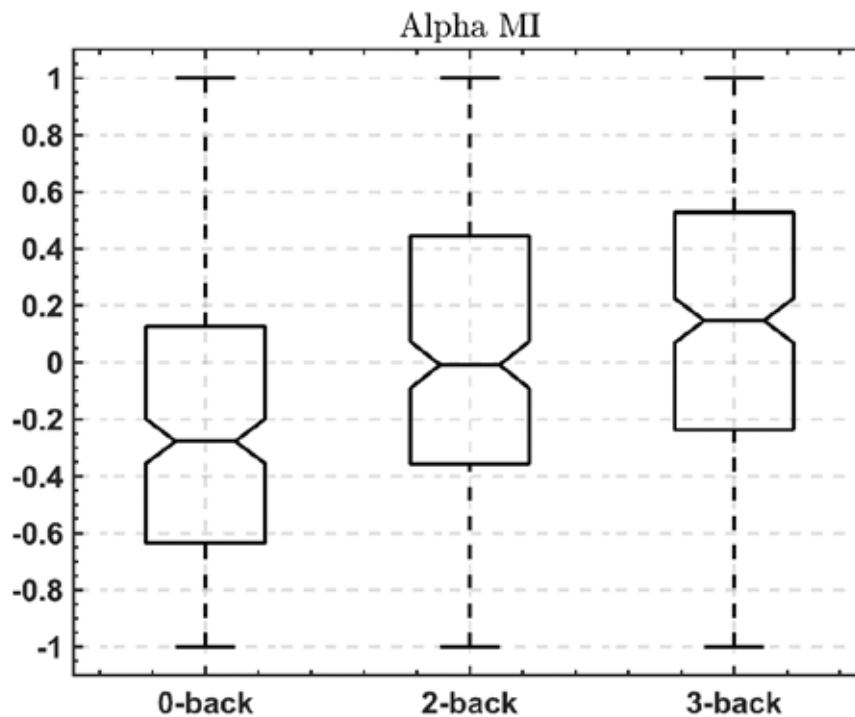


**Figure 5-12** A sample of significant test to represent the difference among three-level workload with the purpose of selecting limited numbers of features. (p-value<0.0001)

## 5.7 References

- Aghajani, H., Garbey, M., Omurtag, A., 2017.  
 . Front. Hum. Neurosci. 11, 359.  
<https://doi.org/10.3389/fnhum.2017.00359>
- Alhudhaif, A., 2021.  
 . PeerJ  
 Comput. Sci. 7, 1–24. <https://doi.org/10.7717/PEERJ-CS.537>
- Anderson, E.W., Potter, K.C., Matzen, L.E., Shepherd, J.F., Preston, G.A., Silva, C.T., 2011.  
 , in: Computer Graphics Forum. pp. 791–800.
- Asgher, U., Ahmad, R., Naseer, N., Ayaz, Y., Khan, M.J., Amjad, M.K., 2019.

. IEEE Access 7, 143250–143262. <https://doi.org/10.1109/ACCESS.2019.2944965>

Bashivan, P., Yeasin, M., Bidelman, G.M., 2016.

, in: 2015 IEEE Signal Processing in Medicine and Biology Symposium - Proceedings. Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/SPMB.2015.7405422>

Blankertz, B., Tangermann, M., Vidaurre, C., Fazli, S., Sannelli, C., Haufe, S., Maeder, C., Ramsey, L.E., Sturm, I., Curio, G., others, 2010.

. Front. Neurosci. 4, 198.

Blinowska, K.J., Rakowski, F., Kaminski, M., De Vico Fallani, F., Del Percio, C., Lizio, R., Babiloni, C., 2017.

. Clin. Neurophysiol. 128, 667–680. <https://doi.org/10.1016/j.clinph.2016.10.002>

Brouwer, A.M., Hogervorst, M.A., Van Erp, J.B.F., Heffelaar, T., Zimmerman, P.H., Oostenveld, R., 2012.

. J. Neural Eng. 9. <https://doi.org/10.1088/1741-2560/9/4/045008>

Cao, J., Grajcar, K., Shan, X., Zhao, Y., Zou, J., Chen, L., Li, Z., Grunewald, R., Zis, P., Marco, M. De, Unwin, Z., Blackburn, D., Sarrigiannis, P.G., 2021.

. Biomed. Signal Process. Control 67, 102554. <https://doi.org/10.1016/j.bspc.2021.102554>

Cao, J., Zhao, Y., Shan, X., Wei, H., Guo, Y., Chen, L., Erkoyuncu, J.A., Sarrigiannis, P.G., 2022.

. Hum. Brain Mapp. 43, 860–879. <https://doi.org/10.1002/hbm.25683>

Chu, H., Cao, Y., Jiang, J., Yang, J., Huang, M., Li, Q., Jiang, C., Jiao, X., n.d.  
0–17.

Han, W., Gao, L., Wu, J., Pelowski, M., Liu, T., 2019.

. Brain  
Cogn. 136, 103613. <https://doi.org/10.1016/j.bandc.2019.103613>

Herff, C., Heger, D., Fortmann, O., Hennrich, J., Putze, F., Schultz, T., 2014.

. Front. Hum. Neurosci. 7, 935. <https://doi.org/10.3389/fnhum.2013.00935>

Herold, F., Wiegel, P., Scholkmann, F., Müller, N.G., 2018.

. J. Clin. Med. 7,  
466.

Hussain, I., Hossain, A., Jany, R., Bari, A., Uddin, M., Raihan, A., Kamal, M.,  
Ku, Y., Kim, J., 2022.

. Sensors 22, 3079.

Hussain, I., Park, S.J., 2021.

. Brain Sci. 11.  
<https://doi.org/10.3390/brainsci11070900>

Hussain, I., Young, S., Kim, C.H., Benjamin, H.C.M., Park, S.J., 2021.

.  
Sensors 21, 1–16. <https://doi.org/10.3390/s21051896>

Kakkos, I., Dimitrakopoulos, G.N., Sun, Yi, Yuan, J., Matsopoulos, G.K.,  
Bezerianos, A., Sun, Yu, 2021.

. IEEE J. Biomed. Heal. Informatics 25, 3824–3833.  
<https://doi.org/10.1109/JBHI.2021.3085131>

Karran, A.J., Demazure, T., Leger, P.M., Labonte-LeMoyne, E., Senecal, S.,  
Fredette, M., Babin, G., 2019.

. *Front. Hum. Neurosci.* 13, 1–20.  
<https://doi.org/10.3389/fnhum.2019.00393>

Khan, M.J., Hong, M.J., Hong, K.-S., 2014.  
. *Front. Hum. Neurosci.* 8, 244.

Kuanar, S., Athitsos, V., Pradhan, N., Mishra, A., Rao, K.R., 2018.  
, in: *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*. Institute of Electrical and Electronics Engineers Inc., pp. 2576–2580. <https://doi.org/10.1109/ICASSP.2018.8462243>

Lim, L.G., Ung, W.C., Chan, Y.L., Lu, C.K., Sutoko, S., Funane, T., Kiguchi, M., Tang, T.B., 2020.  
. *IEEE Trans. Neural Syst. Rehabil. Eng.* 28, 2367–2376. <https://doi.org/10.1109/TNSRE.2020.3026991>

Liu, Y., Ayaz, H., Shewokis, P.A., 2016.  
. *Brain-Computer Interfaces* 4, 175–185.  
<https://doi.org/10.1080/2326263X.2017.1304020>

Lotte, F., Bougrain, L., Cichocki, A., Clerc, M., Congedo, M., Rakotomamonjy, A., Yger, F., 2018.  
. *J. Neural Eng.* 15.  
<https://doi.org/10.1088/1741-2552/aab2f2>

Maimon, N.B., Molcho, L., Intrator, N., Lamy, D., 2020.  
. *arXiv*.

Mandrick, K., Peysakhovich, V., Rémy, F., Lepron, E., Causse, M., 2016.  
. *Biol. Psychol.* 121, 62–73.  
<https://doi.org/10.1016/j.biopsycho.2016.10.002>

Midha, S., Maior, H.A., Wilson, M.L., Sharples, S., 2021. . Int. J. Hum. Comput. Stud. 147, 102580. <https://doi.org/10.1016/j.ijhcs.2020.102580>

Mormann, F., Lehnertz, K., David, P., Elger, C.E., 2000. 144, 358–369.

Mühl, C., Jeunet, C., Lotte, F., 2014. . Front. Neurosci. 8, 114. <https://doi.org/10.3389/fnins.2014.00114>

Nguyen, T., Ahn, S., Jang, H., Jun, S.C., Kim, J.G., 2017. . Sci. Rep. 7, 1–10. <https://doi.org/10.1038/srep43933>

Pei, Z., Wang, H., Bezerianos, A., Li, J., 2021. . IEEE Trans. Instrum. Meas. 70. <https://doi.org/10.1109/TIM.2020.3019849>

Pfeifer, M.D., Scholkmann, F., Labruyère, R., 2018. . Front. Hum. Neurosci. 11, 641.

Pinti, P., Tachtsidis, I., Hamilton, A., Hirsch, J., Aichelburg, C., Gilbert, S., Burgess, P.W., 2020. . Ann. N. Y. Acad. Sci. 1464, 5.

Putze, F., Herff, C., Tremmel, C., Schultz, T., Krusienski, D.J., 2019. , in: Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS. Institute of Electrical and Electronics Engineers Inc., pp. 3103–3106. <https://doi.org/10.1109/EMBC.2019.8856386>

Pyrzowski, J., Sieminski, M., Sarnowska, A., Jedrzejczak, J., Nyka, W.M., 2015.

. Sci. Rep. 5, 1–10. <https://doi.org/10.1038/srep16230>

Saadati, M., Nelson, J., Ayaz, H., 2020.

, in:

Advances in Intelligent Systems and Computing. Springer Verlag, pp. 209–220.

[https://doi.org/10.1007/978-3-030-20473-0\\_21](https://doi.org/10.1007/978-3-030-20473-0_21)

Saadati, M., Nelson, J., Ayaz, H., 2019.

, in: IEEE International Workshop on Machine Learning for Signal Processing, MLSP. IEEE Computer Society.

<https://doi.org/10.1109/MLSP.2019.8918861>

Shan, X., Huo, S., Yang, L., Cao, J., Zou, J., Chen, L., Sarrigiannis, P.G., Zhao, Y., 2021.

. IEEE Trans. Neural Syst. Rehabil. Eng. 29, 841–851. <https://doi.org/10.1109/TNSRE.2021.3076311>

Shin, J., Von Lüthmann, A., Kim, D.-W., Mehnert, J., Hwang, H.-J., Müller, K.-R., 2018.

. Sci. data 5, 1–16.

So, W.K.Y., Wong, S.W.H., Mak, J.N., Chan, R.H.M., 2017.

. PLoS One 12, e0174949. <https://doi.org/10.1371/journal.pone.0174949>

Zaman, S., Md. Rabiul Islam, S., 2021.

. Int. J. Image, Graph. Signal Process. 13, 1–13. <https://doi.org/10.5815/ijigsp.2021.05.01>

Zhang, J., Li, S., 2017.

. Cogn. Technol. Work 19. <https://doi.org/10.1007/s10111-017-0430-6>

Zhao, Yifan, Zhao, Yitian, Durongbhan, P., Chen, L., Liu, J., Billings, S.A., Zis, P., Unwin, Z.C., De Marco, M., Venneri, A., Blackburn, D.J., Sarrigiannis, P.G., 2020.

. IEEE Trans. Med. Imaging 39, 1571–1581.  
<https://doi.org/10.1109/TMI.2019.2953584>

## 6 Effective Connectivity Estimation Using Attention-based Convolutional Neural Networks

This Chapter is based a paper under preparation, 'Imaging of Effective Connectivity Network Using Causality Detection Attention-based Convolutional Neural Networks (CD-ABCNN)'.

### 6.1 Introduction

Effective brain connectivity (EBC) refers to the causal influence of one brain region on another, that is, how one area of the brain affects the activity of another area (Cao et al., 2022; Zarghami and Friston, 2020). It is a key concept in neuroscience and neuroimaging, and it has been used to study a wide range of brain functions, including perception, attention, memory, and decision-making (Babiloni et al., 2003; Chai et al., 2019; Maggioni et al., 2021; Shaw et al., 2019). Effective connectivity can be estimated using various neuroimaging techniques, such as functional magnetic resonance imaging (fMRI) and magnetoencephalography (MEG), as well as electroencephalography (EEG) (He et al., 2019; Yuan Wang et al., 2018; Wei et al., 2021; Zou and Ji, 2021). In recent decades, advances in these and other imaging techniques have allowed researchers to delve deeper into the brain's connectivity patterns and how they relate to behavior, cognition, and disease. The study of effective connectivity networks (ECNs) has provided new insights into the workings of the brain and has implications for our understanding of brain disorders such as Alzheimer's disease (AD), Parkinson's disease (PD), mild cognitive impairment (MCI) and epilepsy (Blinowska et al., 2017; Maggioni et al., 2021; McBride et al., 2015; Zhou et al., 2021). Therefore, finding an effective estimation for ECN has excellent clinical and humanistic relevance.

Model-based methods are techniques used to estimate effective connectivity in the brain by fitting mathematical models to brain imaging data. These models describe the relationships between different brain regions and can provide a more nuanced understanding of the causal interactions between brain regions (Cao et al., 2022; Sakalis, 2011; Salman et al., 2018; Zhao et al., 2012). Some



commonly used model-based methods for effective connectivity include: Dynamic Causal Modelling (DCM) (Brown and Friston, 2012; Herz *et al.*, 2012; Lee *et al.*, 2020; Van de Steen *et al.*, 2019), Structural Equation Modelling (SEM) (Astolfi *et al.*, 2004; Babiloni *et al.*, 2003; Sartori *et al.*, 2012), and Granger causality-related methods (Hasanzadeh *et al.*, 2022; Seth *et al.*, 2015). Most of the Granger causality measures are constructed based on autoregressive models, in which the present samples of the signals are predicted using a linear or nonlinear combination of the past samples (Omidvarnia *et al.*, 2011; van Mierlo *et al.*, 2014; Zhao *et al.*, 2013, 2020). Besides the traditional Granger method, the Error Reduction Ratio–causality (ERR-causality) test was proposed to estimate the time-varying direction and strength of linear or nonlinear causality between two signals as well as their relative time shift (Sarrigiannis *et al.*, 2014; Zhao *et al.*, 2012, 2020). However, Model-based methods for estimating effective connectivity have some limitations and drawbacks: (1) model complexity: model-based methods often require the specification of a complex model, which can be difficult to develop without sufficient knowledge of the target system and can limit the generalizability of the results; (2) model misspecification: if the model used to estimate effective connectivity is not an accurate representation of the underlying brain dynamics, the results may be biased or misleading; (3) Limited flexibility: model-based methods are often inflexible in handling complex data structures or large data sets. This can be a problem when dealing with non-linear relationships or high-dimensional data. These limitations highlight the need for continued development and refinement of model-based methods for effective connectivity and the importance of developing alternative methods (model-free) that can complement or supplement these methods.

Data-driven (non-parametric) methods gradually represent the power of calculation causal intercommunication between the multiple components of brain networks. Unlike parametric-model-based methods, data-driven approaches do not need to build a model before calculating connections. In other words, it could avoid assumptions of an unknown system, improving the efficiency of algorithms to a certain extent. The advancement of deep learning

technology has facilitated the exploration of novel approaches utilizing deep learning methods for the estimating brain effective connectivity networks (ECNs). Such methods include multi-layer perceptron neural networks (Talebi et al., 2018), recurrent neural networks (Yueming Wang et al., 2018), and effective connectivity learning based on spatiotemporal graph convolutional models (Zou and Ji, 2021). However, the current deep-learning methods mostly ignore the distinction between direct and indirect causality, pursuing learning the effective connectivity between each pair of brain areas. In practice, each cortical region causally influences the rest of the brain through different paths. The convolutional neural networks (CNN) are already successfully applied time-series problems, including drilling string vibration (Chen et al., 2020) and multi-sensor based human activity recognition problems (Yang et al., 2015). However, even though time series research is related to EEG brain connectivity problem, the potential of CNN to track ECNs from EEGs has not been explored. To address those limitations and align with the extension of CNN, the attention mechanism is involved to simplify the ECNs by removing indirect effective connectivity.

The present Chapter proposes causality detection attention-based convolutional neural networks (CD-ABCNN) to estimate ECNs with the purpose of efficiently exploring the difference between AD patients and healthy controls (HC) using EEG recordings by constructing a directed causal graph of causal relationships. The attention mechanism significantly increases the interpretability of ECNs. The proposed approach is first validated by the simulated data generated via multivariate autoregressive (MVAR) models and then performed on the practical EEG dataset of ADs and HCs (N=40). With the assistance of novel and advanced visualisation methods, it is found that AD differs from the HC in terms of ECNs, and regional information has been revealed. Overall, this Chapter provides valuable insights into the complex interplay between brain networks and AD and highlights the need for further research to fully understand the mechanisms underlying this neurodegenerative disease.

## 6.2 Methods

### 6.2.1 Causality Detection Attention-based Convolutional Neural Networks (CD-ABCNN)

The CD-ABCNN model comprises a set of independent attention-based convolutional neural networks (CNNs), each having the same architecture but a distinct target time series, and the number of networks ( $N$ ) depends on the number of channels. Figure 6-1 depicts the architecture of a single network  $N_i$ , which aims to predict its target time series,  $X_i$ , by minimizing the loss,  $L$ , between the predicted value,  $\hat{X}_i$ , and the actual values of  $X_i$ . The input to the network comprises an  $N \times T$  dataset,  $X$ , containing  $N$  equal-sized time series. During the training process, the attention mechanism computes attention

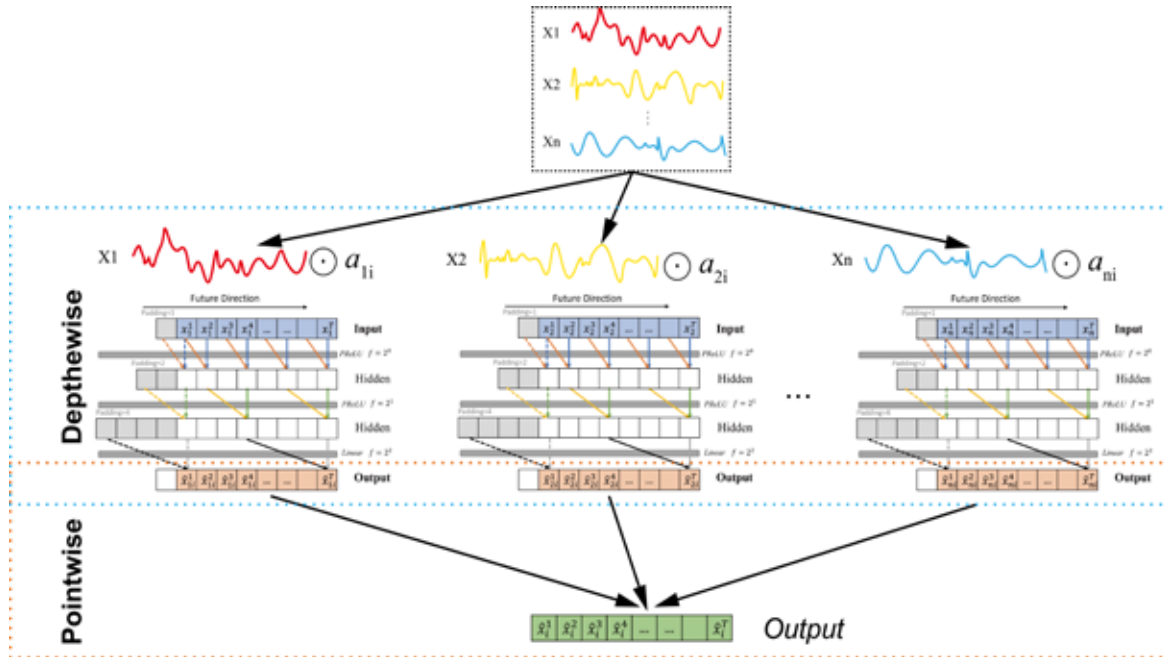


Figure 6-1 The architecture of the proposed Causality Detection Attention-based Convolutional Neural Networks (CD-ABCNN). Here is an example to predict  $x_i$ , with 2 hidden layers ( $L=2$ ), kernel size  $K = 2$  (shown as arrows) and dilation coefficient  $d = 2$ , leading to a receptive field  $R = 8$ . It consists of  $N$  channels, one for each input time series. A PReLU activation function is applied after each convolution. Weights  $W$  are shared across layers, indicated by the identical colours. The  $N$  channels have  $T$  steps,  $L = 1$  hidden layer in the depthwise convolution and  $N \times 2$  kernels with kernel size  $K = 2$ . The attention scores  $a$  are multiplied elementwise with the input time series, followed by an elementwise multiplication with the kernel. In the pointwise convolution, all channel outputs are combined to construct the prediction  $\hat{x}_i$ .

scores,  $a_i$ , indicating where the network focuses its attention while predicting  $X_i$ . As the network employs the attended time series for prediction, this time series is likely to contain information that is useful for prediction, implying a causal relationship between the attended time series and the target time series,  $X_i$ .

The CD-ABCNN incorporates dilated convolutions within its architecture. Dilated convolutions employ a kernel over an area larger than its size by skipping input values with a certain step size  $f$ , known as the dilation factor. The dilation factor increases exponentially with the chosen dilation coefficient  $c$ , such that  $f = c^l$  for layer  $l$ . The stacked dilated convolutions enable the network to operate on a coarser scale without loss of resolution or coverage.

To enable multivariate causal discovery, a one-dimensional depthwise separable architecture is employed in the CD-ABCNN. This architecture keeps the input time series separated and employs depthwise convolutions, where channels are kept separate by applying a different kernel to each input channel, followed by a  $1 \times 1$  pointwise convolution that merges together the resulting output channels. The use of a depthwise separable architecture improves accuracy and convergence speed and allows for a correct interpretation of the relationship between an input time series and the target time series, without mixing the inputs (Nauta et al., 2019).

An attention mechanism is incorporated into the CD-ABCNN architecture to determine where a network focuses when predicting a time series. Each network has its own attention vector  $\mathbf{a}_i = [a_{1i}, a_{2i}, a_{3i} \dots, a_{ni}]$  which is multiplied with the input time series  $X_j$  in network  $N_i$ . Attention score  $a_{ji}$  in  $\mathbf{a}_i$  indicates how much  $N_i$  attends to input time series  $X_j$  for predicting target  $X_i$ . A high value for  $a_{ji}$  in  $\mathbf{a}_i$  suggests that  $X_j$  might cause  $X_i$ , while a low value implies that  $X_j$  is unlikely to be a cause of  $X_i$ . The attention scores are used after training of the networks to identify the potential causes of a target time series. The attention scores can be used to identify potential causal relationships even between time series in the same network. In addition to detecting causal relationships, the proposed CD-ABCNN also determines the time lag between the cause and effect. This is accomplished through the kernel weights  $W$ , which correspond to

a causal input time series  $X_j$  from a network  $N_i$  that predicts the target time series  $X_i$ . The weights  $W$  assigned to each channel output indicate the relative

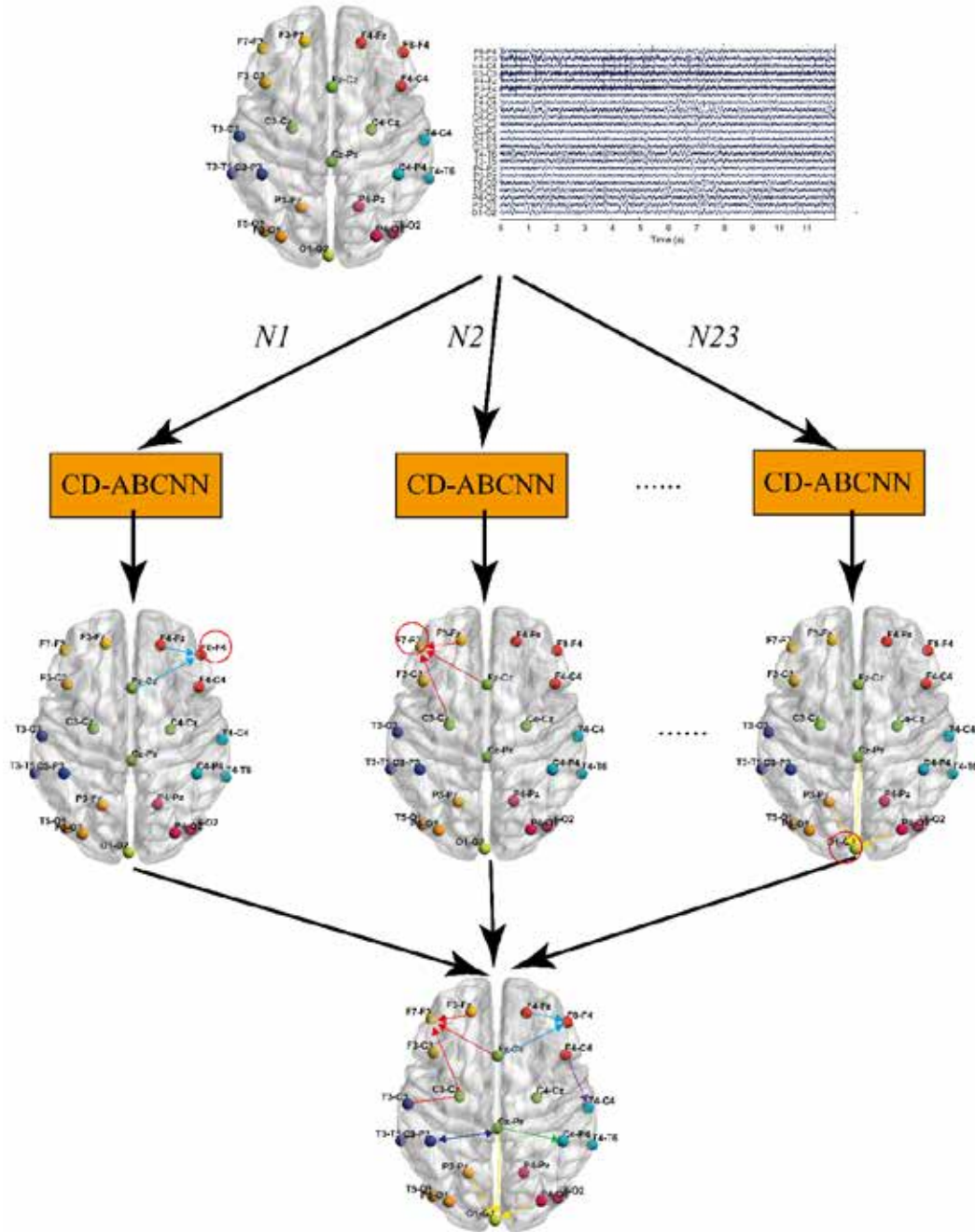


Figure 6-2 Methodology of CD-ABCNN performed on practical 23 channel EEG recordings. CD-ABCNN employs 23 independent CNNs, denoted as  $N1 \dots N23$ , where each network takes each channel of 23-channel EEGs of length  $T$  as input ( $N$  is equivalent to the number of time series in the input data set). The CD-ABCNN constructs a temporal causal graph, namely effective connectivity networks.

significance of each time delay. For example, if the past value  $X_j^5$  of the input  $X_j$  has the highest weights when the network  $N_i$  predicting the future value  $X_i^8$  target time series  $X_i$ , the time delay will be  $8 - 5 = 3$ . In this case, based on the attention mechanism and the learnt time delay, the indirect causality has been removed from the resulting causal graph.

Figure 6-2 illustrates an overview of the proposed CD-ABCNN algorithm when applied to real-world EEG recordings with 23 channels. The methodology involves the use of 23 independent CNN networks, which are denoted as  $N_1, N_2, \dots, N_{23}$ . For the network  $N_i$ , 23-channel EEG recordings were fed into the  $i^{th}$  network to predict  $i^{th}$ -channel EEG. In this case, the output of  $i^{th}$  network is the causal information flows from each of the input 23-channel EEG recordings to the target channel ( $i^{th}$ ). The output of each network represents the causal information flow from each input channel to the target channel. These causal relationships are then combined to construct a directed graph, which represents the ECN. The resulting ECN represents the collective causal relationships between all 23 EEG channels, providing insights into the underlying effective connectivity patterns in the brain.

### 6.2.2 MVAR-model Simulation

A multivariate autoregressive model (MVAR) is a statistical model that describes the relationship between multiple time series variables by representing each variable as a linear function of its own past values as well as the past values of all the other variables in the system (Ghumare et al., 2018; Li et al., 2020). The model is specified by a set of parameters that define the strength and direction of the interdependence among the variables.

For a given time series,  $\mathbf{X}(n) = [x_1(n), x_2(n), \dots, x_m(n)]^T$ , a multivariate autoregressive (MVAR) model of order  $p$  is defined as (Talebi et al., 2018):

$$\mathbf{X}(n) = \begin{bmatrix} x_1(n) \\ x_2(n) \\ \vdots \\ x_m(n) \end{bmatrix} = \sum_{d=1}^p \mathbf{A}_d \begin{bmatrix} x_1(n-d) \\ x_2(n-d) \\ \vdots \\ x_m(n-d) \end{bmatrix} + \begin{bmatrix} \sigma_1(n) \\ \sigma_1(n) \\ \vdots \\ \sigma_m(n) \end{bmatrix} \quad (6-1)$$

where the  $A_d$  is defined as:

$$A_d = \begin{bmatrix} a_{11}^d & a_{1m}^d \\ a_{m1}^d & a_{mm}^d \end{bmatrix}, \quad d = 1, 2, \dots, p \quad (6-2)$$

for  $d = 1, 2, \dots, p$ . The real-valued coefficient  $a_{ij}^d$  is the weighting factor that characterizes the contribution of  $x_j(n-d)$  to  $x_i(n)$ .

### 6.2.3 Real EEG data

The Chapter recruited both healthy control (HC) participants and patients diagnosed with Alzheimer's disease (AD), confirmed through neuropsychology testing and structural and functional MRI scans. The EEG recordings utilized in the Chapter were collected during task-free conditions and were previously employed in other research studies (Durongbhan, 2019; Zhao, 2020). The EEG recordings were acquired using a combination of 10/10 and 10/20 international system of electrode placement with an XLTEK 128-channel headbox (Optima Medical LTD) and were sampled at a rate of 2K Hz (analogue low pass filter at 600Hz) with an earlobe reference. Resting-state EEG recordings were obtained from each participant for 30 minutes, which included distinct 5-minute epochs of eyes open (EO) and eyes closed (EC). The data were categorized based on age group, either below 70 or above 70, and were collected twice for each participant in the two physiological states of EO and EC. Table 6-1 provides an overview of the four groups and their respective composition.

**Table 6-1**

|             | AD | HC |
|-------------|----|----|
| Above 70 EO | 4  | 10 |
| Below 70 EO | 16 | 10 |
| Above 70 EC | 4  | 10 |
| Below 70 EC | 16 | 10 |

The EEGs recorded in this Chapter were reviewed using XLTEK equipment (Optima

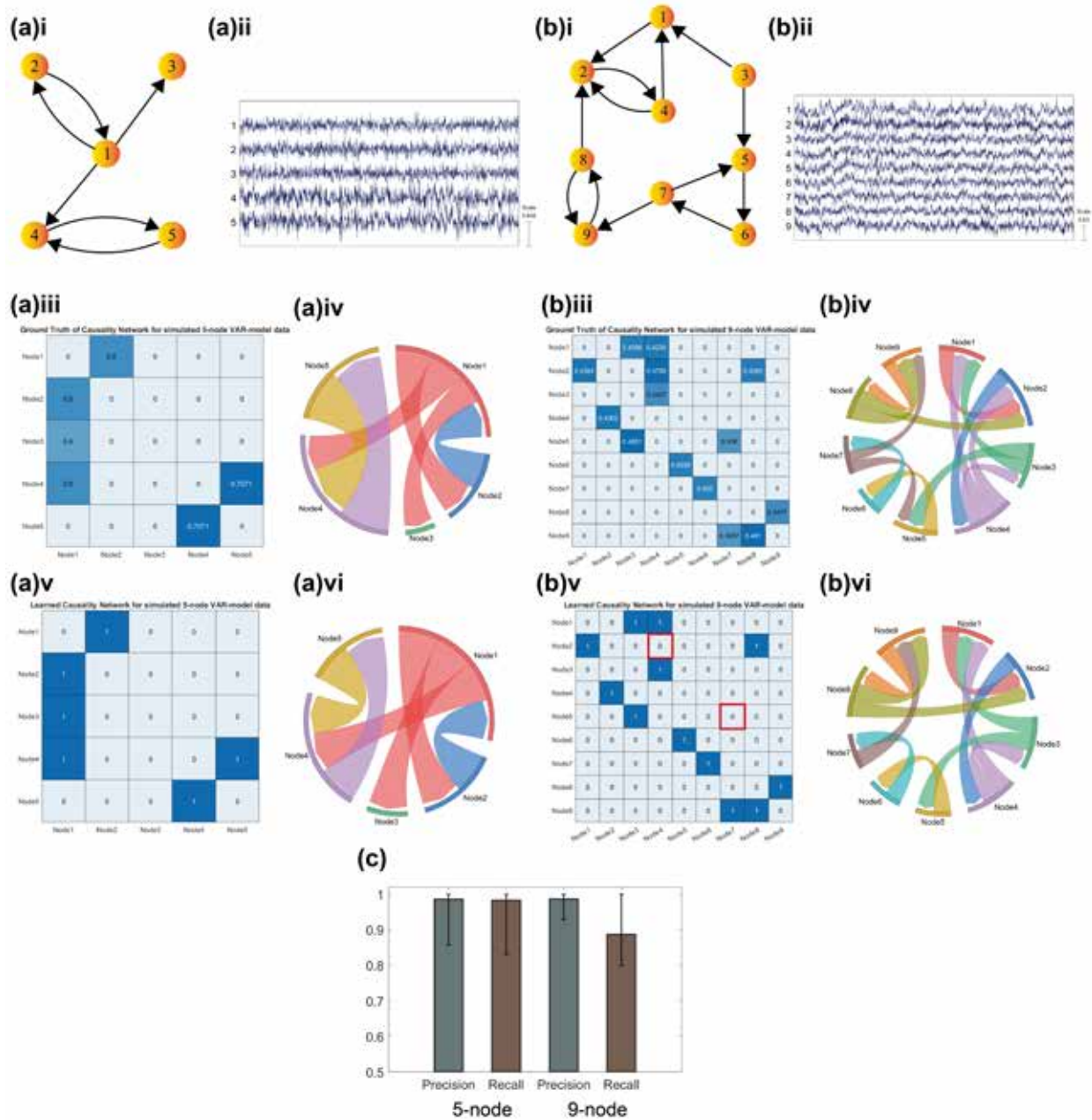


**Figure 6-3** Performance of the proposed CD-ABCNN using the MVAR-model simulated data. (a) the structure of the designed 5-node networks; the simulated data based on 5-node MVAR-simulated networks; and the ground truth of the causal relationship of the designed 5-node networks; and the learnt of the causal relationship of the designed 5-node networks; . (b) the structure of the designed 9-node networks; the simulated data based on 9-node MVAR-simulated networks; and the ground truth of the causal relationship of the designed 9-node networks; and the learnt of the causal relationship of the designed 9-node networks; (c) the performance of the proposed CD-ABCNN on 5-node and 9-node simulated time series.



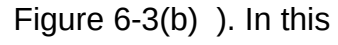
Medical LTD) with bipolar derivations and time-locked video recordings. The available bipolar channels were Fp2-F8, Fp1-F7, F8-F4, F7-F3, F4-C4, F3-C3, F4-FZ, FZ-CZ, F3-FZ, T4-C4, T3-C3, C4-CZ, C3-CZ, CZ-PZ, C4-P4, C3-P3, T4-T6, T3-T5, P4-PZ, P3-PZ, T6-O2, T5-O1, P4-O2, and P3-O1. From each of the 40 healthy control and Alzheimer's disease patients, artifact-free EEG epochs were selected for eyes open and eyes closed conditions, with each epoch lasting 12 seconds. The data pre-processing and export were performed using Spike 2 software (version 8), where a time constant of  $\tau = 0.08s$  (high pass filter at 2 Hz) was applied to attenuate any delta or slower frequency artifacts, such as those generated by eye blinking and movements.

## 6.3 Results

### 6.3.1 Study 1: Simulated Data

The proposed CD-ABCNN has demonstrated robustness and efficiency in estimating causal relationships among nodes in both 5-node and 9-node MVAR-simulated networks. The 5-node model only contained pairwise direct causal information flow (Figure 6-3(a)). The 9-node network (Figure 6-3(b)), designed to include both direct and indirect causality, allowed for an evaluation of the attention mechanisms in the proposed model. Simulated 2000-point time series were generated for both models, and the simulation process was repeated 50 times based on the same MVAR model. Ground truth causality networks were compared to the learned networks for both models. The simulated time series data exhibited clear causal relationships, as illustrated in Figure 6-3(a) and Figure 6-3(b) (ground truth). The learned networks, shown in Figure 6-3(a) and Figure 6-3(b), indicate that the proposed CD-ABCNN was able to accurately estimate pairwise causality for the 5-node network, achieving almost 100% accuracy, recall, and precision (Figure 6-3(c)). This suggests that the attention mechanisms were effective in simplifying the causality networks. For the 9-node network, the precision remained robust (Figure 6-3(c)), which indicates that the proposed framework would not learn wrong causal influence between two channels. The result also suggests that it correctly identified the direct causal influences among nodes without introducing any indirect connections (Figure 6-3(b)).

However, there was a significant decrease in recall (Figure 6-3(c)), due to the model ignored some of t

To be more specific, there are two paths from node 4 to node 2. One is node 4 . In this case, the latter one is recognised as the indirect connectivity, which will be ignored the proposed framework. This observation supports the notion that the proposed attention mechanisms simplify the causality networks, highlighting the most important causal influences. These results suggest that the proposed CD-ABCNN can be a useful tool for investigating causal relationships among nodes in complex systems.

### 6.3.2 Study 2: AD and HC EEG

To reveal the difference between AD and HC in terms of ECN, the proposed CD-ABCNN was performed on EEG recordings of each participant. Figure 6-4 illustrates a comparison between an Alzheimer's Disease (AD) patient and a Healthy Control (HC) participant. By analysing the learned effective connectivity networks of the brain for both participants, it appears that AD may not have a significant impact on the whole brain networks, but it does lead to the loss of several important connections. The heat maps representing the learned pairwise effective connectivity for both participants reveal a similar pattern with some notable differences. For instance, the midline centroparietal area (Cz-Pz) in the HC participant is causally linked to both the left and right centroparietal regions (C3-P3 and C4-P4) with bi-directed effective connectivity, but these connections are missing in the AD patient. This observation is significant because it sheds light on the potential effects of AD on brain networks. The loss of causal connections in specific regions of the brain could have far-reaching consequences for brain function, and may help to explain some of the cognitive and behavioural symptoms associated with AD. The midline centroparietal area is a critical region for attention and working memory (Pezzetta et al., 2023; Pisauro et al., 2017; Salehinejad et al., 2021), and its loss of connection with the left and right centroparietal regions could explain the impairment of these functions in AD patients (Figure 6-4(c) and (f)). Moreover, the finding that AD

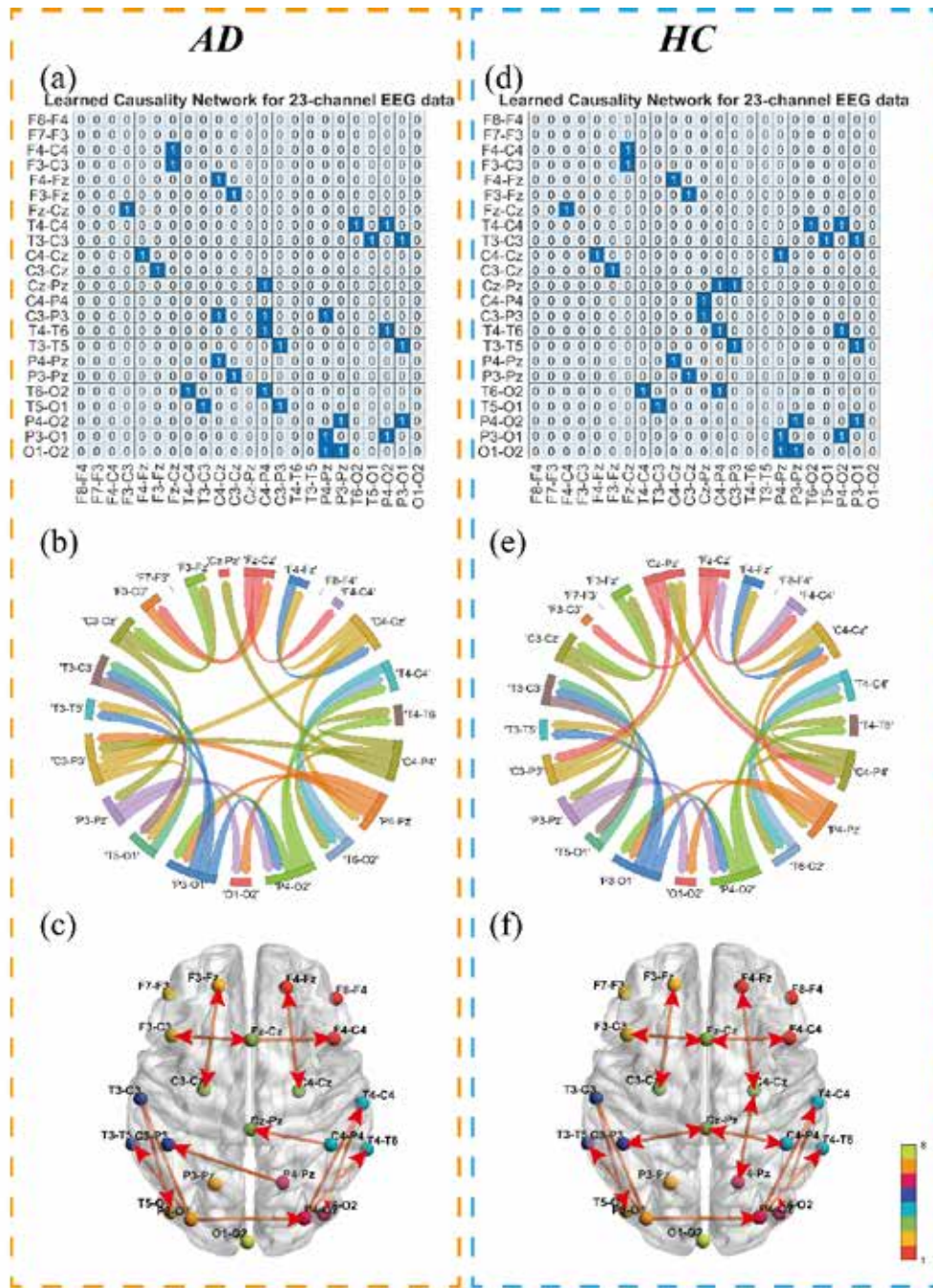


Figure 6-4 The visualisation of the learnt ECNs of one AD (left) and one HC (right). (a) and (d) the heat map of the learnt ECNs; (b) and (e) the chord diagram of the learnt ECNs. The order and locations of channels are designed to correspond the brain areas; (c) and (f) the learnt ECNs with corresponding brain regions. 23 bipolar derivations were created from 8 different brain regions labelled by distinct colour: the left frontal (F7-F3, F3-FZ), right frontal (F8-F4, F4-FZ), left frontocentral, temporo- central, temporal and centroparietal (F3-C3, T3- C3, T3-T5, C3-P3), right frontocentral, temporo- central, temporal and centroparietal (F4-C4, T4-C4, T4-T6, C4-P4), midline frontocentral, centroparietal and left and right parasagittal central regions (Fz- Cz, Cz-Pz, C3- Cz, C4-Cz), left parietal, parieto-occipital and temporo-occipital areas (P3-Pz, P3- O1, T5-O1), right parietal, parieto-occipital and temporo-occipital areas (P4-Pz, P4-O2, T6-O2) and midline occipital region (O1-O2)).

affects specific connections in the brain rather than whole brain networks

suggests that targeted interventions or treatments that restore these connections may be a promising approach. It is essential to identify the specific connections that are most affected in AD and to develop interventions that can promote their restoration. These interventions may include cognitive training, physical exercise, or even brain stimulation techniques such as transcranial magnetic stimulation (TMS) or transcranial direct current stimulation (tDCS). Therefore, this Chapter provides valuable insights into the complex interplay between brain networks and AD and highlights the need for further research to fully understand the mechanisms underlying this devastating disease.

To further statistically compare the HC and AD in terms of learnt ECNs, the ECN of all HC participants ( $N_{HC} = 20$ ) and AD patients ( $N_{AD} = 20$ ) were summed in eyes closed (EC) and eye open (EO) states (Figure 6-5). Similar to previous results obtained from the single participant, the sum of ECNs also indicates a likely pattern of HC and AD in EC and EO state, respectively (Figure 6-5(a), (b), (d) and (e)). The Figure 6-5(c) and (f) illustrate the difference between HC and AD (HC-AD). The result is consistent with previous research findings that suggest that AD patients exhibit altered effective connectivity in the brain, which could be due to neurodegeneration and synaptic dysfunction.

In the EC condition, some pairwise effective connectivity is learnt for all AD patients, such as C3-Cz -Fz, C4- -Fz and Fz- -C3 (yellow-circle area, Figure 6-5(a)) and (b)), while several HC participants seems not have effective connectivity in terms of the mentioned information flows. Furthermore, this area witnesses the most significant difference between HC and AD with value of -6 (Figure 6-5(c)). In the EO condition, the causality from P4-Pz and P3-Pz to O1-O2 is recognised by the proposed CD-ABCNN for all HC participant, while some AD patients loss those connections (green-circle area, Figure 6-5(d) and (e)). This results in the difference value of 5 and 3, respectively (Figure 6-5(f)). However, the most significant difference occurs at channel P3-O1. To be more specific, the most of ADs (15), indicate that there is causal information flow from P3-O1 to T3-C3 under the EO condition while small number of HCs (6) represents this learnt pairwise effective connectivity. In

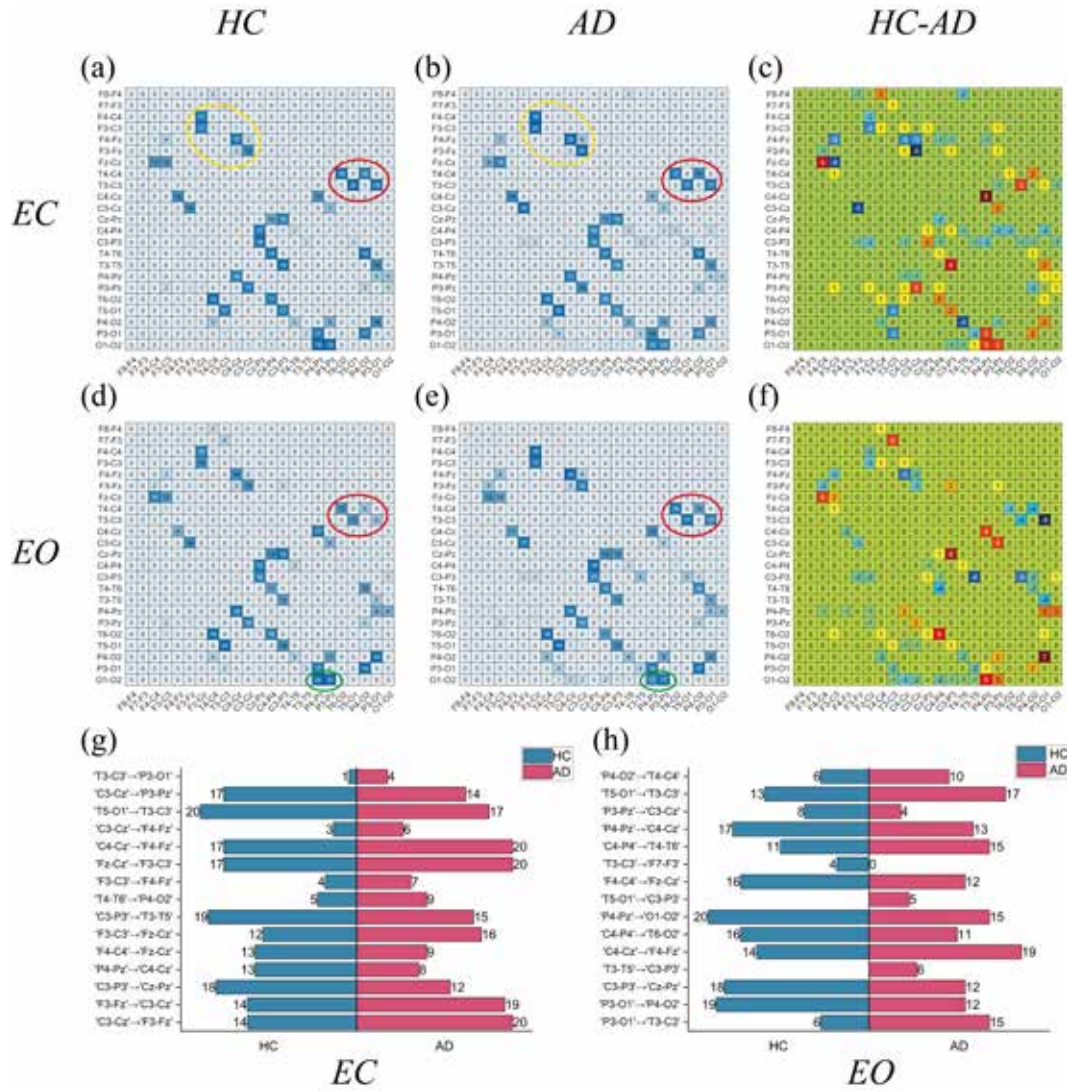


**Figure 6-5** The statistical comparison of AD and HC in terms of the learned ECNs under EO and EC condition, using the proposed CD-ABCNN framework. (a) the sum of ECNs of HC under EC condition; (b) the sum of ECNs of AD under EC condition. The yellow circles indicate specific pairs of channels show learned effective connectivity for all AD patients, but not for all HCs. (c) the heat map of difference between HC and AD under EC condition, namely (a)-(b); (d) the sum of ECNs of HC under EO condition; (e) the sum of ECNs of AD under EO condition. The green circles indicate specific pairs of channels show learned effective connectivity for all HCs, but not for all AD patients. The red circles indicate specific pairs of channels that imply the AD and HC are affected by the eyes' conditions (EO/EC) to varying degrees. (f) the heat map of difference between HC and AD under EO condition, namely (d)-(e); (g) the comparison of AD and HC in terms of top 15 pairs of channels that shows the most significant difference under EC condition; (h) the comparison of AD and HC in terms of top 15 pairs of channels that shows the most significant difference under EO condition. Some combinations have common EEG channels because a bipolar approach is employed (such as F4-C4 and C4-P4), which may result in a misleading high false connectivity between some of the bipolar channels. To avoid this issue, the authors neglected any pair with common EEG electrodes. Therefore, the 46 channel pairs that have this characteristic were not considered.

contrast, almost all HCs (19) is recognised to have directed relationship of P3-



O1 -O2, whereas nearly half of ADs (8) loss this connection.

Besides, the AD and HC are affected by the eyes' conditions (EO/EC) to varying degrees. The red-circle area of ECNs of HC exhibited such a phenomenon that there is a huge decrease in terms of the marked effective connectivity from EC to EO condition. However, the eyes' conditions have few influences on the ECNs of AD at the same area (red circle). This also explain the mentioned findings that this area illustrates the largest gap between HC and AD under EO condition, while there is no obvious difference under EC condition. Another interesting observation is that, for both AD and HC, there is few effective connectivity related to the channel F8-F4 and F7-F3. They do not affect other areas at all and are rarely affected by them.

Finally, Figure 6-5(g) and (h) illustrate the comparison of AD and HC in terms of top 15 pairs of channels that show the most significant difference in EO and EC conditions. The findings showed that the channels varied significantly between the two distinct eyes' conditions, indicating the importance of controlling for the eyes' conditions in any diagnosis and monitoring of AD using ECN-related features.

To reveal the distribution of the learnt ECNs, as well as explore the specific brain regions that contribute to the significant difference, the learnt ECNs for HC and AD under EC and EO are shown in Figure 6-6. The causal relationship that represents the most significant difference between HC and AD is highlighted by the arrows. In the EC condition, the arrow points out the largest difference that more AD patients exhibited the effective connectivity of the left frontal (F3-Fz) and midline centroparietal (C3-Cz) (Figure 6-6(a) and (b)), which are known to be involved in executive functions such as attention and working memory. In the EO condition, the arrow points out the largest difference that more AD patients exhibited the information flow from the left temporo-central (T3-C3) to left parieto-occipital (P3-O1) (Figure 6-6(c) and (d)), which are involved in visual processing and attention. This finding is consistent with previous studies that have reported altered visual processing in AD patients. Interestingly, the effective connectivity among temporal and occipital regions, highlighted by the

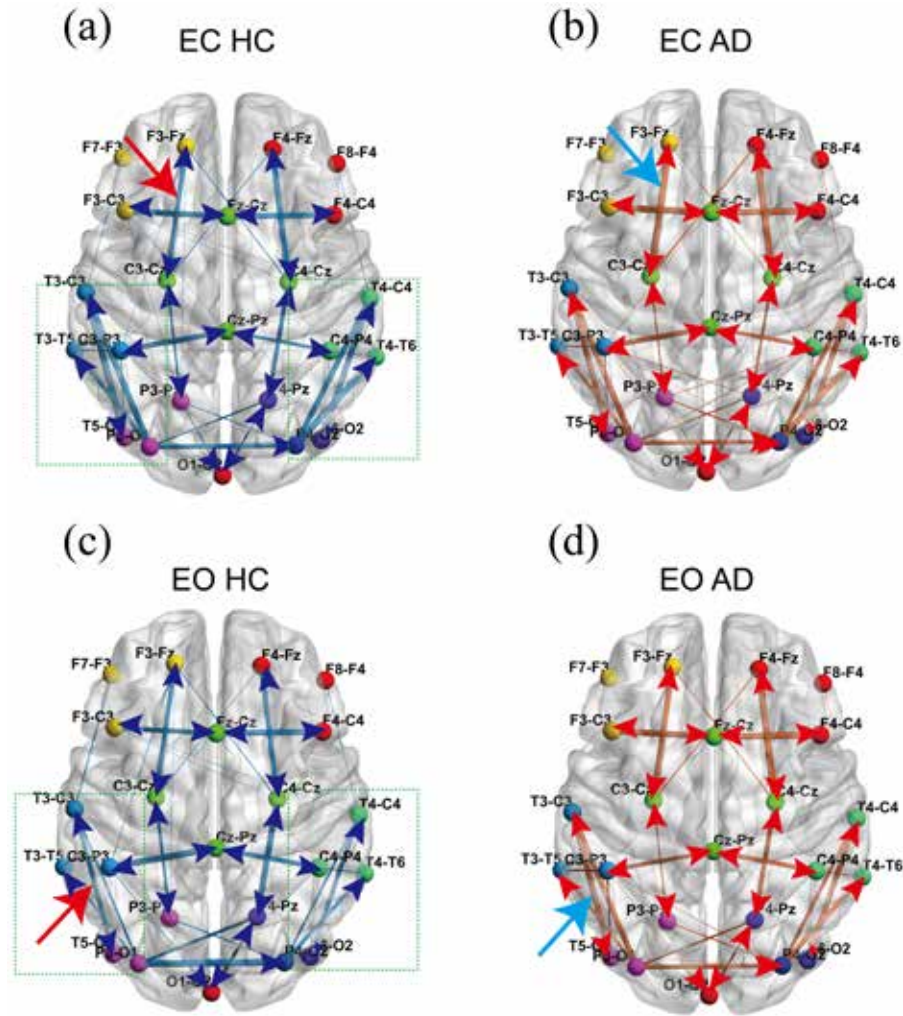


**Figure 6-6** The regional comparison of AD and HC in terms of the learned ECNs under EO and EC condition, using the proposed CD-ABCNN framework. (a) the sum of learnt ECNs of HC under EC condition; (b) the sum of learnt ECNs of AD under EC condition. The arrow points out the bi-directed causal relationship that represent the most significant difference between HC and AD under the EC condition. (c) the sum of learnt ECNs of HC under EO condition; (d) the sum of learnt ECNs of AD under EO condition. . The arrow points out the causal relationship that represent the most significant difference between HC and AD under the EC condition. The width of the lines represents variant numbers of participants. Significant difference between EO and EC for HCs are highlighted by the green rectangles.

green rectangle in Figure 6-6(a) and (c), showed that HCs were more sensitive to the eyes' conditions than ADs. Specifically, there was a dramatical decrease in terms of the causality from P3-O1 to T3-C3 and T3-C5 and from T3-C3 to P3 O1 for the HCs with eyes opening, while ADs patients seem not to obviously respond to the change of eyes' condition in the same areas. This suggests that

AD patients may have difficulty processing visual information, which could contribute to the cognitive deficits observed in this population.

## 6.4 Discussion

The present study has important implications for the diagnosis and monitoring of AD using effective connectivity networks (ECNs). The findings suggest that ECNs can be useful biomarkers for identifying changes in effective connectivity in AD patients. Therefore, ECNs have the potential to be used as a diagnostic and monitoring tool for this population. Several studies also have consistently identified abnormalities of effective connectivity in patients with mild cognitive impairment (MCI) and AD (Blinowska et al., 2017; Sami et al., 2018; Scherr et al., 2021; Xue et al., 2019). Sami et al. (2018) suggested that the use of effective connectivity analysis is highly suitable for studying the widespread effects of neuropathology, as well as the impact of neurodegenerative disease on synaptic and axonal function. It is noteworthy that our study is distinct from previous research as we identified global brain networks using ECN and introduced a unique perspective by eliminating indirect causal relationships. This approach simplified the overall network to a certain extent and significantly improved its interpretability. The Chapter also highlights the importance of controlling the eyes' conditions in ECN-based diagnostic tools to ensure accurate and reliable results. This aspect is in line with prior research. For example, Miraglia et al. (2016) reported that the cortical sources of resting EEG in eyes-closed and eyes-open conditions might vary in MCI and AD patients compared to normal elderly individuals.

Moreover, the observation that specific channels showed obvious differences in effective connectivity in AD and HC groups could be leveraged to develop more targeted interventions for AD patients. Sami et al. (2018) suggested that many neurodegenerative diseases are characterized by regional rather than global pathology. Our findings suggest the AD abnormally behaviour in terms of intercommunication between the left frontal and midline centroparietal regions, as well as left temporo-central and left parieto-occipital. Those cortical areas are highly related to cognitive control, memory and decision-making processes in



human brain (Badre and Nee, 2018; De Jager et al., 2018; Dippel and Beste, 2015; Herding et al., 2019; van Vugt et al., 2019). There are numerous evidences that the AD can result in cognitive decline and affect the ability of memory and decision-making (Durongbhan et al., 2019; Khojaste-Sarakhsi et al., 2022; Kunneman et al., 2017; Mostafavi et al., 2018; Trachsel et al., 2015), which agree with the observations of our study.

However, there are some limitations of this study. First of all, the linear MVAR is used to simulate the EEG signal, whereas non-linear causality is not considered. In the real world, human brain is a complex system and contains non-linear and non-stationary activities. Although the proposed AB-CNN obtains satisfying results in the linear simulated dataset, its generalization and robustness have not been fully explored. In the future, neural mass models may be used to simulate the EEG/MEG data, which have proven especially fruitful for understanding brain rhythms (Byrne et al., 2022). Another drawback is that this Chapter does not consider different frequency bands separately, namely delta (0.5–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–30 Hz) and gamma (30–90 Hz). Moreover, each frequency band is associated with distinct cognitive functions. Therefore, the proposed approaches could be performed on a specific band and the result is worth expecting.

To sum up, the Chapter adds to the growing body of literature on the effective connectivity changes in AD and provides valuable insights into the potential of ECNs as a diagnostic and monitoring tool. In the meantime, EEG is relative low-cost imaging technique with high temporal resolution in the millisecond range (Haufe et al., 2013). Therefore, this Chapter provides possibility of a flexible mobile wearable device to monitor the progress of AD. This Chapter also highlights the importance of effective connectivity analysis in investigating the neural mechanisms underlying AD and sheds light on the importance of controlling for the eyes' conditions in any future diagnostic and monitoring strategies for AD. Further research should explore the potential of effective connectivity analysis in early detection and treatment of AD.

## 6.5 References

Babiloni, F., Cincotti, F., Basilisco, A., Maso, E., Bufano, M., Babiloni, C., Carducci, F., Rossini, P., Cerutti, S., Rubin, D.B.D., 2003.

. Int. IEEE/EMBS Conf. Neural Eng. NER 2003-Janua, 79–82. <https://doi.org/10.1109/CNE.2003.1196760>

Badre, D., Nee, D.E., 2018.

. Trends Cogn. Sci. 22, 170–188. <https://doi.org/10.1016/j.tics.2017.11.005>

Blinowska, K.J., Rakowski, F., Kaminski, M., De Vico Fallani, F., Del Percio, C., Lizio, R., Babiloni, C., 2017.

. Clin. Neurophysiol. 128, 667–680. <https://doi.org/10.1016/j.clinph.2016.10.002>

Cao, J., Zhao, Y., Shan, X., Wei, H., Guo, Y., Chen, L., Erkoyuncu, J.A., Sarrigiannis, P.G., 2022.

. Hum. Brain Mapp. 43, 860–879. <https://doi.org/10.1002/hbm.25683>

Chai, M.T., Amin, H.U., Izhar, L.I., Saad, M.N.M., Abdul Rahman, M., Malik, A.S., Tang, T.B., 2019.

. Front. Neuroinform. 13. <https://doi.org/10.3389/fninf.2019.00066>

Chen, G., Chen, M., Hong, G., Lu, Y., Zhou, B., Gao, Y., 2020.

. Energies 13. <https://doi.org/10.3390/en13040888>

Chen, L., Zhao, Yifan, Zhao, Yitian, 2019. Imaging of nonlinear and dynamic functional brain connectivity based on EEG recordings Imaging of nonlinear and

dynamic functional brain connectivity based on EEG recordings with the application on the diagnosis of brain connectivity based on EEG recording.

De Jager, P.L., Ma, Y., McCabe, C., Xu, J., Vardarajan, B.N., Felsky, D., Klein, H.U., White, C.C., Peters, M.A., Lodgson, B., Nejad, P., Tang, A., Mangravite, L.M., Yu, L., Gaiteri, C., Mostafavi, S., Schneider, J.A., Bennett, D.A., 2018.

. Sci. Data 5, 1–13.  
<https://doi.org/10.1038/sdata.2018.142>

Dippel, G., Beste, C., 2015.

. Nat. Commun. 6.  
<https://doi.org/10.1038/ncomms7587>

Durongbhan, P., Zhao, Yifan, Chen, L., Zis, P., De Marco, M., Unwin, Z.C., Venneri, A., He, X., Li, S., Zhao, Yitian, Blackburn, D.J., Sarrigiannis, P.G., 2019.

. IEEE Trans. Neural Syst. Rehabil. Eng. 27, 826–835. <https://doi.org/10.1109/TNSRE.2019.2909100>

Engineering, M., 2014.

0–7.

Ghumare, E.G., Schrooten, M., Vandenberghe, R., Dupont, P., 2018.

. Brain Topogr. 31, 721–737. <https://doi.org/10.1007/s10548-018-0621-3>

Hasanzadeh, F., Mohebbi, M., Rostami, R., 2022.

. IEEE J. Biomed. Heal. Informatics 26, 2299–2307.  
<https://doi.org/10.1109/JBHI.2021.3138199>

Haufe, S., Nikulin, V. V., Müller, K.R., Nolte, G., 2013.

. Neuroimage 64, 120–133. <https://doi.org/10.1016/j.neuroimage.2012.09.036>

He, B., Astolfi, L., Valdes-Sosa, P.A., Marinazzo, D., Palva, S.O., Benar, C.G., Michel, C.M., Koenig, T., 2019.

. IEEE Trans. Biomed. Eng. 66, 2115–2137.  
<https://doi.org/10.1109/TBME.2019.2913928>

Herding, J., Ludwig, S., von Lautz, A., Spitzer, B., Blankenburg, F., 2019.

. Neuroimage 201, 116011.  
<https://doi.org/10.1016/j.neuroimage.2019.116011>

Khojaste-Sarakhsi, M., Haghighi, S.S., Ghomi, S.M.T.F., Marchiori, E., 2022.  
. Artif. Intell. Med. 130, 102332. <https://doi.org/10.1016/j.artmed.2022.102332>

Kunneman, M., Pel-Littel, R., Bouwman, F.H., Gillissen, F., Schoonenboom, N.S.M., Claus, J.J., van der Flier, W.M., Smets, E.M.A., 2017.

. Alzheimer's Dement. Transl. Res. Clin. Interv. 3, 314–322. <https://doi.org/10.1016/j.trci.2017.04.002>

Li, J., Chan, S.C., Liu, Z., Chang, C., 2020.

. IEEE Access 8, 51230–51245.  
<https://doi.org/10.1109/ACCESS.2020.2979551>

Maggioni, E., Arienti, F., Minella, S., Mameli, F., Borellini, L., Nigro, M., Cogiamanian, F., Bianchi, A.M., Cerutti, S., Barbieri, S., Brambilla, P., Ardolino, G., 2021.

. Front. Aging Neurosci. 13, 1–16.  
<https://doi.org/10.3389/fnagi.2021.657221>

McBride, J., Zhao, X., Munro, N., Jicha, G., Smith, C., Jiang, Y., 2015.

. J. Healthc. Eng. 6, 55–70.  
<https://doi.org/10.1260/2040-2295.6.1.55>

Miraglia, F., Vecchio, F., Bramanti, P., Rossini, P.M., 2016.

. Clin. Neurophysiol. 127, 1261–1268. <https://doi.org/10.1016/j.clinph.2015.07.040>

Mostafavi, S., Gaiteri, C., Sullivan, S.E., White, C.C., Tasaki, S., Xu, J., Taga, M., Klein, H.U., Patrick, E., Komashko, V., McCabe, C., Smith, R., Bradshaw, E.M., Root, D.E., Regev, A., Yu, L., Chibnik, L.B., Schneider, J.A., Young-Pearse, T.L., Bennett, D.A., De Jager, P.L., 2018.

. Nat. Neurosci. 21, 811–819. <https://doi.org/10.1038/s41593-018-0154-9>

Nauta, M., Bucur, D., Seifert, C., 2019.

. Mach. Learn. Knowl. Extr. 1, 312–340. <https://doi.org/10.3390/make1010019>

Omidvarnia, A., Mesbah, M., O'Toole, J.M., Colditz, P., Boashash, B., 2011.

. 7th Int. Work. Syst. Signal Process. their Appl. WoSSPA 2011 179–182. <https://doi.org/10.1109/WOSSPA.2011.5931445>

Pezzetta, R., Ozkan, D.G., Era, V., Tieri, G., Zabberoni, S., Taglieri, S., Costa, A., Peppe, A., Caltagirone, C., Aglioti, S.M., 2023.

. npj Park. Dis. 9, 1–15. <https://doi.org/10.1038/s41531-022-00441-5>

Pisauro, M.A., Fouragnan, E., Retzler, C., Philiastides, M.G., 2017.

. Nat. Commun. 8, 1–9. <https://doi.org/10.1038/ncomms15808>

Sakkalis, V., 2011.

. Comput. Biol. Med. 41, 1110–1117. <https://doi.org/10.1016/j.compbiomed.2011.06.020>

Salehinejad, M.A., Wischnewski, M., Ghanavati, E., Mosayebi-Samani, M., Kuo, M.F., Nitsche, M.A., 2021.

. Nat. Commun. 12.

<https://doi.org/10.1038/s41467-021-24885-0>

Salman, H., Grover, J., Shankar, T., 2018.

2733, 2709–2733.

<https://doi.org/10.1162/NECO>

Sami, S., Williams, N., Hughes, L.E., Cope, T.E., Rittman, T., Coyle-Gilchrist, I.T.S., Henson, R.N., Rowe, J.B., 2018.

. Brain 141, 2500–2510. <https://doi.org/10.1093/brain/awy180>

Sarrigiannis, P.G., Zhao, Y., Wei, H.L., Billings, S.A., Fotheringham, J., Hadjivassiliou, M., 2014.

. Clin. Neurophysiol.

<https://doi.org/10.1016/j.clinph.2013.06.012>

Scherr, M., Utz, L., Tahmasian, M., Pasquini, L., Grothe, M.J., Rauschecker, J.P., Grimmer, T., Drzezga, A., Sorg, C., Riedl, V., 2021.

. Hum. Brain Mapp. 42, 4134–

4143.

Seth, A.K., Barrett, A.B., Barnett, L., 2015.

. J. Neurosci. 35, 3293–3297.

<https://doi.org/10.1523/JNEUROSCI.4399-14.2015>

. Hum. Brain

Mapp. 40, 699–712. <https://doi.org/10.1002/hbm.24446>

Talebi, N., Nasrabadi, A.M., Mohammad-Rezazadeh, I., 2018.

Cogn. Neurodyn. 12, 21–42. <https://doi.org/10.1007/s11571-017-9453-1>

Trachsel, M., Hermann, H., Biller-Andorno, N., 2015.

. Am. J. Alzheimers. Dis. Other Dement. 30, 360–363.  
<https://doi.org/10.1177/1533317514539377>

van Mierlo, P., Papadopoulou, M., Carrette, E., Boon, P., Vandenberghe, S.,  
Vonck, K., Marinazzo, D., 2014.

. Prog.  
Neurobiol. 121, 19–35. <https://doi.org/10.1016/j.pneurobio.2014.06.004>

van Vugt, M.K., Beulen, M.A., Taatgen, N.A., 2019.

. Brain Res. 1715, 1–12.  
<https://doi.org/10.1016/j.brainres.2019.03.008>

Wang, Yueming, Lin, K., Qi, Y., Lian, Q., Feng, S., Wu, Z., Pan, G., 2018.

. IEEE Trans. Biomed. Eng. 65, 1953–1963.  
<https://doi.org/10.1109/TBME.2018.2842769>

Wang, Yuan, Wang, Yao, Lui, Y.W., 2018.

. Neuroimage 178, 385–402. <https://doi.org/10.1016/j.neuroimage.2018.05.042>

Wei, H.T., Francois-Nienaber, A., Deschamps, T., Bellana, B., Hebscher, M.,  
Sivaratnam, G., Zadeh, M., Meltzer, J.A., 2021.

. Neuroimage 242.  
<https://doi.org/10.1016/j.neuroimage.2021.118457>

Xue, J., Guo, H., Gao, Y., Wang, X., Cui, H., Chen, Z., Wang, B., Xiang, J.,  
2019.

Front. Aging Neurosci. 11, 1–15. <https://doi.org/10.3389/fnagi.2019.00326>

Yang, J.B., Nguyen, M.N., San, P.P., Li, X.L., Krishnaswamy, S., 2015.

. IJCAI Int. Jt. Conf. Artif. Intell. 2015-Janua, 3995–4001.

Zarghami, T.S., Friston, K.J., 2020. . Neuroimage 207, 116453. <https://doi.org/10.1016/j.neuroimage.2019.116453>

Zhao, Y., Billings, S.A., Wei, H., He, F., Sarrigiannis, P.G., 2013.

. J. Neurosci. Methods 212, 79–86. <https://doi.org/10.1016/j.jneumeth.2012.09.019>

Zhao, Y., Billings, S.A., Wei, H., Sarrigiannis, P.G., 2012.

. Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys. 86. <https://doi.org/10.1103/PhysRevE.86.051919>

Zhao, Yifan, Zhao, Yitian, Durongbhan, P., Chen, L., Liu, J., Billings, S.A., Zis, P., Unwin, Z.C., De Marco, M., Venneri, A., Blackburn, D.J., Sarrigiannis, P.G., 2020.

. IEEE Trans. Med. Imaging 39, 1571–1581. <https://doi.org/10.1109/TMI.2019.2953584>

Zhou, X., Zhang, Z., Yu, L., Fan, B., Wang, M., Jiang, B., Su, Y., Li, P., Zheng, J., 2021.

. Epilepsy Behav. 124. <https://doi.org/10.1016/j.yebeh.2021.108308>

Zou, A., Ji, J., 2021.

. Proc. - 2021 IEEE Int. Conf. Bioinforma. Biomed. BIBM 2021 284–287. <https://doi.org/10.1109/BIBM52615.2021.9669871>



## 7 The Application of Effective Connectivity on Dementia Classification Supported by Graph Neural Networks

This Chapter is based a journal paper, 'Dementia Classification Using a Graph Neural Network on Imaging of Effective Brain Connectivity', Submitted to Artificial Intelligence in Medicine.

This chapter proposed a framework that introduces a novel directed structure learning GNN (DSL-GNN) to leverage several brain effective connectivity estimations to extract discriminative biomarkers for dementia classification. Different from previous ST-GNN, a novel graph-based framework is proposed to utilise directional features from brain networks, which allows efficiently combining brain effective connectivity estimations, as multivariate features, with univariate biomarkers thanks to the mechanism of GNN. The ST-GNN (Chapter 4) employs functional connectivity and temporal EEG information.

### 7.1 Introduction

Alzheimer's disease (AD) and Parkinson's disease (PD) are two of the most common forms of neurodegenerative diseases, resulting in the loss of memory and cognitive impairments for the former while for the latter motor symptoms dominate the clinical symptomatology prior to the development of significant cognitive deficits typically years later (Amezquita-Sanchez et al., 2019; Borzì et al., 2023; Khojaste-Sarakhsi et al., 2022; Mano et al., 2022; Zhang et al., 2022). They most commonly occur among elderly people over the age of 65, but they may also affect younger individuals (Brookmeyer et al., 2007, 1998). According to the World Health Organisation, more than 55 million people are currently diagnosed with a certain type of dementia. In this case, it has escalated the demand to develop efficient diagnosis methods for AD and PD.

Electroencephalography (EEG) is a non-invasive diagnostic method for studying the bioelectrical function and degeneration of the brain (Chen et al., 2020; Li and Jung, 2021). Compared with other clinical brain imaging techniques, EEG offers higher temporal resolution at a relatively low cost (Ng et al., 2022). The features extracted from EEG can be categorised into two groups: univariate

features that are estimated from a single channel, such as amplitude and power spectrum density (PSD); and multivariate features that estimate the relationship between different regions, namely brain connectivity, which has attracted growing attention in recent years. EEG is well-accepted to estimate brain connectivity, which can reveal complex interconnections among contiguous and remote cortical regions (Cao et al., 2022a, 2021; Tafreshi et al., 2019; van Mierlo et al., 2014). Brain connectivity is believed to help people develop a deeper understanding of various neurological conditions including brain activities and disorders with their underlying mechanisms (Cao et al., 2022b; Goldenberg and Galván, 2015; Moon et al., 2020; Mumtaz et al., 2018; Wen et al., 2022; Xie et al., 2022; Zheng et al., 2022). It can be subdivided into neuroanatomical (or structural), functional and effective connectivity (Cao et al., 2022b). Functional brain connectivity represents the statistical dependencies between regional time series. Effective brain connectivity (EBC) tends to estimate directed causal connections among neural regions (Friston, 2011; He et al., 2013; Seth et al., 2015; Wang et al., 2018; Zarghami and Friston, 2020). Besides the strength of the connection, the directionality of cortical information flow is considered as an additional biomarker (Huang et al., 2015; Supp et al., 2007). The hypothesis of this Chapter is that imaging of EBC could track more differences between AD, PD and healthy controls (HC).

One of the most popular methods for EBC identification and estimation is Granger causality (GC). GC was applied in respect of linear vector autoregressive (VAR) models of stochastic time-series data in the field of econometrics (Granger, 1969). Some modifications were introduced to the original formula by Geweke (1984). Geweke's measures of Granger causality (Geweke-Granger Causality, GGC) were expanded to unconditional causality to include conditional time series components to make it possible to distinguish between direct and indirect influences (Dhamala et al., 2018; Geweke, 1982, 1984; Varsehi and Firoozabadi, 2021). Baccalá and Sameshima (2001) introduced a novel EBC estimation method based on Granger causality - Partial Directed Coherence (PDC). It allows estimation of causality with the PDC method using a Fourier transform of multivariate autoregressive (MVAR) model

coefficients (Luiz A. Baccalá and Sameshima, 2001; Haufe et al., 2011). In other words, it operates in the frequency domain, however, it is not directly related to the power spectrum. Baccala et al. (2007) introduced a modification of that method, called Generalized Partial Directed Coherence (GPDC), to improve PDC's performance in cases of innovations noise that would cause errors in predictive modelling. Another method based on Granger causality called directed transfer function (DTF) was first introduced by Kaminski and Blinowska (1991). It applies GC on MVAR that simultaneously models all signals. Such a model, describing expressively the dataset, is transformed into the frequency domain, which can be changed into the transfer function of the system. Elements of the system indicate the causal influence from input to output in concrete frequency (Babiloni et al., 2016; He et al., 2019). Korzeniewska et al. (2003) noticed that the DTF values may be high in the parts of the spectrum where the relative power of the signal is very small, making a comparison of outflows at different frequencies. This observation led to the introduction of the full frequency Direct Transfer Function (ffDTF). Even though various EBC methods have been developed, there are limited studies to efficiently use the brain networks with directional information flows in clinical applications.

How to effectively use EBC estimations for the research of disease diagnosis is still an open problem. There are two popular strategies, including hand-crafted and imaging features (Cao et al., 2022b). On one hand, Akbarian and Erfanian (2020) employed one value to represent each pair of channels by calculating DTF, PDC, and GPDC, and then performed a binary softmax classifier on the information of EBC networks. Besides, Bae et al. (2017) estimated features of network properties based on Granger causality and successfully designed a support vector machine using the selected features to distinguish between alcoholics and normal subjects. In addition, a study was conducted on EBC for discrimination between Alzheimer's patients and healthy individuals using non-normalized DTF (Blinowska et al., 2017). Although ffDTF was tested in this Chapter, this version of the DTF method provides less accurate results in the classification task, compared with the original DTF. On the other hand, to obtain

imaging features, Saeedi (2021) reconstructed a connectivity image using two connectivity measures (PDC and DTF) and eight frequency bands as the input of deep learning networks. The proposed framework achieved an accuracy of 99.24% in the case of major depressive disorder diagnosis. Besides, Olejarczyk and Jernajczyk (2017) demonstrated the direction of information flow by imaging the strongest connections determined by DTF. It is found that, in patients with schizophrenia, there is a greater number of flows from the occipital to the frontal part of the brain, compared with healthy people. EBC can also play a significant role in the research of PD. Compared to healthy controls, PD patients showed reduced Theta- and Alpha-band EBC among EEG channels, which is estimated by PDC (Maggioni et al., 2021). Those findings suggest that EBC has the potential to reveal abnormal cortical brain activity in patients with various neurological disorders.

To deal with complex brain networks, graph neural network (GNN) has been increasingly popular in recent years (Song et al., 2021). GNN was successfully used to classify AD from health control and mild cognitive impairment (Song et al., 2021) based on the networks estimated by graph theory. It shows that GNN outperforms other machine learning classifiers in terms of classification accuracy. Besides, Zhong (2020) obtained satisfying results by combining GNN and EEG, which considers each EEG channel as a node in the graph. Brain functional connectivity was also employed to work with GNN and it obtained satisfying results in AD classification (Shan et al., 2022). However, the existing research has not employed EBC estimations as the input of GNN. That is to say, the causal relationship has not been considered to represent the difference in dementia-related diseases. Therefore, the effectiveness of the combination of the mentioned EBC and GNN techniques for dementia diagnosis requires further investigation.

This Chapter proposed a framework that introduces a novel directed structure learning GNN (DSL-GNN) to leverage several EBC estimations to extract discriminative biomarkers for dementia classification. The contributions of this Chapter can be summarised into three folds:

A novel graph-based framework is proposed to utilise directional features from brain networks, which allows efficiently combining brain connectivity estimations, as multivariate features, with univariate biomarkers thanks to the mechanism of GNN.

To the best of the authors' knowledge, this is the first time GNN has been applied to EBC in an attempt to distinguish AD, PD and HC. In comparison to existing GNN, the proposed DSL-GNN is capable of dealing with directional information.

Proper visualisation is used to reveal regional information and the difference between AD, PD and HC.

## 7.2 Methods and materials

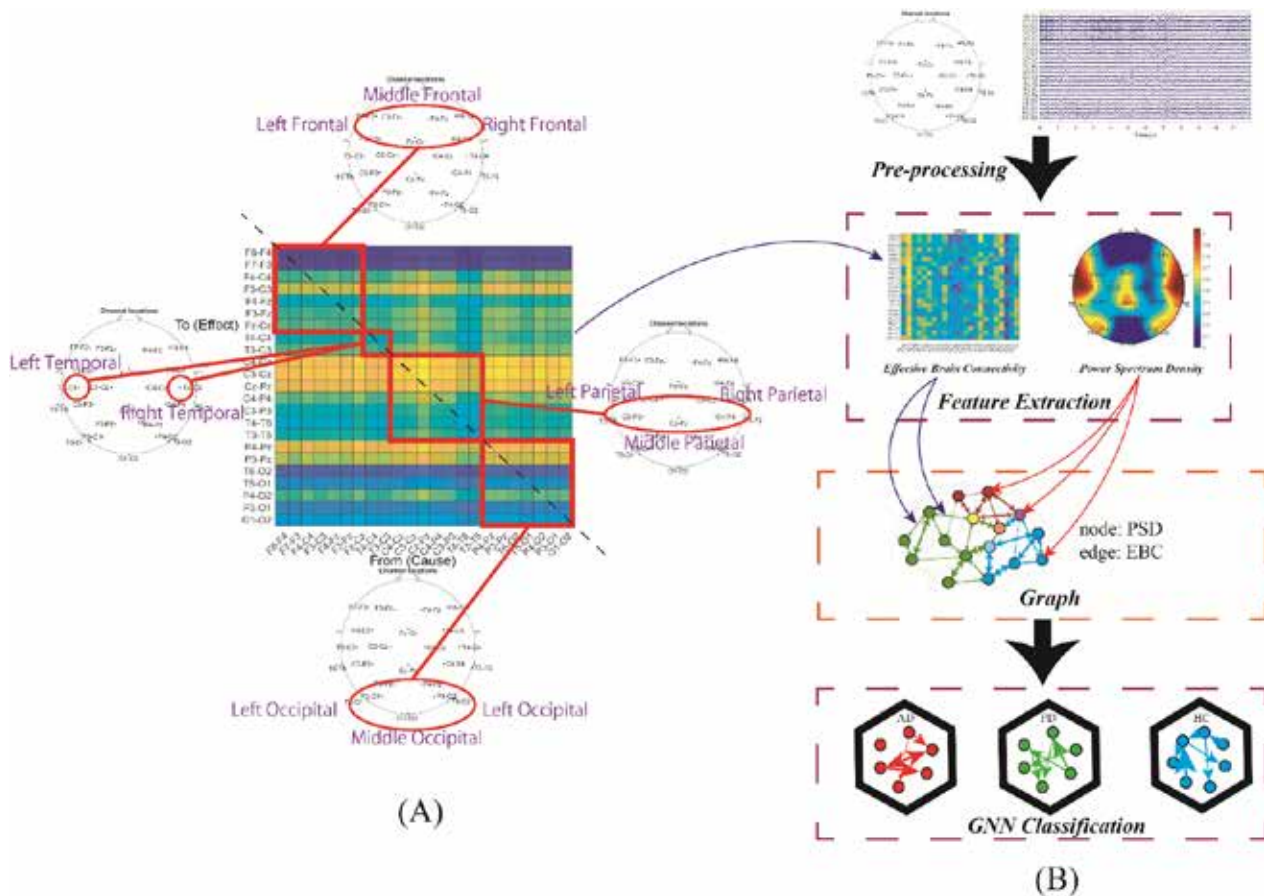


Figure 7-1 (A) Effective brain connectivity visualised by a heat map marked with the corresponding brain regions; (B) the proposed framework includes four main steps: pre-processing, feature extraction (EBC and PSD), graph construction, and DSL-GNN classification. The proposed DSL-GNN employs brain networks as the input, which is constructed by a graph comprised of PSD as nodes and EBC as directional edges.

### **7.2.1 Framework**

In the present Chapter, a novel analytical brain network imaging is constructed by a graph comprised of PSD as nodes and EBC as directional edges. We compared five EBC estimations using EEG recordings: PDC, GPDC, DTF, ffDTF and GGC. Then, a directed structure learning GNN (DSL-GNN) is developed to classify AD, PD and healthy control (HC) based on the EBC networks and frequency response of each channel. The flowchart of the proposed framework is shown in Figure 7-1, including the following steps: pre-processing, feature extraction (EBC and PSD), graph construction, and DSL-GNN classification.

### **7.2.2 Pre-processing**

EEG recordings were filtered with a pass band of 0.1-45 Hz using finite impulse response filtering. To reduce volume conduction effects related to the common reference electrode, 23 bipolar derivations were created from different brain regions the left frontal (F7–F3, F3–FZ), right frontal (F8–F4, F4–FZ), left frontocentral, temporocentral, temporal and centroparietal (F3–C3, T3–C3, T3–T5, C3–P3), right frontocentral, temporocentral, temporal and centroparietal (F4–C4, T4–C4, T4–T6, C4–P4), midline frontocentral, centroparietal and left and right parasagittal central regions (Fz–Cz, Cz–Pz, C3–Cz, C4–Cz), left parietal, parieto-occipital and temporo-occipital areas (P3–Pz, P3–O1, T5–O1), right parietal, parieto-occipital and temporo-occipital areas (P4–Pz, P4–O2, T6–O2) and midline occipital region (O1–O2)). To explore the contribution of different bands, each channel of EEG was filtered into five bands: Delta (<4 Hz), Theta (4-8 Hz), Alpha (8-12 Hz), Beta (12-32 Hz) and Gamma (>32 Hz).

### **7.2.3 Directed Structure Learning Graph Neural Network**

In this Chapter, DSL-GNN has been proposed to learn the directed connectivity from the processed connectivity map and further extract the structural feature to achieve the learning of the high-level representation. The directed graph used in

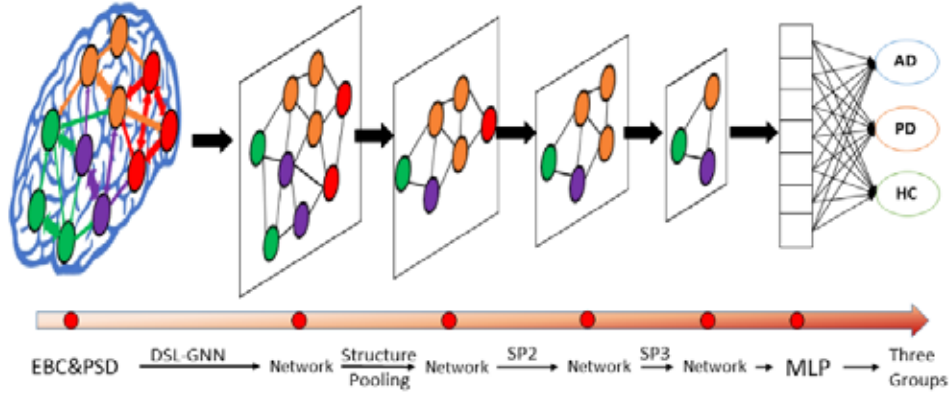


**Figure 7-2 The Flowchart of the proposed DSL-GNN. The input is constructed by a graph comprised of PSD as nodes and EBC as directional edges.**

this Chapter is denoted as  $G = (V, E)$  with  $n$  nodes  $n_i \in V$ , edges  $(v_i, v_j) \in E$ . Its adjacency matrix and feature matrix are denoted as  $A \in \mathbb{R}^{n \times n}$  and  $X \in \mathbb{R}^{n \times c}$ , respectively, where  $n$  is the number of channels which is 23 and  $c$  is the number of features of each node, which is 5 in this Chapter. The conventional Graph Convolutional Network (GCN) (Santerne et al., 2016) based methods cannot be used in the directed graph since it requests the undirected symmetric adjacency matrix. To tackle this, an approximate directed graph Laplacian method (Tong et al., 2020) based on the personalized PageRank has been introduced in this study to encode the directed features. The directed graph Laplacian  $L_{appr}$  in the symmetric normalised format is expressed as:

$$L_{appr} = I - \frac{1}{2} \left( \frac{1}{2} \frac{1}{\pi_{appr}} \frac{1}{\pi_{appr}} + \frac{1}{2} \frac{1}{\pi_{appr}} \frac{1}{\pi_{appr}}^T \frac{1}{2} \frac{1}{\pi_{appr}} \right) \quad (7-1)$$

where  $\frac{1}{\pi_{appr}} = \frac{1}{\|\pi_{appr}\|_1} \text{Diag}(\pi_{appr})$ . The personalized PageRank  $\pi_{appr}$  is a transition matrix  $\pi_{appr}$  to sparsify the Laplacian. It is defined as follows:

$$\pi_{ppr} = \begin{bmatrix} (1 - \frac{1}{n})\tilde{D}^{-1}\tilde{A} & 1^{n \times 1} \\ \frac{1}{n}1^{1 \times n} & 0 \end{bmatrix} \quad (7-2)$$

where  $\tilde{A} = A + I^{n \times n}$ , which is the self-loops added adjacency matrix.  $\tilde{D}(i, i) = \sum_j \tilde{A}(i, j)$  is the diagonal degree matrix.  $\alpha$  is the possibility of the link between node  $n$  and the auxiliary node.  $\pi_{appr}$  is a part of the  $\pi_{ppr}$ , which is the eigenvector of the transition matrix  $\pi_{ppr} \cdot \pi_{ppr} = (\pi_{appr}, \pi)$

Based on Eq. (2), the directed graph convolution is defined as:

$$H = \text{ReLU} \left( \frac{1}{2} \left( \frac{1}{2} \tilde{\sim}_{appr} + \frac{1}{2} \tilde{\sim}_{appr}^T \right) X \right) \quad (7-3)$$

where  $H^{n \times d}$  is the output of the convolution operation with the dimension of  $d$ .  $\text{ReLU}$  is the activation function,  $X$  is the feature matrix and  $\theta^{c \times d}$  is the trainable weight.

Inspired by HGP-SL (Zhang et al., 2019), we further learned the structure representation of the group through graph pooling that could extract the subgraph, which contains the structure information of the graph. If the feature of the node in the graph can be reconstructed by the features of its neighbourhood, such a node can be removed from the subgraph. To achieve that a criterion  $c$  is defined to calculate the score of each node for the  $k$ -th layer. It can be written as

$$c = \|\text{L}_{appr}^k H^k\|_1 \quad (7-4)$$

With the calculated node score, the node can be evaluated in this order. The nodes that have higher scores will be selected in the pooled subgraph. The process can be denoted as follow:

$$\begin{aligned} i &= \text{rank}(c, r \times n^k) \\ \tilde{H}^{k+1} &= H^k(i, :) \\ A^{k+1} &= A^k(i, i) \end{aligned} \quad (7-5)$$



where  $r$  is the pooling ratio and  $rank$  returns the indices of the required node number in the next layer,  $n^{k+1} = r \times n^k$ , based on the score value.

After pooling the graph into a subgraph, the new graph structure needs to be established by calculating the relationship between each pair of nodes. An attention-mechanism-based similarity score function  $S$  has been introduced to achieve this, which can be denoted as:

$$S^{k+1}(i, j) = ReLU(W[H^{k+1}(p, :) || H^{k+1}(q, :)]^T) + \lambda A^{k+1}(p, q) \quad (7-6)$$

where  $W^{1 \times 2d}$  is a trainable weight.  $H^{k+1}(p, :)^{1 \times d}$  and  $H^{k+1}(q, :)^{1 \times d}$  are the  $p$ -th and  $q$ -th row in the matrix  $H^{k+1}$ .  $||$  is the concatenation operation. Specifically,  $A^{k+1}(p, q) = 0$ , if the node  $n_p$  and  $n_q$  are not directly connected.  $\lambda$  is the parameter to adjust the relatively higher score of the direct connection. The similarity score function is further sparsified as  $S_s$  for the convolution operation, which can be expressed as:

$$S_s^{k+1}(i, j) = sparsemax(S^{k+1}(i, j)) \quad (7-7)$$

*sparsemax* is the function that could sparse the matrix with 0 and also provide non-negative and sum-to-one properties.

The structure learning convolution can be expressed as:

$$H^{k+1} = ReLU\left(\frac{1}{2} S_s^k \tilde{H}^k\right) \quad (7-8)$$

As shown in Figure 7-2, the pooling operation repeats 3 times, which provides different levels of structural information in the graph. All these levels of subgraphs need to be considered when learning the representation from the graph. Therefore, mean pooling and max pooling were used to generate a fixed size of the features. Furthermore, the features extracted from subgraphs are aggregated together for the structural representation of the graph. The generation process for the  $k$ -th layer can be expressed as:

$$g(H^k) = ReLU\left(\frac{1}{n^k} \sum_{p=1}^{n^k} H^k(p, :) || \max_{q=1}^{n^k} H^k(:, q)\right) \quad (7-9)$$

The aggregate process can be expressed as:

$$Z = \sum_{i=1}^k g(H^i) \quad (7-10)$$

Finally, a classifier with MLP layers and SoftMax function is employed to achieve the classification. The prediction output  $Y$  can be expressed as:

$$Y = \text{softmax}(\text{MLP}(Z)) \quad (7-11)$$

## 7.2.4 Effective brain connectivity estimation

The brain networks are represented by the EBC estimation performed on the 23-channel EEG recordings, which employs five EBC techniques: PDC, GPDC, DTF, ffDTF and GGC. The EBC of EEG tends to be visualised by a heat map. Figure 7-1(A) illustrates an example of a 23\*23 EBC brain network using a heat map. The colour of each kernel represents the strength of the causality from a horizontal channel to a vertical channel. The red markers indicate the regional information of the proposed visualisation with the carefully designed channel order. The principles of five employed EBC estimators are detailed represented below.

### 7.2.4.1 Granger Causality

Granger causality can estimate effective interactions from time-series data. It implements a statistical, predictive notion of causality whereby causes precede and help predict their effects (Seth et al., 2015). It is defined in both the time and frequency domains, and it allows for the conditioning out of common causal influences. The mathematical formulation of Granger causality can be represented with a bivariate linear autoregressive model of two variables  $X_1$  and

$$X_1(t) = \sum_{j=1}^p A_{11,j} X_1(t-j) + \sum_{j=1}^p A_{12,j} X_2(t-j) + E_1(t) \quad (7-12)$$

$X_2$ :

$$X_2(t) = \sum_{j=1}^p A_{21,j} X_1(t-j) + \sum_{j=1}^p A_{22,j} X_2(t-j) + E_2(t) \quad (7-13)$$

where  $p$  is the maximum number of lagged observations included in the model (the model order), the matrix  $A$  contains the coefficients of the model (i.e. the contributions of each lagged observation to the predicted values of  $X_1(t)$  and  $X_2(t)$ ), and  $E_1$  and  $E_2$  are residuals (prediction errors) for each time series. If the variance of  $E_1$  (or  $E_2$ ) is reduced by the inclusion of the  $X_1(t)$  (or  $X_2(t)$ ) terms in the first (or second) equation, then it is said that  $X_1(t)$  (or  $X_2(t)$ ) Granger-(G)-causes  $X_1$  (or  $X_2$ ). Geweke's measures of Granger causality (Geweke, 1984) expanded the unconditional causality to include conditional time series components to make it possible to distinguish between direct and indirect influences. Besides, GGC can be seen as Granger causality in the frequency domain (Dhamala et al., 2018; Geweke, 1982; Wang et al., 2014).

#### 7.2.4.2 Partial directed coherence

PDC is a method that quantifies the relation between each pair of N-channel signals, avoiding volume conduction by estimating the influences of all other signals (Tafreshi et al., 2019). PDC improves the concept of Partial Coherence by estimating causal influences. This method is estimated on MVAR.

$X(t)$  is a set of estimated signals from N recording channels:

$$X = [x_1(t), x_2(t), \dots, x_N(t)]^T \quad (7-14)$$

The MVAR process is an expressive description of the data set  $X$ :

$$\sum_{r=0}^P A(r)X(t-r) = E(t) \quad (7-15)$$

In this model  $E(t)$  is a zero-mean multivariate uncorrelated white noise vector.  $A(r)$  are the autoregressive coefficients matrix and its elements  $a_{ij}(r)$  indicates the influence of  $x_j(t-r)$  on  $x_i(t)$  and  $P$  represents the border of the model.

PDC from the  $i$ th channel to the  $j$ th channel at the frequency  $f$  is defined as follows:

$$\bar{A}_{ij}(f) = \frac{\bar{A}_{ij}(f)}{\sqrt{\sum_{m=1}^N \bar{A}_{mj}(f) \bar{A}_{mj}(f)}} \quad (7-16)$$

where  $A_{ij}(f)$  is a frequency-domain description of  $a_{ij}(r)$  (Luiz A. Baccalá and Sameshima, 2001; Granger and Aug, 2007)

$$\bar{A}_{ij}(f) = \begin{cases} 1 \sum_{r=1}^P a_{ij}(r) e^{-j2\pi fr}, & \text{if } i = j \\ \sum_{r=1}^P a_{ij}(r) e^{-j2\pi fr}, & \text{otherwise} \end{cases} \quad (7-17)$$

Generalised PDC (GPDC) is defined as follows by normalising (7-16) using standard deviation (Schelter ., 2006; Tafreshi ., 2019)

$$\bar{A}_{ij}(f) = \frac{1/\bar{A}_{ij}(f)}{\sqrt{\sum_{m=1}^N 1/\bar{A}_{mj}(f) \bar{A}_{mj}(f)}} \quad (7-18)$$

#### 7.2.4.3 Directed transfer function

The directed transfer function (DTF) is also based on the concept of Granger causality estimated on MVAR, which models all signals simultaneously (Kamifiski and Bfinowska, 1991).

In this model  $E(t)$  is a zero-mean multivariate uncorrelated white noise vector with  $A(0) = 1$ .  $A(1), A(2), \dots, A(p)$  are coefficients and  $p$  represents the model order. Equation (15) can be transformed into the frequency domain which is defined as:

$$A(f)X(f) = E(f) \quad (7-19)$$

$$A(f) = \sum_{r=0}^p A(r)e^{-j2\pi f \text{ tr}} \quad (7-20)$$

where

$$X(f) = A^{-1}(f)E(f) = H(f)E(f) \quad (7-21)$$

So,  $X(f)$  can be obtained by

$H(f)$  is the transfer function of the system and its element  $H_{ij}(f)$  indicates the causal influence from the  $j^{\text{th}}$  input to the  $i^{\text{th}}$  output at the frequency  $f$ .

$$\gamma_{ij}^2 = |H_{ij}(f)|^2 \quad (7-22)$$

The DTF is defined as:

Normalised DTF is defined as:

$$\gamma_{ij}^2(f) = \frac{|H_{ij}(f)|^2}{\sum_{m=1}^N |H_{im}(f)|^2} \quad (7-23)$$

where  $\gamma_{ij}^2(f)$  describes the influence ratio of the  $j^{\text{th}}$  channel-related cortical area on the  $i^{\text{th}}$  channel-related cortical area with respect to the influence of all estimated cortical signals (Babiloni et al., 2005). In the above formula, the denominator is dependent on frequency, so the spectral characteristics of DTF depend on the frequency properties of all EEG channels, not only the EEG input channel. To avoid this influence, the ffDTF function was introduced. In ffDTF, the denominator is integrated over frequencies. In this way, the spectrum of the estimator ffDTF is not influenced by the frequency characteristics of the denominator (Blinowska et al., 2017).

$$\gamma_{ij}^{-2}(f) = \frac{|H_{ij}(f)|^2}{\int_f \sum_{m=1}^N |H_{im}(f)|^2} \quad (7-24)$$

This is an important difference between DTF and PDC since DTF is normalised for the structure that receives the signal, whereas PDC is normalised for the structure that sends the signal (Luiz A. Baccalá and Sameshima, 2001).

### **7.2.5 Dataset**

Participants were recruited from Sheffield Teaching Hospitals NHS Foundation Trust Neurology clinics. Health controls (HC) were enrolled through educational meetings, word of mouth, family and friends. The Sheffield Teaching Hospital memory clinic provided the majority of these patients. This is a young-onset memory clinic seeing people predominantly aged under 65. The PD cohort was recruited during the Parkinson's UK Roadshow in York in July 2018. The remaining participants were recruited between September 2014 to December 2019. There were sixty participants in this work ( $70y > \text{age} > 48y$ ) and resting state EEG recordings were undertaken from all three cohorts by our research team based at the University of Sheffield. The subjects from all groups were age-matched, HC (12 female/8 male, mean age 61y,  $\pm$  SD 6.7y), AD (8 female/12 male, mean age 60y,  $\pm$  SD 4.4y) and PD (10 female/10 male, mean age 60.25y,  $\pm$  SD 4.96y). The EEG study underwent ethics approval by the Yorkshire and The Humber (Leeds West) Research Ethics Committee (reference number 14/YH/1070). AD patients had their diagnosis confirmed between one month and up to 2 years prior to their EEG recording while they had mild to moderate cognitive deficits, according to their Mini-mental state examination. All patients and controls had brain magnetic resonance imaging (MRI) scans to eliminate other alternative causes of dementia. For the age and gender-matched HC cohort, normal MRI brain scans and cognitive assessments were required before their EEG recordings. The final diagnosis of AD was based on the NINCDS-ADRDA criteria (Dubois *et al.*, 2007). The diagnosis was reached based on a consensus of multidisciplinary evidence, considering clinical history, neurological examination, neuropsychological scores and neuro-radiological findings (Blackburn *et al.*, 2018). In addition, all 20 PD patients had a clinical diagnosis by a consultant with expertise in diagnosing and managing PD. All had normal cognitive function but they were in

different stages of their disease. They also varied in the medication they were on and the symptoms they were experiencing.

EEG recordings were undertaken with an XLTEK 128-channel headbox (Optima Medical LTD) and Ag/AgCL electrodes at a sampling frequency of 2k Hz by implementing a modified 10-10 overlapping a 10-20 international system of electrode placement, with a referential montage (linked earlobe reference). Thirty-minute resting state (task-free – participants were instructed to rest and refrain from thinking anything specific). EEG recordings were obtained from each participant including sustained periods of keeping their eyes closed (EC) alternating with periods during which they kept their eyes open (EO). The recordings obtained were subsequently reviewed by a neurophysiologist - on an XLTEK review station - and for each participant, three 12-second artefact-free mini-epochs of EC and EO were selected. In this Chapter, the EO and EC data are combined for analysis.

### **7.2.6 Classification**

As shown in Figure 7-1, the EBC and PSD features are fed into the proposed DSL-GNN to classify AD, PD and HC. To be more specific, for the Full band, the edges of DSL-GNN are represented by the EBC network ( $23 \times 23$ ) estimated by full-band EEG and the nodes are represented by the PSD values of 23 channels in five separate bands. For each band, the edges are represented by the EBC networks ( $23 \times 23$ ) estimated by corresponding-band EEG and the nodes are represented by the corresponding-band PSD values of 23 channels. There are 19 AD participants, 20 PD participants and 20 HC participants, each of which contains 6 epochs of EEG recording. Therefore, for the AD vs. HC task, the total number of data samples is  $114+120=234$ ; for the AD vs. PD task, the total number of data samples is  $114+120=234$ ; for HC vs. PD task, the total number of data samples is  $120+120=240$ ; and for AD vs. PD vs. HC task, the total number of data samples is  $114+120+120=354$ . The five-fold cross-validation is employed to avoid artificial bias resulting from selecting data and more efficiently utilise the limited data (Jamal et al., 2014). In this case, 80% of

the dataset is used as the training set and the remaining 20% is used as the testing set.

For comparison, a Support Vector Machine (SVM) classifier is employed. To fairly reveal the performance difference, the same features extracted by EBC and PSD are fed into the proposed framework and SVM classifier, namely, for Full band; and

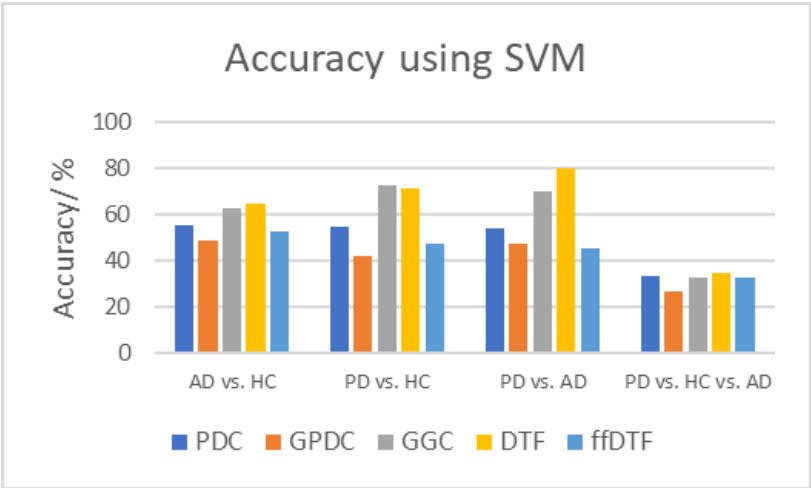


**Figure 7-4 Classification accuracy in Full band using SVM, employing five EBC estimations: PDC, GPDC, GGC, DTF and ffDTF.**

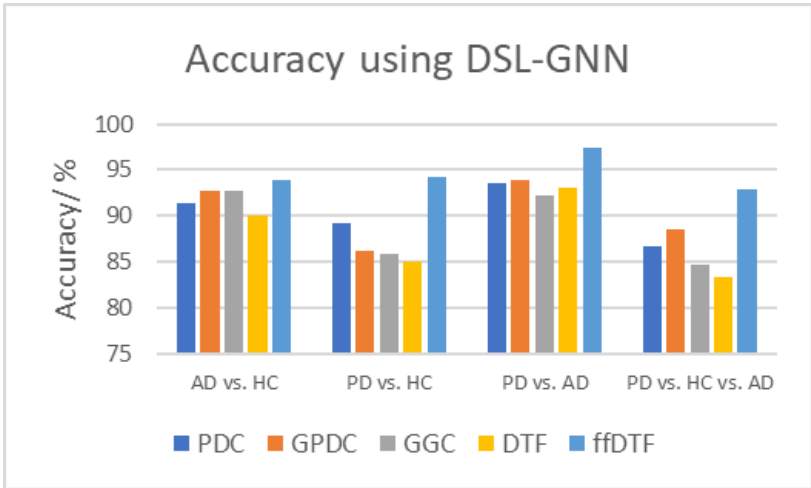


**Figure 7-3 Classification accuracy in Full band using DSL-GNN, employing five EBC estimations: PDC, GPDC, GGC, DTF and ffDTF.**

for each of five separate bands.



## 7.3 Results

### 7.3.1 Classification Performance

In this Chapter, the proposed framework was evaluated in four classification tasks, namely, AD vs. HC, PD vs. HC, PD vs. AD and HC vs. AD vs. PD. The resulting accuracy achieved by SVM and DSL-GNN in the Full band is shown in Figure 7-4Error! Reference source not found. and Figure 7-3 respectively using five EBC estimations individually. Overall, there is a significant improvement in accuracy (at least a 30% increase) for all four tasks using the proposed method, which suggests that DSL-GNN can simultaneously utilise both EBC (multivariate) and PSD (univariate) features more efficiently than SVM. The greatest increase occurred in the three-group discrimination, where DSL-GNN achieved 93.0% while SVM only provided 35% using both ffDTF and PSD features. As observed in Error! Reference source not found., the average performance of SVM in the three-group classification task is close to 33% while that of all other three binary classifications is slightly higher than 50%, suggesting the relatively poor performance of this approach. DSL-GNN is only limited by the worst accuracy among the binary classifications. For example, the results of the three-group classification are extremely close to the PD vs. HC task, due to the accuracies of this task are lowest among the three binary classification tasks. Accuracies of DSL-GNN (96.6%) and SVM (80.0%) for classifying PD vs. AD are highest within the four tasks, which indicates that the difference between PD and AD is the most significant. It should be noted that the EBC estimation with the best performance changes from DTF for SVM to ffDTF for DSL-GNN. Furthermore, ffDTF estimation has obvious superiority over the other four estimations when DSL-GNN is employed, by producing the classification accuracy of over 93% for all four tasks.

### 7.3.2 Band Contribution

To further explore the contribution of each band, the EBC and PSD features from every single band are used for comparison. Table 7-1 and Table 7-2 represent the classification accuracy based on different EBC techniques together with PSD in five separate bands, using SVM and DSL-GNN respectively. It can be found that DSL-GNN using single-band features performs

**Table 7-1 Classification accuracy using SVM in five different bands using five EBC estimations and PSD (%)**

| AD vs. HC           | Delta | 56.88 | 51.26 | 61.53 | 51.26 | 49.13 |
|---------------------|-------|-------|-------|-------|-------|-------|
|                     | Theta | 54.71 | 48.29 | 64.91 | 68.41 | 54.73 |
|                     | Alpha | 52.56 | 42.75 | 64.52 | 68.82 | 55.56 |
|                     | Beta  | 62.40 | 44.01 | 62.38 | 57.72 | 51.30 |
|                     | Gamma | 54.31 | 47.42 | 56.85 | 71.41 | 62.01 |
| PD vs. HC           | Delta | 52.92 | 43.33 | 73.75 | 56.25 | 47.08 |
|                     | Theta | 55.00 | 40.83 | 71.25 | 73.08 | 49.17 |
|                     | Alpha | 57.08 | 43.33 | 70.83 | 74.58 | 46.67 |
|                     | Beta  | 60.83 | 50.00 | 72.92 | 70.00 | 46.25 |
|                     | Gamma | 57.50 | 44.17 | 68.33 | 65.83 | 48.75 |
| PD vs. AD           | Delta | 54.67 | 51.28 | 69.66 | 67.05 | 48.69 |
|                     | Theta | 57.65 | 46.61 | 71.40 | 73.94 | 51.76 |
|                     | Alpha | 58.96 | 51.30 | 68.44 | 66.61 | 56.42 |
|                     | Beta  | 62.77 | 48.30 | 70.09 | 70.88 | 52.16 |
|                     | Gamma | 56.40 | 46.61 | 71.80 | 65.83 | 62.00 |
| PD vs. HC<br>vs. AD | Delta | 32.20 | 30.23 | 34.46 | 32.20 | 31.07 |
|                     | Theta | 31.92 | 26.55 | 37.29 | 31.07 | 31.07 |
|                     | Alpha | 32.20 | 25.99 | 32.20 | 34.75 | 28.53 |
|                     | Beta  | 32.77 | 26.84 | 32.49 | 31.92 | 28.53 |
|                     | Gamma | 30.23 | 29.66 | 33.33 | 39.83 | 32.49 |

much worse than the Full band due to a significant reduction of feature number, while it has smaller influences on SVM. The contribution of 5 bands to the classification outcome for SVM and DSL-GNN differs a lot. Specifically, for the AD vs. HC task, the Theta band performed better than other bands using the DSL-GNN method (about 75%), while SVM provides the highest accuracy in the Gamma band (71.4%) using the DTF technique. For the PD vs. HC task, SVM

**Table 7-2 Classification accuracy using DSL-GNN in five different bands using five EBC estimations and PSD (%)**

| AD vs. HC           | Delta | 62.43 | 62.70 | 60.99 | 61.50 | 70.99 |
|---------------------|-------|-------|-------|-------|-------|-------|
|                     | Theta | 75.23 | 76.43 | 74.37 | 74.37 | 75.60 |
|                     | Alpha | 60.30 | 69.58 | 59.83 | 61.13 | 69.48 |
|                     | Beta  | 57.69 | 54.61 | 58.49 | 57.32 | 61.10 |
|                     | Gamma | 49.63 | 58.59 | 51.63 | 51.70 | 67.84 |
| PD vs. HC           | Delta | 50.00 | 58.75 | 49.17 | 49.17 | 67.08 |
|                     | Theta | 63.75 | 66.25 | 61.67 | 63.33 | 62.92 |
|                     | Alpha | 62.50 | 60.83 | 61.67 | 59.58 | 70.42 |
|                     | Beta  | 50.00 | 54.17 | 50.42 | 50.00 | 63.33 |
|                     | Gamma | 67.92 | 75.83 | 63.33 | 63.75 | 71.25 |
| PD vs. AD           | Delta | 49.63 | 61.17 | 59.18 | 53.81 | 72.40 |
|                     | Theta | 71.29 | 73.86 | 71.69 | 71.22 | 76.90 |
|                     | Alpha | 61.90 | 78.54 | 64.03 | 69.98 | 81.95 |
|                     | Beta  | 55.55 | 59.46 | 51.34 | 49.86 | 65.48 |
|                     | Gamma | 59.39 | 79.81 | 54.78 | 56.52 | 84.99 |
| PD vs. HC<br>vs. AD | Delta | 36.69 | 46.27 | 32.53 | 33.10 | 54.27 |
|                     | Theta | 55.91 | 56.53 | 55.05 | 55.59 | 59.31 |
|                     | Alpha | 45.43 | 53.37 | 44.59 | 42.08 | 52.56 |
|                     | Beta  | 35.30 | 44.42 | 36.71 | 33.75 | 45.19 |
|                     | Gamma | 46.22 | 48.90 | 42.02 | 38.09 | 44.89 |

performs the greatest in the Alpha band (74.58%) using DTF, while DSL-GNN provides the highest accuracy in the Gamma band (75.83%) using GPDC. For the PD vs. AD task, the best performance is provided by DTF in the Theta band (73.94%),

while DSL-GNN achieved the highest accuracy of 84.99% in the Gamma band by employing ffDTF. For the three-group classification, both SVM and DSL-GNN obtained unsatisfactory classification performance, even though DSL-GNN performs better than SVM with about a 15% increase in accuracy. Besides, single-band features provide much worse performance in comparison to Full-band classification with a more than 30% decrease in accuracy, while the influence on binary classification is much less. It seems that the three-group classification task requires much more features than the other three tasks. This

finding may imply that the features and bands that contribute to each binary classification task vary, resulting in the Full band with more information can support better performance in three-group discrimination. In the meantime, the proposed DSL-GNN's ability to deal with a great number of features is necessary, relying on the novel graph analysis of brain networks.

### **7.3.3 Feature Visualisation**

The heat map (Figure 7-5) and topography method (Figure 7-6) are used to visualise the normalised average EBC estimations and PSD values, respectively. In Figure 7-5, it is obvious that the frontal areas are rarely influenced by other regions since the ffDTF value of the first two rows is extremely close to zero for three groups in every band. Compared to HC, the AD and PD groups exhibit different changes. For example, the ffDTF value of the AD group in occipital regions (Effect) increases dramatically, especially in the low-frequency bands (Delta, Theta and Alpha) from 0.5 of HC to 0.8 of AD, which suggests stronger information flow from the whole brain to the occipital area. It should be noted that the biggest change of AD occurs in the Theta band, where the whole brain networks have a significant difference. That explains the highest accuracy of discriminating AD from HC is achieved by DSL-GNN in the Theta band to a certain extent. The left parietal and right occipital of the PD group are more causally influenced by the other regions in all bands, evidenced by the increase of ffDTF value from 0.4 to 0.8. It is worth noting that the bidirectional information flows within the parietal and occipital almost disappear for the PD group, which is significantly different from AD and HC. These findings agree with the classification results of PD vs. HC and PD vs. AD, which suggests the best performance is achieved in the Gamma band.

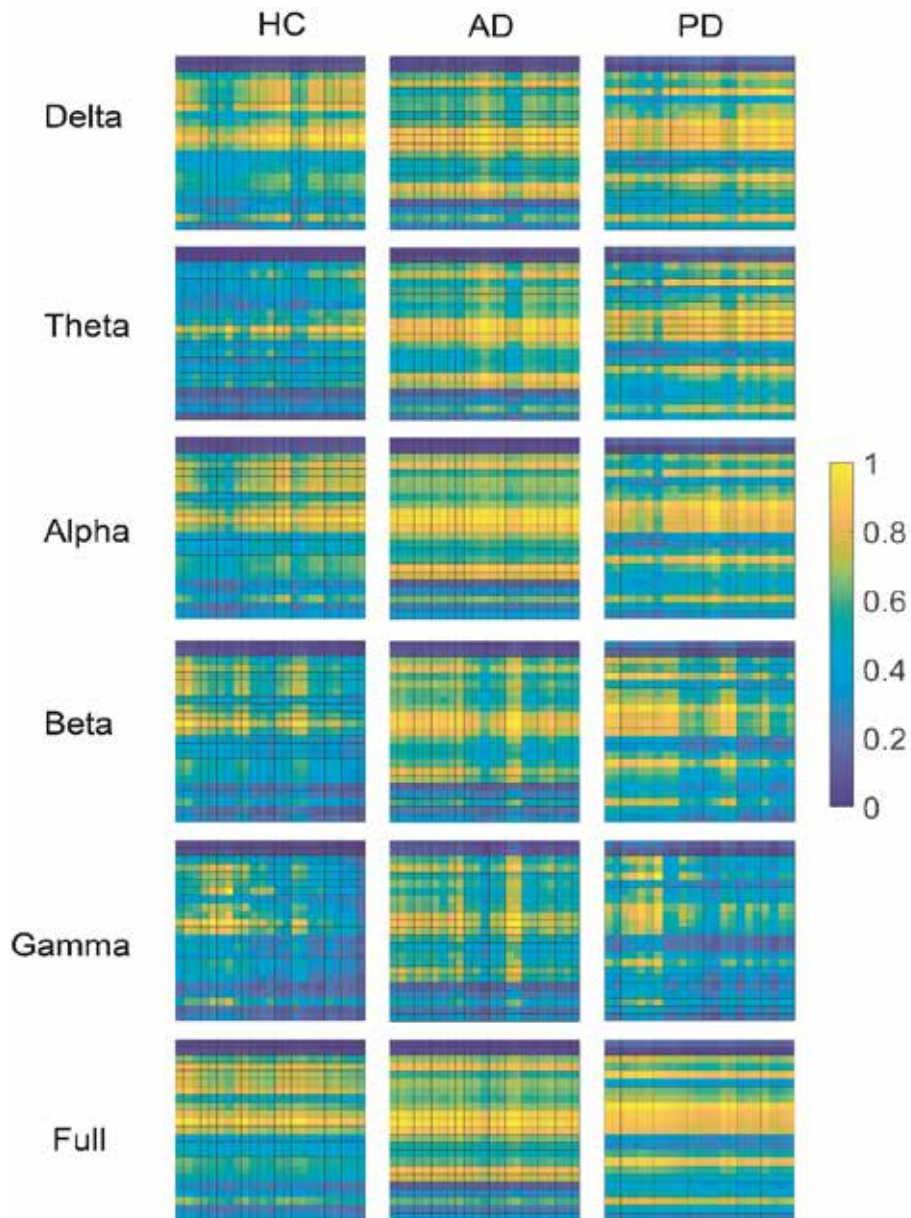


**Figure 7-5 The normalised average fDFTF in different band for HC, AD and PD**

The topographic map is normalised in each band and group individually. That is to say, Figure 7-6 only illustrates the different patterns of PSD distribution and it is meaningless to compare the absolute values in each topography. In Figure 7-6, it is observed that the PSD distribution of PD shows more difference, compared with HC and AD. Specifically, in the low-frequency band, PD has clear higher PSD values in the parietal region of 0.8-0.9, while HC and AD range from 0.3 to 0.7. In the Gamma band, the frontal PSD of PD ( $<0.1$ ) is significantly lower than both AD and HC ( $>0.8$ ). HC and AD are relatively similar

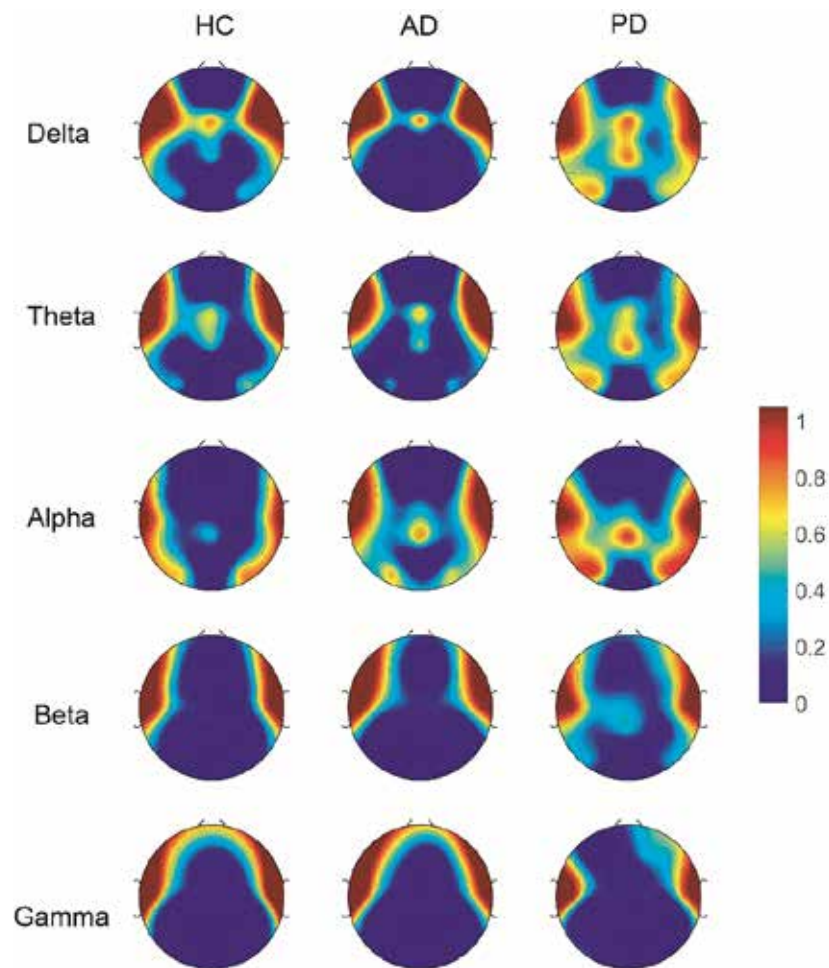


**Figure 7-6 The normalised average PSD in different bands for HC, AD and PD. The PSD distribution of PD shows more difference, compared with HC and AD.**

in almost all bands. The greatest difference occurs in the Alpha band, in which the PSD of HC is 0.3 and AD is 0.7 in the parietal region; and PSD of HC is 0.8 and AD is 0.6 in the occipital areas. Therefore, the bands and areas that contribute the most to AD and PD classification from HC differ, and the degree of difference between AD and HC, and PD and HC varies.

## 7.4 Discussion

The novelty of this Chapter lies in the development of a new Graph Neural Network, called DSL-GNN, to better diagnose AD by integrating both univariate (PSD) and multivariate features (EBC) from EEG recording. The results have proven its effectiveness in improving the classification of HC, AD and PD based on the studied dataset. Table 7-3 compares our proposed method with the

**Table 7-3 Performance of popular features and classifiers that are implemented on the same dataset in the discrimination task of AD or PD**

| Reference                                   | Features                              | Task             | Classifier     | Accuracy     |
|---------------------------------------------|---------------------------------------|------------------|----------------|--------------|
| (Wang et al., 2020)                         | PSD                                   | AD vs. HC        | SVM            | 70.2%        |
| (Nimmy et al., 2019)                        | Mutual information                    | AD vs. HC        | SVM            | 77.23%       |
| (Engels et al., 2015)(Kabbara et al., 2017) | Phase locking value                   | AD vs. HC        | SVM            | 72.11%       |
| (Shan et al., 2022)                         | Wavelet coherence                     | AD vs. HC        | GNN            | 92.3%        |
| (Maggioni et al., 2021)                     | GC                                    | PD vs. HC        | GNN            | 86.2%        |
| (Hassan et al., 2017)                       | PSD +Phase locking value              | PD vs. HC        | SVM            | 82.9%        |
| (Yuvaraj et al., 2016)                      | Correlation+Coherence+Phase lag index | PD vs. HC        | SVM            | 80.2%        |
| (Jeong et al., 2016)                        | Wavelet coherence                     | AD vs. PD vs. HC | LDA            | 40.4%        |
| Our study                                   | <b>EBC+PSD</b>                        | AD vs. HC        | <b>DSL-GNN</b> | <b>94.0%</b> |
|                                             |                                       | PD vs. HC        |                | <b>94.2%</b> |
|                                             |                                       | AD vs. PD        |                | <b>97.4%</b> |
|                                             |                                       | AD vs. PD vs. HC |                | <b>93.0%</b> |

existing methods in terms of the discrimination of AD and PD from HC. The literature review suggests that very limited research focuses on the three-group classification of AD, PD and HC, although several studies obtained good results in binary classification, e.g., AD vs. HC and PD vs. HC. As cognitive dysfunctions can emerge over time in PD patients as the majority develop dementia over a decade (Jeong et al., 2016) there is a clinical need in finding promising biomarkers to explore the differences between AD and PD patients. Our study not only achieved satisfying classification results in distinguishing dementia from HC, but also obtained an accuracy of 93.0% in three-group discrimination. Notably, Jeong et al. (2016) performed linear discriminant analysis (LDA) on the functional connectivity estimated by Wavelet coherence and obtained approximately 80% accuracy in three binary classification tasks. However, the three-group classification was not conducted. The same features and classifier are used to deal with our dataset in the task of AD vs. PD vs. HC and the accuracy is 40.4%, which is quite similar to the results obtained by SVM using EBC in this Chapter. The proposed framework significantly outperforms this method with an accuracy of 93.3%. Besides, Shan et al. (2022) also used Wavelet coherence as the indicator, but GNN was applied to classify AD from HC only. The combination of Wavelet coherence and GNN achieved an accuracy of 92.3%, which is slightly lower than this Chapter. Therefore, the existing research and our findings suggest that the GNN-based classifier has more capabilities in making use of numerous and complex brain network features. In the meantime, the EBC could provide more information than functional connectivity to an extent, as directionality is also used. Relying on the advancement of the proposed DSL-GNN method considering directional edges, it is not surprising that the classification performance is improved dramatically.

The Theta band outperforms other bands in the AD vs. HC task, whichever EBC method is used. Besides, low-frequency bands (0-12 Hz including Delta, Theta and Alpha) perform better than the high-frequency band (12-45 Hz including Beta and Gamma). This finding is also supported by several previous studies (Babiloni et al., 2016; Blinowska et al., 2017; Lehmann et al., 2007), which suggest that the abnormalities of EEG recording at a specific frequency band



(<12 Hz) are easier to be found in AD populations. Different from the AD vs. HC task, the Gamma band contributes the most to the discrimination of PD participants with relatively higher accuracy, which agrees with several recent research studies. For example, Muthuraman (2020) found abnormal cross-frequency coupling at the Gamma band, which was not revealed within the control subject. Besides, Yuvaraj (2016) reported that there are no significant differences between PD patients and HC participants observed in Delta and Theta frequency bands. However, a significant difference is observed in the Beta and Gamma bands. They used four functional brain connectivity methods, including correlation, coherence, phase synchronization index, and bispectrum. Therefore, the band used to diagnose AD and PD from HC is quite different, which may explain the unsatisfying accuracy obtained by individual bands in the three-way classification task (about 50%), whereas DSL-GNN provides a highly promising accuracy of 93% using full-band features.

Although deep learning methods play an important role in the study of dementia diagnosis in recent years (Bi and Wang, 2019; Ieracitano et al., 2020; Shah et al., 2020), effectively making use of numerous features is still challenging. Many studies suggest that more features do not always provide higher classifying accuracy (Jeong et al., 2016; Kakkos et al., 2021). That is to say, the best performance is obtained using a certain percentage of various features. Undeniably, the selection of features would reduce the efficiency of the classifying framework due to it is computationally demanding. I proposed DSL-GNN method demonstrated its superiority in systematically using various brain-network markers since the mechanism of GNN supports a reasonable combination of distinct features. Specifically, the proposed framework constructs a graph using univariate features (PSD) and multivariate features (EBC) and feeds the constructed graph to the classifier. In this case, it is more robust than simply feeding all mixed features into the classifier. It learns the additional interconnection and even the direction among the nodes. That explains the GNN-based method performs significantly better than other machine learning algorithms in this Chapter, especially in the three-group classification task in the Full band where more features are employed.

## 7.5 Conclusion

In this Chapter, a novel DSL-GNN was developed and performed on the EBC estimations and PSD features for AD dementia classification. In comparison to the previous studies on GNN, our proposed approach enhanced the functionality of processing directional information, which builds the basis for more efficiently performing GNN on EBC. Another contribution of this Chapter is creating a new framework for learning univariate and multivariate features simultaneously. The proposed framework and DSL-GNN are validated in four discrimination tasks and our approach exhibited the best performance with the highest accuracy of 94.0% (AD vs. HC), 94.2% (PD vs. HC), 97.4% (AD vs. PD) and 93.0% (AD vs. PD vs. HC). Additionally, the contribution of different bands is verified for these four tasks. Specifically, the most significant band varied, for example, the Theta band for AD vs. HC, the Gamma band for PD vs. HC, the Alpha and Gamma band for PD vs. AD, and the Theta band for AD vs. PD vs. HC. This research empowers the application of GNN in dementia diagnosis based on its significant outcomes. Finally, the proposed framework raises several opportunities and challenges for future research. The proposed framework employs 5 EBC techniques as the multivariate features and PSD as the univariate features, while there are an increasing number of promising brain connectivity and single-channel biomarkers. It can be expected the performance will be further improved if more robust features are utilised.

## 7.6 References

- Akbarian, B., Erfanian, A., 2020. Biomedical Signal Processing and Control A framework for seizure detection using effective connectivity , graph theory , and multi-level modular network. *Biomed. Signal Process. Control* 59, 101878. <https://doi.org/10.1016/j.bspc.2020.101878>
- Amezquita-Sanchez, J.P., Mammone, N., Morabito, F.C., Marino, S., Adeli, H., 2019. A novel methodology for automated differential diagnosis of mild cognitive impairment and the Alzheimer's disease using EEG signals. *J. Neurosci. Methods* 322, 88–95. <https://doi.org/10.1016/j.jneumeth.2019.04.013>

Babiloni, C., Lizio, R., Marzano, N., Capotosto, P., Soricelli, A., Triggiani, A.I., Cordone, S., Gesualdo, L., Del Percio, C., 2016. Brain neural synchronization and functional coupling in Alzheimer's disease as revealed by resting state EEG rhythms. *Int. J. Psychophysiol.* 103, 88–102. <https://doi.org/10.1016/j.ijpsycho.2015.02.008>

Babiloni, F., Cincotti, F., Babiloni, C., Carducci, F., Mattia, D., Astolfi, L., Basilisco, A., Rossini, P.M., Ding, L., Ni, Y., Cheng, J., Christine, K., Sweeney, J., He, B., 2005. Estimation of the cortical functional connectivity with the multimodal integration of high-resolution EEG and fMRI data by directed transfer function 24, 118–131. <https://doi.org/10.1016/j.neuroimage.2004.09.036>

Baccalá, Luiz A., Sameshima, K., 2001. Partial directed coherence: a new concept in neural structure determination. *Biol. Cybern.* 84, 463–474.

Baccalá, Luiz A., Sameshima, K., 2001. Partial directed coherence: A new concept in neural structure determination. *Biol. Cybern.* 84, 463–474. <https://doi.org/10.1007/PL00007990>

Baccala, L.A., Sameshima, K., Takahashi, D.Y., 2007. Generalized Partial Directed Coherence, in: 2007 15th International Conference on Digital Signal Processing. pp. 163–166. <https://doi.org/10.1109/ICDSP.2007.4288544>

Bae, Y., Yoo, B.W., Lee, J.C., Kim, H.C., 2017. Automated network analysis to measure brain effective connectivity estimated from EEG data of patients with alcoholism. *Physiol. Meas.* 38, 759–773. <https://doi.org/10.1088/1361-6579/aa6b4c>

Bi, X., Wang, H., 2019. Early Alzheimer's disease diagnosis based on EEG spectral images using deep learning. *Neural Networks* 114, 119–135.

Blackburn, D.J., Sarrigiannis, P.G., De Marco, M., Zhao, Y., Venneri, A., Lawrence, S., Unwin, Z.C., Blyth, M., Angel, J., Baster, K., Wilkinson, I.D., Bell, S.M., He, F., Wei, H.L., Billings, S.A., Farrow, T.F.D., 2018. 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 Sci.* 8, 1–19. <https://doi.org/10.3390/brainsci8070134>

Blinowska, K.J., Rakowski, F., Kaminski, M., De Vico Fallani, F., Del Percio, C., Lizio, R., Babiloni, C., 2017. Functional and effective brain connectivity for discrimination between Alzheimer's patients and healthy individuals: A study on resting state EEG rhythms. *Clin. Neurophysiol.* 128, 667–680. <https://doi.org/10.1016/j.clinph.2016.10.002>

Borzì, L., Sigcha, L., Rodríguez-Martín, D., Olmo, G., 2023. Real-time detection of freezing of gait in Parkinson's disease using multi-head convolutional neural networks and a single inertial sensor. *Artif. Intell. Med.* 135, 102459. <https://doi.org/10.1016/j.artmed.2022.102459>

Brookmeyer, R., Gray, S., Kawas, C., 1998. Projections of Alzheimer's disease in the United States and the public health impact of delaying disease onset. *Am. J. Public Health* 88, 1337–1342.

Brookmeyer, R., Johnson, E., Ziegler-Graham, K., Arrighi, H.M., 2007. Forecasting the global burden of Alzheimer's disease. *Alzheimer's Dement.* 3, 186–191.

Cao, J., Grajcar, K., Shan, X., Zhao, Y., Zou, J., Chen, L., Li, Z., Grunewald, R., Zis, P., Marco, M. De, Unwin, Z., Blackburn, D., Sarrigiannis, P.G., 2021. Using interictal seizure-free EEG data to recognise patients with epilepsy based on machine learning of brain functional connectivity. *Biomed. Signal Process. Control* 67, 102554. <https://doi.org/10.1016/j.bspc.2021.102554>

Cao, J., Zhao, Y., Shan, X., Blackburn, D., Wei, J., Erkoyuncu, J.A., Chen, L., Sarrigiannis, P.G., 2022a. Ultra-high-resolution time-frequency analysis of EEG to characterise brain functional connectivity with the application in Alzheimer's disease. *J. Neural Eng.* 19, 046034. <https://doi.org/10.1088/1741-2552/ac84ac>

Cao, J., Zhao, Y., Shan, X., Wei, H., Guo, Y., Chen, L., Erkoyuncu, J.A., Sarrigiannis, P.G., 2022b. Brain functional and effective connectivity based on

electroencephalography recordings: A review. *Hum. Brain Mapp.* 43, 860–879.  
<https://doi.org/10.1002/hbm.25683>

Chen, Y., Tang, Y., Wang, C., Liu, X., Zhao, L., Wang, Z., 2020. ADHD classification by dual subspace learning using resting-state functional connectivity. *Artif. Intell. Med.* 103, 101786.  
<https://doi.org/10.1016/j.artmed.2019.101786>

Dhamala, M., Liang, H., Bressler, S.L., Ding, M., 2018. Granger-Geweke causality: Estimation and interpretation. *Neuroimage* 175, 460–463.

Dubois, B., Feldman, H.H., Jacova, C., DeKosky, S.T., Barberger-Gateau, P., Cummings, J., Delacourte, A., Galasko, D., Gauthier, S., Jicha, G., Meguro, K., O'Brien, J., Pasquier, F., Robert, P., Rossor, M., Salloway, S., Stern, Y., Visser, P.J., Scheltens, P., 2007. Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. *Lancet Neurol.* 6, 734–746.  
[https://doi.org/10.1016/S1474-4422\(07\)70178-3](https://doi.org/10.1016/S1474-4422(07)70178-3)

Engels, M.M.A., Stam, C.J., Flier, W.M. Van Der, Scheltens, P., Waal, H. De, Straaten, E.C.W. Van, 2015. Declining functional connectivity and changing hub  
C *Neurol.* 1–8.  
<https://doi.org/10.1186/s12883-015-0400-7>

Friston, K.J., 2011. Functional and Effective Connectivity: A Review. *Brain Connect.* 1, 13–36. <https://doi.org/10.1089/brain.2011.0008>

Geweke, J., 1982. Measurement of Linear Dependence and Feedback between Multiple Time Series. *J. Am. Stat. Assoc.* 77, 304–313.  
<https://doi.org/10.1080/01621459.1982.10477803>

Geweke, J.F., 1984. Measures of conditional linear dependence and feedback between time series. *J. Am. Stat. Assoc.* 79, 907–915.

Goldenberg, D., Galván, A., 2015. The use of functional and effective connectivity techniques to understand the developing brain. *Dev. Cogn. Neurosci.* 12, 155–164. <https://doi.org/10.1016/j.dcn.2015.01.011>

Granger, C.W.J., 1969. Investigating causal relations by econometric models and cross-spectral methods. *Econom. J. Econom. Soc.* 424–438.

Granger, C.W.J., Aug, N., 2007. Investigating Causal Relations by Econometric Models and Cross-spectral Methods 37, 424–438.

Hassan, M., Chaton, L., Benquet, P., Delval, A., Leroy, C., Plomhause, L., Moonen, A.J.H., Duits, A.A., Leentjens, A.F.G., van Kranen-Mastenbroek, V., Defebvre, L., Derambure, P., Wendling, F., Dujardin, K., 2017. Functional connectivity disruptions correlate with cognitive phenotypes in Parkinson's disease. *NeuroImage Clin.* 14, 591–601. <https://doi.org/10.1016/j.nicl.2017.03.002>

Haufe, S., Nikulin, V., Nolte, G., 2011. Identifying brain effective connectivity patterns from EEG: performance of Granger Causality, DTF, PDC and PSI on simulated data. *BMC Neurosci.* 12, 1–2. <https://doi.org/10.1186/1471-2202-12-s1-p141>

He, B., Astolfi, L., Valdes-Sosa, P.A., Marinazzo, D., Palva, S.O., Benar, C.G., Michel, C.M., Koenig, T., 2019. Electrophysiological Brain Connectivity: Theory and Implementation. *IEEE Trans. Biomed. Eng.* 66, 2115–2137. <https://doi.org/10.1109/TBME.2019.2913928>

He, F., Billings, S.A., Wei, H.L., Sarrigiannis, P.G., Zhao, Y., 2013. Spectral analysis for nonstationary and nonlinear systems: A discrete-time-model-based approach. *IEEE Trans. Biomed. Eng.* 60, 2233–2241. <https://doi.org/10.1109/TBME.2013.2252347>

Huang, C.S., Pal, N.R., Chuang, C.H., Lin, C.T., 2015. Identifying changes in EEG information transfer during drowsy driving by transfer entropy. *Front. Hum. Neurosci.* 9, 1–12. <https://doi.org/10.3389/fnhum.2015.00570>

Ieracitano, C., Mammone, N., Hussain, A., Morabito, F.C., 2020. A novel multi-modal machine learning based approach for automatic classification of EEG recordings in dementia. *Neural Networks* 123, 176–190.

Jamal, W., Das, S., Oprescu, I.A., Maharatna, K., Apicella, F., Sicca, F., 2014. Classification of autism spectrum disorder using supervised learning of brain connectivity measures extracted from synchrostates. *J. Neural Eng.* 11. <https://doi.org/10.1088/1741-2560/11/4/046019>

Jeong, D.H., Kim, Y. Do, Song, I.U., Chung, Y.A., Jeong, J., 2016. Wavelet energy and wavelet coherence as eeg biomarkers for the diagnosis of Parkinson's disease-related dementia and Alzheimer's disease. *Entropy* 18. <https://doi.org/10.3390/e18010008>

Kabbara, A., Falou, W. El, Khalil, M., Eid, H., Hassan, M., 2017. A scalp-EEG network-based analysis of Alzheimer's disease patients at rest. *Int. Conf. Adv. Biomed. Eng. ICABME* 2017-Octob, 0–3. <https://doi.org/10.1109/ICABME.2017.8167549>

Kakkos, I., Dimitrakopoulos, G.N., Sun, Yi, Yuan, J., Matsopoulos, G.K., Bezerianos, A., Sun, Yu, 2021. EEG Fingerprints of Task-Independent Mental Workload Discrimination. *IEEE J. Biomed. Heal. Informatics* 25, 3824–3833. <https://doi.org/10.1109/JBHI.2021.3085131>

Kamifiski, M.J., Bfinowska, K.J., 1991. A new method of the description of the information flow in the brain structures. *Biol. Cybern.* 210, 203–210.

Kaminski, M.J., Blinowska, K.J., 1991. A new method of the description of the information flow in the brain structures. *Biol. Cybern.* 65, 203–210.

Khojaste-Sarakhsi, M., Haghighi, S.S., Ghomi, S.M.T.F., Marchiori, E., 2022. Deep learning for Alzheimer's disease diagnosis: A survey. *Artif. Intell. Med.* 130, 102332. <https://doi.org/10.1016/j.artmed.2022.102332>

Korzeniewska, A., M

2003. Determination of information flow direction among brain structures by a modified directed transfer function (dDTF) method. *J. Neurosci. Methods* 125, 195–207.

Lehmann, C., Koenig, T., Jelic, V., Prichep, L., John, R.E., Wahlund, L.O., Dodge, Y., Dierks, T., 2007. Application and comparison of classification

algorithms for recognition of Alzheimer's disease in electrical brain activity (EEG). *J. Neurosci. Methods* 161, 342–350. <https://doi.org/10.1016/j.jneumeth.2006.10.023>

Li, G., Jung, J.J., 2021. Seizure detection from multi-channel EEG using entropy-based dynamic graph embedding. *Artif. Intell. Med.* 122, 102201. <https://doi.org/10.1016/j.artmed.2021.102201>

Maggioni, E., Arienti, F., Minella, S., Mameli, F., Borellini, L., Nigro, M., Cogiamanian, F., Bianchi, A.M., Cerutti, S., Barbieri, S., Brambilla, P., Ardolino, G., 2021. Effective Connectivity During Rest and Music Listening: An EEG Study on Parkinson's Disease. *Front. Aging Neurosci.* 13, 1–16. <https://doi.org/10.3389/fnagi.2021.657221>

Mano, T., Kinugawa, K., Ozaki, M., Kataoka, H., Sugie, K., 2022. Neural synchronization analysis of electroencephalography coherence in patients with Parkinson's disease-related mild cognitive impairment. *Clin. Park. Relat. Disord.* 6, 100140. <https://doi.org/10.1016/j.prdoa.2022.100140>

Moon, S.E., Chen, C.J., Hsieh, C.J., Wang, J.L., Lee, J.S., 2020. Emotional EEG classification using connectivity features and convolutional neural networks. *Neural Networks* 132, 96–107. <https://doi.org/10.1016/j.neunet.2020.08.009>

Mumtaz, W., Saad, M.N. b. M., Kamel, N., Ali, S.S.A., Malik, A.S., 2018. An EEG-based functional connectivity measure for automatic detection of alcohol use disorder. *Artif. Intell. Med.* 84, 79–89. <https://doi.org/10.1016/j.artmed.2017.11.002>

Muthuraman, M., Bange, M., Koirala, N., Ciolac, D., Pintea, B., Glaser, M., Tinkhauser, G., Brown, P., Deuschl, G., Groppa, S., 2020. Cross-frequency coupling between gamma oscillations and deep brain stimulation frequency in Parkinson's disease. *Brain* 143, 3393–3407. <https://doi.org/10.1093/brain/awaa297>



Ng, B., Reh, R.K., Mostafavi, S., 2022. A practical guide to applying machine learning to infant EEG data. *Dev. Cogn. Neurosci.* 54, 101096. <https://doi.org/10.1016/j.dcn.2022.101096>

Nimmy, J.T., Subha, D.P., Menon, R.N., 2019. Mutual information analysis on MCI-AD EEG signal during resting and task conditions. *IEEE Reg. 10 Annu. Int. Conf. Proceedings/TENCON* 2019-Octob, 2295–2299. <https://doi.org/10.1109/TENCON.2019.8929242>

Olejarczyk, E., Jernajczyk, W., 2017. Graph-based analysis of brain connectivity in schizophrenia. *PLoS One* 12, 1–28. <https://doi.org/10.1371/journal.pone.0188629>

Saeedi, A., Saeedi, M., Maghsoudi, A., Shalbaf, A., 2021. Major depressive disorder diagnosis based on effective connectivity in EEG signals: A convolutional neural network and long short-term memory approach. *Cogn. Neurodyn.* 15, 239–252.

Santerne, A., Moutou, C., Tsantaki, M., Bouchy, F., Hébrard, G., Adibekyan, V., Almenara, J.-M., Amard, L., Barros, S.C.C., Boisse, I., Bonomo, A.S., Bruno, G., Courcol, B., Deleuil, M., Demangeon, O., Díaz, R.F., Guillot, T., Havel, M., Montagnier, G., Rajpurohit, A.S., Rey, J., Santos, N.C., 2016. SOPHIE velocimetry of Kepler transit candidates. *Astron. Astrophys.* 587, A64. <https://doi.org/10.1051/0004-6361/201527329>

Schelter, B., Winterhalder, M., Eichler, M., Peifer, M., Hellwig, B., Guschlbauer, B., Lücking, C.H., Dahlhaus, R., Timmer, J., 2006. Testing for directed influences among neural signals using partial directed coherence. *J. Neurosci. Methods* 152, 210–219. <https://doi.org/10.1016/j.jneumeth.2005.09.001>

Seth, A.K., Barrett, A.B., Barnett, L., 2015. Granger causality analysis in neuroscience and neuroimaging. *J. Neurosci.* 35, 3293–3297. <https://doi.org/10.1523/JNEUROSCI.4399-14.2015>

Shah, S.A.A., Zhang, L., Bais, A., 2020. Dynamical system based compact deep hybrid network for classification of Parkinson disease related EEG signals. *Neural Networks* 130, 75–84. <https://doi.org/10.1016/j.neunet.2020.06.018>

Shan, X., Cao, J., Huo, S., Chen, L., Sarrigiannis, P.G., Zhao, Y., 2022. <scp>Spatial–temporal</scp> graph convolutional network for Alzheimer classification based on brain functional connectivity imaging of electroencephalogram. *Hum. Brain Mapp.* 43, 5194–5209. <https://doi.org/10.1002/hbm.25994>

Song, X., Mao, M., Qian, X., 2021. Auto-Metric Graph Neural Network Based on a Meta-Learning Strategy for the Diagnosis of Alzheimer's Disease. *IEEE J. Biomed. Heal. Informatics* 25, 3141–3152. <https://doi.org/10.1109/JBHI.2021.3053568>

Supp, G.G., Schlögl, A., Trujillo-Barreto, N., Müller, M.M., Gruber, T., 2007. Directed cortical information flow during human object recognition: Analyzing induced EEG gamma-band responses in brain's source space. *PLoS One* 2. <https://doi.org/10.1371/journal.pone.0000684>

Tafreshi, T.F., Daliri, M.R., Ghodousi, M., 2019. Functional and effective connectivity based features of EEG signals for object recognition. *Cogn. Neurodyn.* 13, 555–566. <https://doi.org/10.1007/s11571-019-09556-7>

Tong, Z., Liang, Y., Sun, C., Li, X., Rosenblum, D.S., Lim, A., 2020. Digraph inception convolutional networks. *Adv. Neural Inf. Process. Syst.* 2020-Decem, 1–12.

van Mierlo, P., Papadopoulou, M., Carrette, E., Boon, P., Vandenberghe, S., Vonck, K., Marinazzo, D., 2014. Functional brain connectivity from EEG in epilepsy: Seizure prediction and epileptogenic focus localization. *Prog. Neurobiol.* 121, 19–35. <https://doi.org/10.1016/j.pneurobio.2014.06.004>

Varsehi, H., Firoozabadi, S.M.P., 2021. An EEG channel selection method for motor imagery based brain–computer interface and neurofeedback using

- Granger causality. *Neural Networks* 133, 193–206. <https://doi.org/https://doi.org/10.1016/j.neunet.2020.11.002>
- Wang, F., Tian, Y.C., Zhang, X., Hu, F., 2020. Detecting Disorders of Consciousness in Brain Injuries From EEG Connectivity Through Machine Learning. *IEEE Trans. Emerg. Top. Comput. Intell.* 1–11. <https://doi.org/10.1109/TETCI.2020.3032662>
- Wang, H.E., Bénar, C.G., Quilichini, P.P., Friston, K.J., Jirsa, V.K., Bernard, C., 2014. A systematic framework for functional connectivity measures. *Front. Neurosci.* 8, 1–22. <https://doi.org/10.3389/fnins.2014.00405>
- Wang, Yuan, Wang, Yao, Lui, Y.W., 2018. Generalized Recurrent Neural Network accommodating Dynamic Causal Modeling for functional MRI analysis. *Neuroimage* 178, 385–402. <https://doi.org/10.1016/j.neuroimage.2018.05.042>
- Wen, D., Li, R., Jiang, M., Li, J., Liu, Y., Dong, X., Saripan, M.I., Song, H., Han, W., Zhou, Y., 2022. Multi-dimensional conditional mutual information with application on the EEG signal analysis for spatial cognitive ability evaluation. *Neural Networks* 148, 23–36. <https://doi.org/10.1016/j.neunet.2021.12.010>
- Xie, W., Toll, R.T., Nelson, C.A., 2022. EEG functional connectivity analysis in the source space. *Dev. Cogn. Neurosci.* 56, 101119. <https://doi.org/10.1016/j.dcn.2022.101119>
- Yuvaraj, R., Murugappan, M., Acharya, U.R., Adeli, H., Ibrahim, N.M., Mesquita, E., 2016. Brain functional connectivity patterns for emotional state classification in Parkinson's disease patients without dementia. *Behav. Brain Res.* 298, 248–260. <https://doi.org/10.1016/j.bbr.2015.10.036>
- Zarghami, T.S., Friston, K.J., 2020. Dynamic effective connectivity. *Neuroimage* 207, 116453. <https://doi.org/10.1016/j.neuroimage.2019.116453>
- Zhang, W., Han, X., Qiu, S., Li, T., Chu, C., Wang, L., Wang, J., Zhang, Z., Wang, R., Yang, M., Shen, X., Li, Zhen, Bai, L., Li, Zhuo, Zhang, R., Wang, Y., Liu, C., Zhu, X., 2022. Analysis of Brain Functional Network Based on EEG

Signals for Early-Stage Parkinson's Disease Detection. *IEEE Access* 10, 21347–21358. <https://doi.org/10.1109/ACCESS.2022.3150561>

Zhang, Z., Bu, J., Ester, M., Zhang, J., Yao, C., Yu, Z., Wang, C., 2019. Hierarchical Graph Pooling with Structure Learning.

Zheng, R., Feng, Y., Wang, T., Cao, J., Wu, D., Jiang, T., Gao, F., 2022. Scalp EEG functional connection and brain network in infants with West syndrome. *Neural Networks* 153, 76–86. <https://doi.org/10.1016/j.neunet.2022.05.029>

Zhong, P., Wang, D., Miao, C., 2020. EEG-Based Emotion Recognition Using Regularized Graph Neural Networks. *IEEE Trans. Affect. Comput.* 1–1. <https://doi.org/10.1109/taffc.2020.2994159>

## **8 Conclusion**

### **8.1 Contribution to Knowledge**

Many promising approaches related to brain functional and effective connectivity are incessant springing up in recent years and contribute to neurological research. However, there is a lack of comprehensive review in the last five years in terms of estimation, applications, and particularly their associated machine learning approaches. Moreover, numerous researchers developed various visualisation methods to represent functional and effective connectivity, but there is a lack of review on them. To fill those research gaps, I reviewed EEG-based functional and effective connectivity studies undertaken over the last few years, in terms of estimation, visualization, and applications associated with machine learning classifiers. Methods are explored and discussed from various dimensions, such as either linear or nonlinear, parametric or nonparametric, time-based, and frequency- based or time-frequency-based. Then it is followed by a novel review of brain connectivity visualization methods, grouped by Heat Map, data statistics, and Head Map, aiming to explore the variation of connectivity across different brain regions. Finally, the current challenges of related research and a roadmap for future related research are presented. I published this review paper in *Human Brain*

Mapping, titled “Brain functional and effective connectivity based on electroencephalography recordings: A review”.

I propose a novel brain functional connectivity method, Revised Hilbert-Huang Transformation (RHHT), which can better track dynamic and nonlinear brain network information. The effectiveness of the RHHT is evidenced using simulation data and practical EEG recordings. The results of two simulation examples demonstrate that, within a certain statistical confidence level, the proposed novel RHHT-based brain connectivity method outperforms the wavelet-based method in terms of classification accuracy and time-frequency resolution. The result suggests that the proposed method has the potential to better differentiate epileptic patients and healthy controls, and to be translated into a long-overdue biomarker and a diagnostic method for Alzheimer’s disease (AD), benefiting from the enhanced measure of dynamic time-frequency association. The results produce two journal papers. Those findings imply that the proposed RHHT approach could better track dynamic and nonlinear functional connectivity information, paving the way for the development of a novel diagnostic approach. A journal paper has been published by IEEE Transactions on Neural Systems and Rehabilitation Engineering, titled ‘A Revised Hilbert-Huang Transformation to Track Non-stationary Association of Electroencephalography Signals’. Besides, a journal paper has been published in Journal of Neural Engineering, titled ‘Ultra-high-resolution time-frequency analysis of EEG to characterise brain functional connectivity with the application in Alzheimer’s disease.

I proposed a novel framework to diagnose epilepsy based on interictal seizure-free EEG data only using a set of estimates of linear and nonlinear brain functional connectivity. Unlike most of the existing works focusing on either seizure data or single-variate method, I introduced a multi-variate method to characterise sensor level brain functional connectivity from interictal EEG data to identify patients with generalised epilepsy. Despite no interictal abnormalities like IEDs - the empirical diagnostic hallmark of epilepsy - were included in this work, my framework achieved a very high classification accuracy between HC

and EG cohorts, which outperforms the state-of-the-art single channel findings. The findings demonstrate that seizure-free EEG recordings contain invisible information that can be revealed with appropriate methodology to identify patients with generalised epilepsy. Besides, EEG epochs during periods of eyes-open and eyes-closed differ significantly, reflecting the dynamic brain network changes between the two conditions. The results produced a journal paper, which was published by Biomedical Signal Processing and Control, titled “using interictal seizure-free EEG data to recognise patients with epilepsy based on machine learning of brain functional connectivity”.

There is high demand for techniques to estimate human mental workload during some activities for productivity enhancement or accident prevention. Most studies focus on a single physiological sensing modality and use univariate methods to analyse multi-channel EEG data. This paper proposes a new framework that relies on the features of hybrid EEG–fNIRS, supported by machine-learning features to deal with multi-level mental workload classification. Furthermore, instead of the well-used univariate PSD for EEG recording, I propose using bivariate functional brain connectivity (FBC) features in the time and frequency domains of three bands: delta (0.5–4 Hz), theta (4–7 Hz) and alpha (8–15 Hz). It was determined that the best region to assist the discrimination of the mental workload for EEG and fNIRS is different. Specifically, the posterior area performed the best for the EEG in the alpha band and fNIRS had superiority in the right frontal region. Based on those findings, a journal paper was submitted to IEEE Transactions on Cognitive and Developmental Systems, titled ‘Characterising EEG/fNIRS signals for mental workload classification’.

I developed novel GNN-based deep learning methods, including Spatial-Temporal GNN and DSL-GNN. Those two methods focus on dynamic and directional information individually. The performance is better than the traditional machine learning method and prove that the GNN can better use the brain network features. Those findings provide new biomarkers, tools, and frameworks for the diagnosis of dementia. On one hand, different from existing

studies that are based on either topological brain function characteristics or temporal features of EEG, the proposed ST-GCN considers both the adjacency matrix of functional connectivity from multiple EEG channels and corresponding dynamics of signal EEG channel simultaneously. Different from the traditional graph convolutional neural networks, the proposed ST-GCN makes full use of the constrained spatial topology of functional connectivity and the discriminative dynamic temporal information represented by the 1D convolution. The results demonstrate that the proposed method achieves better classification performance than the state-of-the-art methods. This approach can not only help diagnose AD but also better understand the effect of normal ageing on brain network characteristics before we can accurately diagnose the condition based on resting-state EEG. On the other hand, in comparison to the previous studies on GNN, the proposed DSL-GNN approach enhanced the functionality for processing directional information, which builds the basis for more efficiently performing GNN on EBC. Another contribution of this Chapter is the creation of a new framework for applying univariate and multivariate features simultaneously in a classification task. The proposed framework and DSL-GNN are validated in four discrimination tasks and the approach exhibited the best performance, against the existing methods. In a word, this research provides a robust analytical framework to deal with complex brain networks containing causal directional information and implies promising potential in dementia diagnosis. The results produce two journal papers. A journal paper was published by Human Brain Mapping, titled 'Spatial-Temporal Graph Convolutional Network for Alzheimer Classification based on Brain Functional Connectivity Imaging of EEG'. Besides, a journal paper is about to be submitted to Artificial Intelligence in Medicine, titled 'Dementia Classification Using a Graph Neural Network on Imaging of Effective Brain Connectivity'.

The effective connectivity between brain regions provides valuable insight into the underlying neural mechanisms of cognitive processes and disease states. However, the current research more relies on the parametric approaches, which are generally limited by model complexity, model misspecification and data requirements. I proposed causality detection attention-based convolutional

neural networks (CD-ABCNN) to estimate ECNs. The effectiveness of the proposed approach is first validated by the simulated data generated via multivariate autoregressive (MVAR) models. Then, the learned ECNs of AD patients and HC participants were compared to explore the differences between the two groups. The results suggest that AD patients exhibit disrupted connectivity in regions involved in cognitive control, memory and decision-making processes in human brain. The Chapter adds to the growing body of literature on the effective connectivity estimation and highlights the potential usefulness of ECNs in identifying AD-related changes in brain function, providing valuable insights into the potential of ECNs as a diagnostic and monitoring tool for neurodegenerative diseases.

## **8.2 Impact**

Firstly, the review analyses the properties of well-accepted BiFEC and visualisation methods, as well as suggests the appropriate scenario for each one. It suggests potential directions and possible strategies, which will help deepen the understanding of brain functional and effective connectivity from a variety of novel perspectives, as well as how to enhance their availability and reliability. The related research community will benefit from this review by better understanding BiFEC's potential and exploring the identified future opportunities.

Using brain connectivity methods in Alzheimer's disease, epilepsy, Parkinson's disease and workload classification studies obtained better performance compared with traditional methods that are based on a single channel. That is to say, functional and effective connectivity may track more changes from healthy people to patients to some extent. Furthermore, it provides the possibility for doctors and patients to diagnose the disease earlier and more accurately.

The findings indicate that specific regions of the brain can contribute to the diagnosis of epilepsy and dementia disease as well as workload classification based on brain connectivity. By advising the appropriate placement of EEG sensors based on these identified regions, doctors and researchers can more



efficiently and accurately diagnose and classify these neurological disorders. This can ultimately lead to earlier and more effective interventions, improving patient outcomes and reducing the burden on healthcare systems.

The findings suggest that EEG, which is a relatively low-cost imaging technique with high temporal resolution, could potentially be used in a flexible mobile wearable device to monitor the progression of dementia. By utilizing a wearable device, patients can be monitored outside of traditional clinical settings, providing more comprehensive and continuous monitoring of disease progression. This could lead to earlier detection of cognitive decline and allow for more personalized treatment plans, ultimately improving patient outcomes. Additionally, the use of a low-cost imaging technique such as EEG could make this monitoring more accessible and cost-effective for patients and healthcare systems.

### **8.3 Limitation and Future Work**

EEG studies are valuable for understanding brain function, but they have several limitations that must be taken into consideration when interpreting the results. EEG findings can be challenging to interpret, as they are complex and often vary between individuals (Abdullah ., 2010; Hu ., 2011; Jeong ., 2019; Pandey and Seeja, 2022). This can make it challenging to draw clear and universal conclusions about the underlying neural activity. A variety of factors can influence EEG signals, including age, gender, and medications (Asif ., 2019; Hu ., 2011; Yasin ., 2021). This variability can make it difficult to compare results across studies or draw general conclusions about brain function. Similarly, the strength and direction of connectivity between brain regions can vary across individuals and across different experimental conditions (Greene ., 2020; Van Overwalle ., 2019; Zhang ., 2020). This variability can make summarising general conclusions about brain connectivity patterns challenging. The brain is a highly complex system, with multiple levels of organisation and multiple networks that interact with each other. Studying the connectivity between brain regions can be challenging because the relationships between different networks are not fully understood. To overcome

the mentioned limitation, using standardised methods for EEG data acquisition and analysis can help to ensure consistency across studies and increase the reliability of results. This includes standardised electrode placement, data pre-processing techniques, and statistical analysis methods. A large dataset is needed to cover people of different ages, genders and disease conditions to develop and evaluate universal brain connectivity methods (Cao et al., 2022b).

Scalp EEG measures electrical activity on the scalp, which provides limited information about the spatial distribution of neural activity within the brain (Greene et al., 2020; Van Overwalle et al., 2019; Zhang et al., 2020). The scalp-recorded EEG signals resulted from the summed electrical activity of thousands or even millions of neurons, making it difficult to pinpoint the exact location of neural activity. Besides, to reduce volume conduction effects from a common reference, bipolar derivations were used to assess the degree of differences between various pairs of electrodes for two different cohorts of subjects (Cao et al., 2022a; Shan et al., 2022). With this approach—using bipolar pairs of electrodes—the effects of volume conduction are reduced but not eliminated. I recognise that the work is based on a sensor-level scalp EEG analysis, and I do not claim to be able to precisely localise the spatial characteristics underpinning the EEG sensor findings. In this case, one potential future direction for EEG research is to combine it with other imaging techniques, such as fMRI or magnetoencephalography (MEG), to provide better spatial resolution and more comprehensive insights into brain function (Cichy and Oliva, 2020; McIntosh et al., 2021). Combining multiple imaging techniques could help overcome the limitations of individual techniques and provide a complete understanding of brain function. Besides, source localisation is a technique for estimating the location of brain activity underlying EEG signals. Using source localisation methods can help to overcome the spatial resolution limitations of EEG and provide more accurate localisation of brain activity.

Brain connectivity studies could be extended to clinical populations, such as those with neurological or psychiatric disorders, to understand better the neural mechanisms underlying these conditions. These studies could also help identify

potential biomarkers for diagnosis or treatment monitoring. For some of my studies, only a two-group classification (HCs vs AD) has been carried out, without considering preclinical or prodromal populations or staging of disease severity of AD patients. The proposed method has demonstrated effective performance in the case of HCs vs AD, and it may have the potential to deal with the classification of HCs vs MCI vs AD. In the future, I am willing to collect data from MCI, and this three-class classification will be tested in further studies.

Longitudinal studies could investigate how brain connectivity changes over time and how these changes are related to behaviour, cognition, or disease (Filippi et al., 2021, 2020; McQuail et al., 2021). This could provide insights into the dynamic nature of brain networks and help identify critical periods for intervention. Longitudinal studies that track changes in brain connectivity over time could identify critical periods for intervention in neurological or psychiatric disorders. For example, if changes in connectivity occur early in the course of a disease, this could suggest that early intervention is critical to prevent further deterioration of brain function. Besides, studying connectivity changes over time in healthy individuals could improve our understanding of normal brain development and how it is influenced by factors such as genetics, environment, and experience. This could provide insights into optimising brain development and improving cognitive function in children and adolescents.

## REFERENCES

- Abdullah, M.K., Subari, K.S., Loong, J.L.C., Ahmad, N.N., 2010. *Intelligent Control of a Prosthetic Arm*, in: 2010 IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES). IEEE, pp. 303–306. <https://doi.org/10.1109/IECBES.2010.5742249>
- Asif, A., Majid, M., Anwar, S.M., 2019. *Design and Implementation of a Prosthetic Arm*. *Comput. Biol. Med.* 107, 182–196. <https://doi.org/10.1016/j.compbimed.2019.02.015>
- Cao, J., Zhao, Y., Shan, X., Blackburn, D., Wei, J., Erkoyuncu, J.A., Chen, L., Sarrigiannis, P.G., 2022a. *Design and Implementation of a Prosthetic Arm*. *J. Neural Eng.* 19, 046034. <https://doi.org/10.1088/1741-2552/ac84ac>
- Cao, J., Zhao, Y., Shan, X., Wei, H., Guo, Y., Chen, L., Erkoyuncu, J.A., Sarrigiannis, P.G., 2022b. *Design and Implementation of a Prosthetic Arm*. *Hum. Brain Mapp.* 43, 860–879. <https://doi.org/10.1002/hbm.25683>
- Cichy, R.M., Oliva, A., 2020. *Design and Implementation of a Prosthetic Arm*. *Neuron* 107, 772–781. <https://doi.org/10.1016/j.neuron.2020.07.001>
- Filippi, M., Basaia, S., Sarasso, E., Stojkovic, T., Stankovic, I., Fontana, A., Tomic, A., Piramide, N., Stefanova, E., Markovic, V., Kostic, V.S., Agosta, F., 2021. *Design and Implementation of a Prosthetic Arm*. *Mol. Psychiatry* 26, 5429–5440. <https://doi.org/10.1038/s41380-020-0770-0>
- Filippi, M., Canu, E., Donzuso, G., Stojkovic, T., Basaia, S., Stankovic, I., Tomic, A., Markovic, V., Petrovic, I., Stefanova, E., Kostic, V.S., Agosta, F., 2020. *Design and Implementation of a Prosthetic Arm*. *Mov. Disord.* 35, 1987–1998. <https://doi.org/10.1002/mds.28228>

Greene, D.J., Marek, S., Gordon, E.M., Siegel, J.S., Gratton, C., Laumann, T.O., Gilmore, A.W., Berg, J.J., Nguyen, A.L., Dierker, D., Van, A.N., Ortega, M., Newbold, D.J., Hampton, J.M., Nielsen, A.N., McDermott, K.B., Roland, J.L., Norris, S.A., Nelson, S.M., Snyder, A.Z., Schlaggar, B.L., Petersen, S.E., Dosenbach, N.U.F., 2020.

. *Neuron* 105, 742-758.e6.

<https://doi.org/10.1016/j.neuron.2019.11.012>

Hu, B., Liu, Q., Zhao, Q., Qi, Y., Peng, H., 2011.

, in: 2011 IEEE Asia-Pacific Services Computing Conference. IEEE, pp. 436–441.  
<https://doi.org/10.1109/APSCC.2011.87>

Jeong, D., Yoo, S., Yun, J., 2019.

, in: 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR). IEEE, pp. 827–835.  
<https://doi.org/10.1109/VR.2019.8798334>

McIntosh, J.R., Yao, J., Hong, L., Faller, J., Sajda, P., 2021.

. *IEEE Trans. Biomed. Eng.* 68, 78–89. <https://doi.org/10.1109/TBME.2020.3004548>

McQuail, J.A., Dunn, A.R., Stern, Y., Barnes, C.A., Kempermann, G., Rapp, P.R., Kaczorowski, C.C., Foster, T.C., 2021.

*Front. Aging Neurosci.* 12, 1–31. <https://doi.org/10.3389/fnagi.2020.607685>

Pandey, P., Seeja, K.R., 2022.

. *J. King Saud Univ. - Comput. Inf. Sci.* 34, 1730–1738. <https://doi.org/10.1016/j.jksuci.2019.11.003>

Shan, X., Cao, J., Huo, S., Chen, L., Sarrigiannis, P.G., Zhao, Y., 2022.

. Hum. Brain Mapp. 43, 5194–5209.  
<https://doi.org/10.1002/hbm.25994>

Van Overwalle, F., Van de Steen, F., Mariën, P., 2019.

. Cogn. Affect. Behav. Neurosci. 19, 211–  
223. <https://doi.org/10.3758/s13415-018-00659-y>

Yasin, S., Hussain, S.A., Aslan, S., Raza, I., Muzammel, M., Othmani, A., 2021.

. Comput. Methods Programs Biomed. 202.  
<https://doi.org/10.1016/j.cmpb.2021.106007>

Zhang, Y.D., Dong, Z., Wang, S.H., Yu, X., Yao, X., Zhou, Q., Hu, H., Li, M.,  
Jiménez-Mesa, C., Ramirez, J., Martinez, F.J., Gorriz, J.M., 2020.

. Inf. Fusion 64, 149–187. <https://doi.org/10.1016/j.inffus.2020.07.006>

