Imperial College London

IMPERIAL COLLEGE LONDON
DEPARTMENT OF BRAIN SCIENCES

Investigation of the falling asleep dynamics and its closed-loop augmentation strategy

Junheng Li

December 2023

Submitted in part fulfilment of the requirements for the degree of Doctor of Philosophy
in
Clinical Medicine Research

Abstract

Sleep onset insomnia is a common sleep disturbance, particularly in the pathological ageing population such as patients with neurodegenerative diseases and dementia. The bi-directional relationship between sleep disorders and neurodegeneration makes it a novel early-stage treatment venue. Pharmacological treatments for insomnia often cause side effects and fail to show cognitive benefits. Novel neuromodulatory approaches, especially phase-synchronised closed-loop auditory stimulation targeting neural oscillations, have been promising candidates for enhancing deep sleep, but are yet to be explored for augmenting sleep onset.

Standard sleep staging has poor temporal resolution and bad matches to subjective experience, especially for patients with sleep-onset insomnia. Therefore, a better understanding of the falling asleep dynamics outside the sleep staging is necessary for better outcome measurements and closed-loop control. This thesis aims to study the falling asleep brain dynamics measured via electroencephalogram (EEG) on large sleep datasets and test the feasibility of using closed-loop auditory stimulation to augment the process. I first developed a novel feature-based statistical learning approach to explore the EEG dynamics and identified several whole-brain EEG biomarkers for distinguishing the wakefulness and sleep onset states. Combined with the dynamical systems theory, I then built a functional model to describe the continuous falling asleep transition dynamics as a catastrophic bifurcation process. Additionally, I explored the regional brain dynamics and their interactions during the falling asleep process, which provided further insights into sleep initiation.

Lastly, I present a nap study to test the feasibility of phase-synchronised auditory stimulation in augmenting falling asleep, using the EEG biomarkers identified as outcome measurements. I also tested the effect of auditory-intervened naps on memory functions. The results showed a weak positive effect on associative memory. Further refined analyses are necessary to reveal mechanistic insights and guide future experimental explorations for a personalised stimulation design.

Acknowledgements

Research

First of all, I would like to show my greatest gratitude to my mentor Dr. Nir Grossman. It was a great honour to be your student and become a scientist under your guidance. What I learned from you is far beyond scientific knowledge, but the high-level thinking about science, and the dedication towards the truth. It always surprises me how much effort you put into your work, despite being so brilliant, which was the most inspiring stimulant to me. Doing a PhD was a great adventure both physically and mentally, I am grateful for your firm support in my academics as well as personal life.

I extend my gratitude towards my mathematics mentor Prof. Mauricio Barahona. Your brilliant mind always has inspiring novel ideas that push my work to a higher level. Your rectifications in the mathematical details inside the results are vital for someone like me with no formal mathematics training.

Thank you, Dr. Robert Peach, for your constant help on mathematics and computation throughout my PhD (and during my Master's). When I was stuck, you always came to my rescue with the most direct solution. All the work in this thesis would not be possible without your great ideas and extensive support.

Thank you also to Dr. Edward Rhodes, for mentoring and supporting me in conducting experiments. In addition, your 'critical' feedback on my work is important for making it better. I also acknowledge your help in refining data analysis and thesis, as well as the 'few' drinks that you bought me throughout my PhD.

I further thank Prof. Derk-Jan Dijk for your valuable insights concerning sleep neuroscience. Thanks to Dr. Ines Violante, Dr. Danielle Kurtin, Dr. Matteo Vinao-Carl and Dr. Emma-Jane Mallas for your suggestions and support of my work.

Thanks to Tianyu Wei, who helped conduct a lot of explorations of the results presented in Chapter 5. Thanks to Jiulian Gutierrez, Rutvi Patel, Gayathri Mathusuthan and Priti Amin who helped with the nap experiment presented in Chapter 6.

I would like to acknowledge Imperial College London, China Scholarship Council (CSC), and Dr. Nir Grossman for their joint sponsorship of my PhD.

Lastly, I would like to thank all Interventional Systems Neuroscience lab members, who created a wonderful research environment together to work in. Working with you all was a great pleasure and an inspiring experience.

Dataset

Sincere gratitude to Professor Nadia Goesslin and Professor Julie Carrier from the University of Montreal for providing the sleep dataset with cognitive impairment patients.

The Multi-Ethnic Study of Atherosclerosis (MESA) Sleep Ancillary study was funded by NIH-NHLBI Association of Sleep Disorders with Cardiovascular Health Across Ethnic Groups (RO1 HL098433). MESA is supported by NHLBI funded contracts HHSN268201500003I, N01-HC-95159, N01-HC-95160, N01-HC-95161, N01-HC-95162, N01-HC-95163, N01-HC-95164, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168 and N01-HC-95169 from the National Heart, Lung, and Blood Institute, and by cooperative agreements UL1-TR-000040, UL1-TR-001079, and UL1-TR-001420 funded by NCATS. The National Sleep Research Resource was supported by the National Heart, Lung, and Blood Institute (R24 HL114473, 75N92019R002).

Friends

The first special thanks go to one of my best friends, colleague, and 'comrade' during the PhD journey, Dr. Xiaoqi Zhu. Looking back on the past four years, it now seems impossible to have gotten through without your company and consistent support, specifically during the COVID lockdown periods. Without family nearby, you were my only emotional support.

Thanks to my many other friends, Chaoran, Jiabin, Xiaowen, Xingjian, and Waleed, who helped and supported me in various ways. Thanks also to those 'remote' friends who provided important mental values to me, even if we have only met in 'virtual' software. It was wonderful to have you lifting me from time to time, and bringing positivity in my life moving forward.

Family

Last but not least, I thank wholeheartedly my parents, for your consistent and unwavering love, understanding, and support both mentally and financially, to make what I can have today possible. Before the start of this thesis, we had been separated by COVID for 3 years and 9 months. It was an unimaginably difficult time, with the coexistence of heavy pressures of PhD and overwhelming nostalgic affections. It was the hope of reuniting with you that made me persist through. I would like you to know that I love you with all my life.

And to Lulu, my love. It was your existence that brightened my entire life and brought me infinite strength when facing challenges. It was also you who provided me with a feeling of family during this monotonous PhD life. Thank you and love you. 感谢你们的支持。

Declaration of Originality

The Multi-Ethnic Study of Atherosclerosis (MESA) dataset was obtained with permission from the National Sleep Research Resources (NSRR) (https://sleepdata.org/datasets/mesa). The Montreal dataset was obtained from collaborators Prof. Nadia Goesslin and Prof. Julie Carrier from the University of Montreal.

In Chapter 5, the inter-regional dynamics of the falling asleep process are explored and conducted along with the undergraduate project student Tianyu Wei, under my supervision and with my extensive input. All results shown belong to my own original ideas.

In Chapter 6, the experiment was co-designed by me and Dr. Edward Rhodes. The experiment was conducted by me and Dr. Edward Rhodes, along with four Master's project students under our supervision: Jiulian Gutierrez, Rutvi Patel, Gayathri Mathusuthan and Priti Amin. The data analyses for all the results shown were done by myself, with input from Dr. Edward Rhodes for initial data preparation and pre-processing.

I declare herewith, that all other work shown within this thesis is my own unless referenced otherwise, and conforms to the rules and guidelines set out for PhD theses by Imperial College London. Furthermore, I confirm that all data and findings have not been falsified or embellished and that the work has not been used for alternative exam processes.

Copyright

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial 4.0 International Licence (CC BY-NC).

Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose.

When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes.

Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Contents

Acronyms **14**

List of Tables **17**

List of Figures **21**

1 Introduction **27**

1.1 Sleep, insomnia, and its interventions 27

1.1.1 Sleep and sleep disturbances in neurodegenerative diseases 27

1.1.2 Insomnia, and its current treatments 29

1.1.3 The need for novel neuromodulation paradigm in augmenting sleep-onset 30

1.2 Electroencephalogram for the quantification of sleep 32

1.2.1 Electroencephalogram in capturing brain electrophysiology 32

1.2.2 Sleep EEG and standard sleep stagings 34

1.3 The quantification of falling asleep dynamics 36

1.3.1 The falling asleep period 36

1.3.2 Behavioural and physiological changes during falling asleep 37

1.3.3 The EEG brain dynamics during falling asleep 39

1.4 The neurophysiology of wake-sleep regulations 43

1.4.1 The neuronal circuits underlying the wakefulness and sleep 43

1.4.2 Regulations of wakefulness and sleep 45

1.5 The structure of this thesis 47

1.6 Publications statement 48

2 Material and Methods **49**

2.1 Sleep EEG datasets . . . 49

2.1.1 Dataset 1: Multi-Ethnic Study of Atherosclerosis (MESA) - Elderly cohort 49

2.1.2 Dataset 2: the Montreal's sleep EEG dataset with mild cognitive impairment 51

2.2 EEG pre-processing . . . 54

2.2.1 Definition of sleep-onset . . . 54

2.2.2 Preprocessing and extracting sleep-onset EEG . . . 54

2.3 Comprehensive EEG time series feature base for quantitative analysis of the falling asleep process . . . 57

2.3.1 Feature-based EEG analysise . . . 57

2.3.2 Feature-base computation . . . 59

2.3.3 Limitations and extensions . . . 61

2.4 Feature-based statistical learnings . . . 62

2.5 Dynamical systems theory and the brain . . . 63

2.5.1 The basic dynamical system theory . . . 63

2.5.2 The brain as a dynamical system . . . 65

2.6 Closed-loop auditory stimulation paradigms . . . 66

3 The distinctive brain EEG activities at sleep onset **68**

3.1 Introduction . . . 68

3.2 Methods . . . 69

3.2.1 Datasets . . . 69

3.2.2 Feature extractions for bedtime and falling asleep period . . . 70

3.2.3 Classification using linear SVM and leave-one-out cross-validation . . . 70

3.2.4 Hierarchical clustering and feature selection . . . 72

3.2.5 Statistical analysis . . . 73

3.2.6 Regression models . . . 73

3.3 Results . . . 75

3.3.1 EEG features governing sleep onset compared to bedtime . . . 75

3.3.2 The effect of ageing . . . 77

3.3.3 The effect of gender . . . 79

3.3.4 Validation using a second sleep EEG dataset (Montreal dataset) . . . 79

3.3.5 The effects of cognitive impairment on the features at bedtime and sleep onset . . . 83

3.3.6 Relationship to sleep onset latency . . . 84
3.4 Discussion . . . 91
3.4.1 A set of orthogonal EEG features distinguishing sleep onset from wakefulness 91
3.4.2 Sleep onset latency is correlated with and can be predicted by a small subset of EEG features at bedtime . . . 93
3.4.3 Ageing is associated with shorter sleep onset latency and distinct brain dynamics at sleep onset . . . 94
3.4.4 Other general limitations . . . 95

4 Bifurcation brain EEG dynamics during falling asleep 97
4.1 Introduction . . . 97
4.2 Methods . . . 99
4.2.1 Feature dynamic time series during falling asleep . . . 99
4.2.2 Conceptualisation of a dynamical system in feature state space for analysis of bifurcation . . . 101
4.2.3 Theoretical dynamical system model for bifurcation analysis . . . 103
4.2.4 Dimensionality reduction via functional principal component analysis (FPCA) 106
4.2.5 Granger causality analysis in time-series dynamics . . . 107
4.2.6 Critical slowing down and its early warning signals (EWS) prior to bifurcation . . . 108
4.3 Results . . . 111
4.3.1 The brain EEG of falling asleep has a catastrophic bifurcation dynamic . 111
4.3.2 The key dynamical EEG features during falling asleep . . . 114
4.3.3 The key dynamical features govern the falling asleep bifurcation . . . 117
4.3.4 Causal interactions between the key dynamical features . . . 119
4.3.5 Critical slowing down during falling asleep . . . 119
4.4 Discussion . . . 121
4.4.1 The falling asleep transition can be described as a catastrophic bifurcation 122
4.4.2 Using early warning signal (EWS) to predict sleep onset . . . 124
4.4.3 General limitations and future directions . . . 126
4.4.4 Implications . . . 127

5 The regional and inter-regional brain dynamics during falling asleep 128

5.1 Introduction . . . 128

5.2 Methods . . . 129

5.2.1 Analysis of regional differences . . . 129

5.2.2 Development of the feature base for quantifying inter-regional interactions and its implementations . . . 130

5.2.3 The analysis of the inter-regional interaction dynamics . . . 137

5.2.4 Other statistical methods . . . 139

5.3 Results . . . 139

5.3.1 The regional specific EEG dynamic features at sleep onset compared to bedtime . . . 139

5.3.2 The region-specific bifurcation dynamics during falling asleep . . . 143

5.3.3 The inter-regional interactions dynamics at sleep onset compared to bedtime 145

5.3.4 The continuous inter-regional interaction dynamics during falling asleep transition . . . 150

5.3.5 A Granger-causality based framework between regional dynamics and inter-regional interactions during falling asleep . . . 153

5.4 Discussion . . . 154

5.4.1 The regional EEG brain dynamics during falling asleep change quantitatively but not qualitatively . . . 154

5.4.2 Sleep onset is hallmarked by a global increase of information transfer between regions . . . 155

5.4.3 A Granger-causality based framework eluding to a self-enhancing loop between regional and inter-regional dynamics during sleep initiation . . . 157

5.4.4 General limitations and future directions . . . 159

6 Testing sleep onset augmentation via phase-locked auditory stimulation 160

6.1 Introduction . . . 160

6.2 Methods . . . 161

6.2.1 Participants recruitment . . . 161

6.2.2 Experimental design . . . 162

6.2.3 Data analysis . . . 168

6.3 Results . . . 176

6.3.1 Demographics . . . 176

6.3.2 Phase-locking accuracy . . . 177

6.3.3 The effect of auditory stimulation on nap . . . 178

6.3.4 Effect of auditory-intervened nap on memory performances . . . 183

6.4 Discussion . . . 185
6.4.1 Reflections on experimental results . . . 185
6.4.2 Potential personalised stimulation design in augmenting sleep onset . . . 188
6.4.3 Implications and other potential work . . . 189

7 Conclusions and discussion **191**
7.1 Summary of the results . . . 191
7.2 Novel contributions and implications . . . 193
7.3 Outlook . . . 194

Bibliography **196**

Appendices **229**

A List of EEG time-series feature base used in the analysis **230**
A.1 List of conventional EEG analysis features . . . 230
A.2 List of CATCH-22 highly comparative time-series features (Reused from Lubba et al. (2019)) . . . 232
A.3 List of inter-regional interaction features between EEG channels . . . 233

B Supplementary Materials **236**
B.1 Sleep onset latency distributions before exclusion for MESA and Montreal datasets 236
B.2 Examplary EEG cardiac signature trace . . . 237
B.3 Examplary EEG power spectrum before and after differentiation filtering . . . 237
B.4 List of features in the hierarchical order in Figure 3.2 and the explanations of the key representative features selected . . . 238
B.5 GLM statistics of all features at sleep onset to predict age, and explanations of the significant features in the model . . . 240
B.6 GLM statistics of all features at bedtime to predict age . . . 243
B.7 Key feature statistics at bedtime and sleep onset for male and female participants in Figure 3.6 . . . 245
B.8 List of features in the hierarchical order in Figure 3.7 . . . 245
B.9 List of features in the hierarchical order in Figure 3.9a and b . . . 247
B.10 Statistics of the feature distributions in Normal and MCI participants at bedtime and sleep onset (for Figure 3.10 and 3.11) . . . 250
B.11 List of features in the hierarchical order in Figure 3.15 . . . 251
B.12 The GLM statistics of all features to predict sleep onset latency at bedtime . . . 253

B.13 Confounding test for sleep state space variable and feature state space velocity (aligned to half sleep onset for each participant) . . . 255

B.14 Sleep state space variable and feature state space velocity comparison between fast (< 40 minutes) and slow (> 60 minutes) falling asleep groups . . . 256

B.15 List of features in the hierarchical order in Figure 4.7 . . . 256

B.16 Confounding test for feature dynamic time-series in Figure 4.7 . . . 259

B.17 Similarity score between the first two FPCs and all 47 features, and explanation of the selected representative features . . . 259

B.18 Similarity between the feature time series to the control variable time series $c(t)$ 262

B.19 Table of p-values of the pair-wised Granger-causality test between the four key dynamical feature time series . . . 264

B.20 Group averaged early warning signal for a non falling asleep transition period as controlled testing . . . 265

B.21 The EWS results on exemplary participants and statistical methods for individual-level analysis . . . 265

B.22 An exploration of the sleep state variable between healthy and MCI cohorts in the Montreal dataset . . . 266

B.23 Classification accuracy between bedtime and sleep onset for each region separately and compared to global accuracy . . . 267

B.24 Hierchical clustering results for each EEG region separately in the MESA dataset 269

B.25 The features added and statistics for the classification in Figure 5.5 . . . 274

B.26 The theoretical model fitting on the s time series dynamics of each region separately 275

B.27 The continuous feature dynamics for the three different regions in the MESA dataset . . . 276

B.28 The distribution statistics of the two key feature types at bedtime and sleep onset between all region pairs (for Figure 5.9) . . . 278

B.29 Hierarchical clustering of all the inter-regional interaction features without combining into categories in the MESA dataset . . . 280

B.30 The distribution statistics of all the TE features at bedtime and sleep onset between all region pairs in the Montreal dataset (for Figure 5.11) . . . 281

B.31 Results for the re-referenced Montreal dataset and comparison to the previous results . . . 282

B.32 The continuous inter-regional interactional dynamics for all features during falling asleep . . . 284

B.33 The continuous inter-regional interactional feature dynamics plot for some specific feature categories . . . 284

B.34 A zoom-in of specific regional band power dynamics for illustration . . . 289

B.35 The p-value tables for the Granger causality test between all the regional and inter-regional FPC dynamics . . . 290

B.36 Phase histogram at stimulation onset for all 32 channels 291

B.37 Phase histogram at stimulation onset for Fz channel in the single-channel phase-locking EEG system 293

B.38 Feature-based classification between audio and control conditions in the last 3 minutes of nap 294

B.39 The model statistics for FNAT performance comparison between audio and control groups in Figure 6.11 294

C Permissions to reuse figures **297**

Acronyms

AASM	American Academy of Sleep Medicine.
AD	Alzheimer's Disease.
ANOVA	Analysis of Variance.
ARAS	Ascending Reticular Activating System.
CATCH-22	CAnonical Time series CHaracteristics.
CLMM	Cumulative Link Mixed Model.
CRT	Choice Response Task.
DMN	Default Mode Network.
DTF	Directed Transfer Function.
DTW	Dynamical Time Warping.
ECG	Electrocardiogram.
ecHT	End-point Corrected Hilbert Transform.
EDF	European Data Format.
EEG	Electroencephalogram.
EMG	Electromyography.
EOG	Electrooculography.
ERP	Event Related Potential.

ESS Epworth Sleepiness Score.

EWS Early Warning Signals.

fMRI functional Magnetic Resonance Imaging.

FNAT Face-Name Association Task.

FPCA Functional Principal Component Analysis.

GC Granger Causality.

GLM Generalised Linear Model.

GLME Generalised Linear Mixed-Effect Model.

HCTSA Highly Comparative Time Series Analysis.

KSS Karolinska Sleepiness Scale.

MCI Mild Cognitive Impairment.

MESA Multi-Ethnic Study of Atherosclerosis.

MI Mutual Information.

NREM Non-Rapid Eye Movement.

NSRR National Sleep Research Resources.

OSA Obstructive Sleep Apnea.

PAC Phase Amplitude Coupling.

PD Parkinson's Disease.

PDC Partial Directed Coherence.

PSD Power Spectral Density.

PSG Polysmonography.

PSQI Pittsburgh Sleep Quality Index.

REM Rapid Eye Movement.

RMS Root Mean Square.

SEM Slow Eye Movement.

SVM Support Vector Machine.

SWS Slow-Wave Sleep.

TE Transfer Entropy.

VAR Vector Auto-Regression.

List of Tables

3.1 List of significant features at sleep onset to predict age in GLM 78

3.2 List of significant features at bedtime to predict sleep onset latency in GLM . . . 89

5.1 List of all the key features selected from each region separately in the MESA dataset. 142

A.1 List of conventional EEG analysis features 230

A.2 List of Catch-22 time-series features 232

A.3 List of inter-regional interaction features between EEG channels 233

B.1 List of features in the hierarchical order in Figure 3.2 238

B.2 List of features in the hierarchical order in Figure 3.2 239

B.3 GLM statistics of all features at sleep onset to predict age 240

B.4 Explanantions of the significant features in the GLM model to predict age at sleep onset 242

B.5 GLM statistics of all features at bedtime to predict age 243

B.6 Key feature statistics at bedtime and sleep onset for male and female participants in Figure 3.6 245

B.7 List of features in the hierarchical order in Figure 3.7 246

B.8 List of features in the hierarchical order in Figure 3.9a 247

B.9 List of features in the hierarchical order in Figure 3.9b 249

B.10 Statistics of the feature distributions in Normal and MCI participants at bedtime and sleep onset (for Figure 3.10 and 3.11) 251

B.11 List of features in the hierarchical order in Figure 3.15. 251

B.12 GLM statistics of all features at bedtime to predict sleep onset latency 253

B.13 Feature names list in the hierarchical order in Figure 4.7 256

B.14 Similarity score between the first two FPCs and all 47 features 259

B.15 Similarity between the feature time series to the control variable time series $c(t)$ 262

B.16 Table of p-values of the pair-wised Granger-causality test between the four key dynamical feature time series 264

B.17 Single-region classification accuracy between bedtime and sleep-onset in the MESA dataset 267

B.18 List of features in the hierarchical order in Figure 5.4. 270

B.19 The features added into the classification model in Figure 5.5 at each step, along with the Accuracy and F-score at that step. 274

B.20 The FPC similarity scores between all the features and the FPCs identified from the feature dynamics in each region, respectively. 276

B.21 The distribution statistics of the two key feature types at bedtime and sleep onset between all region pairs (for Figure 5.9). 278

B.22 The distribution statistics of all the TE features at bedtime and sleep onset between all region pairs in the Montreal dataset (for Figure 5.11). 281

B.23 Table of classification accuracies for each inter-regional feature type in MESA and Montreal datasets (Raw and Re-referenced). 282

B.24 Table of the inter-regional FPC similarity scores for all the TE (6), DTF (30), and power band ratio (30) features (between all region pairs and directions). . . 286

B.25 The p-value tables for the Granger causality test between all pairs of the regional and inter-regional FPC dynamics. 290

B.26 The phase distribution statistics at stimulation onset for all 32 EEG channels. . . 291

B.27 The post hoc contrasts in the percentage of response in each response type between audio and control groups. 295

B.28 The GLME model statistics for the response time (RT). 295

B.29 The post hoc contrasts of the response time (RT) in each response type between audio and control groups. 295

B.30 The CLMM model statistics for prediction of the confidence. 296

B.31 The post hoc contrasts of the confidence in each response type between audio and control groups. 296

B.32 The group-level statistics for the response type percentage, response time, and confidence. 296

C.1 List of reused figures and permissions. 297

List of Figures

1.1 Demonstration of standard sleep staging and corresponding EEG properties. . . 35

2.1 The EEG montage of the MESA dataset. 50

2.2 Age distribution grouped by gender in the MESA dataset. 50

2.3 Demographics detail for Montreal dataset. 52

2.4 Overview of the pre-processing pipeline. 55

2.5 The sleep onset latency distribution for MESA and Montreal. 55

2.6 The phase-locking auditory stimulation paradigm. 67

3.1 Illustration of SVM classification in a 2D space. 72

3.2 Hierarchical clustering of the features that distinguish bedtime and sleep-onset. . 75

3.3 Classification accuracy and F-scores using all features and key features only. . . . 76

3.4 Linear correlations between the four key features and ages. 77

3.5 Linear correlations between the significant features from the GLM model and age. 78

3.6 Differences of the four key features between genders. 80

3.7 Hierarchical clustering of the features that distinguish bedtime and sleep-onset in the Montreal dataset. 81

3.8 Classification accuracy and F-scores using all features and key features only in the Montreal dataset. 81

3.9 Hierarchical clustering of the features that distinguish bedtime and sleep-onset for Normal and MCI participants separately. 82

3.10 Differences in the six key features between Normal and MCI participants at bedtime. 84

3.11 Differences in the 6 key features between Normal and MCI participants at sleep onset. . . . 85

3.12 Effects of age, gender and cognitive impairment on sleep-onset latency. . . . 86

3.13 Correlation between key features and sleep-onset latency at bedtime and sleep onset. . . . 86

3.14 Division of the fast and slow participant cohorts based on sleep-onset latencies. . . 88

3.15 Hierarchical clustering of the features that distinguish fast and slow latencies cohort at bedtime. . . . 88

3.16 Correlation between the significant features in GLM and sleep-onset latency at bedtime. . . . 90

3.17 Comparison of the predicted sleep onset latency from the regression model and the actual latency. . . . 90

4.1 Transformation of EEG signals into the feature state space. . . . 100

4.2 Conceptualisation of falling asleep in feature state space as a dynamical system. . 102

4.3 Bifurcation diagram of the theoretical model. . . . 104

4.4 Concept of critical slowing down and its early warning signals. . . . 109

4.5 The dynamics of sleep state variable can be described as a bifurcation process. . 112

4.6 The dynamics of the feature state space variable. . . . 113

4.7 The feature dynamic time series during the falling asleep transition. . . . 114

4.8 FPCA analysis in identifying key feature dynamic patterns and representative features. . . . 115

4.9 Transition dynamics in the 2D feature state space between each pair of the four key dynamical features. . . . 117

4.10 The sleep state variable as a function of the four key dynamical features as control parameters. . . . 118

4.11 Granger causal interactions between the four key feature dynamical features. . . 119

4.12 The critical slowing down dynamics in the sleep state variable during the falling asleep transition. . . . 120

5.1 Visual illustration of the computation of mutual information and transfer entropy .136

5.2 Illustration of feature type-wise hierarchical clustering approach. 138

5.3 Distribution of the accuracy difference between the global and each EEG channel for all features. 140

5.4 Hierarchical clustering of key features that distinguish between bedtime and sleep-onset in Posterior region (Cz-Oz channel). 141

5.5 Classification accuracy between bedtime and sleep onset using a combination of the key features from the 3 different regions in the MESA dataset. 142

5.6 The sleep state variable dynamics in three regions separately, with the potential analysis. 144

5.7 The FPCA analysis for the feature dynamics of the three different regions in the MESA dataset. 145

5.8 Hierarchical clustering of the inter-regional interaction features (arranged in each category) that distinguish bedtime and sleep-onset in the MESA dataset. 146

5.9 The distribution of the features in the two key inter-regional feature types at bedtime and sleep-onset in the MESA dataset. 147

5.10 Hierarchical clustering of the inter-regional interaction features (arranged in each category) that distinguish bedtime and sleep-onset in the Montreal dataset. . . . 148

5.11 The transfer entropy distribution differences between bedtime and sleep-onset in the Montreal dataset with 3 regions (Frontal Fz, Central Cz and Parietal Pz). . . 149

5.12 The sleep state variable dynamics in the interactional feature state space. 150

5.13 Functional Principal component analysis (FPCA) for extracting key feature dynamical patterns for inter-regional features. 151

5.14 The Granger causality between all pairs of regional and inter-regional FPC dynamics. 153

6.1 The entire experimental set-up and procedure. 162

6.2 The design of the Delayed Free-Recall task. 164

6.3 The design of Face-Name Association Task (FNAT) interleaved with Choice Response Task. 166

6.4 The setup of the two EEG systems. 167

6.5 Demographics details of the participants contrasted between two conditions. . . . 176

6.6 Phase-locking accuracy across the brain in the 32-channel EEG system. 178

6.7 Differences between audio and control in sleep stages and KSS tests. 179

6.8 Comparison of the four key feature changes on the whole brain between audio and control across the nap period. 181

6.9 The key feature time series during nap normalised to baseline and compared between audio and control. 182

6.10 Results comparison in the Delayed Free-Recall task performance between audio and control group. 183

6.11 Results comparison in FNAT performance between audio and control groups. . . 184

B.1 The sleep onset latency distribution for MESA and Montreal. 236

B.2 An example EEG cardiac signature trace. 237

B.3 Examplary EEG power spectrum before and after differentiation filtering. 237

B.4 Confounding test for sleep state space variable and feature state space velocity aligned to half sleep onset. 255

B.5 Sleep state space variable and feature state space velocity comparison between fast ($<$ 40 minutes) and slow ($>$ 60 minutes) falling asleep groups. 256

B.6 Feature time series aligned to half of sleep onset point as a confounding test. . . 259

B.7 Early warning signal in the sleep state variable for a control period without falling asleep transition. 265

B.8 The EWS results for three exemplary participants. 266

B.9 The dynamics of sleep state variable compared between healthy and MCI cohorts in the Montreal dataset. 267

B.10 The hierarchical clustering result of the features for classifying bedtime and sleep onset state in the Frontal region. 272

B.11 The distribution comparison between bedtime and sleep onset of the key features selected from the Frontal region. 272

B.12 The hierarchical clustering result of the features for classifying bedtime and sleep onset state in the Right-Central region. 273

B.13 The distribution comparison between bedtime and sleep onset of the key features selected from the Right-Central region. 274

B.14 The theoretical model fitting on the s time series dynamics of each region separately. 275

B.15 The continuous feature dynamics for the 3 different regions in the MESA dataset. 276

B.16 The hierarchical clustering of the 60 top-performing interactional features in the MESA dataset. 280

B.17 The specific inter-regional features distribution differences between bedtime and sleep-onset (selected from hierarchical clustering). 281

B.18 Hierarchical clustering of the inter-regional interaction features (arranged in each category) that distinguish bedtime and sleep-onset in the re-referenced Montreal dataset (Fz-Cz, Cz-Pz, C4). 282

B.19 The continuous inter-regional interactional dynamics for all features during falling asleep. 284

B.20 The continuous feature dynamics of the six TE features in the MESA dataset. . . 285

B.21 The continuous feature dynamics of the thirty DTF features in the MESA dataset. 285

B.22 The continuous feature dynamics of the thirty cross-channel power band ratio features in the MESA dataset. 286

B.23 The Frontal delta band dynamics and the alpha dynamics in the Right-Central and Posterior region. 289

B.24 Phase histogram at stimulation onset for all 32 channels. 291

B.25 Phase histogram at stimulation onset for Fz channel in the single-channel phase-locking EEG system. 293

B.26 Feature-based classification between audio and control conditions in the last 3 minutes of nap. 294

1. Introduction

1.1. Sleep, insomnia, and its interventions

1.1.1. Sleep and sleep disturbances in neurodegenerative diseases

Sleep is normally defined as a recurring reversible state of reduced or even loss of cognitive responses, sensory actions, and active physical interactions with the external world (Bear et al., 2020). It is a universal biological phenomenon and an indispensable biological necessity for the preservation of human well-being, like food or air (Grandner and Fernandez, 2021). We spend, on average, one-third of our lifetime in sleep, so why is sleep so important? It has been demonstrated that sleep plays a crucial role in various physiological processes. One of the most well-established functionalities of sleep is the formation and consolidation of long-term memory on both neuro-behavioural levels and immunological memories, especially during slow-wave sleep (SWS) (Born et al., 2006; Rasch and Born, 2013). Normal sleep patterns have also served as an important regulation for normal body physiology, such as hormone levels, blood pressure, and other metabolisms (Allada and Bass, 2021). Additionally, sleep is also important for clearing the metabolic waste products and neurotoxins (such as β-Amyloid) in the brain via regulations of the flow of interstitial fluid and cerebrospinal fluid, thus improving cognitive function and brain health (Lewis, 2021; Yankova et al., 2021). In the situation of sleep deprivation (lack of sleep), normal cognitive functions and mood will be impaired (Dijk and Landolt, 2019), as sleep has been found to play a crucial role in regulating emotional processing and maintaining emotional well-being (Goldstein and Walker, 2014). In sum, deviant sleep patterns are closely associated with impacted life quality and a lot of diseases (Zisapel, 2007).

In the reverse way, these factors are also important in the regulation of normal sleep. For instance, disruptions in mood and normal body health could significantly impact the quality of sleep (Stewart et al., 2020; Konjarski et al., 2018). Specifically, sleep disturbances have been reciprocally related to many mental diseases (such as depression or anxiety) (Alvaro et al., 2013; Anderson and Bradley, 2013), and the cognitive declines and the pathologies of many neurodegenerative diseases and dementia (Naismith and Mowszowski, 2018; Anderson and Bradley,

2013). Formally, sleep disturbances encompass several different symptoms, including insomnia (difficulty in initiating and maintaining sleep), hypersomnia (e.g., excessive daytime sleepiness, narcolepsy), and parasomnia (such as nightmares) (Cormier, 1990).

Compared to young people, the normal ageing population will have reduced capability in initiating and maintaining sleep. Specifically, elder people tend to fall asleep more slowly, have more fragmented and shallower sleep, and have shortened sleep duration (Mander et al., 2017). There will also be significant changes in the micro-structures of sleep patterns in elderly people (such as slow-waves and spindles in electroencephalograms), which are closely related to the disrupted functionalities and benefits provided by the sleep (Mander et al., 2017). In neurodegenerative dementia patients (or, say, pathological ageing), sleep disturbances are more pervasive (Bliwise, 1993). Different neurodegenerative diseases have different pathologies, and thus different potential symptoms of sleep disturbances (Raggi and Ferri, 2010; Naismith et al., 2011). For example, Parkinson's disease (PD) belongs to the synucleinopathies and is characterised by the loss of dopaminergic neurons and the spread of α-synuclein and ubiquitin proteins. Rapid eye movement (REM) sleep behaviour disorder (RBD) is a type of sleep disorder that is specifically related to PD and often treated as early-stage hallmarks of PD pathology and PD-related dementia (Lin and Chen, 2018). One of the most well-studied neurodegenerative diseases is Alzheimer's disease (AD). AD belongs to tauopathies and is characterised by the accumulation of β-Amyloid proteins and tau tangles in the brain. AD patients suffer from many different sleep disorders, such as the specific sundowning syndrome: abnormal agitation or distress causing disruptions of sleep (especially in the evening). The AD pathologies are closely related to sleep disturbances. The accumulation of β-Amyloid is closely associated with increased sleep fragmentation and loss of non-rapid-eye-movement (NREM) sleep in both mice models (Roh et al., 2012) and human (Mander et al., 2015). The regional accumulation of tau tangles (especially in the sleep-regulating areas such as locus coeruleus) also causes early-stage sleep abnormalities before neurodegeneration (Holth et al., 2017). Specifically, this interaction between neurodegenerative dementia and sleep disturbances has been shown to be bi-directional. It has been found that β-Amyloid level is regulated by the wake-sleep patterns, and sleep deprivation could lead to the accumulation of the neurotoxin (Kang et al., 2009). The β-Amyloid level increases with wake time due to the synaptic activities in the neuronal firing, which would then be cleaned up mostly during NREM sleep (Xie et al., 2013) (the waste-cleaning function of sleep mentioned previously). Disrupted wake-sleep cycles, shallower slow-wave sleep (SWS), and sleep fragmentation could then lead to the failure of neurotoxin clearance, and as a vicious cycle, the accumulation leads to worse sleep disturbances and more increased neuronal activities, pushing deeper the disease progression (Ju et al., 2014). In a longitudinal view, even prior to symptoms of cognitive declines or other pathologies, increased sleep disruptions are causally related to higher risk factors for AD (Lim et al., 2013; Shi et al., 2018). Early treatment of sleep disorders has also been proven to delay the progression of cognitive impairment (Osorio et al., 2015). These bi-directional causal

linkages make sleep an important potential early-stage treatment venue to prevent or change the disease progression of neurodegenerative dementia (Mander et al., 2016).

1.1.2. Insomnia, and its current treatments

Among the many sleep disturbances, insomnia is the most common one, which affects around 30% of the population (Roth, 2007). Insomnia is normally defined as the difficulty in falling asleep, maintaining sleep (many fragmentations and shallower sleep), and early-morning awakenings. Clinically, the diagnosis of insomnia will classify patients into primary insomnia (insomnia is the main symptom not associated with any other disorders), or secondary insomnia caused by external drugs or neuropsychiatric disorders. (Roth and Roehrs, 2003). Whichever type, insomnia disorder itself is a vicious cycle. Insomnia normally causes disrupted daytime functionalities and accounts for a lot of mental disorders. In the reverse way, these mental disorders are also important risk factors for sleep disruptions (Anderson and Bradley, 2013). In general, the causes of insomnia can be complex. For instance, the previously mentioned neuropsychiatric disorders and neurodegenerative diseases could lead to insomnia. There are other causes, such as sleep-disordered breathing (Lavie, 2007) or behavioural disorders (Roth and Roehrs, 2003). Contrary to being drowsy and non-responsive, insomnia patients tend to be hyperaroused, with increased beta electroencephalogram (EEG) activities and lower delta band (slow frequency) activities, indicating a shallower sleep state (Marzano et al., 2008). The hyperaroused sleep state is also manifested in the incapability of maintaining the sleep states (Wei et al., 2017), as well as the phenomenon of paradoxical insomnia, where the participants remain consciously awake while brain EEG patterns show another way (Rezaie et al., 2018; Manconi et al., 2010). The neurophysiological mechanisms behind insomnia disorders are also complex, and not fully studied. For instance, disrupted circadian phases (such as in night-shift workers) could lead to insomnia. Recent research has proposed that the vulnerability of insomnia depends more on brain circuits governing emotions rather than those governing wake-sleep regulations (Van Someren, 2021).

Linking to the previously mentioned benefits and functionalities of sleep, for insomnia patients, better initiation and maintenance of the sleep state are more important for the potential benefits to occur. As discussed previously, sleep is an important early-stage intervention target for neurodegenerative diseases and dementia. The current treatment methods for insomnia include pharmacological treatments and non-pharmacological approaches. Pharmacological treatments include γ-aminobutyric acid (GABA) inhibition-enhancing drugs such as benzodiazepine or non-benzodiazepine hypnotics, or other anti-depressants and antipsychotics (Dauvilliers, 2007). However, most of the drugs have significant side effects, no cognitive benefits, and even deteriorate disease progressions. For instance, benzodiazepine hypnotics have been shown to cause amnestic effects and worse memory performances as in cognitive impairment (Fava, 1996; Hall-Porter et al., 2014). Non-benzodiazepine drugs such as zopiclone, although remaining as the first-line

insomnia treatment drugs, have also been shown to cause amnestic effects (Fava, 1996; Wisden et al., 2019), and even increased risk of dementia (Shih et al., 2015). Furthermore, most of these GABAergic-based drugs can cause drug resistance, dependence, and rebound after stopping (Wisden et al., 2019). Mechanistically, these drugs normally disrupt NREM sleep rather than enhance it and cause changes in neuronal oscillations countering memory consolidation (Mander et al., 2016).

There are many non-pharmacological approaches for treating insomnia disorders. A very simple, non-invasive, but effective treatment technique is called cognitive behavioural therapy for chronic insomnia (CBT-I). By changing sleep hygiene, thoughts, and behaviour, normally insomnia issues can be relieved. However, it is unclear whether CBT-I could have memory benefits and changes in disease progression (Mander et al., 2016). Novel neuromodulation approaches tend to be a more promising candidate for treating insomnia disorders, especially NREM sleep enhancement, by targeting the neuronal oscillations directly. In the seminal study by Marshall et al. (Marshall et al., 2006), they applied a transcranial electrical stimulation at 0.75 Hertz (Hz), and successfully entrained the slow-wave activities and enhanced overnight declarative memory performances. Further successful replications have also shown its efficacy in elderly adults (Westerberg et al., 2015). Of course, there are failed replications (Eggert et al., 2013), which actually underscore the complexity and unknown mechanisms behind the modulation of brain rhythms.

In recent years, as an even less invasive neuromodulation technique, auditory stimulation has become a promising sleep aid for the treatment of sleep disorders (Yoon and Baek, 2022). Compared to electrical stimulation, auditory stimulation is safer, easier to apply, and requires less training for the patients to use. In this thesis, the use of auditory stimulation in targeting sleep-onset insomnia (difficulty in falling asleep) will be the main focus and will be further discussed.

1.1.3. The need for novel neuromodulation paradigm in augmenting sleep-onset

The auditory stimulation in targeting the brain oscillations has been widely applied and tested in sleep neuromodulation and treating sleep disorders. The most successful application by far is the use of auditory stimulation in boosting slow-wave activities during deep NREM sleep. Ngo et al. applied auditory stimulation (50 milliseconds of pink noise) phase synchronised to the up-phase of the slow-wave oscillations in EEG could boost slow-wave activities, and enhance overnight memory performances (Ngo et al., 2013). There are many successful replications of this result (Ong et al., 2016), still, there are failed attempts as well (Cox et al., 2014; Weigenand et al., 2016). Mechanistically, the nonlemniscal pathway is activated during sleep via auditory stimulation, alerting the brain to external events, and causing widespread depolarisation of neurons across the cortex followed by hyperpolarisation, similar to the slow-wave activities

(Bellesi et al., 2014). This activation of synchronisation leads to a broader distribution of slow-wave activities, which were believed to be generated in a local pattern (Bellesi et al., 2014; Siclari et al., 2014). Under this mechanistic framework, there are many important factors to consider for effective stimulation, which further underlines its complexity. One of the most important aspects could be the stimulation timing (the phases of the neuronal oscillations to which the stimulation is synchronised). For instance, open-loop stimulation has been shown with no memory benefits (Weigenand et al., 2016), and targeting the wrong phases (down-phase of slow waves) could lead to reversed effects in both slow-wave activities and memory functions (Fattinger et al., 2017). This will be further discussed in Section 2.6, where the phase-locking paradigm will also be introduced. Aside from the timing, the type of audio tones used is also important. Research by far used many different types of audio tones, including coloured noise (white or pink), natural sounds (rain, for instance), or even music (Yoon and Baek, 2022). Most of the studies used pink noise in deep sleep stimulation, which has a broadband spectrum, to elicit a more pronounced response. The nonlemniscal pathway is very susceptible to habituation under repeated stimulation due to the fact that the original purpose of this pathway is to detect abnormalities in an unconscious situation (Bellesi et al., 2014). Interestingly, pink noise has also been shown to be robust in effects (Debellemanière et al., 2022). Other factors include the sound volumes (stimulation intensity) and sound frequency. The volumes should be strong enough to take effects on the ongoing brain oscillations, but also not too strong to elicit arousals, as the nonlemniscal pathway comprises the arousal-promoting systems (Bellesi et al., 2014). Volumes are also suggested to be adjusted in real-time according to sleep depth to achieve maximal effects (Esfahani et al., 2023). Similarly, the sound frequency is suggested to vary over time, to avoid the habituation effects. Most importantly, there does not seem to be a 'one for all' solution as the optimal stimulation. Schneider et al. (Schneider et al., 2020) tried to replicate the slow-wave enhancement results in elderly participants but failed, implying the need to further understand the nature of the stimulation, as well as the need to develop a more 'personalised' stimulation paradigm.

Focusing on insomnia disorder, and specifically sleep-onset insomnia, there are very few studies that studied the effects of inducing sleep onset (Forquer and Johnson, 2007; Messineo et al., 2017; Kawada and Suzuki, 1993; Montgomery-Downs et al., 2010; Nasari et al., 2018). However, most of these studies used different types of sounds and volumes: 60-70 decibels (dB) white noise (Forquer and Johnson, 2007), 60 dB pink noise (Kawada and Suzuki, 1993), broadband sounds (Messineo et al., 2017; Montgomery-Downs et al., 2010), or natural sounds (Nasari et al., 2018). Although they showed a significant decrease in the sleep onset latency, some studies are severely underpowered given a very small amount of participants ($N = 4$ for (Forquer and Johnson, 2007; Kawada and Suzuki, 1993). The results are further confounded by the distinct and potentially biased sleep onset latencies measurements, such as sleep questionnaires or surveys (Forquer and Johnson, 2007; Nasari et al., 2018), or even just signs of immobility (Montgomery-Downs et al.,

2010). In general, these differences make all these studies difficult to compare and generalise. There are also negative results, for instance, Daum et al. showed that the effect on latencies (to stage N1 or N2) highly depends on personal responsiveness to stimulation (Daum et al., 1988). There has been no study by far that systematically explores the timing and other stimulation factors (such as target frequency of neuronal oscillation) in the effect of inducing sleep-onset. Importantly, sound stimulations can disrupt the brain rhythms underpinning the transition to sleep thereby impeding falling asleep. Following the path of the investigations in the slow-wave sleep enhancement, phase-synchronised auditory stimulation will be applied to augment sleep onset (or facilitate falling asleep), targeting neuronal oscillations directly. There have been preliminary studies supporting the rationale of using phase-locked auditory stimulation to modulate vigilance. This will be further discussed and introduced in Chapter 6, where the experiment using auditory stimulation to boost sleep onset will be shown as well.

Currently, the most commonly used outcome measurement is based on sleep staging (see introduction below). It can only be a statistical quantification but not a subjective experience, especially in insomnia patients (see previous). In addition, to develop a closed-loop 'personalised' stimulation, there needs to be a deeper understanding of the brain dynamics during the falling asleep transition. Therefore, this thesis starts by exploring systematically the brain dynamics (quantified via EEG) during falling asleep: the dynamics biomarkers that best distinguish the wake and sleep states (Chapter 3 and Chapter 5), and functional models in describing the transition dynamics (Chapter 4). These understanding and tools will provide a solid ground for further developing a 'personalised' system based on the specific information from the brain directly.

1.2. Electroencephalogram for the quantification of sleep

1.2.1. Electroencephalogram in capturing brain electrophysiology

The brain is an incredibly intricate system, composed of billions of neurons with complex structures and interconnections. There are many different levels of brain activity measurements, ranging from the activity of individual neurons (patch-clamp) to the activity of extracellular activities (such as local field potential) and whole-brain imaging techniques. The electroencephalogram (EEG) is one of the whole-brain imaging techniques. EEG measurement is normally done by attaching flat metal electrodes to the scalp, to detect local electrical activities of the cortical neurons. As one can see, EEG only measures cortical activities, with no knowledge about deep brain activities. Briefly, the electric potential that EEG records comes from the synaptic activities (excitatory or inhibitory potentials) of the regional cortical pyramidal cells (Olejniczak, 2006). The polarity of the EEG signal is determined by the excitatory (or inhibitory) synaptic

inputs at the apical (or basal) dendrites of the pyramidal cells. Of course, what EEG is measuring is actually the sum of all the electrical activities underlying the electrode. There are many determinants for the magnitude of the EEG signals. For instance, when the cells have long somatodendritic axes and are aligned with each other, their electrical currents will have a stronger dipole and sum with each other (compare Purkinje cells and pyramidal cells, for instance). The other key factor is the synchronisation of the neuronal firings. When the neurons fire simultaneously, their electrical currents sum up to a larger EEG amplitude. One example would be the slow-wave activity during NREM sleep, as mentioned previously. Another example would be the EEG dynamics during the seizure, where the neurons are highly synchronised during which the patient will be unconscious. In sum, the different oscillation patterns and frequencies represent important information about underlying neuronal dynamics.

The first EEG recording was done by Hans Berger (Berger, 1929). His name is also well-known for the Berger effect, when people close their eyes, the EEG pattern will show synchronised alpha activity (around 10 Hz), especially at the occipital region. To measure EEG activity, the electrodes need to be placed on the relevant locations on the scalp. Currently, the standard for the electrode placement follows the 10-20 system (Homan, 1988). The electrode names are dependent on the cortical lobes they are measuring: Frontal, Parietal, Occipital and Temporal, and the numbers denote the left (odd) or right (even) side. There are 21 electrodes in the original 10-20 system, but nowadays, there are many other systems with many more electrodes, called high-density EEG recordings. The absolute EEG amplitude is extremely small (only around 10-20 microvolts); therefore, to get high-quality EEG data, the signal amplifier and the recording environment need to be well controlled. In the actual EEG recording setups, there will be a ground electrode (to put the amplifier and body at the same potential level) and a reference electrode (relative to which the EEG activity is measured).

The most standard way to analyse the EEG signals is the power spectrum analysis, which generally analyses the oscillation power at different frequencies. The oscillation activities are generally divided into different frequency bands: delta band (0.5-4 Hz), theta band (4-8 Hz), alpha band (8-12 Hz), beta band (12-30 Hz), and gamma band (above 30 Hz). Different frequency bands correlate to different brain states and functionalities. The Berger effect was one example previously. Low-frequency band activities normally correspond to drowsiness; for instance, the slow-wave activities belong to the delta band. Higher frequencies, such as beta or gamma, are normally related to high-level cognitive functions in the brain. The abnormalities in these different band activities are also important diagnostic criteria for many diseases (Schomer and Lopes da Silva, 2017). The previously mentioned higher EEG beta-band power at sleep onset for insomnia patients is a quick example. With the advance of mathematics and signal processing, there are now many other advanced analysis methods for the EEG signals as time series data, for instance, the many nonlinear metrics used in this thesis (introduced later). These different methodologies can contribute to a better mechanistic understanding of brain function, as well

as more sensitivity as biomarkers to detect different diseases.

Compared to other whole-brain imaging methods of brain activities, such as functional magnetic resonance imaging (fMRI), EEG is much easier to set up, has a lower cost, and a much higher temporal resolution. However, fMRI has much greater spatial resolution compared to EEG, as it is capable of detecting hemodynamic responses in deep brain structures. Modern high-density EEG recordings combined with a source-locational analysis (Srinivasan et al., 2006) has, to some extent, remedied this drawback, but the algorithmic reconstruction of the sources of activities still cannot compare to the actual three-dimensional imaging by the fMRI.

1.2.2. Sleep EEG and standard sleep stagings

EEG has long been used to study sleep state changes in the brain. The study by Loomis et al. (Loomis et al., 1937) provided the very first descriptions of the brain dynamics at different stages of sleep. They classified the sleep into five different states from A to E: from alpha trains to 'null' (low voltage recording) and to 14-per-second (Hz) waves, now known as spindles, and to 'random' patterns. Dement and Kleitman did some minor revision in 1957 (Dement and Kleitman, 1957b), and spindles were marked as the start of sleep stage 2, where stages A and B (from the previous study) were combined into stage 1, which is very similar to the criterion nowadays. In the same year, they also identified the cyclic pattern of NREM and REM sleep across the night, and found that eye movements in sleep are related to dreaming (Dement and Kleitman, 1957a). The current standards of sleep stagings all stand on these solid research foundations.

There are two major staging criteria used nowadays: the Rechtschaffen and Kales (Rechtschaffen and Kales, 1968), and the American Academy of Sleep Medicine (AASM) (Silber et al., 2007), with the latter one dominant now, and used in this thesis. The AASM sleep staging criterion is illustrated in Figure 1.1. When people close their eyes to go to sleep, the brain EEG patterns show alpha trains (recall the Berger effect). During the process of falling asleep, the brain rhythms slow down gradually, showing the theta waves (stage N1) and then delta waves in stage N2 sleep (where 'N' stands for 'NREM'). The start of stage N2 is normally defined as the appearance of K-complexes and spindles, two important EEG micro-events during sleep. Sleep spindles refer to a type of waxing and waning fast oscillation (12-15 Hz) that normally happens in light NREM sleep (stage N2 mostly). Spindles have been shown to be modulated by the up-phase of the slow-wave activities and provide important time frames for the information transfer from the hippocampus to the neocortex (Mölle et al., 2009; Staresina et al., 2015). The density of spindles and the functional coupling between spindles and slow-wave activities have also been shown to be correlated with memory performances (Helfrich et al., 2018). The K-complex is a brain wave pattern characterized by a brief high-voltage spike that lasts for longer than 0.5 seconds. K-complexes have been proposed to be a specific type of slow waves

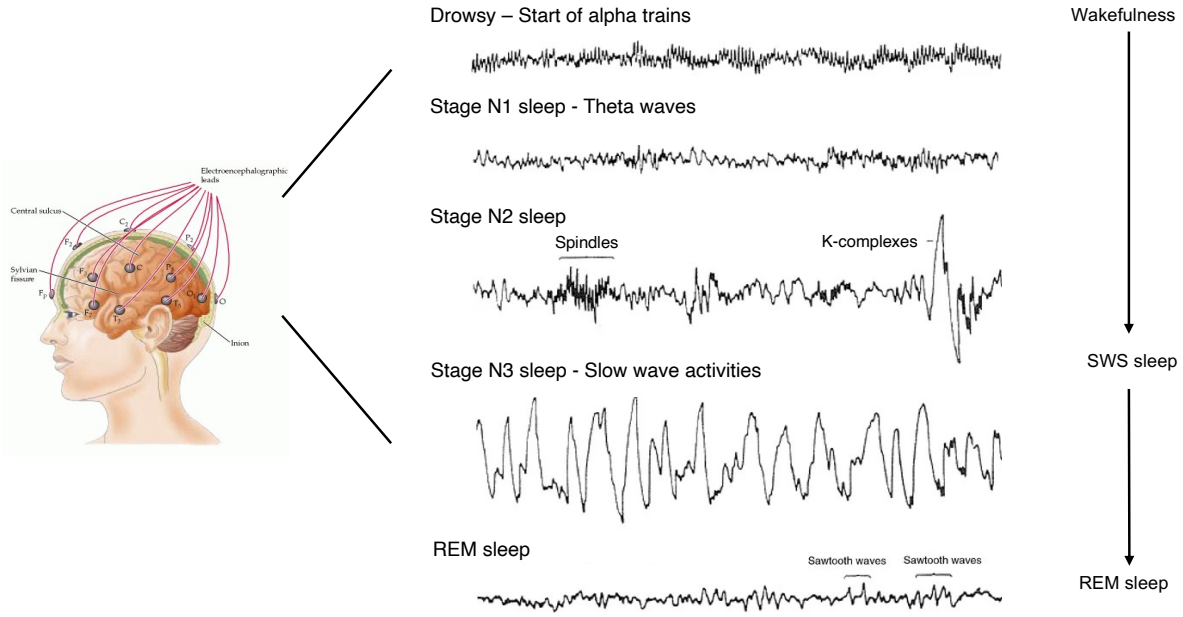


Figure 1.1: Demonstration of standard sleep staging and corresponding EEG dynamic properties. The different EEG patterns are associated with different stages of sleep. The start of alpha trains in EEG implies drowsiness. With sleep going deeper, theta waves will occur (stage N1), with the occurrence of spindles and K-complexes hallmark stage N2, which is also widely accepted as the onset of sleep. Stage N3 is represented by large-amplitude slow-wave activities, this is also the deepest NREM sleep. REM sleep pattern consists of sawtooth theta waves and looks like the awake EEG. Figure adapted from Wright (2009) with the kind permission from the publisher Springer Nature.

that occur in NREM sleep that further promote cortical synchronisation (Halász, 2016). As the sleep goes deeper, the cortex starts to synchronise, displaying a large amplitude, but slow-frequency slow-wave activities (normally 0.5-2 Hz, stage N3 sleep). All the previous stages are called NREM sleep in general. The REM sleep stage is the most intriguing stage during the entire night, where the EEG patterns resemble that of wakefulness, along with the suppression of muscle tones and the presence of rapid eye movements. REM sleep can also be further classified into phasic and tonic states, representing different physiological processes in the brain such as alertness or information processing (Simor et al., 2020). REM sleep normally happens later in the night (the second half of nocturnal sleep), and it alternates with NREM sleep. In the AASM criterion, REM sleep is marked as stage 5. However, it is not the 'deepest' sleep state, rather, it is closer to wakefulness, as it is a special sleep stage with cortical activations (Halász et al., 2004). This is also the reason that in the hypnogram (sleep stages plotted over time across the night), the REM stage is just below the wake stage. In this thesis, the falling asleep period (from wakefulness to the start of sleep) will be the focus to study and discuss.

In clinical applications, sleep staging is done with additional information from electrocardiogram (ECG), electromyography (EMG) and electrooculography (EOG), called polysomnography (PSG) as a whole. Nowadays, there are also breathing measurements (such as nasal airflow) for the detection of sleep apnea. These measurements contribute to a better sensitivity in mark-

ing the sleep stages and detections of sleep disorders. For instance, during the falling asleep transition, the heart rate and breathing will slow down, with slow eye movements in the EOG (see later for further details). The start of REM sleep can be better judged from the flat line in EMG and rapid eye movements in the EOG. There are also specific EEG montages recommended for PSG measurements, for optimal detection of those specific EEG micro-events (Iber, 2007; Silber et al., 2007). For instance, spindle activities are best detected in the Central region, while K-complexes and delta-band oscillations are better recorded in the Frontal region.

1.3. The quantification of falling asleep dynamics

The previous section briefly introduced the 'golden standard' for the quantification of sleep. So how is the falling asleep transition happening in the brain? From the perspective of the golden standard, it seems that the brain 'jumps' discretely into sleep from 0 to 1 and then from 1 to 2. However, from personal experience, this process is a smooth 'continuous' transition rather than a jump. Indeed, it has been widely acknowledged that the falling asleep transition (aka, the sleep-onset process) is a continuous dynamical process in the brain, accompanied by ceased behavioural responses and different physiological changes (Ogilvie, 2001). Still, one can easily distinguish whether a person is widely alert or sound asleep, but the intermediate process between them has been troubling scientists over the years as a 'no man's land' separating the two states (Merica and Fortune, 2004). In this section, the falling asleep transition dynamics will be briefly reviewed, from the perspective of biomarkers outside the brain (behavioural and physiological changes), as well as the brain dynamics quantified via EEG.

1.3.1. The falling asleep period

Before discussing the falling asleep dynamics, it is important to define what is the falling asleep period. The most crucial aspect of defining this period is to find the sleep onset point. The exact moment of sleep onset has been elusive, due to the lack of human feedback during this transition. Currently, the sleep onset is normally defined as the start of stage N2 sleep, with the appearance of sleep spindles and K-complexes as undoubtful moments of falling asleep (De Gennaro et al., 2001b,a). Compared to stage N2, stage N1 is more likely a transient state with mixed features of both wakefulness and sleep and with the least inter-rater consistency (Rosenberg and Van Hout, 2013). Using a data-driven approach to decode the intrinsic states during sleep from the functional magnetic resonance imaging (fMRI) data, stage N1 was also found to be the most ambiguous and heterogeneous state (Stevner et al., 2019; Rué-Queralt et al., 2021).

Still, this does not mean that stage N2 pinpoints the sleep onset perfectly. Sleep staging is a highly constrained method due to its low temporal resolutions (one value every 30 seconds).

Specifically, the sleep stage of one specific 30s epoch is determined by the dominant pattern. That is to say, if 49% of this epoch shows a wakefulness pattern, while the remaining 51% show a stage N2 pattern, this 30s epoch will be staged N2 sleep rather than wake, and biasing the judgements of actual sleep onset. Besides, as previously mentioned, the sleep stages determined from the PSG patterns correspond badly to the subjective experiences of sleep onset (Rezaie et al., 2018; Yang et al., 2010). Thus, there have been many studies trying to break out the sleep stage framework and use data-driven approaches to identify representations of intrinsic brain dynamic states in sleep. For instance, some studies transform the EEG data into a Fourier-based state space using power spectrum analysis (Schoch et al., 2017; Hermans et al., 2022). Using advanced machine learning methods, such as topic models or auto-encoders, the EEG dynamics in sleep can be transformed into clusters, or non-categorical latent states (Hermans et al., 2022). Similar approaches have also been applied to the fMRI data, as mentioned previously. These approaches are not perfect, as most of them suffer from sensitivity in hyperparameters and difficulty in interpretation for clinical or practical uses.

Focusing on the falling asleep period, aside from the endpoint, many studies would also define a starting point for this period. The most commonly used starting point is the start of alpha trains (Hori et al., 1994; Siclari et al., 2014). To deal with the inconsistency of the period length between subjects, some studies would just take a uniform time period before the start of stage N2 sleep. Defining a starting point (such as drowsiness) might be good for getting a homogeneous starting state, but one might lose the chance to get insights into the transition of the pre-sleep state as a key transient path (a preparation stage) towards actually falling asleep transition (Strauss et al., 2022; Kinreich et al., 2014). In this thesis, the falling asleep period will be defined from bedtime (the start of the PSG recording) to sleep onset (defined as the first 1-minute of stage N2 sleep), further details will be discussed in the following chapters.

1.3.2. Behavioural and physiological changes during falling asleep

Intuitively, when talking about sleep and falling asleep, the impression is that the person lies down, stops initiating any movements, and stops responding to external stimuli below a certain arousal threshold. That definition is primitive, but it was the original way researchers used to study the transition from wake to sleep, along with EEG dynamics (Davis et al., 1937, 1938; Flanigan, 1972; Liberson and Liberson, 1966). There are normally two types of behavioural measurements used, one active and another passive. Active behavioural responses are normally quantified as responses to external stimuli, for example, auditory stimulations (Ogilvie and Wilkinson, 1984; Ogilvie et al., 1989) or photic stimulation (Liberson and Liberson, 1966), where, as expected, the response times lengthen as the transition goes on. The convergence of these behavioural metrics with the gradual process of loss of consciousness while falling asleep provides additional value for studying corresponding EEG brain dynamic features (Strauss et al., 2022;

Comsa et al., 2019), or the prediction of sleep onset probability (Prerau et al., 2014). Passive behavioural measurements include actigraphy (Hauri and Wisbey, 1992; Martin and Hakim, 2011), which has now become a standard for sleep metrics. Another passive measurement is the continuous pressure system, and nowadays, combined with EMG, it is used to measure the loss of muscle control along with the falling asleep process.

There are many physiological changes accompanying the falling asleep transition, along with brain dynamic changes. There is a significant decrease in respiratory rate during falling asleep, accompanied by alpha to theta transition in the EEG patterns (Kay et al., 1994; Worsnop et al., 1998). This reduced ventilation is due to the change in upper airway resistance and respiratory pumps (Kay et al., 1994; Worsnop et al., 1998), and fundamentally, due to the central neural control mechanisms (Naifeh and Kamiya, 1981). Heart rates also slow down during this process, accompanied by a decrease in body temperature (Okamoto-mizuno et al., 2008). The changes in cardiovascular activities provide a window for studying the changes in the autonomic systems. Using the power spectrum analysis in the heart rate variability, it was identified that this cardiovascular change arises from the shift to parasympathetic dominance during falling asleep (Trinder et al., 2001; Shinar et al., 2006).

Eye movement changes have been the most well-studied physiological mechanisms in sleep. In 1955, Kleitman and Aserinsky used EOG, an electrical recording of eye movements, to study the eye movement changes during sleep, while all previous studies observed by manually raising the eyelids (Aserinsky and Kleitman, 1955). In their study, they identified two types of eye movements in sleep, one is the famous rapid eye movement in REM sleep. Another is the slow eye movements (SEM). During the process of falling asleep, SEM increases steadily and disappears when slow-wave activities appear. SEM activity is also found to be positively correlated to sleep onset latencies and subjective drowsiness levels (Torsvall and Åkerstedt, 1988; Åkerstedt and Gillberg, 1990; De Gennaro et al., 2005). This is further validated in EEG power spectrums, where increases in SEM are paralleled by increases in slow frequency EEG powers (De Gennaro et al., 2000; Marzano et al., 2007). Therefore, SEM has been a useful biomarker for sleep staging or prevention of sleep-related accidents (Shin et al., 2011).

Another important physiological marker is body temperature, and it has been known that body temperature decreases while falling asleep. Heat transfer from the core body to distal areas (hand or foot) to reduce the core body temperature has been an important physiological step for facilitating the falling asleep transition (Kräuchi et al., 1999, 2000). On the contrary, increasing core body temperature disrupts sleep (Fletcher et al., 1999). It has also been shown that skin and body temperature can be modulated and manipulated to facilitate sleep and improve sleep depth (Raymann et al., 2005, 2008). The fundamental mechanism behind this is related to the bidirectional relationship between the body temperature and related neuronal areas that control sleep and circadian rhythms in the hypothalamus, that the body and skin temperature provide

additional signalling pathways for regulating sleep patterns (Van Someren, 2000, 2004).

Most of these indexes have been incorporated into the PSG and are important clinical tools for aiding sleep staging and studying sleep disorders. Fundamentally, these changes are all external reflections of the internal neuronal regulation process and handovers to the parasympathetic controls during the falling asleep process. Besides, all these pieces of evidence show that the falling asleep transition is potentially a snapshot of larger circadian and homeostasis patterns behind the wake-sleep regulations.

1.3.3. The EEG brain dynamics during falling asleep

The sleep staging provided very poor temporal resolution in quantifying this process. As an extension, the most well-known fine gradation of the falling asleep period is Hori's nine-stage system (Hori et al., 1994). Hori used 5-second epochs and divided the sleep onset process from standard stage wakefulness to stage 2 into finer-grade 9 sub-stages. EEG starts from the continuous alpha train (stage 1) to alpha diminishing (stage 2-3); then the EEG starts to flatten (Hori stage 4), with occurrences of ripples in stage 5; the appearance of vertex sharp waves (a shorter version of K-complexes that happen before stage N2 sleep normally) with incomplete spindle traces hallmarks stage 6-7; and finally, complete spindles is the stage 9, where they defined it as definitive sleep. This finer staging provides insights about the detailed EEG pattern changes during this process and also provides the framework to be combined with other modern EEG analysis methodologies (spectrum, coherence or others) in analyzing the brain dynamics during sleep onset. However, practically, doing sleep staging on 5-second epochs is very labour-demanding. EEG patterns are also likely to be unstable during the short 5-second epochs, making scoring even more difficult and confusing (Ogilvie, 2001).

Most of the key findings of the falling asleep brain dynamics come from the EEG power spectrum analysis. According to the sleep stages, during the falling asleep transition, the EEG rhythms generally slow down, with increases in the low-frequency powers and decreases in high-frequency powers. Upon conducting a more comprehensive spatial-temporal quantitative EEG analysis, it was found that different brain areas exhibit distinct dynamic patterns in the EEG power bands and display varying connectivity based on coherence analysis. Again, Tadao Hori and his team initiated the field with an extensive exploration of the topological and temporal changes of the different EEG frequency bands during the falling asleep process. In 1985, Hori used the spatial-temporal EEG power spectrum analysis on five midline EEG electrodes to study the changes in EEG dynamics from 10 minutes before stage 1 sleep to 30 minutes after (Hori, 1985). He found that the theta band activity started to increase in the frontal region (Fz) first, and the occipital area (Oz) showed the slowest increase of delta activity. This was the first direct observation that the EEG power dynamics are not uniformly changing during the falling asleep period across the brain. A series of different studies followed this path. Wright et

al. (Wright et al., 1995) studied the EEG power changes across the brain in the first and last minute form stage 1 to 2 transition and found that the midline areas showed the most significant changes. Using a coherence analysis, they also found a breakdown of alpha coherence between inter-hemisphere regions. In recent studies, it was found that this breakdown of alpha coherence actually corresponds to the loss of behavioural responsiveness (Comsa et al., 2019).

Starting from 1997, Hori and his team published a series of papers describing fully the topographical and temporal dynamics of different EEG frequency band activities as well as the coherence changes (Tanaka et al., 1997; Morikawa et al., 1997; Tanaka et al., 2000, 1999; Morikawa et al., 2002), mostly based on the 9-stage system they developed. Most of these findings laid important foundations in the current understanding of falling asleep brain dynamics. It was found that the anterior-central brain areas started to show synchronisation and increased activity in low-frequency bands (delta and theta) in the early stages of falling asleep (prior to Hori's stage 5), and the increases were more clear after the appearance of vertex sharp waves (Morikawa et al., 1997; Tanaka et al., 2000). On the other hand, the alpha band switched from posterior dominance to anterior between Hori's stages 3 to 7 (Tanaka et al., 1997). However, notice that they had a different definition of the alpha frequency band, and they divided the alpha band into three sub-bands for study based on a principal component analysis. They further extended their analysis into the sigma band (the spindle frequency range) and found a global increase of both coherence and power in this frequency band after the stage N1 sleep (Morikawa et al., 2002). These early studies provided important insights that sleep initiation involves different local processes and starts before stage N1 sleep.

From the year 2000, De Gennaro and his colleagues contributed to the most significant part of the picture of falling asleep EEG dynamics. De Gennaro et al. in 2001 conducted a single-Hz EEG power analysis on a 4-minute period before stage N1, a 5-minute period after stage N2, and a 4-minute period between the two stages (De Gennaro et al., 2001a). The delta and theta bands increased monotonically during the process of sleep initiation (before stage N2), with the alpha and beta bands displaying a monotonic decrease. However, they found that after the start of stage N2, there was a reversed change in the alpha band (8-11 Hz). This difference implies a change in the functional meaning of the alpha band activities after sleep onset, where it was indeed confirmed that while occipital alpha correlates to drowsiness, the frontal alpha activity after sleep initiation is related to sleep maintaining process (Pivik and Harman, 1995). Recall from previous that the alpha-band activity shifted from occipital dominance to anterior regions during sleep onset (Tanaka et al., 1997), and it was also confirmed in many later results that there is a posterior-to-anterior-spreading of alpha-band activity prior to sleep onset (De Gennaro et al., 2001b; Gennaro et al., 2004). From their analyses, they also confirmed that stage N2 is a more solid boundary as sleep onset, and stage N1 is more likely a transient period where the synchronisation is still building up and alpha-activity is shifting (De Gennaro et al., 2001a,b).

In 2004, De Gennaro et al. (Gennaro et al., 2004) applied a novel approach called the directed transfer function, to analyse the information transfer during the falling asleep process, while previously, all connectivity was based on spectrum coherence with no directional information. In this study, they found the significant role of the frontal regions in the spread of synchronisation patterns during the falling asleep period, and this anterior-posterior functional connection was enhanced in conditions of sleep deprivation (Gennaro et al., 2005). Indeed, the key role of the frontal region has later been confirmed in several other studies. Frontal regions have naturally more homeostasis pressure (D' Ambrosio et al., 2019), and display more sleep needs (Ferrara et al., 2002), have stronger delta-band activities (especially slow-wave activities) compared to other regions prior to sleep (Tanaka et al., 2000; Siclari et al., 2014; Marzano et al., 2013), where these analyses further show that it also plays the imperative role as a stabiliser in sleep initiation.

Recent analyses of EEG data during the falling asleep transition focus more on the neurophysiological and functional insights, thanks to the source locational analysis that can help identify current sources of EEG data. Siclari et al. (Siclari et al., 2014) analysed the falling asleep period from alpha trains to the start of slow-wave activities and found two types of slow-wave activities, with different cortical sources and functional meanings. At an earlier stage, a bottom-up synchronisation (similar to the function of K-complexes) results in the type-I slow waves, which involve a large frontal-medial region. At a late stage, when sleep is probably established, there will be a different slow wave with a smaller amplitude but more density across the entire cortex, as a sign of global synchronisation. They also discovered that spindles were sparse in the beginning, while during the transition, spindle frequency dropped and spread to more cortical regions. Guerrero and Achermann (Fernandez Guerrero and Achermann, 2018) further used effective connectivity metrics to study deeper brain structures and found the anterior-posterior decoupling of the default mode network (DMN) while falling asleep. Specifically, the posterior cingulate cortex (as a central node of DMN) plays an important functional role by information transfer to the other DMN nodes. This was later verified in the fMRI study by Stevner (Stevner et al., 2019) that the mean increase of activation in the DMN seems to be an essential path towards sleep onset transition. Aside from the spectrum power itself, the changes in the oscillation rhythms (local frequency peaks in the spectrum) have also been widely studied recently. Marzano et al. (in De Gennaro's team) (Marzano et al., 2013) systematically studied the power and oscillation changes in each frequency bin before and after sleep onset. They verified many of their previous findings, for instance, the anterior dominance of the slow-wave activities, early-stage synchronisation in low-frequency bands in the anterior regions, and the shift of alpha from posterior to anterior. In addition, they identified that the occipital region shifted from alpha dominance to theta. Oscillation patterns also change significantly with their power. Beta oscillations decreased gradually along with their power reduction, while delta power and oscillations both showed increases. Sigma band oscillation increased (due to the appearance of spindles) and

showed a central dominance. Achermann et al. (Achermann et al., 2019) used a different method to detect the oscillatory events (using an autoregressive model to detect oscillation damping) and identified similar findings in general.

One important and interesting phenomenon during the falling asleep process is the hypnagogic experience or dream-like mentation. In Hori's study (Hori et al., 1994), it was found that the hypnagogic experience is highest in Hori's stage 5, after alpha flattening and before the appearance of vertex sharp waves. Functionally, the hypnagogic events are related to the activation of the DMN (as described previously), which is further related to the increased metabolic activity in the visual cortex (occipital area) during the transition process (Kjaer et al., 2002). The visual images during the falling asleep period can even be decoded from the occipital fMRI dynamics (Horikawa et al., 2013). This functional relevance could also probably explain the later slow-wave dominance in the posterior regions (Gorgoni et al., 2019). Further intracranial EEG recordings also found that the earlier deactivation of the thalamus than the cortex resulted in a still-excited cortex during the falling asleep transition, and thus the hallucinations (Magnin et al., 2010).

What can we get from these studies as a whole? Firstly, the falling asleep transition is not uniform both temporally and spatially, with different frequency bands showing different cortical dominance, cortical sources, oscillation patterns, and significant cortical connectivity changes and information transfer patterns. For a more detailed review, see Gorgoni et al. (2019, 2020). Secondly, although lots of these studies showed complex spatial-temporal EEG power band dynamics, and defined different 'falling asleep periods' (in both duration and definition) to study, some general patterns can still be observed. The frontal region starts to synchronise first, with the highest low-frequency activity (indicating the sleep pressure) and serves an important functional role in synchronising the entire cortex. The alpha activity shifts from occipital to frontal, serving a sleep maintenance role. Occipital areas shift from alpha dominance to theta activity, which is functionally linked to hypnagogic events and DMN activations. Still, one important challenge within this field of study is to link the EEG-based findings to underpinning neuronal controls. The wake-sleep transitions are regulated by deeper brain structures (coming next), but EEG only captures the superficial cortical activities. In Chapter 4, a dynamical system framework based on underpinning neurobiology of wake-sleep regulations will be proposed to capture the changes of different EEG feature dynamics as a whole.

The previous dynamics discussed all belong to a 'normal' falling asleep transition. Another path of the study revolves around how sleep deprivation could change the dynamics. In sleep-deprived participants, there would be higher sleep homeostasis forces, which could thus lead to a different transition pattern and provide more mechanistic insights about the process. This is not reviewed here. Fundamentally, the transition pattern is a more intense and accelerated version of the normal one (Gorgoni et al., 2020). There are, of course, more questions in this regard. For

instance, children display a different synchronisation pattern during the transition (Spiess et al., 2018). Genders also play a role, and girls were found to be more susceptible to sleep initiation problems (de Zambotti et al., 2020). Ageing is another big factor, as it has been shown that older adults reach a relatively shallower sleep onset state, and are thus more susceptible to sleep fragmentation (Gorgoni et al., 2021). More questions can be asked about different sleep disorders, such as narcolepsy and insomnia, or other different neurological diseases, such as AD and dementia (Ogilvie, 2001). These are by far, not fully studied, and they will not be further reviewed here as they are outside the general study scope of this thesis. In Chapter 3, some of these factors will be studied with a novel EEG feature-based approach.

Taking a broader view, the previous studies do not capture the full picture of the EEG brain dynamics, as all of them are based on power spectrum analysis. There is some other literature using event-related potential (ERP) analysis to probe the loss of consciousness during the falling asleep transition. Specifically, an odd-ball stimulation paradigm was very commonly used to analyse the mismatch negativity (the difference in the ERP when presented a deviant stimuli), and during the falling asleep transition, the mismatch negativity will gradually diminish (Harsh et al., 1994; Nittono et al., 2001), implying the underlying changes of the cognitive processing. Aside from that, there is an emerging field of nonlinear EEG analysis, which could all provide novel mechanistic insights into different aspects of underpinning dynamics and novel biomarker potentials (Pritchard and Duke, 1992; Stam, 2005). For instance, it has been found that the sleeping brain has a lower dimension and lower complexity structures, with lower entropy and reduced fractal dimension (Ma et al., 2018; Shen et al., 2003). In this thesis, many nonlinear EEG dynamic features will be considered to study the wake-sleep transition dynamics as a whole.

Fundamentally, EEG is a limited method to measure the brain, only reflecting cortical activities. There have also been fMRI (Stevner et al., 2019) or positron emission tomography (PET) (Maquet, 2000) studies for the hemodynamic changes during the falling asleep transition. They are not further reviewed or discussed here, as from the perspective of this thesis, EEG is the most feasible way to measure brain activities for neuromodulation purposes.

1.4. The neurophysiology of wake-sleep regulations

1.4.1. The neuronal circuits underlying the wakefulness and sleep

Previously, the falling asleep dynamics were discussed on the level of EEG only. How is this transition supported by the underlying neuronal circuit interactions? Compared to the time spent in the NREM sleep bouts, the actual transition time into the NREM sleep only takes less than 1% of the duration (Saper et al., 2010). In this section, I will shortly present the current knowledge of the neuronal circuits underpinning the wakefulness and sleep state, and how the

transition is happening on the circuit level. For a more comprehensive neurobiological review, see Saper et al. (2005, 2010); Sulaman et al. (2023).

Wakefulness is supported by a wide range of neuronal populations distributed across the brain. Currently, the framework is mostly based on the well-known ascending reticular activating system (ARAS). The wakefulness-supporting circuits have two main pathways, one through the thalamic relay neurons and the other directly into the cortex (Saper et al., 2001). The first path contains the ascending pathway from the brainstem towards the thalamus, which activates the thalamus reticular nucleus and thalamic relay neurons and transfers information towards the cortex. The major input comes from the cholinergic nuclei: the pedunculopontine and laterodorsal tegmental nuclei (PPT and LDT), whose neurotransmitter is acetylcholine (ACh). Interestingly, the PPT and LDT also play an important role in the regulation of REM sleep. The role of the thalamus in supporting arousal is not clear, but the thalamic reticular nucleus has been found to be crucial in selective attention, which controls the key gateway in the communication between the thalamus relay neurons and the cortex (McCormick, 1989; Saper et al., 2005). The second path activates a wide range of neurons in the lateral hypothalamus, basal forebrain and the cortex, without the thalamus in the middle. The major inputs come from the monoaminergic neuron populations at the mesopontine level in the upper brainstem and caudal hypothalamus, including the locus coeruleus (norepinephrine, NA), the dorsal and median raphe (serotonin, or 5-HT), the ventral periaqueductal grey matter (dopamine), and the tuberomammillary neurons (histamine). Contrary to the cholinergic nuclei, neurons from these nuclei are only active during wakefulness and fall silent in NREM and REM sleep. In this pathway, the inputs from the brainstem structures to the cortex are augmented via the basal forebrain structures, especially the orexin (or hypocretin) and the melanin-concentrating hormone neurons inside the lateral hypothalamus. The basal forebrain structures play a more important role than just relaying the messages, especially the orexin neurons in the posterior-lateral hypothalamus. The orexin neurons are also shown to be a critical stabiliser in the wake-sleep regulation system, by suppressing REM sleep while sustaining wakefulness (Saper et al., 2005, 2010). Lesions of these orexin neurons will cause narcolepsy (Hara et al., 2001), whose major symptoms are very short sleep-onset latencies, with initial sleep periods containing REM sleep. It remains an open question why there are so many neuron groups supporting wakefulness. Potential reasons could be that in the presence of lesions in one structure, the others could compensate and sustain normal wake-sleep functions (Posner, 2007). These different structures are also important in the regulation of normal behavioural functions and sensory inputs during wakefulness. One interesting theory proposed recently is that the cortex is defaulted to synchronised slow oscillations (Bandarabadi et al., 2020), while these arousal systems stimulate the cortex to disrupt the synchronisation and support cognitive actions (Sulaman et al., 2023).

The NREM sleep-promoting nuclei mainly reside inside the lateral hypothalamus, following the findings by Von Economo in early 1930. The most important and well-studied structure is the

ventrolateral preoptic area (VLPO). VLPO neurons contain mostly the GABA and galanin inhibitory neurotransmitters and innervate massively the ASAR system components (Sherin et al., 1998). The VLPO neurons fire actively during NREM sleep but remain silent in wakefulness and REM. The lesions in VLPO could significantly decrease the amount of sleep. The previously mentioned GABAergic-enhancing drugs mostly function by enhancing the inhibition of the VLPO to those wake-promoting nuclei. Similar to wakefulness, there are also many other sub-structures in the thalamus and basal forebrain that help maintain NREM sleep (Sulaman et al., 2023). There are also specific REM-regulatory neuron populations (such as the PPT and LDT discussed previously), but they fall outside the scope of this thesis and will not be further discussed.

How are the wake-sleep transitions happening? It was found that there is a mutual inhibitory relationship between VLPO and many of the wake-promoting structures. Based on this mutual inhibition, Saper et al. proposed a flip-flop switch model (Saper et al., 2010). For instance, during the wakefulness-to-sleep transition, the wake-promoting neurons will cease firing gradually, while the sleep-promoting neurons (VLPO) will gain advantages and start to fire. The increase in the VLPO neuron firings will thus cause a stronger inhibition of those wake-promoting systems, and, thus, a rapid collapse towards the sleep state. This is the fundamental reason behind the 'quick' state transitions, which are much shorter than the sleep state itself. In the individual neurons, the changes in firing take less than one second (Takahashi et al., 2010), but of course, to the size of the whole brain, this would normally take several seconds or up to minutes of transition. This can also be observed in the EEG dynamic feature state space (Diniz Behn et al., 2010), where the two states show clear-cut EEG power features in two clusters. Based on these neurobiological insights, several mathematical models (e.g., using neuron mass models) have been built, illustrating that the wake-sleep transition is a bifurcation process. This will be further extended in Chapter 4, as currently, there is no empirical evidence, especially in EEG data, demonstrating this bifurcation dynamic pattern. Inside the flip-flop model, there is an additional component, with the orexin neurons as the stabilising hand on the two sides. Without orexin neurons, it was found that the switches become unstable and cause several sleep disorders (Saper et al., 2010).

1.4.2. Regulations of wakefulness and sleep

The two-process regulation model

The switch presented previously simply serves the 'functionality' for the transition, rather than initiation. The VLPO neurons do not initiate the firings themselves, but only when the external inhibition from the wake-promoting neurons can be overcome by its internal firings. Currently, the most widely-used wake-sleep regulation model is the two-process model (Borbély et al., 2016),

that sleep is regulated by the homeostasis forces and circadian rhythms. The homeostasis forces have been considered the actual 'accumulating' force that helps initiate the transition into sleep. During wakefulness, several neurochemical interactions and cellular metabolism will produce 'wastes', which serve as the internal forces that will cause sleep transitions. Current knowledge normally associates the homeostasis drives with adenosine, a by-product of cell metabolic activities, as the internal biomarker. Functionally, adenosine works on the excitatory A2a receptors of the meningeal cells in the VLPO, leading to stronger inhibition of the arousal systems (Scammell et al., 2001). The caffeine that we normally consume daily also works via the blockage of A2a receptors (Huang et al., 2005). Another well-established hallmark of the homeostasis force is the slow-wave activity in EEG (Borb and Achermann, 1999); The amount of slow-wave activity increases with the time spent awake.

The other important regulatory factor is the circadian rhythm. One day is 24 hours, and people's lives revolve around this rhythm. The master clock in the human brain is found inside the hypothalamus as well, which is the suprachiasmatic nucleus (SCN). SCN adjusts its internal clock by responding to external light input in the retina and by the release of a hormone called melatonin from the pineal gland during darkness (Saper et al., 2005). Injections of melatonin could entrain the circadian rhythms (Cassone et al., 1986), and it is currently used as a short-term treatment of insomnia. The SCN does not directly function on the wake-sleep regulatory systems, but rather, it outputs to an intermediate called the subparaventricular zone. From this subparaventricular zone, it further extends to the dorsomedial nucleus of the hypothalamus. The dorsomedial nucleus functions as an amplifier and projects to VLPO for sleep regulations. Aside from sleep, a lot of other body metabolisms are also regulated by this circadian rhythm. The dorsomedial nucleus also projects widely to other locations for these regulations, such as the medial preoptic area (for body temperature), and many different hormone-control areas (Saper et al., 2005). The 3-stage controlling (with intermediates between SCN and VLPO) ensures the potential to shift the physiological rhythms to the actual environment and other factors (Saper et al., 2005). In the previous section, it was mentioned that by modulating the body temperature, sleep can also be induced via the pathways to the circadian rhythms. Ensuring a healthy and normal circadian rhythm is thus important in many of these aspects, and disrupted wake-sleep patterns could lead to disruptions in many of these physiological cycles, and act back to disrupt sleep initiation. As Sulaman et al. proposed (Sulaman et al., 2023), sleep initiation is a gradual de-arousal process, where at the usual time, with the same stereotypic behaviour, one can entrain the body's rhythms and facilitate sleep initiation by the body's own regulations to reduce wake-promoting tones.

Local sleep regulations

Contrary to prevailing theories that view sleep as a global state, the concept of local regulation of wakefulness and sleep offers a fresh perspective that could elucidate various sleep phenomena. A prime example of this local regulation in nature is the bottlenose dolphin, which exhibits hemisphere-wise sleep: one hemisphere can rest while the other remains awake. Sleep deprivation in one hemisphere does not impact the other, suggesting independent regulatory processes in each hemisphere (Oleksenko et al., 1992). In a seminal study on rats, Vyazovskiy et al. (Vyazovskiy et al., 2011) discovered that cortical slow-wave activities can occur locally, with the entire cortex simultaneously exhibiting both wakefulness and sleep states. Similar coexistence of wake-like and sleep-like EEG patterns has been observed in humans (Nobili et al., 2011). Crucially, these local slow-wave activities, driven by local synaptic changes and cellular demands, are associated with learning functions and memory benefits (Huber et al., 2004; Murphy et al., 2011). These findings suggest that sleep could be regulated based on local demands. Krueger et al. (Krueger et al., 2008) proposed that sleep is not a bottom-up control from the brainstem but a fundamental property of neuronal assemblies (or cortical columns), driven in a use-dependent manner. Sleep is influenced by regional sleep needs, determined by regional cellular activities and synaptic changes. Homeostasis and circadian regulations could also function locally, with global state transitions driven by accumulating areas of 'local sleep patterns' (D' Ambrosio et al., 2019).

The theory of local sleep regulation could provide insights into various sleep phenomena and abnormalities. For instance, the spatial-temporal differences in EEG patterns across different brain regions during the transition to sleep discussed previously could be explained by this theory. The frontalisation of slow-wave activities reflects a higher sleep pressure in the frontal regions (Ferrara et al., 2002), due to intrinsically different activation inputs (D' Ambrosio et al., 2019). Sleep state transition is not an all-or-none process, but rather a fluid process with a mixture of both states existing across brain areas (Sarasso et al., 2014). Insomnia patients may have more locally awake regions, even when polysomnography shows general sleep patterns (Ståle-sen Ramfjord et al., 2020). Paradoxical insomnia could be explained by excessive local activity that maintains consciousness. Local sleepiness could also potentially cause cognitive impairment resulting from sleep deprivation or sleep inertia (a brief period of cognitive impairment after waking up). While this theory is promising, it is still in its early stages. For a more detailed review, refer to Krueger et al. (2008, 2019).

1.5. The structure of this thesis

The overarching goal of this thesis is to develop the foundations for a closed-loop neuromodulation intervention for sleep-onset insomnia to prevent or slow dementia.

Chapter 1 (this chapter) summarizes the key background knowledge and knowledge gaps relevant to the thesis. It includes the fundamentals of sleep EEG, the current approaches to quantify the falling asleep dynamics, the neurophysiology behind the wake-sleep regulations, etc.

Chapter 2 summarizes the key methodologies used in the thesis. It describes the sleep EEG datasets, the principles of the feature-based statistical learning approach (used in Chapter 3) and the dynamical systems computational approach (used in Chapter 4), as well as the phase-locked auditory stimulation paradigm (used in Chapter 6).

Chapter 3 deploys a feature-based statistical learning approach to existing sleep-EEG datasets to identify the most informative EEG features that distinguish sleep onset from bedtime wakefulness. It then explores the dependency of those sleep onset features on demographic factors and cognitive impairment.

Chapter 4 deploys a dynamical systems approach to show that the brain EEG transition from wakefulness to sleep can be accurately modelled in the feature state space as a catastrophic bifurcation dynamic.

Chapter 5 extends the results in Chapter 3 and Chapter 4 to examine differences in the falling asleep EEG dynamics between brain regions and the interaction between them. It uses a Granger-causality method to derive a mechanistic insight into the sleep initiation dynamics.

Chapter 6 tests the feasibility of augmenting the falling asleep process using phase-locked auditory stimulation in a cohort of healthy young adults. The study was a sham-controlled day nap and included pre-post nap cognitive tests. The falling asleep process is quantified using the insights generated in Chapter 3.

1.6. Publications statement

The outcomes of the thesis are in preparation for publication. Specifically, one paper will report the catastrophic bifurcation description of the falling asleep process (Chapter 4 and some of Chapter 5). A second paper will report the most informative EEG features that distinguish sleep onset from bedtime wakefulness (Chapter 3 and some of Chapter 5). A third paper will report the feasibility testing of using phase-locked auditory stimulation to promote daytime sleep (Chapter 6) together with an earlier study establishing feasibility in awake eyes-closed state (not included in the thesis).

2. Material and Methods

2.1. Sleep EEG datasets

2.1.1. Dataset 1: Multi-Ethnic Study of Atherosclerosis (MESA) - Elderly cohort

The Multi-Ethnic Study of Atherosclerosis (MESA) dataset (Chen et al., 2015; Zhang et al., 2018) was obtained with permission from National Sleep Research Resources (NSRR) (https://sleepdata.org/datasets/mesa). The study was originally aimed as a longitudinal study of the sleep factors that could lead to cardiovascular diseases in middle-aged to elderly cohorts. There are no disease records or notes for the participants of this dataset, and participant recruitment was not limited to cardiovascular disease patients. Thus, the participants are generally heterogeneous and can be considered 'healthy' elderly in general.

All data were collected as part of research protocols that were approved by the local institutional review board at each institution; written, informed consent was obtained from each individual before participation. The participants were generally from community-based samples, and the dataset originally contained 2055 subjects with whole-night PSG recordings. According to NSRR, the PSG collection was done at home using the Compumedics Somte System (Compumedics Ltd., Abbotsford, Australia). The PSG data are available in European Data Format (EDF) files along with their annotation files. The sleep stages were scored in 30-s epochs using the American Academy of Sleep Medicine (AASM) (Iber, 2007; Silber et al., 2007) criterion and other annotations of sleep events (such as arousals and sleep apnea) are also provided. The PSG data contains 3 bipolar linked EEG channels (Fz-Cz, Cz-Oz, C4-M1, see (Homan, 1988) for 10-20 EEG system) (see Figure 2.1), with a sampling rate of 256 Hertz (Hz). This is also one of the standard montages recommended in AASM criterion (Silber et al., 2007). The ECG data was also extracted along with the EEG data, sampled at 256 Hz as well. The demographic information (gender and age) was extracted for every participant with available PSG data. For each EEG channel, a signal quality index was provided (ranging from 1 to 5, with 5 being the best quality), which was also extracted. The signal quality index measures the percentage of good-quality data in the overnight recording, and this index will be used later to exclude

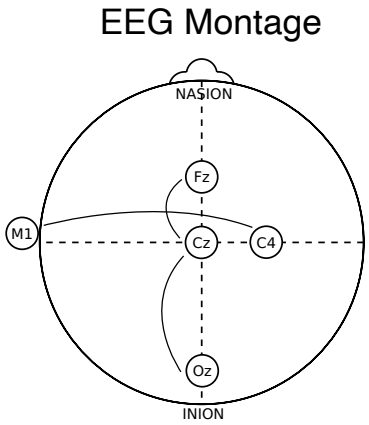


Figure 2.1: The EEG montage of the MESA dataset. The montage contains three linked EEG channels: Fz-Cz, Cz-Oz and C4-M1, with relative locations marked in the plot.

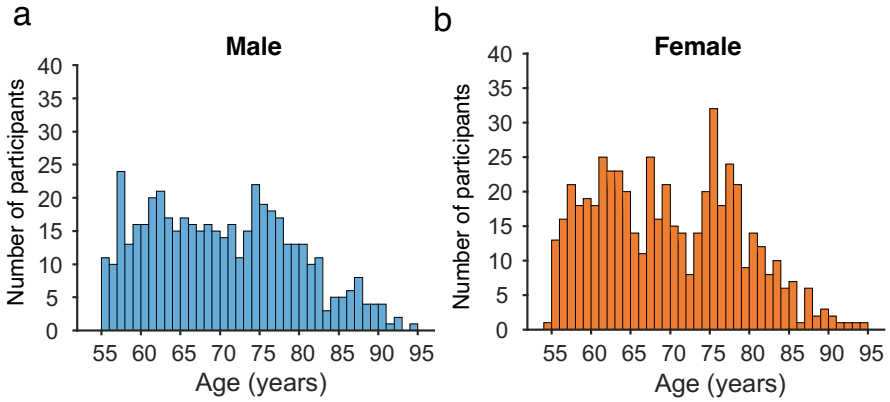


Figure 2.2: Age distribution grouped by gender in the MESA dataset. (**a**) for male and (**b**) for female particpants.

channels with bad overall quality.

After reading both the PSG data (EEG and ECG) and sleep annotations file, the PSG data was firstly ensured to align with the sleep scoring. Then the sleep scoring for each participant is checked through to ensure that: a. The sleep scoring starts from stage 0 (wake), and b. the first continuous 1-minute stage N2 sleep as the sleep onset (see later section) can be found. Participants who did not meet these criteria were excluded from further analysis. This resulted in a total of 1651 participants left. The sleep onset latency was then calculated as the time from the scoring start to the marked sleep onset. The latencies for all participants ranged from 1-minute to over 6 hours, with mean $\pm$ st.d. 81.72 ± 54.96 minutes (see Appendix Figure B.1a). This is a very wide range, and it does not make sense to study the participants who did not have a sleep onset for 6 hours. In this thesis, after consulting with sleep experts, I excluded all participants with sleep onset latency below 3 minutes or above 90 minutes. See the next section for a detailed explanation and discussion of the sleep onset definition and latency exclusions.

After this initial data cleaning, there are $N = 1011$ participants left for analysis.

The participant age distribution is shown in Figure 2.2, separated by gender (blue for male and orange for female). The dataset included $N = 477$ males with a mean age of 69.81 ± 9.21 st.d. years and $N = 534$ females with ages 69.28 ± 8.96 years. The Kruskal-Wallis test between the age distribution of males and females shows no significance (p-value = 0.45). It shows that the MESA dataset includes a middle-aged to elderly cohort.

2.1.2. Dataset 2: the Montreal's sleep EEG dataset with mild cognitive impairment

Mild Cognitive Impairment (MCI) is a neurological condition that involves a gradual decline in cognitive abilities faster than normal ageing but is not as severe as dementia (Gauthier et al., 2006). People with MCI may experience difficulties in one or more of these cognitive domains: attention and speed processing, executive functions, visual and verbal episodic learning and memory, visuospatial abilities, and language. However, these challenges do not significantly interfere with their daily functioning or independence. MCI is generally considered a transient state between normal ageing and more severe cognitive dysfunctions. MCI can generally be categorised into amnestic (aMCI) or non-amnestic (naMCI) subtypes, depending on whether the memory function is impaired (aMCI) or not (naMCI). As illustrated in Section 1.1, sleep disturbances are closely associated with cognitive decline, and the relationship is bi-directional. Therefore, it is also important to study the falling asleep dynamics of cognitively impaired patients and see how they differ from healthy controls. This understanding will be important for further developing the neuromodulation system for sleep disturbance intervention.

The Montreal's sleep EEG dataset (later called the Montreal dataset) was obtained from collaborators Professor Nadia Goesslin and Professor Julie Carrier from the University of Montreal, Canada. The dataset was previously used in another study published by their team (Gagnon et al., 2018) for studying cognitive tests in obstructive sleep apnea (OSA) patients. Each participant recruited first completed the sleep questionnaires and then had an overnight laboratory-based PSG recording. The neuropsychological assessment and cognitive tests were administered the next morning. The diagnosis criterion of MCI can be found in the paper and was based on Petersen and Morris (2005): an objective cognitive impairment (score ≤ 1.5 st.d.) in one or more cognitive domains (see previous) that could not be explained by current medications but preserved the capability of life independence. The cognitive impairments were grouped into aMCI or naMCI and single or multiple-domain impairment. In this thesis, I included the amnestic or non-amnestic MCI labelling (aMCI or naMCI), regardless of the number of domains affected. The sleep stages for this dataset were also scored using the AASM criterion with 30s-length epochs.

The dataset included full PSG sleep data from 128 participants. The PSG data contains 17-

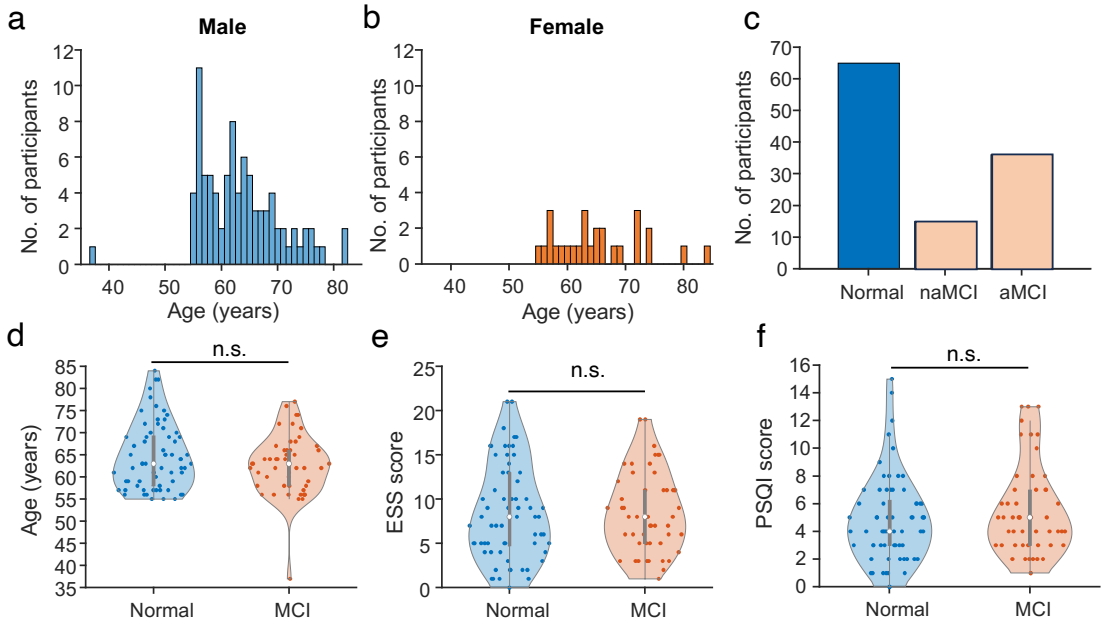


Figure 2.3: Demographics detail for Montreal dataset. The age distribution histogram of the Montreal dataset grouped by male (**a**) and female (**b**). **c**. The distribution of mild cognitive impairment (MCI) patients or normal healthy participants. naMCI: non-amnestic MCI; aMCI: amnestic MCI. **d**. Age distribution for Normal (blue) and MCI (red) participants; n.s. non significance using the Kruskal-Wallis distribution test. **e**. The Epworth sleepiness score (ESS) distribution for Normal (blue) and MCI (red) participants; n.s. non significance using a two-sample, two-sided t-test. **f**. The Pittsburgh Sleep Quality Index (PSQI) distribution for Normal (blue) and MCI (red) participants; n.s. non significance using a two-sample, two-sided t-test.

channel EEG recordings (F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T3, T4, T5, T6, Fz, Cz, Pz of the 10-20 system, all linked-ear referenced) sampled at 256 Hz, and ECG recording was sampled at 512 Hz. There were 7 participants with a different EEG system (montage and sampling rate) that were excluded. The demographics information (age, gender, neuropsychiatric diagnosis for MCI) was extracted, along with the sleep questionnaire scores available: Epworth sleepiness score (ESS) for measuring the level of subjective sleepiness (Johns, 1991) and Pittsburgh Sleep Quality Index (PSQI) for measuring subjective sleep qualities and even for sleep function screenings (Buysse et al., 1989; Mollayeva et al., 2016). The ESS questionnaire contains eight questions, with each one having a scale from 0 (no chance of dozing) to 3 (highly likely to doze off). The highest score is 24, implying a higher subjective sleepiness. The PSQI questionnaire measures the subjective sleep quality in seven sleep domains (sleep latency, sleep duration, habitual sleep efficiency, sleep disturbances, use of sleeping medication, and daytime dysfunction) over the past month. Each domain has a scale from 0-3, with a total score of 21, and higher scores imply a worse subjective sleep quality over the past month. Note again that the sleep questionnaires are done prior to the PSG recording.

Similar to the MESA dataset, the sleep scoring was checked for each participant to ensure that the participant started from stage 0 and had a valid sleep onset as defined. The EEG data is

then aligned to the sleep scoring (with unscored EEG epochs in the beginning removed). Then, participants with sleep onset latencies outside the range of 3-90 minutes were excluded. The raw sleep onset latency distribution can be seen in Appendix Figure B.1b, with mean ± st.d. 25.34 ± 23.90 minutes. After cleaning, there are $N = 116$ participants left for later analysis.

The demographic information and comparison between normal and MCI participants are shown in Figure 2.3. The age distribution for male and female are shown separately in Figure 2.3a,b, with $N = 89$ males and ages 63.31 ± 7.24 (mean ± st.d.) years, and $N = 27$ females and ages 65.72 ± 7.49 years. The MCI diagnosis result is shown in Figure 2.3c, indicating 65 Normal, 15 naMCI and 36 aMCI participants separately. Since the number for aMCI and naMCI is different from Normal participants, I later considered them as MCI patients ($N = 51$ in total) and compared them against the Normal participants.

The age distribution between Normal and MCI participants is shown in Figure 2.3d, where normal participants have ages 64.69 ± 7.57 (mean ± st.d.) years, and MCI patients 62.69 ± 6.82 years. The Kruskal-Wallis distribution test between them shows no significance with a p-value of 0.35. Similarly, the sleep questionnaire scores (ESS and PSQI) are compared in Figure 2.3e,f. The ESS scores for Normal and MCI participants are 8.45 ± 5.40 (mean ± st.d.) and 8.57 ± 4.49, respectively, with two-sample, two-sided t-test between them showing no significance (p-value 0.90). The ESS maximum ESS score is 21, but the mean score is within the range of normal daytime sleepiness (Doneh, 2015). The PSQI scores for Normal and MCI participants are 4.83 ± 3.13 and 5.49 ± 3.27, respectively, with a two-sample, two-sided t-test showing no significance (p-value 0.27). There is an age outlier in the MCI cohort (with an age of 37 years old), who was therefore excluded from all analysis later. After all these exclusions, there are in total $N = 115$ participants for all later analyses, with 65 Normal and 50 MCI participants. Overall, the Montreal dataset has a similar age distribution as the MESA dataset but with participants who suffer from cognitive impairments. In addition, it includes more EEG channels supporting richer spatial analysis of brain dynamics.

OSA (along with other sleep breathing disorders) is an important risk factor for cognitive decline and further dementia (Leng et al., 2017; Osorio et al., 2015). The important aspect could be that OSA patients normally have excessive daytime sleepiness (Lal et al., 2021; Seneviratne and Puvanendran, 2004), which could affect the sleep-onset latency (Fong et al., 2005). Given the original purpose of this dataset was to detect cognitive impairment with or without OSA, therefore both Normal and MCI cohorts have patients with OSA symptoms (Gagnon et al., 2018). Moreover, the ESS score distributions between these two cohorts are also not significantly different (Figure 2.3e), and thus were not further considered a confounding factor in the analysis.

2.2. EEG pre-processing

2.2.1. Definition of sleep-onset

In this thesis, sleep onset was conservatively defined as the first 1-minute of stage N2 sleep (which means 2 continuous sleep scoring of N2 with epoch size 30s). This is a conservative definition, given that we want to ensure that the participant is definitely asleep at our definition point. Stage N2 sleep has been proposed by many previous literatures as a safe boundary between wakefulness and sleep (De Gennaro et al., 2001b,a), where the appearance of spindles and K-complexes most safely imply definite sleep (Ogilvie, 2001). While undeniably the transition from wakefulness to sleep stage N1 involves some level of EEG changes and specifically physiological changes (Ogilvie and Wilkinson, 1984), stage N1 sleep is more likely to be a transient state during the falling asleep transition, with heterogeneous features of both wakefulness and definite sleep states (Rué-Queralt et al., 2021; Stevner et al., 2019). Stage N1 sleep has also been demonstrated to have the lowest inter-rater consensus (Somaskandhan et al., 2022).

The falling asleep period was then defined as the start of the recording (called bedtime) and up to the point of sleep onset. Similarly, the sleep onset latency was defined as the time between the recording start (here designated as 'bedtime') and the sleep-onset time. The sleep-onset latency was defined as the time from bedtime to the sleep-onset. Given that the sleep scoring was done at 30s-length epochs, the time resolution for the sleep onset latency was 30s. Notice that this definition of the sleep-onset period, as well as latency, is different to some of the previous literature (see Section 1.3.1 for detailed literature discussions), as in some cases the 'starting point' is defined as the onset of alpha trains in EEG.

2.2.2. Preprocessing and extracting sleep-onset EEG

The entire pre-processing pipeline can be viewed in Figure 2.4. Both datasets have first been checked through to ensure that all participants remained have complete data needed (demographics, overnight PSG recordings, sleep annotations).

The data reading and pre-processing were then done in MATLAB (MathWorks Inc.). As mentioned before in the previous section, the sleep scorings for each participant were checked to ensure a valid falling asleep period (from bedtime to sleep onset) exists. The participants with a sleep onset latency outside the range of [3,90] minutes were excluded, as mentioned. Participants with shorter sleep latencies have risks of narcolepsy or missing recordings before sleep onset, while participants with longer sleep latencies have risks of lacking stage N2 or significant fragmented sleep. Of course, it would be interesting to explore these participants on their own, but it was outside the scope of this thesis. The sleep onset latency distribution for both datasets

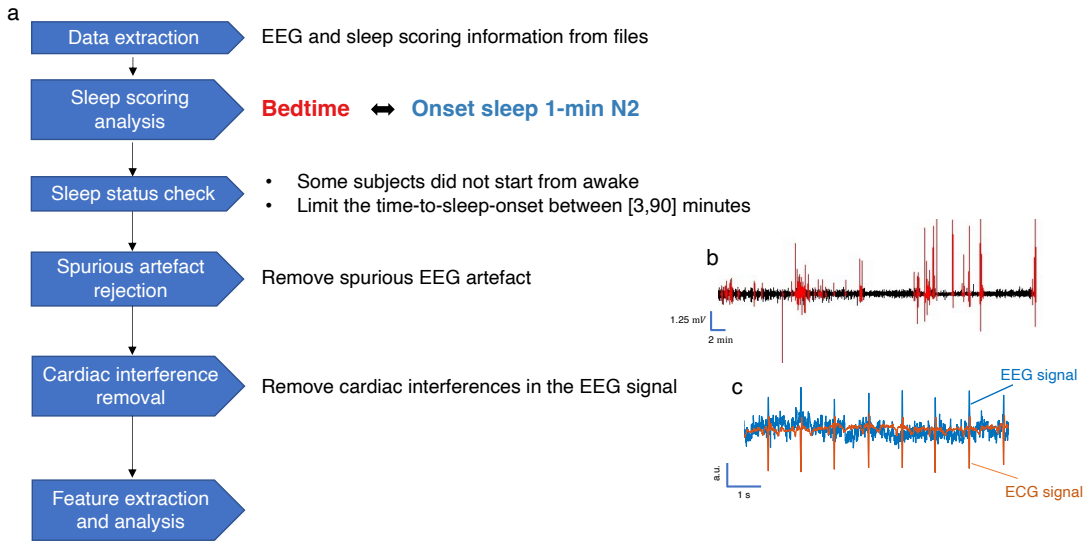


Figure 2.4: Overview of the pre-processing pipeline. **a**. The pre-processing pipeline is shown in a hierarchical order. After extracting the data and necessary information, the sleep scoring is analysed to extract the sleep onset as defined. The sleep-onset latency is then analysed, and participants within the latency range [3,90] minutes are kept. Then, two-step artefact rejections are done: the spurious EEG artefacts rejection and cardiac interference removal. The EEG signals during the sleep-onset period are then ready for analysis. **b**. Exemplary EEG signal, with red parts indicating artefacts. **c**. Exemplary EEG signal (blue) aligned with ECG signal (red), where the cardiac interference can be visualised. The ECG signal is re-scaled to show along with EEG data, thus the amplitude is an arbitrary unit (a.u.).

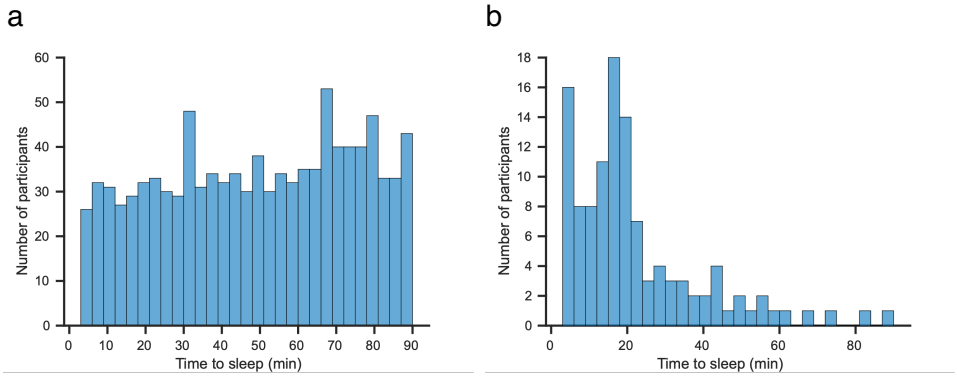


Figure 2.5: The sleep onset latency distribution for MESA (a) and Montreal (b) datasets. The distribution is shown after excluding participants outside the 3-90-minute range.

after cleaning is shown in Figure 2.5. For MESA, the remaining participants have a latency of 48.83 ± 25.01 minutes, and for Montreal participants, 21.88 ± 17.35 minutes.

After all the cleaning and exclusion (see also the previous section), there were $N = 1011$ participants left for the MESA dataset and $N = 115$ left for the Montreal dataset. After this point, there will be no more systematic exclusion of participants; the only potential reason for further exclusion will be due to EEG artefacts and analytic needs, which will be explained where applied. Before running feature-based analysis on EEG data, the EEG will be further cleaned and

explained below.

EEG movement artefact rejections

The typical EEG amplitudes are normally within the range of a few tens (e.g., 10-20) microvolts (μV). It can be corrupted by different sources of artefacts including eye blinks (especially in frontal regions), muscle movements (body movements or even just chin muscles), electrical noise (from the power sources of the recording devices), and even cardiac interferences (Urigüen and Garcia-Zapirain, 2015). The electrical noise from the power sources (50 Hz in the UK or 60 Hz in the US) is normally suppressed while recording the EEG via a notch filter. In the sleep EEG data, the blinking artefacts are minimal since the participants closed their eyes during sleep. Hence, in this step, I aim to simply remove EEG epochs with spurious movement artefacts (for example, arbitrarily large EEG amplitudes) to ensure general data quality. The cardiac inferences will be discussed in the next part.

There would normally be many spurious large artefacts, as expected in the overnight recordings. Firstly, using the EEG signal quality index provided (in the MESA dataset only), all EEG channels with signal quality smaller than 4 were excluded from the analysis. To further exclude the spurious EEG artefacts, I computed the root mean square (RMS) values (which are useful metrics for signal amplitude variability) for each EEG channel, across epochs (30s length, same as sleep scoring) over the entire EEG recording, using the following formula:

$$RMS = \sqrt{\frac{1}{N}\sum_{n=1}^{N}|x_n|^2}, \tag{2.1}$$

where N is the total number of samples in each EEG epoch, and x_n is the amplitude of the EEG samples. The median RMS values of all epochs were then calculated, and epochs with RMS values above 2.5 units of the median of all epochs were marked as artifactual and excluded from analysis. This process was repeated twice to ensure sufficient cleaning of the spurious peaks in the EEG signal. An example outcome after cleaning can be found in Figure 2.4b (on the right side of 'Spurious artefact rejection'), where the red part implies the artefact detected and rejected.

The RMS-based rejection was applied for all the EEG channels (the Montreal dataset has 17 channels, but not all were used later). Artifactual epochs were marked as 'Not a number' (NaN). To ensure consistency in the time domain between channels, if an epoch was identified as an artefact in one of the channels, then the epoch was set to NaN in all the channels.

The RMS-based method for artefact rejection was used throughout this thesis in different scenarios (different epoch lengths, artefact thresholds, and different EEG periods). The exact

parameters are specified where applied. There are other approaches for EEG artefact rejection or cleaning methodologies (see reviews Urigüen and Garcia-Zapirain (2015); Jiang et al. (2019).

EEG cardiac interference removals

When visually observing the EEG signals in the MESA dataset, I found significant cardiac interferences, which might confound our findings about brain dynamics (see Figure 2.4c on the right side of 'Cardiac interference removal'. To remove cardiac interferences, I applied the same method as described in (Purcell et al., 2017). Using existing ECG data, the R-peaks are extracted using the Pan-Tompkins method (Pan and Tompkins, 1985). For each EEG channel, for the entire period, I extracted 2s EEG signal aligned to each R-peak detected, and then averaged over all 2s EEG segments to get a characteristic EEG cardiac signature per channel per participant (similar to the concept of cardiac event-related potential). Then, this EEG cardiac signature was aligned back to each R peak and subtracted from EEG data per channel. More details and evaluation of this method can be found in the paper (Purcell et al., 2017). An example EEG cardiac signature trace is shown in Appendix B.2. The Montreal' s dataset did not have any obvious cardiac traces. Therefore, this method was only applied in the pre-processing of the MESA dataset.

2.3. Comprehensive EEG time series feature base for quantitative analysis of the falling asleep process

2.3.1. Feature-based EEG analysise

Time series is a type of repeated measurement of the same quantity over time used across many different fields (finance, biology, etc.) (Fulcher and Jones, 2017). Different time series data from different subjects will exhibit different properties, such as being stationary or nonstationary. The dynamics of a time series can provide important insights into the underlying characteristics of the measured quantity. Time series analysis itself is a very big topic, with different time series categories, different analytic approaches (based on its category or properties) and different analytic aims (prediction or pattern searching) (Fulcher et al., 2013; Hamilton, 2020). For example, financial or economic time series data (such as stock prices) are typically non-stationary and have a great repertoire of mathematical models for causal inferences and predictions (Geweke, 1984; De Gooijer and Hyndman, 2006). EEG is a specific type of time series data which measures the underlying cortical neuronal activity over time. Many time series analysis approaches can be applied to EEG signals (Gersch and Yonemoto, 1977) and gain mechanistic insights into

neuronal mechanisms and disease diagnosis such as epilepsy (Lehnertz, 1999) and Parkinson's disease (Janjarasjitt et al., 2008).

Here I aim to develop a feature base to study the EEG time series during the falling asleep transition. The most typical and widely used approach is the spectrum-based analysis. EEG power spectrum changes at sleep onset have been extensively explored in previous literature (Gorgoni et al., 2020; Ogilvie, 2001). It has been well established that the EEG spectrum shifts to lower rhythms during falling asleep, with increases in delta and theta band power and decreases in alpha and beta band power. The theta to alpha band power ratio, for example, has often been extensively used as an index of drowsiness (Jap et al., 2009; Lal and Craig, 2001) and related to cognitive functions (Zawiślak-Fornagiel et al., 2023). Therefore, the basic features included are these power band features and their ratios. Their oscillation peaks are also considered here, which show the actual oscillation rhythm and change at sleep onset (Achermann et al., 2019; Gorgoni et al., 2020). Another important EEG analytic approach is the cross-frequency coupling (Canolty and Knight, 2010). It has been applied in feature-based sleep stages classification and proved to provide extra value and information (Dimitriadis et al., 2021; Sanders et al., 2014), but its specific dynamics and mechanistic insights on the falling asleep are barely studied, and therefore added as part of the feature base. I also added a temporal coherence measurement, which can capture the internal neuronal oscillator pattern, as in our earlier study in essential tremor patients (Schreglmann et al., 2021). This is different from other coherence metrics between different brain areas, which quantifies the synchronisations (Leocani and Comi, 1999), as this is location-agnostic. Refer to Appendix A.1 for an overview of all the features evaluated in different frequency bands (in total, 25 features).

All of the previous analytic approaches are Fourier-based. As introduced before, EEG is a time series data, and there is a massive repertoire of analytic features outside the normal spectrum analysis for it, which could provide interesting unknown insights into neural dynamics. Nonlinear dynamical metrics (Lyapunov exponent, scaling exponents, etc.), for instance, have provided great mechanistic insights into sleep regulations (Acharya U. et al., 2005; Stam, 2005). It would be impossible to enlist all available time-series analytic approaches, but comprehensive time series analysis tools have been collected and aggregated in previous work called highly comparative time series analysis (HCTSA) toolbox (Fulcher et al., 2013; Fulcher and Jones, 2017). Although powerful, this collection has over 7000 features, which are overly computationally costly. Therefore, I resorted to a reduced version of it, which contains only the 22 most representative ones with retained distinguishability, called canonical time series characteristics (CATCH-22, see Lubba et al. (2019)). Refer to Appendix A.2 for all the CATCH-22 features.

In total, the feature base included 47 features. It was used to systematically identify the key EEG signatures that distinguish sleep onset from wakefulness and the transition dynamics for tracking the falling asleep process.

2.3.2. Feature-base computation

There are many important aspects to consider in computing the EEG time series features. For instance, the sampling rate and data length could significantly impact the outcomes (especially the Fourier-based features). The CATCH-22 features are generally more robust to these parameters but keeping them consistent is important for comparisons. The dataset itself limits the sampling rate, but the choice of EEG epoch lengths for evaluating features is crucial for the outcomes and interpretation of results. The EEG epoch length cannot be too short because the power spectrum estimation would be distorted in the low-frequency range, and the frequency resolution would be limited. On the other hand, if the epoch is too long, the assumption of stationarity may not be valid (Kaplan et al., 2005) and the computational cost could increase dramatically. In this thesis, the EEG features were computed in 6s epochs (unless otherwise specified), ensuring a 'quasi-stationary' assumption and sufficient power spectrum resolution.

The feature evaluation was done in MATLAB. The EEG signals of the sleep EEG period were preprocessed as described in Section 2.2, and then divided into 6s epochs and features were computed as described below.

Power band features

The EEG signal was first bandpass filtered between [0.1,32] Hz range. Since the 1/f spectral noise (a power-law decay of EEG power with frequency) could confound band power calculations (especially for low-frequency ranges), therefore, I applied a differentiation filter to remove the 1/f frequency component as in (Demanuele et al., 2007). The technical details and quantification of the filter can be referred to in the paper, and it was shown to be great for distinguishing low-frequency components. Many other sleep toolboxes use this method, such as Luna. In this thesis, the differentiation filter was done using MATLAB function `diff`, and one exemplary power-spectrum result is shown in Appendix B.3. There are more advanced methods for extracting this 1/f component (such as Donoghue et al. (2020)), and this was not used here to avoid over-complicating the pipeline and interpretations. The power spectral density (PSD) was then estimated using Thomson' s multitaper approach (MATLAB function `pmtm`), with 4 tapers and the number of fast Fourier transform points the same as the sampling frequency (i.e., 256). The PSD was then normalised to the total spectrum power so that the band power is a ratio compared to the entire spectrum. The band powers of frequency bands were computed by summing the normalised PSD of all the frequency bins within that band. The band powers were then log-transformed. The band power ratios were computed by the quotient of two normalised band powers and then log-transformed the ratio. The peak frequency in each frequency band was defined as the frequency bin with maximum normalised power. Four frequency bands are used in the analysis: delta band 1 to 4 Hz, theta band 4 to 8 Hz, alpha band 8 to 12 Hz and

beta band 12 to 30 Hz.

Cross-frequency coupling features

Cross-frequency coupling is normally measured as phase-amplitude coupling (PAC), which measures the coupling strength between the phases of lower frequency bands and the amplitudes of higher frequency bands. A differentiation filter was applied to the input EEG epoch before further evaluation for removing the 1/f component (same as before). The PAC was then calculated by following methods in (Samiee and Baillet, 2017). Briefly, for the input EEG epoch, we designed a zero-phase, even-order finite impulse response filter to filter the signal into the four respective bands (same as before). I then applied the Hilbert transform to calculate the phase and amplitude for each band. Phase cycles were then extracted (from $-\pi$ to π is a cycle), and for each cycle, the modulation phase and amplitude are put into a complex signal:

$$z(t) = A_{f1}(t) \times e^{\varphi_{f2}(t)}, \tag{2.2}$$

where A_{f1} is the amplitude from higher frequency bands, and φ_{f2} is the phase from lower frequency bands. The PAC was then quantified by summing $z(t)$ over all phase cycles and then normalising to the average amplitude A_{f1} in each cycle. The result PAC value is between 0 and 1, where closer to 1 implies stronger coupling. The PAC is thus calculated for each frequency band pair (low to high).

Temporal coherence features

The differentiation filter was still applied first to each EEG epoch data to ensure the same spectrum property. The temporal coherence vs frequency for each EEG epoch was quantified using the magnitude squared coherence algorithm across 1s segments using MATLAB' s `mscohere` function. Briefly, the EEG epoch is divided into several 1s segments, where the magnitude squared coherence between each pair of them is calculated, and then the coherence spectrums are averaged. For each frequency band, the feature value is calculated as the sum of the values of all the frequency bins within that band. The temporal coherence features are evaluated for those four frequency bands specified before.

CATCH-22 features

Catch-22 (Lubba et al., 2019) is an openly available toolbox (`https://github.com/DynamicsAndNeuralSystems/catch22`) implemented in both MATLAB and Python. The features include both linear and nonlinear metrics and many other time series analysis methods. The

descriptions of all the features are given in the original paper. In this thesis, I only interpret the key features relevant to sleep onset. In each EEG epoch, I first applied the differentiation filter the same as before to ensure the same time series property. Then, the 22 features were calculated using the provided functions.

2.3.3. Limitations and extensions

The feature base is a comprehensive approach to capture the brain dynamic changes during the falling asleep process. There are many other sleep EEG features on slow wave oscillations or spindles, but they normally appear after sleep onset and thus were not included here. The EEG power band features were calculated up to the beta band (i.e., 30Hz), as typically sleep EEG does not have high-frequency gamma-band activity associated with high-level cognitive activity (Jensen et al., 2007). Gamma band activity might appear again during slow wave sleep (Valderrama et al., 2012), but slow wave sleep only occurs after sleep onset as well.

The feature base is, of course, not perfect. The features are evaluated on single-channel EEG signals, where the features would only represent regional dynamics. As previously described in Section 1.3, falling asleep happens both temporally and spatially across the brain areas. Therefore, later in Chapter 5, there will be further development of inter-regional interaction feature base to capture more of the spatial dynamics.

The CATCH-22 features are sometimes difficult to interpret in the aspect of brain dynamics, as many of those time series analytic approaches come from other disciplines. Besides, as a child of HCTSA, CATCH-22 preserves the general performance, but some specific features might be prevented from discovery. Many nonlinear features such as complexity, entropy or fractal dimensions have immense values in quantifying state transitions and brain dynamic structural changes (Stam, 2005); For example, in sleep, it has been demonstrated that the sleeping brain shows lower dimensional chaotic structures and these nonlinear measurements can be used in classifying different sleep stages (Acharya U. et al., 2005; Acharya et al., 2015). Mediano et al. also showed that complexity metrics provide important insights into cognition and consciousness (Mediano et al., 2023). These specific features, which are well-researched in sleep and neuroscientific contexts, could have stronger interpretability, but it is impossible to find matches in CATCH-22, although the latter claims to best capture and represent the nonlinear features in HCTSA.

In the feature computation, the 1/f component (aka., the arrhythmic component)was removed from the EEG signal using the differentiation filter. However, the 1/f component of EEG captures important aperiodic dynamics in the brain and may contain important information about underlying neuronal electrophysiology relating to information processing and cognition (Voytek et al., 2015; Gyurkovics et al., 2022). This scale-free component (power-law decay) has

also been demonstrated to be capable of distinguishing different sleep stages and anaesthetic states (Lendner et al., 2020; Colombo et al., 2019). Of course, there will be more such relevant features which are not fully reviewed here. In this thesis, only CATCH-22 is applied, but extensions to the feature base with the above metrics are done in a Master's project I supervised, where many metrics mentioned above were found to have a good predictive value of sleep-onset (Ilina, 2023).

Another limitation lies in that the parameters of those feature computation metrics are fixed into the pipeline, but surely, specific time series structures could have specific parameters that best capture their dynamics (especially in nonlinear EEG features, such as entropy or information Pritchard and Duke (1992); Micheloyannis et al. (1998)). This is not feasible here given that one wants interpretability across huge amounts of data and samples. Still, most of the features are robust enough to capture the statistical properties of that specific type of feature, and one could dive into that category of feature to explore different parameters for better interpretations and mechanistic insights if that feature is useful and informative.

2.4. Feature-based statistical learnings

Feature-based statistical learning is a powerful approach in machine learning that involves identifying and extracting relevant features from data to make accurate predictions or classifications. The extraction of key features can also lead to a better mechanistic understanding of the system studied (such as the brain in this case). One important aspect is the feature base used, which, as discussed above, should be sufficiently good to describe the properties of the system (resolved before). The other aspect is the statistical learning methods. As the name specified, statistical learning uses the statistical properties of the features (distributions, etc.) and applies different machine learning techniques to make inferences (about underlying systems or relationships) or make predictions when observing new data (James et al., 2021a).

Given a set of features, the task of statistical learning is to learn the best fit for:

$$Y = f(X) + \epsilon, \tag{2.3}$$

where Y is the target predictor, X is the feature (can be high dimensional), ϵ is the error term. The key aim of statistical learning is to estimate f, the so-called target function, to get the optimal performance for estimating Y using different machine learning techniques. The predictor Y can be categorical (for example, sleep or wake) or a continuous variable (the time to sleep-onset). Based on the predictor type, the estimation of f can generally be called classification (for categorical variables) or regression (for continuous variables). The estimation techniques can be parametric (with important assumptions about the data, such as normality) or non-

parametric. With a clear target variable Y, the learning is classified as supervised, where unsupervised learning can also be applied to learn a general pattern of the input data without a target predictor (for instance, clustering). The most important aspect of estimating f is to minimise the error ϵ (in mathematical terms, the optimisation problem). Optimisation is a big topic, and many different methodologies exist based on the estimation methods and predictor types but they are not discussed here.

Statistical learning is a very broad topic as well and reviewing all available methodologies is outside the scope of this thesis. One of the key tasks in this thesis was to classify wakefulness and sleep-onset states to identify the key distinguishing features. In this case, I typically used a linear support vector machine (SVM) (Cortes and Vapnik, 1995). I also used regression, for example, to predict sleep-onset latency. Essentially, as mentioned in the beginning, all these techniques are applied to make predictions, and then extract important features and interpret them for a mechanistic understanding of the falling asleep brain dynamics.

One important problem which arises only very recently is the dimensionality reduction of the features X. With increasing numbers of analytic methods, the number of features also increases dramatically. In fact, the process from HCTSA to CATCH-22 is already a typical process of dimensionality reduction, which minimises the redundancy (removing statistically similar features) but preserves the performance. For the feature base developed before (47 in total), I also applied a key feature selection procedure using hierarchical clustering based on feature similarities (Liu et al., 2011). The same methodology has been applied in tremor time series features to identify predictive features of stimulation response for mechanistic insights (Schreglmann et al., 2021). The exact algorithmic and implementation details of these methods will be introduced in Chapter 3.

2.5. Dynamical systems theory and the brain

2.5.1. The basic dynamical system theory

Dynamical systems theory is a branch of mathematics that studies the behaviour of systems that change over time, which provides an important framework for understanding the interaction of complex systems. A dynamical system is a system where its system state changes with time following a specific rule that can be described as a function (Ott, 2002). A very basic simple example is:

$$\frac{dx}{dt} = f(x), \tag{2.4}$$

where x is the system state, and f is the function that governs the rule of how the system state evolves over time. The example equation in (2.4) is a first-order (highest derivative order),

ordinary (derivative only to one variable), dynamical system. The variable x can be high dimensional, which means for each state variable, there will be a differential equation governing its dynamics. Based on the property of function f, the system can be linear (such as $f(x) = x$) or nonlinear (containing higher orders, for example, x^2). Aside from the linearity or nonlinearity of the system, another important property of a dynamical system is the perseverance of the energy (Stam, 2005). Dissipative systems will lose or exchange energy during the dynamic process, while conservative systems do not. Almost all real-world systems will be dissipative, as there are numerous energy exchanges in almost all processes. Finally, the system can be deterministic or stochastic (if f is a noise term unrelated to x). There are many more complicated systems with partial differential equations and higher-order derivatives, which are not discussed here.

As the equation shows, the system dynamics over time are governed by the function f. However, the differential equation is very often non-solvable via integration. Therefore, the system's dynamics are normally quantified via the state space phase plane. In a multi-dimensional system, the system state can be represented as a point in the multi-dimensional space at each time point, and the evolvement of the state space over time is the trajectory of the system dynamics that can imply important system properties. An important aspect is that when one observes the system for long enough, for most dissipative systems, there will be a 'regular pattern' in the trajectory dynamics called an attractor to which the system converges. For example, if a normal pendulum is released from a point in a 2-D plane, after a sufficiently long time, the pendulum will stop upright in the middle (a stable point attractor). In a complex nonlinear system, more complicated attractors will appear, such as limited cycles or, more interestingly, chaotic behaviour (see examples in (Feldman, 2019)). The attractor is the fundamental property of a dynamical system, and there is rigorous mathematics to quantify the different properties of an attractor (Lyapunov exponent, fractal dimensions, etc.), and the dynamical systems are classed mostly based on the attractor properties.

The attractor can be called a stable state, which means once the system dynamics get into this attractor, it will not get out again unless it is disturbed by external forces. A system can have multiple attractors at the same time, and an attractor can appear or disappear when different system parameters are taken. These parameters are often called control parameters, and when changing a control parameter causes significant changes in the system's attractors (such as the disappearance of one attractor), a bifurcation (also called phase transition) happens. Bifurcation can be observed in many real-world systems. The Hodgkin-Huxley neuronal models will undergo Hopf bifurcation (into a limited cycle) when the input stimuli (as a control parameter) becomes large enough (Wang et al., 2007a), and in ecological systems, when the harvesting rate (the control parameter) becomes too high, the species will go extinct (catastrophic bifurcation) (May, 1977).

The system dynamics can be fairly easily analysed with an explicit differential equation such

as (2.4). However, for most complex real-world systems (such as the brain), deriving such equations is impossible. Still, one can observe the temporal dynamics of a system (a time series), such as EEG measurements. Then, from a top-down approach, one can infer the system or attractor properties from the time series data. There is a plethora of analytic approaches for this called nonlinear time-series analysis, such as time-delayed embedding to re-construct the attractor shapes or Lyapunov exponents to estimate the chaotic properties (Parlitz, 1998). These approaches can be applied to EEG data to infer nonlinear properties of the underlying brain networks (Stam, 2005; Pritchard and Duke, 1992, 1995), for example, those nonlinear features mentioned previously 2.3.

2.5.2. The brain as a dynamical system

What type of dynamical system is the brain? Given that the brain is a biological system, and given its complexity level, it should be a nonlinear dissipative system. Indeed, the brain can be considered a complex, nonlinear, high-dimensional dynamical system (Hancock et al., 2022; Stam, 2005). It contains rich temporal dynamics and intricately designed connections inside to achieve different functionalities. From the level of individual neuron spikes to a nucleus (collections of a bunch of neurons), and to the system-level oscillations, each level of resolution has its own dynamics (called a functional source) and can be considered a dynamical system. The functional sources are interconnected with each other (functional connectivity). This connection can happen at different levels, for instance, the inhibitory and excitatory connections between different neurons or nuclei. The system dynamics observed from EEG, for example, are thus the complex dynamics from a functionally coupled network of underlying dynamical systems in the regional cortex, where the cortical dynamics are further driven by other inter-connected cortical areas and deeper brain structures. The proposed nonlinear dynamical system framework of the brain by Stam (Stam, 2005) well demonstrated that the brain is a network of well-structured, weakly interconnected dynamical systems that form a self-organised criticality. This critical interconnected system can be adjusted by changing the connectivity (such as inhibitory or excitatory changes) to achieve quick adjustments to the brain to get into different brain states (e.g., cognitive tasks, sleep) and, finally, to affect the behavioural outputs.

Indeed, the nonlinear patterns from the EEG time series dynamics can imply different brain states and even identify diseases (Stam, 2005). One well-studied example is the epilepsy. When a seizure happens, the EEG is high-amplitude and highly synchronised, which implies a disruption of the underlying neuronal network (and thus the behavioural states) (Lehnertz, 2008). The EEG pattern will display a low-dimensional structure, implying a reduced 'dynamical complexity' and disrupted capability of the brain to do normal functions (Stam, 2005). The sleeping brain also displays a reduced dimension and complexity, where the brain rhythms slow down. In Alzheimer's disease, the weak nonlinear connectivity structure was disrupted, leading to more

noise-like brain dynamics, with reduced complexity structures (Jeong, 2004). These nonlinear metrics have also been proven to have values in disease diagnosis (Pritchard et al., 1991; Stam et al., 1996), and correlate to cognitive functions and behavioural changes that progress with diseases (Stam, 2014).

The important property of the brain organised as a critical dynamical system is the quick phase transitions (bifurcations) it can make. For instance, the transition of seizure onset can be considered as a bifurcation (Lopes da Silva et al., 2003). Wake-sleep transitions are clearly another type of critical transition that can happen quickly inside the brain. The critical transition or bifurcation, as mentioned before, is the disappearance of one state attractor and the appearance of another in the system dynamics. By conceptualizing these brain state transitions as critical transitions, there are many fundamental properties of the system dynamics that can be used to predict the tipping point of the transition, such as critical slowing down (Scheffer et al., 2009, 2012). Predicting such transitions in the brain could help pre-warning (such as seizures) or develop closed-loop systems (for seizure prevention or sleep tracking). The critical slowing down has been proven to be of tremendous value in seizure forecast (Maturana et al., 2020). This thesis examines whether the falling asleep process is a critical transition, in Chapter 4.

2.6. Closed-loop auditory stimulation paradigms

Auditory stimulation is an easy-to-apply, non-invasive stimulation technique which is widely applied in sleep neuromodulation (Capezuti et al., 2022). Acoustic stimulation synchronised to the up-phase of slow-wave activities during deep non-rapid-eye-movement (NREM) sleep has achieved extensive success in boosting slow-wave sleep and enhancing cognitive performances (especially memory) after (Bellesi et al., 2014; Ngo et al., 2013) (see review in Section 1.1). The targetted phase is important, as it has been demonstrated that down-phase stimulation could, in turn, disrupt slow-wave sleep, and even interfere with the memory consolidation process (Ngo et al., 2013; Fattinger et al., 2017). The EEG oscillations (phases and amplitudes) come from the underlying neuronal firing patterns and excitatory or inhibitory relationships (Klimesch et al., 2007; Hebron et al., 2023), therefore, targeting different phases of the activities could result in different effects on the cortical oscillations (entrainment or disruption) (Sousouri et al., 2022). The optimal timing to deliver the stimulation to achieve different aims could be different, and could also change with age (Navarrete et al., 2020), emphasizing the need to personalise the target phase. Therefore, to achieve this personalised closed-loop auditory stimulation, there needs to be a closed-loop system to track the phases of EEG oscillations and deliver audio outputs in real time.

The phase-locked stimulation paradigm is shown in Figure 2.6. To compute the instantaneous phase of the neural oscillation in real-time accurately, one has to overcome Gibb' s phenomenon,

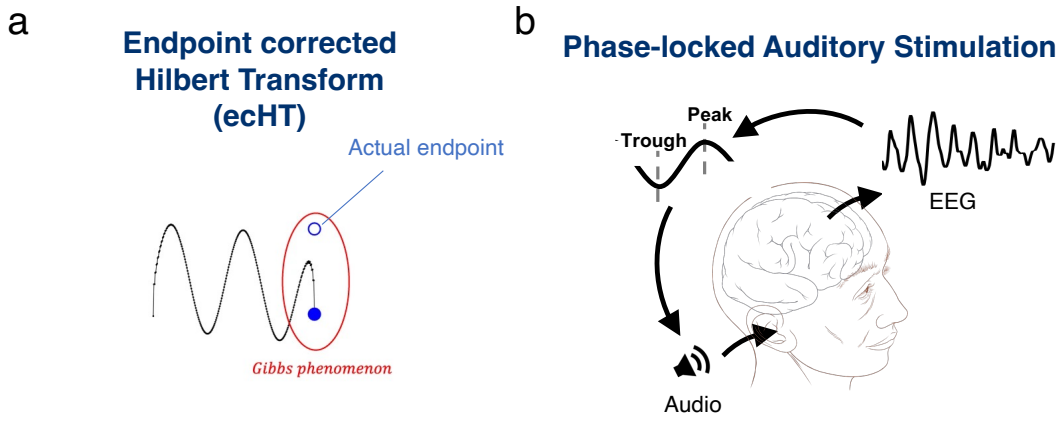


Figure 2.6: The phase-locking auditory stimulation paradigm. a. Concept of end-point corrected Hilbert transform (ecHT) algorithm, which corrects the Gibbs phenomenon and estimates real-time phases correctly. The solid blue dot is the wrong endpoint due to the Gibbs phenomenon, where the non-filled blue point shows the end-point corrected (actual) phase position. Figure adapted from Schreglmann et al. (2021) under Creative Commons Attribution 4.0 International License. **b**. Concept of phase-locking auditory stimulation. EEG is recorded with its phases extracted in real-time using ecHT algorithm. The auditory stimulation is then delivered in synchronisation to the desired phase (peak or trough shown in the plot).

which distorts the ends of the Hilbert transform. The distortion originates in the discrete Fourier transform (part of the Hilbert transform algorithm) of a finite signal that its ends are not continuous and differentiable (Hewitt and Hewitt, 1979). To mitigate this problem, our group developed the end-point corrected Hilbert transform (ecHT) algorithm (Schreglmann et al., 2021) (Figure 2.6a), where a causal band-pass filter is applied to remove the extra distortion and ensure correct phase estimation at the end-point. The phase-locking accuracy was proven in the original paper. In the actual hardware set-up, a 2nd order Butterworth filter is applied.

The closed-loop acoustic stimulation was set up as in Figure 2.6b. The ecHT algorithm estimates the instantaneous phase at the target stimulation frequency in real time. Then, via the hardware control, the acoustic output is delivered at the target phase (for example peak or trough in the figure). The audio output (sound types, volumes) can also be controlled in the device to deliver desired outputs. Of course, there would be hardware limitations in the target phase and estimation, for example, the firmware iteration time will limit the phase resolution. In Chapter 6, this paradigm will be applied to deliver phase-locked auditory stimulation targeting low-frequency oscillation, in an attempt to augment the sleep onset, where exact implementation details will be further introduced.

3. The distinctive brain EEG activities at sleep onset

3.1. Introduction

The brain EEG dynamics during falling asleep have been studied before showing global slowing down of brain rhythms associated with an increase of power in slow frequency ranges and a decrease of power in higher frequency bands (see review in Section 1.3). The EEG of sleep onset has also been associated with changes in nonlinear features. For example, Acharya et al. showed that sleep onset is linked to changes in the Lyapunov exponent and correlation dimensions (Acharya et al., 2015). However, the generalisation of these early results has been limited by using specific features (or a feature type) and using small datasets. Thus, the most informative EEG features that differentiate sleep onset from wakefulness remain unknown.

The richness of the datasets also allows us to ask further questions that had been rarely explored previously. Demographic factors such as age and gender could significantly impact sleep, but they are barely studied for the falling asleep process. The ageing population normally has more difficulty initialising sleep (Li et al., 2022), but the exact EEG dynamics changes have not been studied. Most studies focused on the deep sleep process, specifically the reduction of slow wave activity and disrupted homeostatic regulations (Campos-Beltrán and Marshall, 2021; Landolt et al., 1996; Carrier et al., 2001; Landolt and Borbély, 2001), which are then related to disrupted cognitive functions (Mander et al., 2013, 2014), but the changes in sleep onset process are unclear. The study by Gorgoni (Gorgoni et al., 2021) provided the very first analysis of the sleep onset dynamics of older adults, which showed that older people have reduced delta band activity and a higher arousal level throughout (delta-to-beta ratio), which relates to a longer sleep onset latency and more sleep fragmentations. Still, the analysis was restricted by the methodology (EEG power spectrum only), the sample sizes, and poorly controlled conditions between young and old adults (see the paper). The question is equivalent to gender, which has not even been well-studied. Aside from these two, the Montreal dataset provides the potential to study how cognitive impairment impacts the EEG features at bedtime and sleep onset. As MCI patients suffer more from the difficulty of sleep initiation, the analysis can surely provide

better insights into it. The strongest point is that by using feature-based statistical learning, one can get a more overall understanding of the key EEG dynamic differences in these questions, while most previous studies only took EEG power spectrum analysis.

An exploratory question is regarding the sleep onset latency. The sleep onset latency can be affected by many different factors, for instance, the drowsiness state of the participant or specific diseases (obstructive sleep apnea, insomnia, narcolepsy, etc.). It would be very interesting to explore if there would be any EEG signatures that change with the actual latency from bedtime to sleep onset and even predict the latency based on EEG features at bedtime. Diaz et al. have demonstrated that resting state EEG contains biomarkers specifically in the theta oscillations that correlate with the sleep onset latency, although they defined the sleep onset as two consecutive stage N1 sleep (Diaz et al., 2016). Here again, the feature-based statistical learning will provide a more comprehensive insight into whether the latency can be predicted and what EEG dynamics are relevant to it.

In this chapter, I aim to address the knowledge gap of the most informative features by classifying the brain EEG at sleep onset using a comprehensive feature base, including all the common sleep EEG band features and more modern nonlinear features, and then clustering correlated features to discover an orthogonal set of features. I use a large open-source sleep EEG dataset (including more than a thousand subjects) and then validate the findings in an independent dataset received from a collaborating sleep lab in Montreal (the largest sleep clinic in Canada). I also examine whether the brain activities at sleep onset (and bedtime) are affected by the age or gender of the participants and whether there are differences associated with MCI. Finally, I explore the relationship between brain EEG activities at bedtime and sleep onset latency. By discovering the key EEG features differentiating the sleep onset state from wakefulness, I aspire to provide functional biomarkers for the future development of closed-loop neuromodulation (the overarching goal of the thesis). Given the specific interest in preventing and slowing sleep-dependent dementia drivers, I focus the investigation on elderly people and people with MCI.

3.2. Methods

3.2.1. Datasets

The investigation was performed using two datasets, i.e., the MESA dataset ($N = 1011$ subjects) and the Montreal dataset ($N = 115$) described in Section 2.1. The participants' EEG signals were extracted and pre-processed as described in Section 2.2. In the case of the MESA dataset, the three EEG channels, i.e., Fz-Cz, Cz-Oz and C4-M1, were extracted and in the case of the Montreal dataset, the three channels Fz, Cz and Pz (linked-ear referenced) were extracted.

In the MESA dataset, EEG channels with a quality index smaller than 4 were excluded (an EEG channel quality index was provided in the dataset). Thus, subsets of participants had only two channels or one channel. Yet, since the analysis examines global changes, i.e., the feature values across the EEG channels were considered equally, and not expected to confound the results. In total, the Fz-Cz, Cz-Oz, and C4-M1 channels included recordings from 900, 870, and 870 subjects, respectively. In the Montreal dataset, no quality index was provided, and hence, all three channels were initially extracted.

3.2.2. Feature extractions for bedtime and falling asleep period

The EEG was divided into 6s epochs. Epochs with artefacts were detected using the RMS criterion (30s epoch, 2.5 units of median threshold) and cardiac interference approach as described in Section 2.2. Epochs with artefacts were rejected by replacing the EEG values with NaNs.

Then, one minute (min) EEG segment at sleep onset (i.e., the first minute of continuous stage N2 sleep) and one minute EEG segment at bedtime (i.e., the first recording minute) were extracted. If the first epoch in the segment was a rejected epoch due to an artefact, then the extracted segment started at the next non-artefact. Next, four representative 60-second (s) artefact-free epochs were extracted from each 1 min EEG segment. The epochs were randomly sampled, and to further ensure data quality, all extracted epochs had RMS values below 3 units of the median RMS value of the entire 1-minute segment. Next, the values of the 47 features (see Section 2.3) were computed in each of the four representative 6 s epochs at bedtime and sleep onset and were labelled as 'bedtime' and 'sleep-onset' accordingly. Finally, the values of each feature were normalized to [0,1] using a Sigmoid function. The normalisation was done to each feature, across all EEG epochs of both bedtime and sleep onset from all participants.

3.2.3. Classification using linear SVM and leave-one-out cross-validation

Basic SVM theory

Support vector machine (SVM) is a widely used statistical learning method for supervised classification tasks (Cortes and Vapnik, 1995). As a statistical learning method (see Section 2.4), the aim of SVM is to learn an optimal hyperplane in the high-dimensional feature space that best separates different classes (labels). Taking a two-class classification problem as an example, for each sample, there would be a corresponding point in the feature space $\mathbf{x}$ (a vector of all the feature values of this sample) and a corresponding label $y = \pm 1$ indicating the class. SVM will then train a hyperplane $f(\mathbf{x}) = \mathbf{wx} - \mathbf{b}$ which satisfies $\mathbf{wx} - \mathbf{b} = 0$ trying to separate these two groups. To maximise the separability, the margin from the point to the hyperplane has to be maximised, where the margin can be calculated as $\frac{2}{\|\mathbf{w}\|}$. This is then equivalent to minimising

the $\|\mathbf{w}\|$, which is a convex optimisation problem that can be solved using many mathematical algorithms and will not be discussed here.

An exemplary 2-D feature space example is shown in Figure 3.1, where the hyperplane here is a line. For any sample point with feature vector $\mathbf{x}$, when $f(\mathbf{x}) \geq 1$, the classifier will classify it to be class $y = 1$, and vice versa. The points that help construct the optimal hyperplane will be called the support vectors, and that is where the name comes from. One obvious limitation of SVM is that it only works well when the data are linearly separable. This assumption is often not true, where kernel tricks can be applied to transform the current feature space into a different shape, which might become more linearly separable (Boser et al., 1992). Commonly used kernels include polynomial kernels or Gaussian radial basis function kernels. SVM can also be applied to multi-class problems, but essentially, it will be subdivided into several binary classification questions (Duan and Keerthi, 2005). This thesis only uses binary SVM and hence multi-class SVM is not further discussed.

Implementations

In this thesis, SVM was used in two-class classification problems with linear kernels. The aim was not to achieve 100% classification accuracy but instead to learn the fundamental statistical properties of the features that separate the two classes (i.e., bedtime wakefulness and sleep onset).

SVM was implemented in MATLAB using the function `fitcsvm`. To prevent overfitting, a leave-one-out cross-validation approach was taken (James et al., 2021b). The leave-one-out approach can also prevent the model from learning participant-specific information and thus lead to overfitting. Briefly, for a k-fold cross-validation, the participants were equally subdivided into k groups. Then, the leave-one-out approach took $(k-1)$ groups to train the SVM classifier and used the remaining group for testing validation. This was done k times, with each group validated once. The final accuracy was the average accuracy of all the k folds.

Throughout the thesis, I used $k = 10$ unless otherwise specified. The classification performance was evaluated using the balanced accuracy:

$$acc = \frac{precision + recall}{2}, \tag{3.1}$$

where $precision = TP/(TP + FP)$ and $recall = TP/(TP + FN)$, with TP as true positive, FP as false positive, and FN as false negative in the typical machine learning denotations. In most cases, the SVM classification used only 1 feature to evaluate its performance. In some cases, the features would be aggregated together to train the model, and the performance was evaluated using both the balanced accuracy above and the F-score, which is the harmonic mean

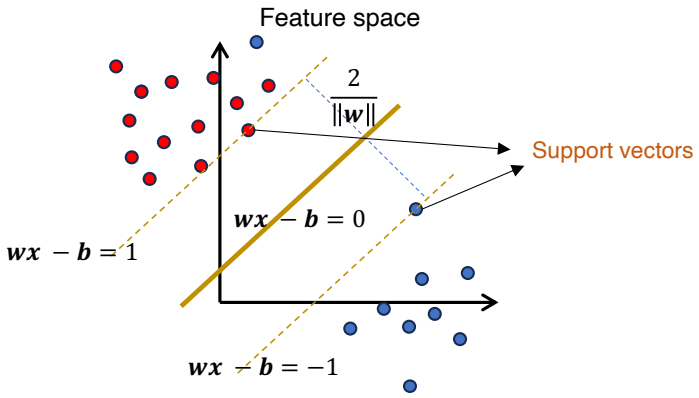


Figure 3.1: Illustration of SVM classification in a 2D space. Each point represents a sample in the feature space ($\mathbf{x}$, with blue or red indicating the class it belongs to. SVM aims to learn an optimal hyperplane $\mathbf{wx} - \mathbf{b} = 0$ (the solid yellow line in this 2D example), to separate these two groups best by maximising the margin from the support vectors (marked in the figure) to the hyperplane $\frac{2}{\|\mathbf{w}\|}$.

of precision and recall:

$$F\text{-}score = \frac{2 \times precision \times recall}{precision + recall}. \tag{3.2}$$

Both the accuracy and F-score have a maximum value of 1. The higher the values, the better the model performance.

3.2.4. Hierarchical clustering and feature selection

Dimensionality reduction in feature space is crucial for discovering the key governing features. One way to achieve this is hierarchical clustering (Park, 2013). Hierarchical clustering is a specific clustering method that will return the clusters in a multi-level hierarchy, where features from lower-level clusters will join higher-level clusters. This allows a strong degree of freedom to choose the clustering level that one would like to use. As in other clustering methods, the distance metric is a crucial element that affects the interpretation of the clustering results. To select key informative features and reduce feature space redundancy, the distance metric should be based on the similarity measurements of different features. Correlation measurements are simple and good to use in this case, of course, there are other more advanced methods, such as mutual information (Liu et al., 2011).

In this chapter, hierarchical clustering was done based on the Pearson correlation distance between the features:

$$d_{ij} = 1 - \rho_{ij}, \tag{3.3}$$

where d_{ij} is the Pearson correlation distance, and ρ_{ij} is the Pearson correlation coefficient between feature i and j. Essentially, it means that the more correlated two features are, the closer

they will be in clustering. Hierarchical clustering was done in MATLAB using function `linkage`, with the cluster distance metric as the average distance between all pairs of features in any two clusters. The clustering threshold was chosen to be 0.5 (unless otherwise specified), which means that the features in each cluster have a similarity (correlation coefficient) of $\rho_{ij} > 0.5$. To select the representative feature of each cluster, each cluster' s centroid (average of all features in the cluster) was computed, and the feature that was closest to this centroid was taken as the representative one for this cluster. This procedure was done only on the top 40 features (out of 47) with the highest SVM classification accuracy.

3.2.5. Statistical analysis

To test differences in the mean values of the normalized features between the two cohorts, a double-sided two-sample t-test was used. Although the t-test only works under the assumption of a normal (i.e., Gaussian) distribution, for large enough sample numbers (and normalised values), this is acceptable (King and Eckersley, 2019). To test the relationship between the features and continuous variables such as age, linear Pearson correlation was used, outputting a correlation coefficient (ranges from [-1,1]) and a p-value indicating significance. To test differences in the mean values of non-normalised variables such as age and sleep onset latency between two cohorts the non-parametric Kruskal-Wallis test was used. The p-value threshold was normally 0.05 for all these statistical tests unless otherwise specified or Bonferroni corrected for multiple comparisons (Sedgwick, 2012).

3.2.6. Regression models

Basic regression analysis theory

The SVM described above solves a binary classification but cannot estimate continuous variables such as age or sleep onset latency. In this case, regression models can be used. Regression models belong to the statistical learning method (see Section 2.4), with the only difference to SVM that now Y is a continuous variable. The aim of regression is thus to learn an optimal function $f(X)$ that can best predict the variable. In this chapter, only linear regression models were used for exploratory and easy interpretation purposes, although f can also be a nonlinear function to uncover nonlinear interactions between features and predictors. Notice that some 'regression' models can solve classification questions, such as logistic regression. The logistic regression model would be further introduced and applied in Chapter 6.

The problem that a linear regression model is trying to solve is similar to Equation (2.3), where f will be a linear function:

$$Y = \beta X + \epsilon. \tag{3.4}$$

Here Y is the response variable, X is a matrix of the predictors or features, and β is the coefficients to be learnt. Linear regression models take a linear combination of the predictors and try to predict the response variable. To train the model and get optimal coefficients, the aim is to minimise the error term ϵ (also called residuals), and again, this comes to optimisation algorithms (not described here). Still, how this optimisation is done affects the model types. Normal linear regression models assume that ϵ or the response variable Y follows a conditional Gaussian distribution, whereas generalised linear models (GLM) disregard this assumption and allow more types of residual distribution to be fitted. After training the GLM, the p-value infers whether the model is significant (i.e., whether the features can have significant linear prediction relationships to response variables). The model can have a p-value for each predictor so that one can identify which ones are significant, thus doing feature selection and dimensionality reduction as well. GLM can also describe nonlinear relationships between predictors and responses, which will not be used here.

GLM is a great tool for statistical inferences but not actual prediction tasks. One potential issue of regression models is that the features could be highly correlated to each other (especially for large feature space dimensions). This correlation will cause the covariance matrix of X to be rank deficient (with non-positive eigenvalues), where the loss function for optimisation is ill-posed to solve (Hastie, 2020). Regularisation is normally needed in model training to solve this issue. One most commonly used is the Ridge regularisation (Hoerl and Kennard, 1981). Another problem, similar to all other machine learning problems, is the overfitting of the regression model. Thus, cross-validation is needed to be applied.

The GLM can be further extended to a generalised mixed-effect model (GLME). GLME aims to solve similar questions to GLM but with additional considerations such as grouping of the samples. There are two components in a GLME model, the fixed effects (the predictors that one wants to test), and the random effects (potential inter-dependence of the samples). For instance, if there are 200 samples that come from 20 different participants, with each participant contributing 10 samples. The 10 samples from the same participant are not independent samples. The GLME adds this to the model with a grouping capability of the samples as a random effect. GLME is further introduced and applied in later chapters.

Implementations

In this chapter, the GLM was used for statistical inference of linear relationships between the features and response variables such as age or sleep onset latency assuming a normal distribution of the response variables. GLM was fitted using MATLAB function `fitglm`. For the prediction of sleep onset, a more rigorous model was applied with Ridge regularisation and cross-validation, using MATLAB function `fitrlinear`. The Ridge regularisation automatically reduces the feature dimensionality to prevent overfitting.

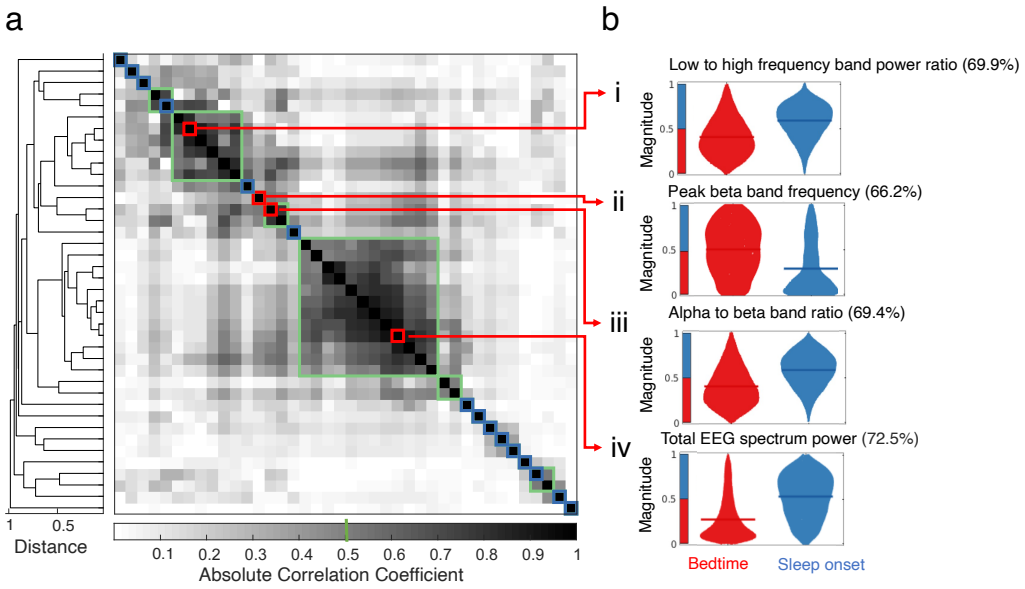


Figure 3.2: Hierarchical clustering of the features that distinguish bedtime and sleep-onset. **a**. The top 40 features that best classify wakefulness (bedtime) and sleep-onset are chosen and hierarchically clustered based on the absolute correlation coefficient with a clustering threshold of 0.5. The features are re-arranged according to the clustering, with the dendrogram shown on the left; Green box, outlines of a feature cluster; Blue square, central feature of a cluster; Red square, central feature with mean classification accuracy in that cluster above 65% and chosen as a key governing feature. The grey colour indicates the Pearson correlation coefficients between each pair of features, with the colour bar shown at the bottom. **b**. The distribution of the four key features identified from (**a**): the low-to-high band power ratio (**i**), the peak beta-band frequency (**ii**), the alpha-to-beta band ratio (**iii**), and the total EEG spectrum power (**iv**). Red indicates bedtime and blue indicates sleep-onset.

3.3. Results

3.3.1. EEG features governing sleep onset compared to bedtime

I first set out to identify the EEG features that best distinguish the brain activity at sleep onset state from the brain activity at bedtime using the MESA dataset and the feature-based statistical learning approach. I computed the feature values of the epochs in the 1-minute bedtime and 1-minute sleep onset segments. Then, I used SVM for each feature to classify bedtime and sleep onset, with a 10-fold leave-one-out cross-validation. Finally, I used a hierarchical cluster of 40 features with the highest accuracy to identify the most informative features.

The results are shown in Figure 3.2a. It shows 20 feature clusters with four clusters having a mean classification accuracy higher than 65%. See Appendix B.4 for the mean accuracy of all 40 features and the mean accuracy of each cluster. Figure 3.2a shows the values of the features at the centre of those four clusters during bedtime (red) and sleep onset (blue) together with their classification accuracies (See Appendix B.4 for a detailed explanation and origins of these features). The features were: the low-to-high frequency band power ratio (i.e., the ratio of

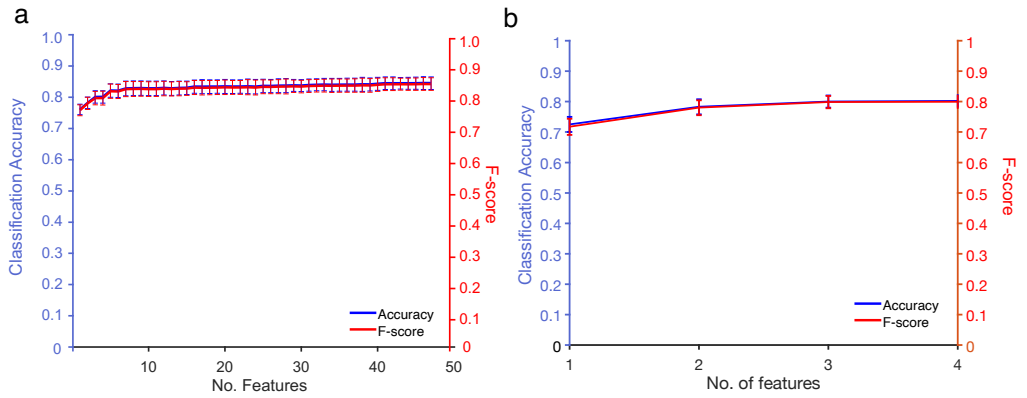


Figure 3.3: Classification accuracy and F-scores using all features and key features only. **a**. Classification accuracy (blue) and F-score (red) between bedtime and sleep-onset as a function of features used (in total 47 features). Shown are the mean ± st.d. for 10-fold leave-one-out cross-validations of the SVM model. **b**. Same as shown in (**a**) but now only with the 4 key features identified in Figure 3.2.

spectral power in the delta and theta bands to the alpha and beta bands, 69.9% accuracy), the peak beta band frequency (the frequency bin with the largest PSD within the beta frequency band, 66.2% accuracy), the alpha-to-beta band power ratio (69.4% accuracy), and the total EEG spectrum power (i.e., the power between 0 and 25 Hz relative to the power across the spectrum, 72.5% accuracy). The result suggests that sleep onset was best described by an increase in the low-to-high frequency band power ratio, a reduction in the beta band frequency, and an increase in the alpha-to-beta band power ratio and the total EEG power. Other features commonly reported in the context of sleep onset such as the theta-to-alpha band power ratio and theta power, were captured within the clusters of alpha-to-beta band power ratio and the low-to-high power band ratio. In addition, many of the canonical time-series features with both linear and nonlinear metrics were captured in the total EEG spectrum power (a Catch-22 feature) cluster. Interestingly, the changes in the peak beta band frequency feature were orthogonal to other feature changes (i.e., an independent cluster). I call these key features the 'governing features' later in the text.

Using all the features (i.e., 47 features) arranged from highest to lowest classification accuracy, the SVM model achieved a high accuracy of 85% (F-score 0.84) (Figure 3.3a). Note that in the figure, the features were first ranked based on their classification performance and then inputted into the model sequentially. When only the governing features (i.e., 4 features) were used, the SVM model also achieved a high accuracy of 80% (F-score 0.8) (Figure 3.3b). This result suggests that there was a lot of redundancy among the features and that by focusing on the four governing features one could already achieve accurate distinguishing of the sleep onset state. The results also demonstrate the utility of the hierarchical clustering method for feature selection.

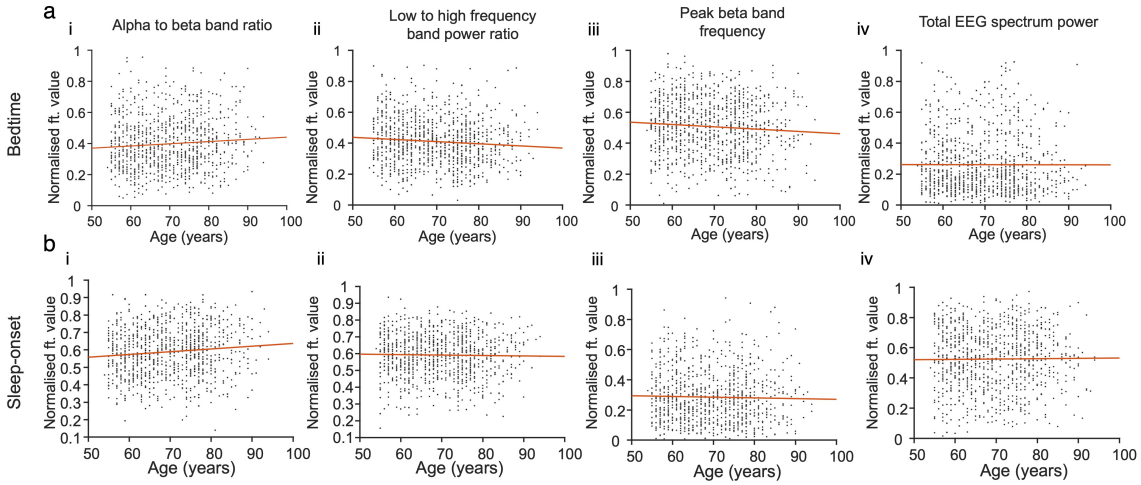


Figure 3.4: Linear correlations between the four key features and ages. a. Correlation between age and the normalised values of 4 key governing features at bedtime: (i) Alpha to beta band power ratio, (ii) low to high frequency band power ratio, (iii) peak beta band frequency and (iv) total EEG spectrum power. The red line shows the polynomial fit (regression line). **b**. Similar to (**a**) but now showing feature values at sleep onset. The Pearson correlation stats for bedtime (**a**): (i) coefficient $m = 0.08$, $p = 0.02$; (ii) coefficient $m = -0.08$, $p = 0.01$; (iii) coefficient $m = -0.08$, $p = 0.02$; (iv) coefficient $m = 1.85 \times 10^{-5}$, $p = 1$; For (**b**): (i) coefficient $m = 0.10$, $p = 0.001$; (ii) coefficient $m = -0.02$, $p = 0.56$; (iii) coefficient $m = -0.03$, $p = 0.43$; (iv) coefficient $m = 0.01$, $p = 0.76$.

3.3.2. The effect of ageing

After identifying the EEG features that best distinguish brain activity at sleep onset, I explored whether those features are affected by age by computing the correlation (using the Pearson correlation test) between the subjects' feature values and their age. The results are summarized in Figure 3.4. It shows that at bedtime, ageing was associated with a higher alpha-to-beta power ratio (coefficient $m = 0.08$, $p = 0.02$) but a lower low-to-high power band ratio ($m = -0.08$, $p = 0.01$) and peak beta band frequency ($m = -0.08$, $p = 0.02$). At sleep onset, ageing was associated with a higher alpha-to-beta power ratio (coefficient $m = 0.10$, $p = 0.001$). The correlation between sleep onset alpha-to-beta power ratio and ageing survived a Bonferroni correction of the eight comparisons (corrected p-value threshold 0.00625). The results suggest that elderly people had overall higher alpha-to-beta power that is sustained at sleep onset. The interpretation of the findings is however limited, especially for bedtime features. This will be discussed further later.

Next, I used GLM to explore whether there is a relationship between all the EEG features and ageing. I focused on the sleep onset period since at bedtime brain activity can be influenced by many factors apart from age. Before fitting the model, the feature values from all samples of the same participant (four samples at sleep onset) were averaged. I found that the GLM model was significant with F-stat 3.15, R-squared 0.14 and $p < 0.0005$. The significant features are listed in Table 3.1. It shows that at sleep onset, ageing was associated with lower values of Beta

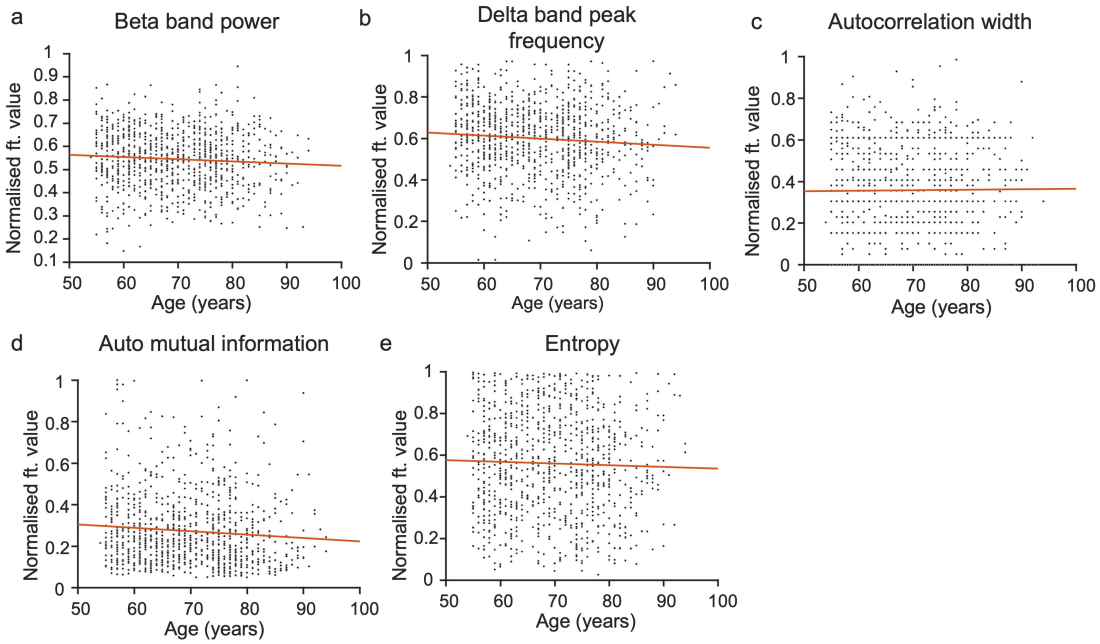


Figure 3.5: Linear correlations between the significant features from the GLM model and age. a-e. Linear correlation between age and the significant features in the GLM model. (**a**) beta band power, (**b**) delta band peak frequency, (**c**) autocorrelation width, (**d**) auto mutual information, and (**e**) entropy. Red line, linear polynomial fit; (**a**) coefficient $m = -0.07$, $p = 0.03$ for Pearson correlation test, (**b**) $m = -0.08$, $p = 0.01$, (**c**) $m = 0.01$, $p = 0.80$, (**d**) $m = -0.08$, $p = 0.01$, and (**e**) $m = -0.03$, $p = 0.34$.

band power, Delta band peak frequency, Autocorrelation width, Auto mutual information, and Entropy, and with higher values of Low to high-frequency band power ratio (a governing feature shown before). Refer to Appendix B.5 for full GLM statistics of all features and explanations of the six features that are significant in the model. One of the features (the 'low to high frequency band power ratio') was the same as one of the four key governing features before.

Table 3.1: List of significant features at sleep onset to predict age in GLM, along with their statistics. 't-Stat' stands for t-statistics.

Feature name	**Estimated Coefficient**	**Standard Error Estimate**	**t-Stat**	**p-value**
Beta band power	-45.33	13.59	-3.34	0.001
Low to high frequency band power ratio	51.37	22.15	2.32	0.021
Delta band peak frequency	-6.86	2.37	-2.90	0.004
Autocorrelation width	-8.97	3.75	-2.39	0.017
Auto mutual information	-12.89	2.82	-4.57	5.58×10^{-6}

Continued on next page

Table 3.1: List of significant features at sleep onset to predict age in GLM, along with their statistics. 't-Stat' stands for t-statistics. (Continued)

Entropy	-19.95	6.07	-3.28	0.001

The GLM showed that at sleep onset, a linear combination of these six features was correlated with age. To test whether each individual feature is correlated with age, I repeated the Pearson correlation analysis to all the features apart from the low to high-frequency band power ratio, which was already computed, see Figure 3.4bii. The results are summarized in Figure 3.5. The results confirmed that ageing was associated with reduced beta band power, delta band peak frequency, and auto-mutual information at sleep onset. Overall, the results suggest that sleep onset of older people was associated with slower brain activity and stronger noise-like patterns with less repetition (Escudero et al., 2009). It is interesting to note that although a linear combination of all these features could predict ages, not every feature was correlated to age.

Finally, I deployed the GLM to test the effect of ageing on bedtime EEG. The fitted model had an F-stat of 1.64, an R-squared value of 0.079 and a p-value of 0.005. The full model statistics and statistics for each feature can be found in Appendix B.6. The six features shown before were not significant in this model, and there were only a few features just below the significance line (p-value 0.05), which is not further interpreted here.

3.3.3. The effect of gender

Next, I tested, using a two-sample t-test, whether the governing EEG features differ between male and female participants. The results are shown in Figure 3.6.

I found that female participants had higher total EEG spectrum power ($p = 0.04$) and beta band frequency ($p = 0.02$) at bedtime. However, the interpretation is limited since the results did not survive correction for multiple comparisons. The full statistics are given in Appendix B.7.

To explore the effect of gender across all the EEG features, I used the classification followed by hierarchical clustering as before but now for 'male' and 'female' labels. I did not find features with an accuracy larger than 55%, suggesting no differences between genders.

3.3.4. Validation using a second sleep EEG dataset (Montreal dataset)

To test the generalization of the results obtained with the MESA dataset, I repeated the analysis with the Montreal dataset. As mentioned previously, the Montreal dataset contains 17 channels, but only three channels, Fz, Cz and Pz, were taken here, which measure approximately the same

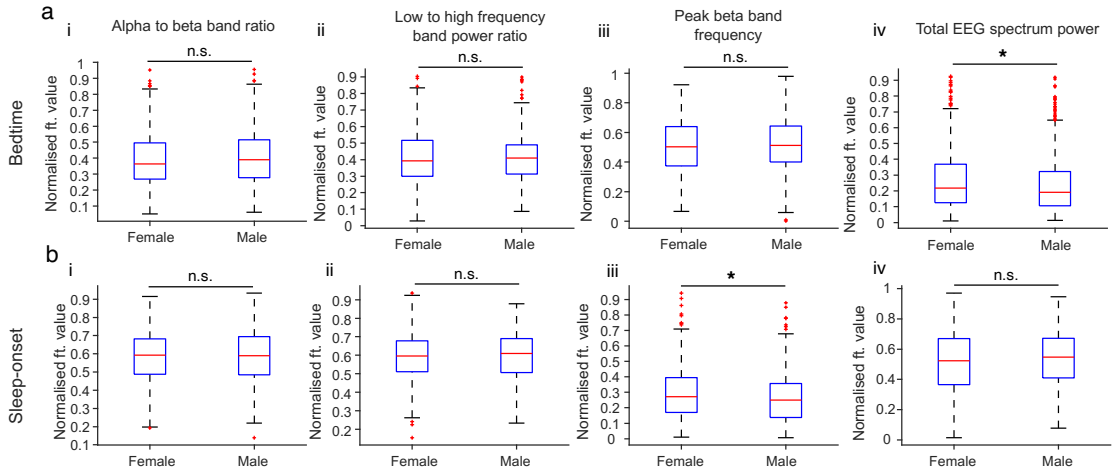


Figure 3.6: Differences of the four key features between genders. a. Differences of the 4 key governing features between female and male participants at bedtime; (i) alpha to beta band ratio, (ii) low to high band power ratio, (iii) peak beta band frequency and (iv) total EEG spectrum power. **b**. Similar to (**a**) but now for features at sleep onset. '*' implies $p < 0.05$ using a double-sided two-sample t-test; n.s. non-significant. The full distribution statistics can be checked in Appendix B.7. The red horizontal line shows the median values, with the box showing the upper and lower 25% quantile lines. Red plus markers show the outliers, who are more than 1.5 interquartile range (distance from top to bottom) from the top and bottom of the box.

cortical areas as in the MESA (frontal, central and parietal/occipital). Figure 3.7 shows the EEG features that best distinguish the brain activity at sleep onset state from the brain activity at bedtime in this dataset. As in the MESA dataset, there were four feature clusters with a mean classification accuracy higher than 65%. The corresponding governing features were the low to high frequency band power ratio (accuracy 80.6%) and the peak beta band frequency (accuracy 72.6%) as in the MESA dataset, and the theta band peak frequency (accuracy 68.6%) and the delta band peak frequency (accuracy 72.6%), that different from the MESA dataset, Figure 3.7b. In comparison to bedtime, brain activity at sleep onset was associated with an increase in the low to high-frequency band power ratio and a decrease in the peak beta band frequency as in the MESA dataset and with lower theta frequency and higher delta frequency.

The total EEG spectrum power found to govern the sleep onset in the MESA dataset but not the Montreal dataset had a high accuracy of 69.9% in the Montreal dataset. However, since the feature occurred inside a large cluster with a mean cluster accuracy smaller than 65% it was not picked up. The alpha to beta band power ratio was found to govern sleep onset in the MESA dataset, but not the Montreal dataset had a lower accuracy of 61.9% in the Montreal dataset. It was not linked to other key feature clusters. The peak delta and theta band frequencies found to govern the sleep onset in the Montreal dataset but not the Mesa dataset, had accuracies of 64.79% and 53.81%, respectively, in the MESA dataset. Refer to Appendix B.8 for a table of all the features listed in the hierarchical order as in the figure, along with their accuracies.

As in the MESA dataset, classifying with all the features (i.e., 47 features arranged from highest

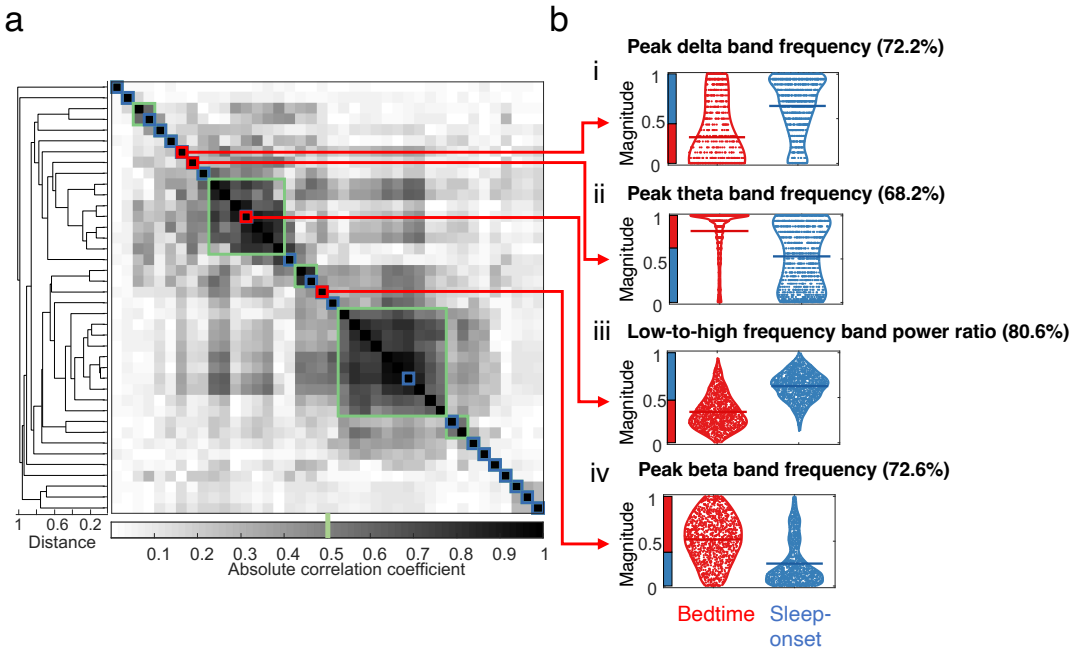


Figure 3.7: Hierarchical clustering of the features that distinguish bedtime and sleep-onset in the Montreal dataset. a. The top 40 features that best classify wakefulness (bedtime) and sleep-onset are chosen and hierarchically clustered based on the absolute correlation coefficient with a clustering threshold of 0.5. The features are re-arranged according to the clustering, with the dendrogram shown on the left; Green box, outlines of a feature cluster; Blue square, central feature of a cluster; Red square, central feature with mean classification accuracy in that cluster above 65% and chosen as a key governing feature. The Pearson correlation coefficients between pairs of features are shown in grey-scale, with the colour bar at the bottom. **b**. The distribution of the four key features identified from (**a**): peak delta band frequency (**i**), the peak theta band frequency (**ii**), the low to high frequency band power ratio (**iii**), and peak beta band frequency (**iv**). Red shows bedtime, and blue shows sleep-onset.

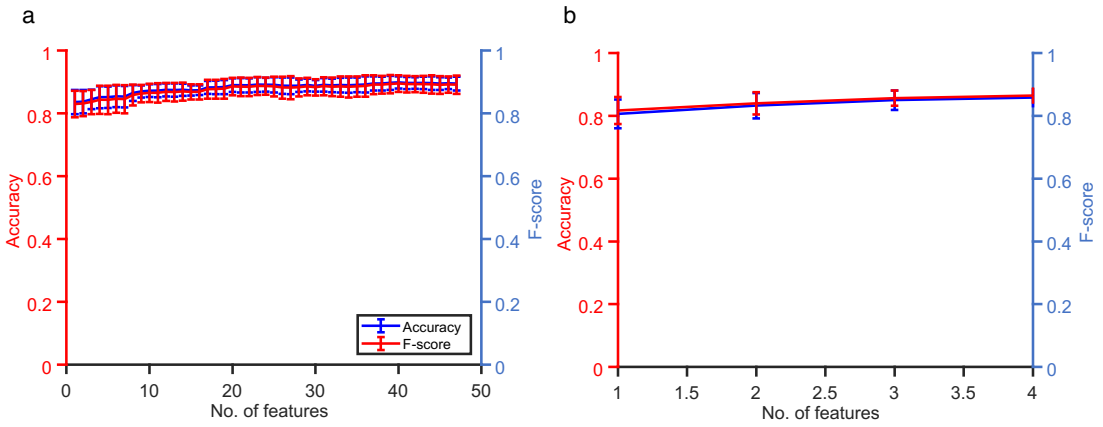


Figure 3.8: Classification accuracy and F-scores using all features and key features only in the Montreal dataset. a. Classification accuracy (blue) and F-score (red) between bedtime and sleep-onset as a function of features used (in total 47 features) in the Montreal dataset. Shown are the mean ± st.d. for 10-fold leave-one-out cross-validations. **b**. Same as shown in (**a**) but now only with the 4 key features identified in Figure 3.7.

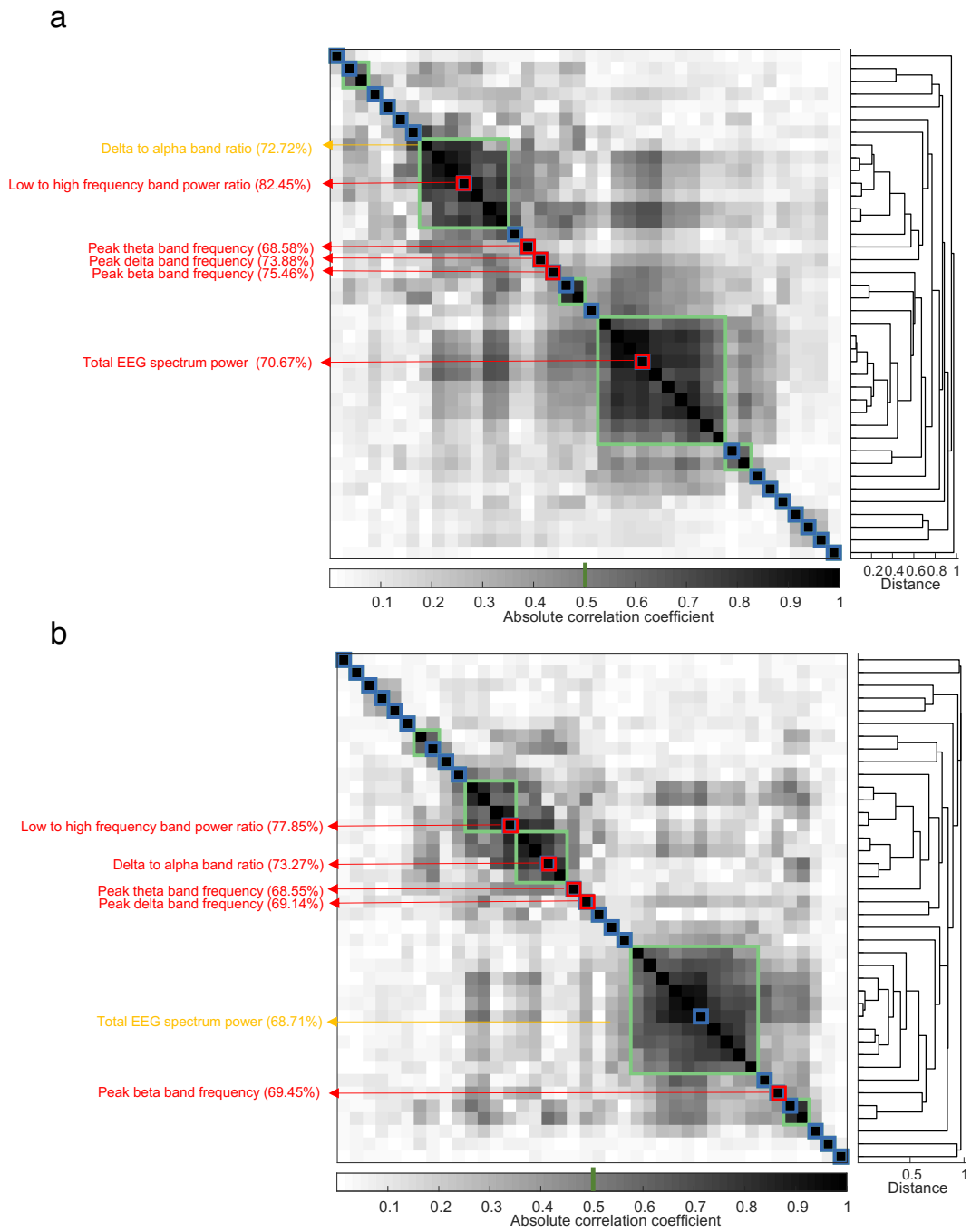


Figure 3.9: Hierarchical clustering of the features that distinguish bedtime and sleep-onset for Normal and MCI participants separately. (**a**) for Normal and (**b**) for MCI participants; Hierarchical clustering is done based on Pearson correlation distances with cluster threshold 0.5; Dendrograms of the clustering are shown on the right; Green box, outlines of a feature cluster; Blue square, centroid feature of a cluster; The Pearson correlation coefficients between pairs of features are shown in grey-scale with colour bars at the bottom. Red square and red features, central feature with mean classification accuracy in that cluster above 65% and chosen as a key governing feature. Yellow features are the key features identified in another cohort but not in this current cohort. The relevant feature names, along with their classification accuracy, are shown on the left side.

to lowest classification accuracy), the SVM model achieved a very high accuracy (89%, f-score 0.89) (Figure 3.8a). Similar accuracy could be achieved by using only the governing features (86%, f-score 0.86) (Figure 3.8b). As a quick test, the same four key features from the MESA dataset were used here in the Montreal dataset to classify between the two states, and the model reached 83% accuracy (F-score 0.84), which was also higher than that in the MESA dataset.

Overall, the replication of the analysis in the Montreal dataset confirmed that the brain activity at sleep onset could be accurately distinguished from the brain activity at bedtime using a small subset of features, aka governing. The differences in the governing features and accuracy between the datasets could have originated from the fact that the Montreal cohort includes subjects with MCI. To explore the effect of MCI, I repeated the analysis of the Montreal dataset but separately for the normal and MCI participants. The results are shown in Figure 3.9. It shows that the sleep onset in healthy participants was governed by the same four features found when the whole dataset was analyzed (i.e., the low to high band power ratio, the peak delta, theta and beta band frequencies) plus an additional feature, i.e., the total EEG spectrum power (Figure 3.9a). The sleep onset in participants with MCI was governed by the same four features found when the whole dataset was analyzed (i.e., the low to high band power ratio, the peak delta, theta and beta band frequencies), plus a different additional feature, i.e., the delta to alpha band ratio (Figure 3.9b). The delta-to-alpha band ratio also had high accuracy (72.72% in the healthy group but was correlated with the low-to-high band power ratio feature (yellow in Figure 3.9a). The total EEG spectrum power also had a high accuracy (68.71%) in the MCI group but was embedded in the cluster with a low mean accuracy (yellow in Figure 3.9b). The tables of the list of features in the hierarchical order for each one of them can be found in Appendix B.9.

3.3.5. The effects of cognitive impairment on the features at bedtime and sleep onset

To further investigate the effect of cognitive impairment on sleep EEG activity, I directly compared the governing features between the normal and MCI groups at bedtime and sleep onset. The results are shown in Figure 3.10 for bedtime and Figure 3.11 for sleep onset. The full statistics of the two figures can be found in Appendix B.10.

It reveals that at bedtime, the participants with MCI had higher low-to-high frequency band power ratio and peak delta band frequency (both increase at sleep onset compared to bedtime), but lower peak beta band frequency (which decreases at sleep onset compared to bedtime). At sleep onset, the participants with MCI had a higher peak beta band frequency but a lower total EEG spectrum power (which increases at sleep onset compared to bedtime). Together these results demonstrate that people with MCI had brain activities that are closer to sleep when going to sleep and shallower sleep activities at its onset. Earlier studies have shown that people with MCI have reduced vigilance during wakefulness (Jaussent et al., 2012). The results

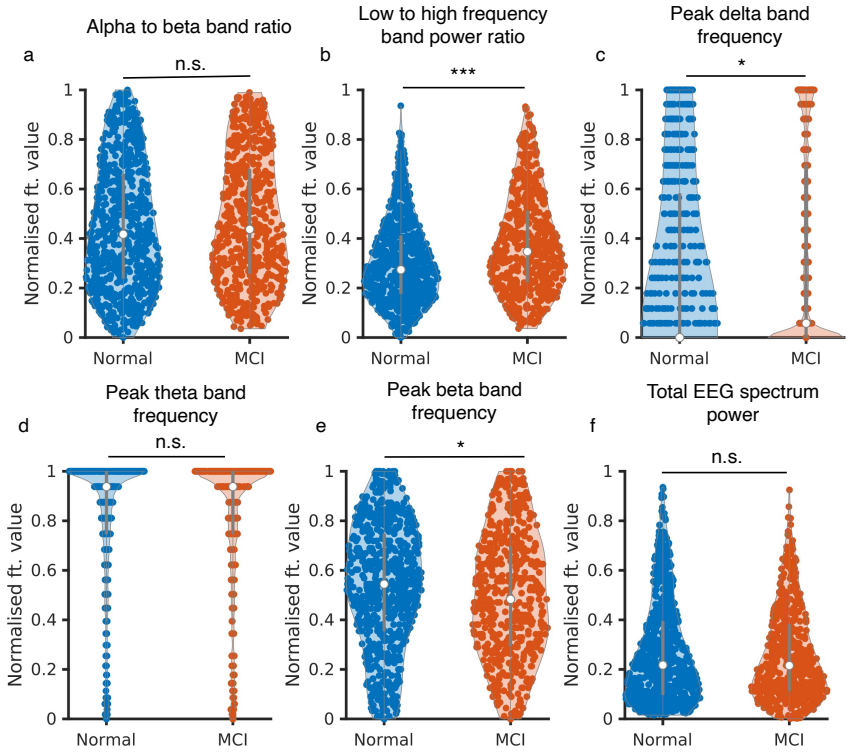


Figure 3.10: Differences in the six key features between Normal (blue) and MCI (red) participants at bedtime. **a** Alpha to beta band ratio; **b** low to high frequency band power ratio; (**c**) peak delta band frequency; **d** peak theta band frequency; **e** peak beta band frequency; **f** total EEG spectrum power. *** implies $p < 0.0005$ using double-sided two-sample t-test; * implies $p < 0.05$. Violins show the kernel density plot (external shape), median values (white non-filled circle) and upper and lower 25% quantiles (grey bar).

reported here raise the possibility that reduced vigilance during wakefulness could be associated with the ability to initiate sleep.

To test whether other features could differentiate the brain activity of the MCI participants from healthy participants, I tested classification between the Normal and MCI labels. I did not find features that could classify the brain activity of the MCI participants with sufficient accuracy (all feature accuracies were smaller than 60%) at both bedtime and sleep onset.

3.3.6. Relationship to sleep onset latency

In this section, I aim to explore whether the time participants required to fall asleep, i.e., the sleep onset latency relates to the features of their brain EEG. I first tested the relationship between the participants' sleep onset latency and their demographics (age, gender, cognitive impairment). I then tested the relationship between the participants' sleep onset latency and the value of the EEG features governing sleep onset. Finally, I tested whether the sleep onset can be predicted via a regression using all the feature base and demographic information. I used

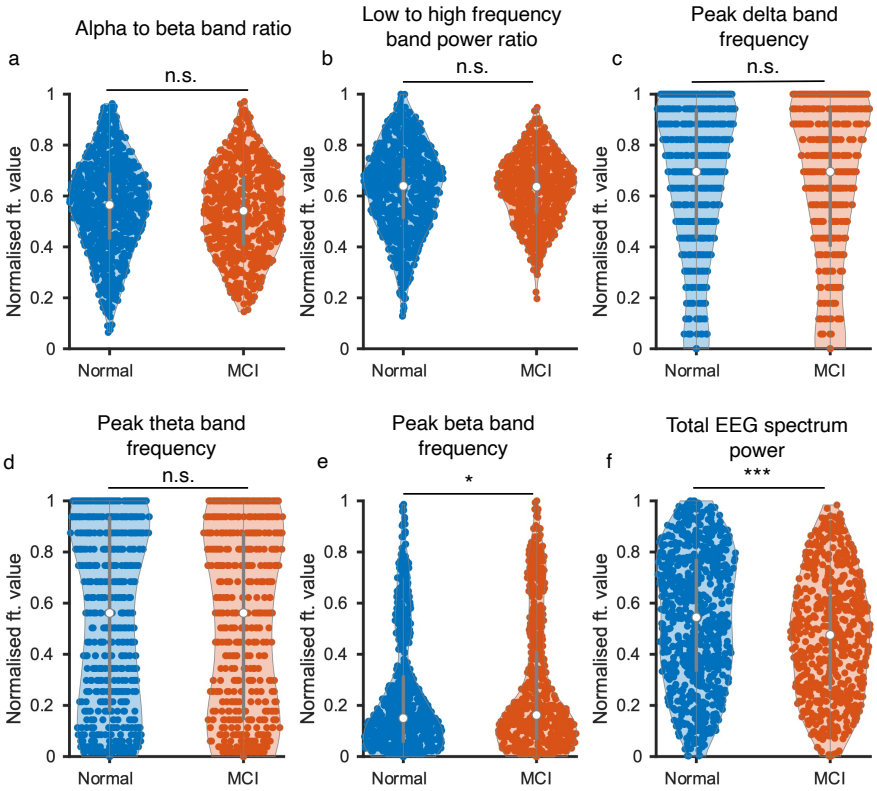


Figure 3.11: Differences in the 6 key features between Normal and MCI participants at sleep onset. Same as in Figure 3.10 but now for feature values at sleep onset. *** implies $p < 0.0005$ using double-sided two-sample t-test; * implies $p < 0.05$. Violins show the kernel density plot (external shape), median values (white non-filled circle) and upper and lower 25% quantiles (grey bar).

the MESA dataset in all the analyses apart from the analysis that tested the effect of cognitive impairment. In all the analyses later, the features of the EEG mini-epoch samples from the same participant were averaged to get representative feature values for that participant (at bedtime or sleep onset separately).

Effects of age, gender and cognitive impairment on sleep onset latency

Figure 3.12 shows the relationship between sleep onset latencies, and age, gender, and cognitive impairments. I found that the participants' sleep onset latencies were slightly negatively correlated with their age (coefficient -0.06, p= 0.049), Figure 3.12a. The sleep onset latencies of female participants were slightly longer than male participants (female: 50.42 ± 25.40 minutes mean ± st.d.; male: 47.05 ± 24.48 minutes; $p = 0.03$), Figure 3.12b. The sleep onset latencies of participants with MCI were longer than those of healthy participants (MCI: 24.20 ± 17.02 minutes; healthy: 20.28 ± 17.63 minutes; $p = 0.048$; Montreal dataset), Figure 3.12c. It is worth noting that the Montreal datasets had a lower latency range than the MESA dataset.

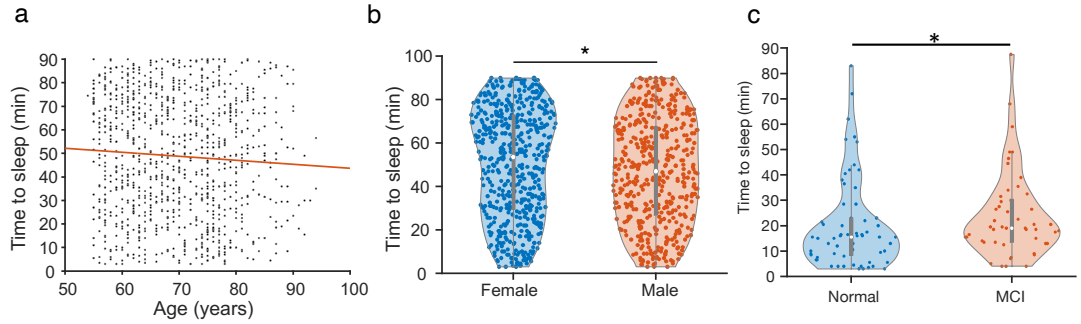


Figure 3.12: Effects of age, gender and cognitive impairment on sleep-onset latency. **a**. The sleep-onset latency vs ages. Red-line, the linear regression with coefficient -0.06 and p-value 0.049 for Pearson correlation test. **b**. The distribution of sleep-onset latency between female and male participants. Significance is tested using the Kruskal-Wallis test; $*$ indicates $p < 0.05$. **c**. The distribution of sleep-onset latency between Normal and MCI participants in the Montreal dataset. $*$ indicates $p < 0.05$, Kruskal-Wallis test. Violins show the kernel density plot (external shape), median values (white non-filled circle) and upper and lower 25% quantiles (grey bar).

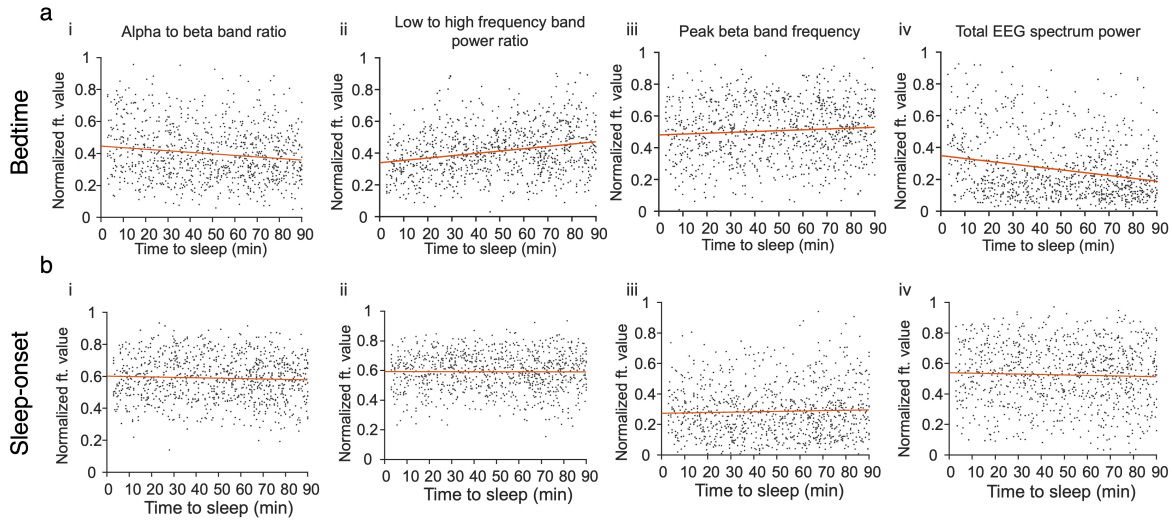


Figure 3.13: Correlation between the four key features and sleep-onset latency at bedtime (a) and sleep onset (b). The four features are alpha to beta band ratio (i), low to high frequency band power ratio (ii), peak beta band frequency (iii), and total EEG spectrum power (iv) separately. Red line, the polynomial fit (regression line). For (**a**), (i) line slope $m = -0.14$, $p < 0.0005$ Pearson correlation test; (ii) line slope $m = 0.25$, $p < 0.0005$; (iii) line slope $m = 0.08$, $p = 0.02$; (iv) line slope $m = -0.24$, $p < 0.0005$; For (**b**), (i) line slope $m = -0.05$, $p = 0.13$; (ii) line slope $m = -0.003$, $p = 0.92$; (iii) line slope $m = 0.04$, $p = 0.24$; (iv) line slope $m = -0.04$, $p = 0.26$;

Effect of governing EEG features on sleep latency

Figure 3.13 shows the correlation between sleep onset latency and the EEG features found to govern sleep onset. I found that the participants' sleep onset latencies were correlated with the bedtime but not sleep onset values of the features governing sleep onset (i.e., the alpha to beta power ratio, the low to high frequency band power ratio, the peak beta band frequency, and

the total EEG spectrum power). Specifically, longer sleep onset latencies were associated with lower alpha to beta band ratio ($m = -0.14$, $p < 0.0005$) and total EEG spectrum ($m = -0.24$, $p < 0.0005$) at bedtime (both increase at sleep onset compared to bedtime), but higher low to high frequency band power ratio ($m = 0.25$, $p < 0.0005$) and peak beta band frequency ($m = 0.08$, $p = 0.02$) at bedtime (increase and decrease at sleep onset compared to bedtime, respectively). However, the correlation with the peak beta frequency did not survive Bonferroni correction for the eight comparisons (p-value threshold 0.0065).

These results demonstrate that, no matter how long it took for the participants to fall asleep, their sleep onset state (based on the features) was not affected by it. The negative correlation of the sleep onset latency with the bedtime' s alpha to beta band ratio and total EEG spectrum suggests that the more similar the brain activities to sleep the shorter the sleep latency. However, the positive correlation with the low to high frequency band power ratio implies that it is not a simple relationship.

Prediction sleep latency from bedtime EEG

Could sleep onset latency be predicted from the EEG features at bedtime? To test the feasibility of predicting sleep onset latency, I first used the feature-based learning approach to test whether fast and slow sleepers can be classified based on their bedtime EEG. Thus, I extracted the 20 percent fastest and 20 percent slowest participants (Figure 3.14). The fast sleeper group had a sleep onset latency smaller than 23 minutes ($N = 203$ participants, 98 males). The slow sleeper group had a sleep onset latency larger than 74 minutes ($N = 205$ participants, 84 males). The fast sleeper group was slightly older than the slow sleeper group (fast sleepers: 70.38 ± 8.94 years; slow sleepers: 68.65 ± 8.98 years; $p = 0.046$, Kruskal-Wallis test), Figure 3.14, consistent with Figure 3.12a.

The classification outcome is shown in Figure 3.15. I found three feature clusters with a mean accuracy of approximately 60% (the rest of the clusters had a lower mean accuracy): total EEG spectrum power, lower delta to alpha band power ratio, and delta band power (Figure 3.15b). The key feature selection criterion was adjusted here since the general classification accuracy was not very good for all features. See Appendix B.11 for the feature list in the order of hierarchical clustering. Participants with faster sleep latencies had higher total EEG spectrum power (increase at sleep onset compared to bedtime), but lower delta to alpha band power ratio and delta band power (increase at sleep onset compared to bedtime).

The association of fast sleep latency with higher total EEG power is consistent with the analysis shown in Figure 3.13aiv, suggesting the closer the bedtime values to sleep onset ones, the faster the transition. However, the association of fast sleep latency with lower delta to alpha band power ratio and delta band power suggested a spring-like phenomenon, i.e., the further the

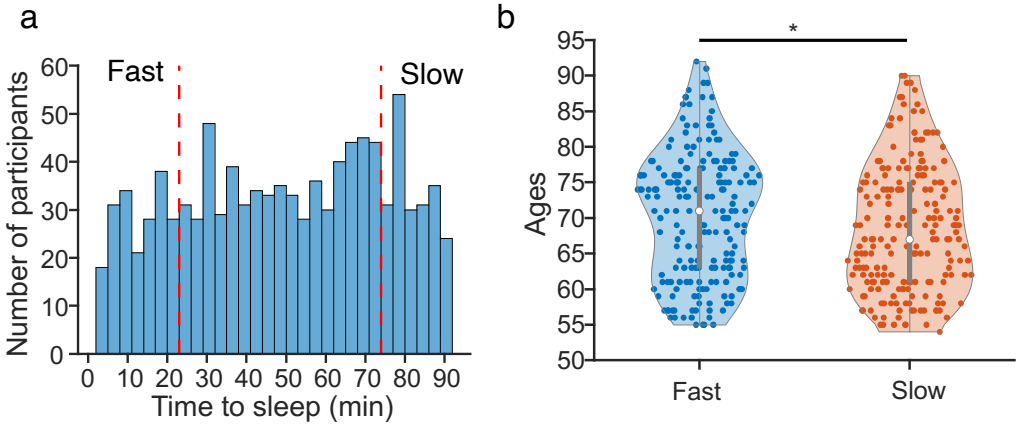


Figure 3.14: Division of the fast and slow participant cohorts based on sleep-onset latencies. a. Histogram of the sleep-onset latency for the remaining N=1009 participants within ranges [3,90] minutes. The two vertical dashed lines are the upper and lower 20-percentile dividing lines. **b**. The age distribution for the fast (N=203) and slow (N=205) cohort, and * indicates $p < 0.05$ for the Kruskal-Wallis test. Violins show the kernel density plot (external shape), median values (white non-filled circle) and upper and lower 25% quantiles (grey bar).

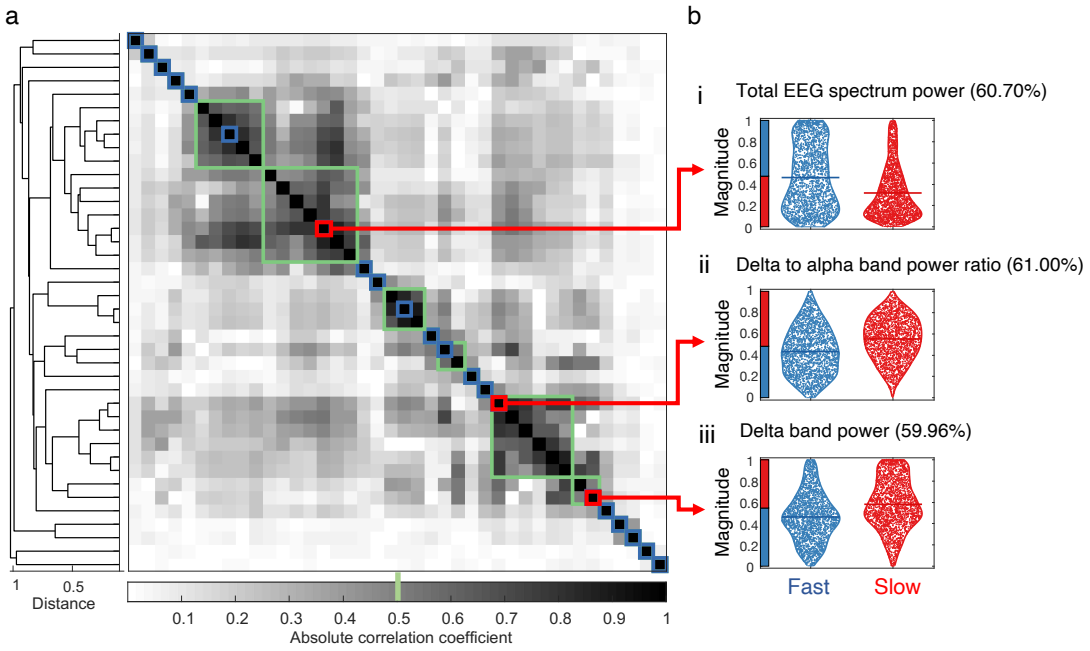


Figure 3.15: Hierarchical clustering of the features that distinguish fast and slow latencies cohort at bedtime. a. Hierarchical clustering of the top 40 features that best distinguish between fast and slow. The top 40 features that best classify wakefulness (bedtime) and sleep-onset are chosen and hierarchically clustered based on the absolute correlation coefficient with a clustering threshold of 0.5. The features are re-arranged according to the clustering, with the dendrogram shown on the left; Blue squares indicate cluster centres; Green boxes indicate features that are clustered together; Red squares indicate the key features selected. The Pearson correlation coefficients between pairs of features are shown in grey scales, with the colour bar at the bottom. **b**. The key features selected along with their accuracies and distribution plots for fast (blue) and slow (red) cohorts: (i) the total EEG spectrum power; (ii) the delta to alpha band power ratio; (iii) the delta band power.

bedtime values to sleep onset ones, the faster the transition. Using all features, the classification (leave-one-out linear SVM) accuracy was 63.5% (F-score 0.61). Using only the three governing features, the accuracy was similar i.e., 63.7% (F-score 0.59). Note, that although the accuracy is not high, it was significantly larger than a chance level (50%).

After validating that the EEG features at bedtime can be classified between fast and slow sleepers, I used GLM to test whether the sleep latency could be predicted from the EEG features at bedtime. The GLM was applied using all features (at bedtime) as predictors and the latencies as the response variable, with the Gaussian distribution assumption (as no other clear distribution properties described well the data). I found that the model was significant with an F-stat of 3.66, an R-squared of 0.16, and $p < 0.0005$. The model had four significant features: the delta band power, the theta band power, the theta to beta band power ratio, and the total EEG spectrum power (Table 3.2). The full GLM statistics for all features can be found in Appendix B.12. Note, that the delta band power and the total EEG power were consistent with the previous classification analysis.

Table 3.2: List of significant features at bedtime to predict sleep onset latency in GLM, along with their statistics. 't-Stat' stands for t-statistics.

Names	**Estimate Coefficient**	**Standard Error Estimate**	**t-Stat**	**p-value**
Delta band power	117.25	59.27	1.98	0.048
Theta band power	-166.32	68.92	-2.41	0.016
Theta to beta band ratio	179.12	72.02	2.49	0.013
Total EEG spectrum power	64.99	32.85	1.98	0.048

The outcomes of the slow vs fast classification analysis and the GLM suggest that the sleep onset latency was associated with five features: bedtime power at the delta and theta bands, the power ratios between the delta and alpha bands and the theta and beta bands, and the total EEG spectrum power. The correlations between these features and the sleep onset latencies are shown in Figure 3.16 except for the total EEG spectrum power, which was shown in Figure 3.13aiv. It shows that the delta band power and delta to alpha band ratio had a positive correlation to the sleep onset latency. The correlation between the theta band power and the theta to beta band ratio and the sleep latency was not significant. These results support the spring-like phenomenon, i.e., the lower the absolute or relative delta power at bedtime the faster the sleep latency.

Next, I tested whether the sleep onset latencies can be predicted using the previously identified key features, combined with the knowledge of age and gender, using the Ridge regression model. To avoid biasing the model, the age values were z-score normalised before fitting into the model

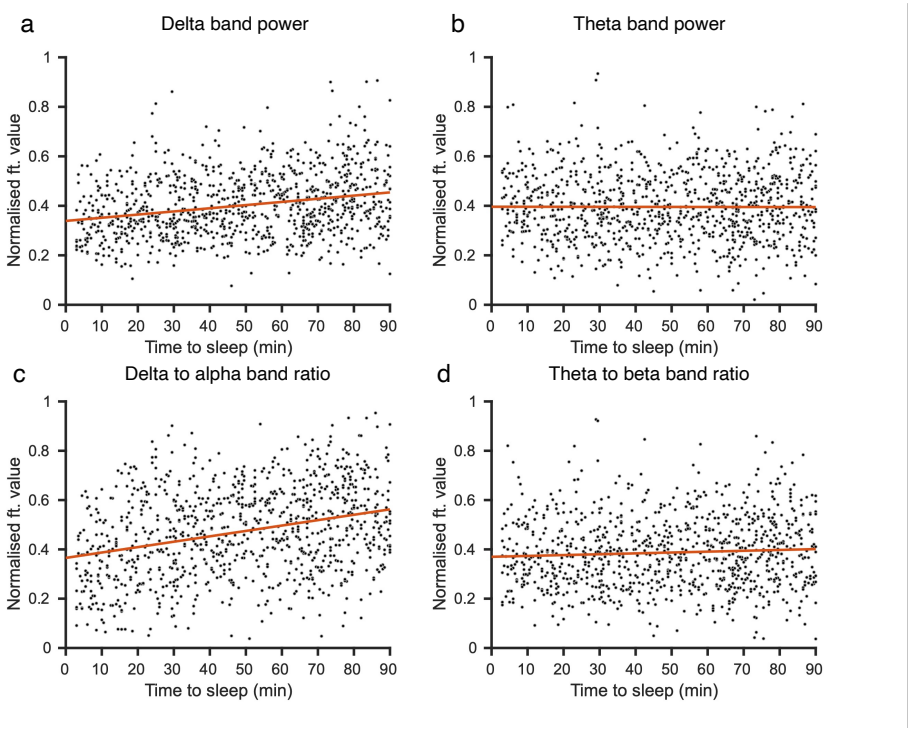


Figure 3.16: Correlation between the significant features in GLM and sleep-onset latency at bedtime. The three features are delta band power (**a**), theta band power (**b**), delta to alpha band ratio (**c**) and theta to beta band ratio (**d**) separately. Red line, the polynomial fit regression line. (**a**), line slope $m = 0.25$, $p < 0.0005$ Pearson correlation test; (**b**), $m = -0.003$, $p = 0.93$; (**c**), $m = 0.31$, $p < 0.0005$ and (**d**), $m = -0.06$, $p = 0.06$.

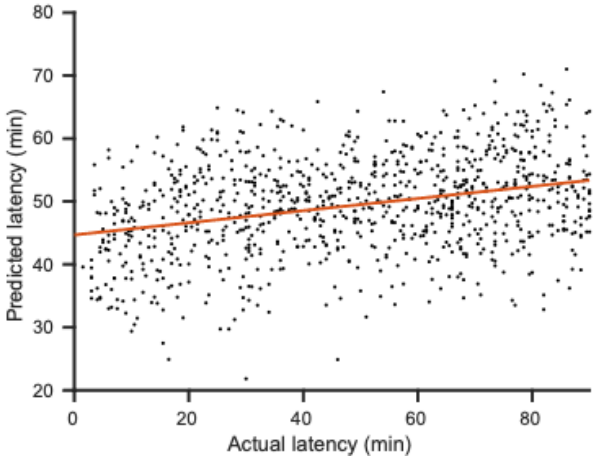


Figure 3.17: Comparison of the predicted sleep onset latency from the regression model and the actual latency. Red line, linear polynomial fit line. The correlation coefficient is 0.32 with the p-value smaller than 0.0005 using the Pearson correlation test.

(as all other features vary from 0 to 1). I found that using the five governing EEG features, age and gender, the prediction root mean squared error was 23.8 minutes (10-fold leave-one-out cross-validation), Figure 3.17. Using all the EEG features, age and gender, the prediction root

mean squared error was similar (23.7 minutes).

Dependency of sleep onset EEG dynamics on sleep latency

Are the EEG features at sleep onset dependent on the sleep latency? Classification of the fast and slow sleepers using the EEG features at sleep onset was not accurate (highest feature accuracy was smaller than 55%, accuracy using all features was 52.5%, F-score 0.52). Similarly, using GLM sleep latency could not be predicted using the EEG features at sleep onset (an F-stat of 1.09, R-squared of 0.05 and p-value of 0.32). These results demonstrate that the brain activity at sleep onset was independent of the latency to sleep onset.

3.4. Discussion

In this chapter, I used a feature-based statistical learning approach to identify the EEG features that distinguish the brain EEG activity at sleep onset from bedtime in older adults and explored the effect of demographic factors such as age and gender as well as MCI. In contrast to earlier studies, I used a comprehensive feature base and statistical learning including hierarchical feature clustering in two large independent sleep EEG datasets (including more than a thousand subjects) to systematically find the most informative features.

3.4.1. A set of orthogonal EEG features distinguishing sleep onset from wakefulness

I showed that brain activity at sleep onset can be reliably and accurately distinguished from bedtime by changes in a small set of four EEG features. Specifically, sleep onset is associated with an increase in the alpha-to-beta band ratio, low-to-high band power ratio and total EEG spectrum power, and a decrease in the peak beta band frequency. The results suggest that these feature changes best capture the changes in brain activity at the initiation of sleep across individuals. Contextualizing past literature: Earlier studies suggested that sleep onset could be described by features such as delta or theta band activities Merica and Fortune (2004); Achermann et al. (2019), theta to alpha band ratio (Jap et al., 2009; Lal and Craig, 2001), or power band analysis in general (De Gennaro et al., 2001a). I found that these features are embedded in feature clusters represented by the low-to-high frequency band power ratio and the alpha-to-beta band power ratio. The low-to-high frequency band power ratio found in my investigation was previously proposed by Slipher et al. to capture the consciousness state (Slipher et al., 2014). The other features, however, have not been previously reported to be useful in capturing drowsiness or sleep. The increase in total EEG spectrum power is expected given the larger oscillations at sleep. However, it is interesting that this feature represents a

cluster with many nonlinear features. Later analyses in this thesis further confirm the power of this feature. Ogilvie et al.'s finding long ago hallmarked such an all-band power increase as the key implication of transitioning into behavioural sleep (Ogilvie et al., 1991).

In comparison to earlier studies, here I used a more comprehensive feature base and large datasets. The feature base included all the common EEG features and the CATCH-22 features that were found efficient for different types of time series data (selected from more than seven thousand features). I then deployed a hierarchical feature clustering to identify the most informative features accounting for redundancy in feature information. To test the generality of the results, I repeated the analysis in an independent dataset.

Broad implications: We used to assess the onset of sleep by the development of delta waves and micro-events such as K-complex and spindles. (see the AASM Silber et al. (2007) and the Hori' s nine-stage system Hori et al. (1994)). This chapter's findings suggest an additional strategy to assess sleep onset based on four orthogonal EEG features, aka sleep onset EEG biomarkers. The sleep onset EEG biomarkers will assist future development and sleep-promoting therapies, including closed-loop neuromodulation. The results also demonstrate the utility of the feature-based statistical learning pipeline, and how it can be applied to better capture and understand the underpinning neural signatures behind a process, such as the transition into sleep. With a specifically designed feature base, the same pipeline can be applied to many other questions, such as the underpinning brain dynamics differences between diseased brains, key functional biomarkers behind the neural cognition process, or perturbation biomarkers informing neuro-intervention effects.

There are of course several limitations of the study as well. The two datasets used in this chapter provide independent validation of the main outcomes. However, differences in the dataset might increase the variance in the analysis. For example, the two datasets used different electrode montages and differed in health conditions (the Montreal dataset included patients with MCI). Importantly, the sleep onset latency distributions differ between the datasets (see results in Section 2.2). This means that the bedtime state for the MESA could be different to that of the MCI, leading to different features found. As some literature proposed, there could be a 'pre-sleep-onset' state (Merica and Fortune, 2004; Kinreich et al., 2014), where the brain would enter an intermittent state before the transition started. It could be that the bedtime dynamics for the Montreal dataset participants were closer to the 'intermittent state'. This will be further discussed in Chapter 4 about the continuous transition dynamics, where it was found that the theta band peak frequency (as well as delta) had a bursting type of dynamic before sleep onset and then went back after sleep onset.

The wakefulness state during bedtime (i.e., the first minute of recording) is heterogeneous. Differences in the starting of the EEG recording and/or the participant's sleepiness at bedtime might increase the variance in the analysis. The only certainty about the bedtime state is that

the participants were awake, while many other factors could confound the actual brain stats (e.g., moods, thinking). On the contrary, sleep onset dynamics could be of more homogeneity.

The EEG feature base used in this chapter was comprehensive but, in principle, other features not considered here could also be informative in describing the EEG at sleep onset. The analysis in this chapter focused on global changes in temporal dynamics of brain activity but did not examine the spatial dynamics, i.e., inter-regional changes and interactions. I explore the spatial features in Chapter 5. The analysis of this chapter focused on the brain EEG differences at the sleep onset state compared to the bedtime wakefulness state, but it did not examine the transition between the states (i.e., the falling asleep process). I investigate the transition from wakefulness to sleep in Chapter 4.

3.4.2. Sleep onset latency is correlated with and can be predicted by a small subset of EEG features at bedtime

I found that the falling asleep time, i.e., sleep latency, is associated with a small set of three brain EEG features at bedtime, i.e., the total EEG spectrum, the delta band power, and the delta to alpha band ratio (all features increase at sleep onset compared to bedtime). It is also found that the sleep onset EEG dynamics are independent of the actual latency that leads to it. There is not much research that systematically studies the relationship between EEG brain dynamics and sleep onset latency, but many studies do show that resting state EEG (similar to the relaxed bedtime state) has predictive value in the sleep onset latency (Diaz et al., 2016; Kovrov et al., 2018).

We used to think that the more similar the bedtime brain activity to sleep, the faster the transition to sleep. Some of these features support this view, such as the total EEG spectrum power and the alpha to beta band power ratio (larger values indicate shorter sleep latency). However, other features (especially the delta power) cast questions on this point, suggesting an alternative, coil spring-like phenomenon (Yamagata et al., 2021). This is also illustrated in the comparison between healthy normal participants and MCI patients. When comparing the Normal and MCI participants, it was found that the MCI participants were closer to sleep onset state at bedtime (higher low to high band power ratio, higher peak delta band frequency and lower peak beta band frequency), but they reached a much shallower sleep onset state (higher peak beta band frequency and lower total EEG spectrum power at sleep onset). However, the MCI participants had a significantly higher sleep onset latency than normal participants. Relating to the 'spring theory' in sleep proposed by Yamagata et al. (Yamagata et al., 2021), intense wakefulness could compress the 'spring' more and lead to a stronger sleep rebound. This would imply that the further the features at bedtime to sleep onset, the stronger the drive in the sleep initiation. As a further comment on the findings in MCI patients, MCI is very often related to obstructive sleep apnea and excessive daytime sleepiness. Although this was not

demonstrated in the ESS scores, the feature dynamics captured this very well. The shallower sleep onset state compared to Normal participants implies a more aroused asleep state and could be related to sleep fragmentations in cognitively impaired patients (Gorgoni et al., 2021).

The finding is limited by the specificity of the feature base and database, and by the heterogeneity of the bedtime brain EEG as described above. In addition, the finding was not validated in multiple datasets, not at a single subject level. Future studies using multiple recordings from the same individuals could test the subject-level prediction of this approach. It also relates to the definition of the falling asleep period (or sleep onset period) and latency. The actual 'transition' into sleep could be much shorter than from bedtime to sleep onset, which means that the brain dynamics stay stagnant until the actual transition occurs. Therefore, the model would lose functionality if the bedtime dynamics stay anywhere outside the transition zone. Further, the actual temporal dynamics of the transition were not considered. The prediction was based on the assumption that the features change monotonically and linearly while falling asleep, which might not be true. Relating to the previously identified contradictory feature results (feature changes to sleep latencies), the nonlinear feature dynamics could be one explanation for it as well (see Chapter 4 later).

3.4.3. Ageing is associated with shorter sleep onset latency and distinct brain dynamics at sleep onset

I found a negative correlation between sleep onset latency and age, suggesting older people fall asleep faster. In addition, older people had a higher alpha-to-beta power ratio (feature increases at sleep onset compared to bedtime) at sleep onset. This result contradicts earlier findings that older people have longer sleep onset latency and have a more aroused sleep state (Li et al., 2022; Gorgoni et al., 2021; Mander et al., 2017). Firstly, the population in the MESA dataset is already in the elderly range and does not capture a full age spectrum. There are no 'young' participants to compare with, as in the other studies. Other factors include the differences in the sleep onset definition. I conservatively used the first continuous minute of stage N2 while often the first epoch of N1 or N2 is used. In addition, it could be associated with the differences in the dataset.

A general slowing down of the brain dynamics at the sleep onset state was found (negative correlations in beta band power and delta band peak frequency). This corresponds to the typical understanding that ageing populations have a slowing of the brain rhythms (Dauwels et al., 2011). Ageing brains also have disrupted weak nonlinear structure (especially in cognitive impairment and neurodegenerative disease), which results in a decreased complexity but a more noise-like and less nonlinear EEG structure Jeong (2004); Stam (2005); Delbeuck et al. (2003). This corresponds to the current finding that the auto mutual information (which measures the signal's nonlinear structures) decreases with age. These findings generally imply that the

typical ageing brain characteristics can also be observed at the sleep onset state. Of course, the correlations between age and these features were weak (and did not survive Bonferroni corrections), and it could be possible that GLM was biased due to the different number of samples for each age number. However, the results show that the behaviourally homogeneous state does have intrinsic dynamic differences (given the six features from GLM) related to ageing, and can be potentially extended for further phenotyping of the participants.

The general limitations listed above pertain to the specificity of the feature base and database and to the heterogeneity of the bedtime brain EEG applied. In addition, the finding was not validated in multiple datasets.

3.4.4. Other general limitations

The statistical metrics used in this chapter were mostly intuitive and simple, like linear kernel SVM and GLM models, along with pre-defined thresholds for choosing features. One could argue why not use more sophisticated models for better accuracy or predictions. More complex models mean less interpretable results. Deep learning models are great, but 'what' is learnt in the model is hard to interpret. The simple models do not provide great sensitivity, but combined with the feature base, they can provide significant statistical inferences about the important features, leading to an understanding of the basic principles and mechanistic interpretations. With the basic principles ready (e.g., in sleep onset latency, there are important EEG dynamics at bedtime to predict it, and it is independent of the sleep onset state), one can then deploy more complicated models to get more accurate predictions and make into engineering applications. Sometimes, it could be possible that the simple metrics cannot distinguish well (for instance, in the cases of gender, no features classified with above 55% accuracy), and it is possible to deploy more complicated models and comprehensive statistical tests (such as surrogate testings) to extract key features for more insights.

That being said, the statistical inferences in huge heterogeneous datasets are great for understanding general mechanistic principles, but it could still be beneficial to conduct more rigorous, well-controlled studies to gain better insights into the variabilities within and between subjects. In predicting the sleep onset latency, for instance, it would be better to phenotype the participants based on their bedtime EEG characteristics or other information (e.g., demographics), and for each type, there could be an independent prediction model.

There are many other questions regarding the bedtime and sleep onset EEG dynamics outside the demographics listed in this chapter. The participants studied in this chapter did not include those with extremely short or long sleep onset latencies (3-90 minutes range), and as previously mentioned, they could be studied themselves specifically, as narcolepsy or insomnia are both a kind of 'diseased' sleep initiation problem. With other datasets, it would be good to study

how different disease patterns change these dynamics and gain more insights into them (e.g., neurodegeneration, depression, brain injury), and provide more diagnostic values (Ogilvie, 2001).

4. Bifurcation brain EEG dynamics during falling asleep

4.1. Introduction

In the previous chapter, I investigated brain EEG at sleep onset, identifying the features that best distinguish the state from bedtime wakefulness. In this chapter, I aim to investigate the dynamics of the brain EEG that governs the falling asleep process, i.e., the transition from the wakefulness state to the sleep onset state.

Early EEG studies suggest that the falling asleep process is continuous and involves gradual changes in brain activity (Ogilvie, 2001; Merica and Fortune, 2004; Gorgoni et al., 2019) accompanied by behavioural and physiological changes (Ogilvie and Wilkinson, 1984; Ogilvie, 2001). The changes in EEG include the generation of alpha band bursts followed by the emergence of Vertex sharp waves and spindles (Hori et al., 1994) as well as spatiotemporal patterns across frequency bands (De Gennaro et al., 2001a; Gorgoni et al., 2020) (see Section 1.3 for detailed discussions). Most of the analyses, however, are bound to the sleep stages, to which most of their conclusions (in the EEG spectral changes) are statistically linked. The distinct dynamical patterns of different power bands (how they change temporally and locally, how they interact and link to other regions) are complex and hard to generalise as a whole. A study by Merica and Fortune (Merica and Fortune, 2003) provided an interesting insight that the transition to sleep involves distinct quantitative regional differences (e.g., different band-power changes), but is uniformly controlled by the brainstem temporally.

Conceiving the brain as a complex dynamical system (see Section 2.5), the wake-sleep transition can be considered as a phase transition. Underpinned by neurobiology, the wake-sleep transitions are governed by mutual inhibition of wake-promoting nuclei and sleep-promoting ones, with external regulations of the circadian rhythms and homeostatic forces (Saper et al., 2010; Sulaman et al., 2023). This mutual inhibition makes the actual wake-sleep transition very rapid compared to the entire sleep duration, with underlying neuronal firing dynamics switching from wake-promoting ones to sleep-promoting ones (Saper et al., 2010; Takahashi et al., 2010). Several

conceptual models have been developed to model this regulation, for example, the well-known two-process model (Borbély et al., 2016), or the flip-slop switch model (Saper et al., 2001). The two models well delineate the macroscopic dynamics of the wake-sleep regulation process, where the brain switches abruptly between these two states. This 'switching' process can be conceived as a type of bifurcation (or critical transition), with bi-modal brain dynamic patterns and inaccessibility between these two states (de Mooij et al., 2020). Still, the microscopic dynamics in the timescale of seconds or minutes cannot be described in the model. Based on neurological theories, several neurophysiological models (especially neural mass models) have been developed and demonstrated that the wake-sleep transition is theoretically a bifurcation process (Phillips and Robinson, 2007; Robinson et al., 2011; Yang et al., 2016). These models provided valuable mechanistic insights into the neuronal dynamics during the wake-sleep transition and the theoretical foundation for describing the dynamics as a bifurcation.

However, to date, there has not been experimental evidence that the human falling asleep process is a bifurcation dynamic. In the neurophysiological models, there are many simplified assumptions, for instance, the neuronal firings of a nucleus are summarised into one differential equation, with simplified inputs and outputs between different nuclei as one inhibitory or excitatory parameter. There are also numerous brain nuclei not included in the models, and the complex interactions are simplified. The dynamics of the model can also be easily observed in the firing outputs of the nuclei in the model. There are thus many difficulties in demonstrating such dynamics in actual brain EEG data: the quantification of the high-dimensional system state variables, the measurement of the underlying sleep control parameters that drive the bifurcation process, and the unclear governing law of the brain as a complex dynamical system itself. One exception is a recent study by de Mooij et al. (de Mooij et al., 2020) who examined a bifurcation dynamic in the sleep EEG of two subjects. The study did not find bifurcation but the small number of subjects and the small EEG feature base (a subset of power spectrum features) limited the investigation.

What are the benefits of describing the brain state transition with a bifurcation dynamic (i.e., a critical transition)? One potential benefit is an empirical definition of sleep onset outside the framework of sleep staging (de Mooij et al., 2020). A second potential benefit is the ability to track and predict the sleep onset transition as was previously shown in epileptic seizure onset prediction (Stam, 2005; Yuan et al., 2012; Lopes da Silva et al., 2003; McSharry et al., 2003). One important characteristic of a bifurcation dynamic is the critical slowing down (Scheffer et al., 2009, 2012). A critical slowing down implies that when a system is close to a bifurcation (i.e., the disappearance of an old state and the emergence of a new state), the system is more easily perturbed (due to a weaker state attractor) resulting in a larger variance and higher autocorrelation due to small perturbations. These so-called 'early warning signals' (EWS) can help predict the critical transition before it happens. A critical slowing down has been widely observed in dynamical systems such as ecological systems (before species extinction, van Nes

and Scheffer (2005)) or financial markets (before the financial crisis, May et al. (2008)). It was also been used in the prediction of epileptic seizures (de Mooij et al., 2020).

In this chapter, I test whether the changes in the brain activity during the falling asleep process have a bifurcation dynamic by examining the wake-sleep trajectory in the feature state space. The state space analytic approach was previously applied in other studies as well, for instance, Behn et al. (Diniz Behn et al., 2010) used EEG band power features to delineate the wake and sleep states as two distinct clusters in the feature space (bimodality as previously mentioned). Here, I use the more comprehensive feature base as described in Chapter 3 and the large MESA open-source sleep EEG dataset (including more than a thousand subjects). I then use dimensionality reduction strategies to identify the EEG features dynamics that underpin the bifurcation dynamic. Finally, I explore whether EWS could predict the participants' transition to sleep. By developing a model that represents the dynamic by which the brain EEG changes from wakefulness to sleep onset, I aspire to support the development of closed-loop neuromodulation intervention for sleep onset insomnia (the overarching goal of the thesis).

4.2. Methods

In this section, I will introduce the main methodologies used in this chapter, specifically, the conceptualisation of the feature state space, the bifurcation model used, governing feature dynamics analysis, and the critical slowing down theories. Most of the analyses were done on the group level dynamics, with only early warning signals done on individual levels. Some post hoc analyses or statistical tests based on observations of the data are introduced along with the results.

4.2.1. Feature dynamic time series during falling asleep

Transformation of the EEG into feature time series dynamics

The first step was to transform the EEG time series during the falling asleep transition into the feature space. This step can be visualised as in Figure 4.1. The same feature base was used (47 features). To construct the continuous feature time series, the entire falling asleep period for every participant (from bedtime to sleep onset) was taken, with an additional 10 minutes (min) taken after sleep onset. The EEG period was then divided into several 6-second (s) mini-epochs with 50% overlapping (3 s). Unavoidably, there were artefacts inside this period of EEG data. Thus, to exclude artifactual mini-epochs, a similar RMS-based method (see Section 2.2) was applied, but here the RMS values were evaluated on every 6s mini-epochs and epochs with RMS values above 2.5 the median of all were excluded (marked as NaNs). Similar to previously,

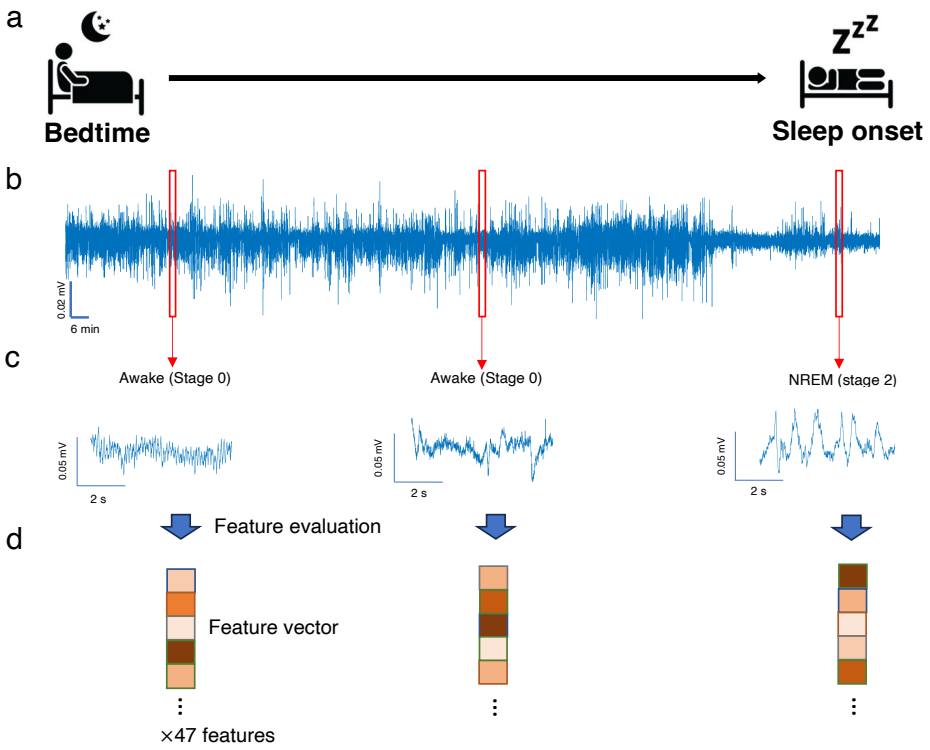


Figure 4.1: Transformation of EEG signals into the feature state space. (**a**) illustrates that the EEG time series during the falling asleep period (from bedtime to sleep onset) are taken with exemplary EEG traces in (**b**). Then, for each 6s EEG mini-epoch (exemplary traces in (**c**)), all 47 features are extracted (in (**d**)) and thus concatenated into a multi-dimensional feature dynamic time series.

artefacts were marked consistently across all 3 EEG channels, which means that if any mini-epochs were marked as artefacts in any channel, the mini-epochs at the same timestamp in the other channels were marked as NaNs as well. Then, for every clean mini-epoch, all 47 features were extracted. By doing so, it is conceptualised that for every mini-epoch, the features represent their dynamics at that specific timestamp. By concatenating all the features at all timestamps, the multi-dimensional feature time series was derived (sampled at 3s intervals). This was done for every EEG channel (3 linked channels in total) of each participant. The feature vectors between different channels were then averaged to get the global brain dynamics, and for each participant, there was a 47-dimensional feature time series capturing the global dynamics while falling asleep (imagine as a 47 by T dimension feature matrix, where N is the number of participants and T is the number of timestamps). Each feature dimension can be considered to capture the dynamics of one aspect of the system (the brain), and more than just Fourier-based ones as well.

Before doing further analysis, the features were z-score normalised globally. This means that for each feature, the feature values from all timestamps of all participants were used to calculate the mean value and standard deviation, and then all feature values were z-score normalised to

them. By doing so, different feature dimensions are comparable to each other without relative magnitude differences and also without destroying inter-subject differences. All later analyses used the z-score normalised feature time series unless otherwise specified.

Group-averaged individual feature dynamic time-series

For each participant, there was a 47-dimensional feature time series matrix capturing their full falling asleep EEG dynamics (from bedtime to 10 minutes after the marked sleep onset). However, as one can see, different participants had different lengths of falling asleep periods (different latencies). Besides, there were missing values (artifactual epochs) inside each time series which further complicates the story. To study the group-level dynamics, it is therefore important to align the data and keep consistent sample numbers at each timestamp.

The most common method used in past literature to deal with the inconsistent period length is to divide the entire period into equal percentiles, and average within each percentile to get representative dynamics (Aeschbach and Borbély, 1993; Marzano et al., 2013). Although they normally defined a starting point of their falling asleep period with some level of certainty that the 'transition' might already started, this method will surely distort, especially the nonlinear temporal structures of the original dynamics. Therefore, to get the average feature time series dynamics, the feature time series from all participants were firstly aligned to their sleep onset point. Then for each participant, their entire feature time series was normalised to their sleep onset by subtracting the median of their 1-minute sleep onset values (since the features were already globally z-score normalised). The feature time series dynamics were then the dynamic changes relative to their sleep onset respectively. Since participants had different sleep onset latency (3-90 minutes in the MESA dataset after exclusion), the number of data samples decreased with time going further from the aligning point. To ensure correct statistics, a minimum of $N = 200$ data samples were demanded for each timestamp, and thus the dynamics before -73.1 minutes (minus sign means before sleep onset) were ignored. Then for each timestamp, I bootstrapped all data samples available to ensure that only 200 samples were used for evaluating the mean feature value. This was done for all 47 features separately. The bootstrapping was done randomly five times to ensure sufficient data representation at each timestamp. I found that the average group-level dynamics was stable with $N = 200$. I defined the first epoch of the 1-minute sleep onset as time zero and focused on 60 min before sleep onset.

4.2.2. Conceptualisation of a dynamical system in feature state space for analysis of bifurcation

Figure 4.2 illustrates the concept of a dynamical system in the feature space state for the discussion of bifurcation. In the high-dimensional feature state space (or a manifold), the EEG

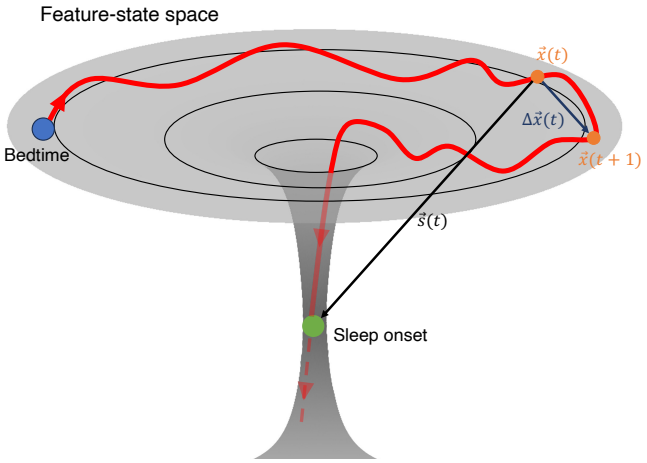


Figure 4.2: A conceptualisation plot of the falling asleep transition in feature state space from bedtime to sleep-onset. On the trajectory of falling asleep (red curve), $\vec{x}(t)$ represents the feature vector (feature state variable) at time t, where the vector from $\vec{x}(t)$ to $\vec{x}(t+1)$ is the state-space velocity $\Delta\vec{x}(t)$. The state variable might go through a manifold in the state space rather than going directly to the sleep-onset point (green point), therefore we define another variable $\vec{s}(t)$, which is the direct vector from $\vec{x}(t)$ to sleep onset point in the feature space (the absolute value is, therefore, the Euclidean distance $s(t)$).

brain dynamics start from the bedtime state (blue point) and move (red curve) till reaching a tipping point of the bifurcation dynamic in which it is attracted (pulled) towards the sleep onset point (green). The trajectory is illustrative only, and different participants could follow different paths. The state space vector $\vec{x}(t)$ is the instantaneous location in the feature space (relative to the starting coordinates). The time derivative of $\vec{x}(t)$ is the state space velocity vector denoted as $\Delta\vec{x}(t)$. Note that the state variable could go through a complex manifold before reaching the sleep tipping point. The sleep state vector $\vec{s}(t)$ is the vector between the instantaneous location x(t) and the sleep onset coordinates. The size of the sleep state vector $\vec{s}(t)$ (i.e., denoted as a scaler $s(t)$): $s(t) = \|\vec{s}(t)\|$ is the Euclidean distance to the sleep onset centroid. This variable is later called the sleep state variable, which captures the dynamics towards the sleep onset.

Evaluations of the variables and group-averaged dynamics

For each participant, the globally z-score normalised, 47-dimensional feature time series were used. The $\vec{x}(t)$ was then the feature vector at timestamp t. The feature state space velocity vector was evaluated by direct subtraction. The absolute velocity (or distance of change) was thus the Euclidean distance from $\vec{x}(t)$ to $\vec{x}(t+1)$ and denoted as the vector norm of $\Delta\vec{x}(t)$ as $\|\Delta\vec{x}(t)\|$. To evaluate the vector $\vec{s}(t)$, the sleep onset point vector needs to be defined. To ensure smooth data and avoid abrupt zeros in the final evaluation of $s(t)$, the median vector of the 1-minute sleep onset and 10-minute post-asleep period (11-minute period in total) was taken as the representative centroid of the sleep-onset state. $\vec{s}(t)$ was then calculated as the subtraction of the sleep onset centroid vector to the current state vector, with the Euclidean distance as the

norm of $\vec{s}(t)$. All these evaluations were done in MATLAB with Euclidean distance evaluated using function `pdist2`.

The previous actions were done for each participant separately, and here, aside from the feature time series, there was a feature state space velocity time series $\|\Delta\vec{x}(t)\|$ and a distance to sleep onset time series $s(t)$. To get the group-averaged dynamics of these two variables, the same bootstrapping was applied here, where for each timestamp, there were $N = 200$ samples for averaging, to get the averaged time series. This bootstrapping was done five times as well. As a side note, the random bootstrapping indexes were kept fixed across different analyses (with random number generators having fixed seeds). Similarly, time series data was truncated to start from 60 minutes before sleep onset to 10 minutes after. Later, the time series $s(t)$ and $\|\Delta\vec{x}(t)\|$ will just be denoted as the s and $\|\Delta\vec{x}\|$ for convenience.

4.2.3. Theoretical dynamical system model for bifurcation analysis

Description of the model

The basic dynamical system theory has been introduced in Section 2.5. To systematically explain and test whether the falling asleep transition dynamics can be modelled as a bifurcation, it needs to be compared and fitted to a theoretical model. Here, I chose a catastrophic bifurcation model (there are other types of bifurcation models). The model chosen here came from a well-studied ecological system describing the catastrophic outcomes of over-exploitation (May, 1977; Noy-Meir, 1975):

$$\frac{dx}{dt} = \left(rx(1 - \frac{x}{K}) - c\frac{x^2}{(x^2 + h^2)} \right), \tag{4.1}$$

where x denotes the system state, c denotes the control parameter of the system. r, K, h are fixed constants related to the system's nature: r is the growth rate of the resource biomass (reproduction rate), K is the ecosystem's maximum carrying capacity (maximum population without external interventions due to limited natural resources), and h is the half-saturation constant (fixed constant when r is half of the maximum value). The detailed interpretations of these constants will not be further discussed here, as they were developed in the context of ecological harvesting systems. Conceptually, the gradual increase in the grazing rate (c) leading to the catastrophic bifurcation in the system could simulate the dynamics of the falling asleep transition, during which there would be a gradual accumulation of sleep pressure as well as decreased wake tones. These constants of course will determine the shapes of the temporal dynamics and the optimal values will be fitted from the empirical data, but the exact meanings in this context are by far unexplainable. The model used in the current context is purely descriptive, to simulate the bifurcation dynamics of the falling asleep transition, and it could be a different dynamical system with similar bifurcation behaviour. The important point is that

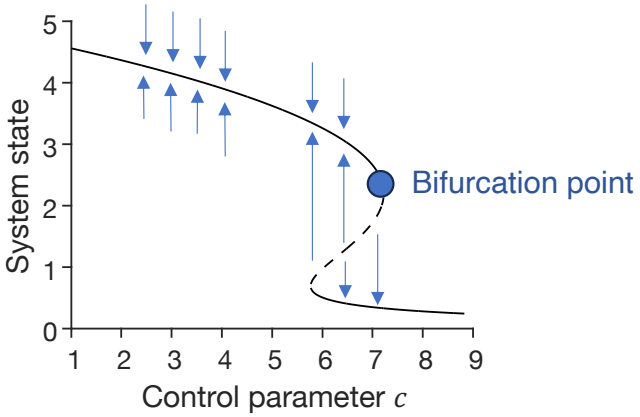


Figure 4.3: A simulation of the theoretical model demonstrating bifurcation. Key parameters: $r = 5.73$, $K = 4.73$, $h = 0.58$, $m = 0.11$ (Optimal parameters fitted later to sleep state variable s). Solid lines indicate stable system states (attractors), while dashed lines indicate unstable ones, with arrows indicating the direction of attraction. The bifurcation point (with critical values) is marked as a blue solid point in the figure.

this model provides an intuitive understanding of this process, and the dynamical patterns are widely generalisable across different bifurcation models (Maturana et al., 2020).

The bifurcation dynamics are controlled by the control parameter c, i.e., the gazing rate in the original model. Higher grazing rates lead to a reduction of resource biomass x, and when it reaches the critical threshold, the system goes through a catastrophic bifurcation resulting in species extinction without the capability to return to the stable biomass state. The model simulation can be seen in Figure 4.3 (the plot is typically called a bifurcation diagram), illustrating how the dynamics of the system state x change along with control parameter c. This type of bifurcation is called 'catastrophic' since when c crosses the critical point, the resources will catastrophically die out. 'Catastrophic' also implies that when the bifurcation happens, the system state cannot immediately return to its previous stable state by reducing c back. This is called the hysteresis phenomenon, where the transition to and back different system state attractors across the bifurcation point do not follow the same path, and there are time delays before one can transit back to the previous state (a transition barrier). This is intuitive in the context of species, when people harvest hard enough and the species is going to die out, they are not going to return to normal even if people stop by the next day.

In the wake-sleep transition context, it implies that if the falling asleep transition is catastrophic, then the reversed waking-up process is not straightforward and that a return will follow a different path. The phenomenon of sleep inertia immediately after waking up, where the cognitive performances are much worse than the normal wakefulness state (Ferrara and De Gennaro, 2000; Tassi et al., 2006; Lovato and Lack, 2010), is the direct evidence of hysteresis. Such hysteresis in the wake-sleep transition has also been shown in neural modelling (Phillips and Robinson, 2007; Robinson et al., 2011; Yang et al., 2016). Hwang et al. demonstrated the same pattern

in the transition to and from anaesthetic states (Hwang et al., 2012). Using functional magnetic resonance imaging (fMRI) data, Stevner et al. showed that wakefulness-after-sleep is an independent state from the normal wakefulness state before sleep (Stevner et al., 2019).

For the bifurcation to happen in the model, the change of the control parameter is the most important underlying drive. To fit the model into the falling asleep transition dynamics and simulate the bifurcation dynamics, the control parameter is assumed to increase linearly over time:

$$\frac{dc}{dt} = m, \tag{4.2}$$

where m is a constant controlling how sharply it increases. It can be conceived as during the falling asleep transition, the sleep forces increase linearly until the bifurcation threshold, where the wakefulness state will collapse into the sleep state.

Model fitting and control parameters

The model fitting to the empirical data (the sleep state variable s) was done in MATLAB. The two ordinary differential equations 4.1 and 4.2 were solved using MATLAB function `ode45`, with initial state x equivalent to the starting point of s ($t = 1$), and initial $c = 1$. The optimal parameters r, K, h and m were then found based on minimised root mean squared error, using MATLAB function `fminsearch` to search the optimal values. The model performance was then quantified via the R-squared value: $R^2 = 1 - RSS/TSS$, where RSS is the total sum of squares of residuals, and TSS is the total sum of squared raw values. The higher the R-squared, the more variance the model is capturing, and thus the better fitted the model is.

The most important value in the fitting is m, which determines how fast the control parameter is changing. The time series of the control parameter is thus $c(t) = mt + 1$, and the similarity between $c(t)$ and feature time series was analysed to see which feature best represents the control parameter in the system. To test the similarity between $c(t)$ and the 47 group-averaged feature time series (see previous), two metrics were used: the normalised inner product and the dynamical time warping (DTW) distance. The absolute normalised inner product (similarity score) between the two time series is given by:

$$a_i = \left| \frac{c(t) \cdot x_i(t)}{\|c(t)\| \times \|x_i(t)\|} \right|, \tag{4.3}$$

where a_i is the similarity score for the i-th feature, and $x_i(t)$ is its feature time series. The score is between 0 and 1, where closer to 1 indicates higher similarity. The inner product of the two time series captures the similarity at exactly the same time points. DTW, however, is a better metric which captures the same dynamical pattern which might occur at different times or speeds for two time series (Berndt and Clifford, 1994). In mathematical words, the

DTW algorithm nonlinearly stretches the signals until the Euclidean distance between them (the sum of every timestamp) is minimised. Thus, DTW is more sensitive in capturing the same dynamical patterns in the time series. The DTW was implemented using MATLAB function `dtw`. The feature with the maximum similarity score and minimum DTW distance was then taken as the most representative feature as the control parameter of the system. The empirical bifurcation diagram between the sleep state variable s and the control parameter (the feature dynamics) was then plotted to observe the bifurcation dynamics.

4.2.4. Dimensionality reduction via functional principal component analysis (FPCA)

Each feature time series measures a specific dimension of the brain dynamics, so it would be interesting to test causal interactions between these different dimensions. 47 features were very high-dimensional, therefore Functional principal component analysis (FPCA) was applied for such a dimensionality reduction. FPCA is a specific method of functional data analysis (Wang et al., 2016), which is fundamentally a principal component analysis but for functional data observed continuously in time or space. FPCA analysis uses a combination of different basis functions (linear, polynomial or splines) to simulate the time series dynamics and identify intrinsic patterns of the time series, which can be used for reducing dimensionality (finding key patterns) or comparison analysis. FPCA has the mathematical foundations from the Karhunen-Loève transform (Maccone, 2009), which decomposes a signal into a linear combination of a series of orthogonal bases (derived from the covariance matrix of the signal). FPCA is an extension of this transformation into discrete data.

FPCA was implemented using the existing Python toolbox 'scikit-fda' (`https://fda.readthedocs.io/en/latest/index.html#`). To identify the specific dynamical patterns of all 47 features during falling asleep, the FPCA was applied to the group-level averaged feature time series before the sleep onset point (see previous method, the additional 10-minute post-asleep period excluded here), with an L2 regularization (also called the Ridge regularisation) applied. Similar to normal principal component analysis, the selection of the functional principal components (FPC) was based on the variance explained. To map the FPCs back to the features (which features best represent the FPC dynamics), the similarity scores between the FPCs and the feature time series were computed as a normalised inner product between them:

$$score = \frac{x_{ft} \cdot x_{FPC}}{\|x_{ft}\| \times \|x_{FPC}\|}, \tag{4.4}$$

where x_{ft} is the feature time series, and x_{FPC} is the FPC time series. The score ranges from -1 to 1 without taking absolute values, since this similarity can both be positive and negative. The features with maximum and minimum scores were selected for each FPC.

The selected features were the features with the most representative temporal dynamics that

capture the falling asleep transition process, therefore, their dynamics were plotted out in 2-D feature state space to observe the transition (between each pair of them). The averaged and global z-score normalised feature time series were used. To simplify the plot and visualise the directions and trajectories of the dynamic changes, the median point of every 30s segment (10 feature samples at 3s sampling intervals) was taken and scattered on the figure, with the direction vector as the median vector from all data samples in the current 30s median to those in the next one. Therefore in the plot, one could observe the dynamical changes at a 30s interval, along with the averaged directions of change.

4.2.5. Granger causality analysis in time-series dynamics

As specified before, FPCA was applied to reduce the dimensionality of the feature time series to find key representative patterns of changes during falling asleep. It is thus interesting to test and infer the causal interactions between these different dimensions of changes, which can provide more mechanistic insights into the brain's dynamical interactions, and identify 'sources' of change for tracking this process or closed-loop interventions. Here, I used Granger causality (GC) to test the causal interactions. GC tests whether the past of one time series is useful in forecasting the future or another. Of course, there are many other methods to test different causal inferences, for instance, the transfer entropy, although they measure fundamentally similar properties (sharing information and future predictions) (Barnett et al., 2009).

Given that most of the feature time series were non-stationary and cointegrated (with long-term correlations between them), standard GC testing approaches directly using the vector auto-regression (VAR) model were not valid. Therefore, the Toda-Yamamoto approach (Toda and Yamamoto, 1995) was applied to deal with causality testing in nonstationary and cointegrated data. For the given set of feature time series (K features), the exact steps for testing follow (see the original paper as well as the detailed tutorial given by Giles (Giles, 2011)):

1. Test each feature time series for their order of integration using both Kwiatkowski–Phillips–Schmidt–Shin (KPSS) tests and the Anderson-Darling test with a p-value threshold of 0.05; The order of integration means the minimum number of differentiations needed to make the time series stationary in covariance measurements. Note down the maximum order of integration of all K time series as m.
2. Set up a VAR model using MATLAB function `varm` on the K feature time series. Akaike information criterion was then used to select the optimal model order p_0. The Akaike information criterion quantifies the information loss of the model, and smaller values indicate a better model.
3. To make sure that the VAR model was well specified, the residual autocorrelation for each feature time series was tested using Ljung-Box Q-test (MATLAB function `lbqtest`), and

model order p_0 was increased until there was no significant residual autocorrelation (p-value threshold 0.05) for every time series. Denote the additional lags required as p_1.

4. The cointegration between time series was then tested to see whether the VAR model was stable at the current time series level. The Johansen cointegration test was used, which converted the VAR model into a vector error-correction (VEC) model and tested the traces or maximum eigenvalues of the error-correction matrix. If the error correction matrix of the VEC model is full rank, then the K time series must have cointegration relationships, and the VAR model is stable at this level (otherwise, further differentiation of the time series is required before fitting the VAR model). The existence of cointegration also implies that there must be causalities between two or more time series in the model.
5. Finally, the pair-wise Granger causality was tested between each pair of the feature time series (conditioned on remaining time series), using MATLAB' s function `gctest`, where the final VAR model order is $(m + p_0 + p_1)$. The final statistics excluded the extra m lags in the model as they were counted for time series integration.

Then between each pair of feature time series, there was a p-value from the test indicating Granger causality. I took a p-value threshold of 0.025 as the statistical significance level (corrected for each pair of tests) and the significant causal relationships were plotted out in arrows from one to the other. The test here was conducted only test pair-wise causality for easy interpretation to start with. One could conduct tests for more complicated block-wise interactions to test causality from a bunch of time series to another bunch.

Aside from this method, there are many other existing toolboxes for doing the GC test. One example is the MVGC toolbox (Barnett and Seth, 2014), which can test causality for multiple realisations of each time series process (each feature here). The GC can also be done on the frequency domain to test causality in a specific frequency band. The theories behind the testing and different tools are outside the scope here, with the method above easy and transparent to use in inferring basic Granger causal interactions.

4.2.6. Critical slowing down and its early warning signals (EWS) prior to bifurcation

The basic critical slowing down theory

As outlined before, when the dynamical system is close to a catastrophic bifurcation, there will be a critical slowing down. The critical slowing down is hallmarked by more vulnerability (less stability) to external perturbations (Figure 4.4ci), where the system dynamics will have increased autocorrelation and variance (Figure 4.4cii and ciii). The underpinning reason behind these dynamics is the change in the system's attractor landscape (Figure 4.4b), where closer to bifurcation indicates a shallower state attractor. The exact mathematical proof will not be

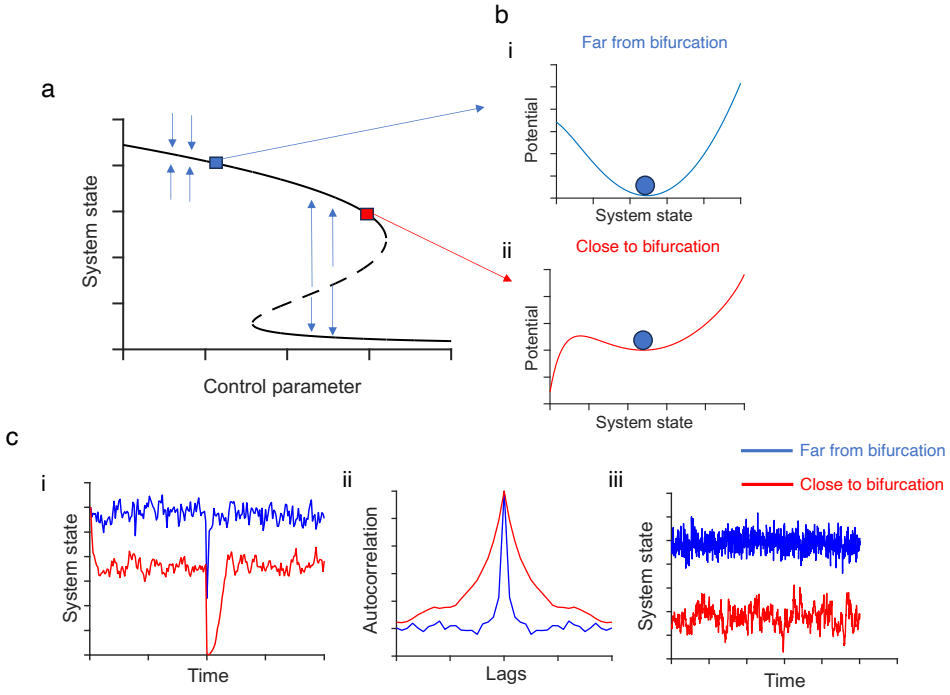


Figure 4.4: Illustration of critical slowing down prior to bifurcation and its early warning signals (EWS). a. The theoretical bifurcation diagram of the same system as in Equation (4.1). When the system is far from bifurcation (blue square), the system state attractor will be deep and stable (**bi**), while when close to bifurcation (red square), the attractor will become shallow and almost gone with new one emergent (**bii**). **c**. Comparison of system dynamics between far from bifurcation state (blue) and close to bifurcation state (red). It can be seen that when the system state is closer to bifurcation, it will be more vulnerable to perturbations (i), with increased autocorrelation (ii), and increased system state fluctuations (iii). The simulations are done in MATLAB, the axes units are arbitrary here for illustrative purposes.

explained here, but can be found in Scheffer et al. (2009). The model shown in the figure is the same as the previous one, with the landscape approximated via the 15th-order Taylor expansion of the differential equation.

It is important to note that, although critical slowing down is a very generic indicator for a lot of critical transitions (catastrophic bifurcation is simply one type), there are exceptions in natural systems (Scheffer et al., 2009). For instance, the previously discussed increased autocorrelation and variance are based on the fact that the system is driven towards bifurcation at a moderate speed (with changes in the underlying control parameter). However, it is possible that the system suffers from a very huge perturbation with catastrophic outcomes where the system state cannot return to the original attractor. Maturana et al. have demonstrated that this is a potential path to epilepsy already (Maturana et al., 2020). Aside from critical slowing down, when the system is under highly stochastic noise, the system dynamics could flicker towards another state (called flickering) (Scheffer et al., 2012; Dakos et al., 2013). Flickering means that the system state fluctuates back and forth between two states (due to large stochastic noise effects) until being trapped stably in one of the states. Statistically, the same properties (increased autocorrelation

and variance) still apply in such systems, as fundamentally, it is still a critical transition. The general bifurcation theory is a huge topic that will not be covered here, but in principle, in highly connected homogenous systems such as the brain, critical transitions are expected and critical slowing down should be empirical indicators in those transitions (Scheffer et al., 2012).

Computation of EWS in brain EEG during falling asleep

The critical slowing down theories are mostly illustrated in simple and theoretical systems such as the harvesting model used. In natural systems, the inferences about upcoming critical transitions could only be made by computing the early warning signals (such as autocorrelation and variance) in the system's time series observed. For example, in the harvesting system, the actual time series that describes the system state could only be the resources biomass recorded over time. Aside from autocorrelation and variance, there is a rich plethora of methodologies that can be applied to help detect the tipping point in the time series for 'early warning' of the upcoming bifurcations. For instance, the detrended fluctuation analysis (Peng et al., 1995), or the model-based state space auto-regressive analysis (Ives and Dakos, 2012), see Dakos et al. (2012) for a detailed review of all the methodologies. In this chapter, only autocorrelation and variance were computed on the feature time series during falling asleep as they are the two most solid and straightforward signs (Scheffer et al., 2009).

There are many challenges in using EWS to detect tipping points in empirical data and systems, for instance, the statistical difficulty in determining the exact point of critical transition, and the technical difficulties of different parameters and pre-processing of the data. Detecting falling asleep could be more difficult due to the lack of ground truth in actual sleep onset transition aside from the sleep stages, which are insufficient and inaccurate in both temporal resolution and subjective experiences (Hermans et al., 2019). More discussions will follow after the results.

In this chapter, the main aim is to empirically demonstrate that critical slowing down exists in the feature state space dynamics during the falling asleep transition at the group level, which could be further extended to the prediction of the transition prior at the individual level. For each participant, the evaluation of the EWSs was done on the sleep state variable time series s, which was the variable that shows the bifurcation dynamics. Only 20 minutes before the sleep onset was taken for this evaluation, based on the empirical observation that the brain dynamics remained almost unchanged before this point, and it was sufficiently long to observe the increase of autocorrelation and variance. The post-asleep period was removed here as it would confound the final result. Before the actual computation, the participants were further filtered through, to ensure good data quality in the last 20 minutes. As mentioned previously, artifactual epochs will leave NaNs in the feature time series. Hence, the participants with more than 15 continuous NaNs (45s continuous artifactual EEG data), with total artifactual epochs taking more than 10 percent of all, and with a falling period of shorter than 20 minutes, were excluded from further

EWS analysis. After exclusion, there were in total $N = 210$ participants left for this analysis. The two EWSs, autocorrelation and variance, were then evaluated on the 20-minute sleep state variable time series s for each participant.

The implementation of EWS computation followed in general Dakos et al. (2012). Before the evaluation, the missing values (NaNs) were interpolated using the median values of the samples around it (median filtering). The evaluations of autocorrelation and variance were done in a running-window approach with a window size of 5 minutes (100 feature time samples), with gaps between windows the same as the feature time series intervals (3s). This means that whenever a new feature sample comes, the autocorrelation and variance would be evaluated again in the current window. In each running window, the 5-minute feature time series were linearly detrended (using MATLAB function `detrend`) and the residuals were taken for calculating autocorrelation (using MATLAB function `autocorr`) and variance (using MATLAB function `std`). This detrending is important to avoid spurious increases in autocorrelation and variance due to changes in the mean values of the time series. After evaluations, there was an autocorrelation and variance time series returned for each participant's sleep state variable time series (which are 5 minutes shorter due to the first running window). To statistically test whether the autocorrelation and variance increased with time, the linear mixed effect (LME) model was used with the EWSs as the response variable and time as the predictor. Random effects of the participant were considered in the model, where time indexes were grouped by participant ID. The model statistics were evaluated using the analysis of variance (ANOVA), where the coefficient and p-value of the 'Time' variable would be indicative of whether the linear increase was significant over time. The EWSs were also compared to sleep stages, where they were averaged across all participants, and the changes were tested against the mean value of the first 3 minutes to see at which timestamp they started to increase significantly.

4.3. Results

4.3.1. The brain EEG of falling asleep has a catastrophic bifurcation dynamic

I first examined the group-level change in the sleep state variable $s(t)$, i.e., the instantaneous Euclidean distance to sleep onset centroid. The result is shown in Figure 4.5a. Note that the s time series was shifted downward to make the minimum value to be 0, and the time series was smoothed in a moving-median window of size 20 data samples for better illustration. It reveals that the distance to the sleep onset centroid does not monotonically decrease but instead drops rapidly 10-12 minutes before the sleep onset, independent of the sleep onset latency (right-tailed two-sample t-test to test with Bonferroni correction for multiple timestamps). The sharp drop in the sleep state variable resembles a bifurcation dynamic. To directly test whether the change

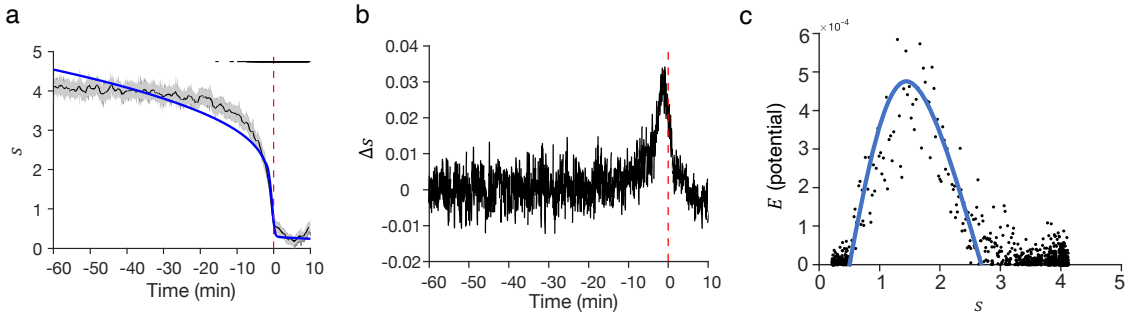


Figure 4.5: The dynamics of the sleep state variable can be described as a bifurcation process. **a**. The group-averaged temporal dynamics of sleep state variable s, shown are mean ± standard error of the mean (s.e.m.) across 200 samples. The blue line is the theoretical model with parameters optimally fitted to the empirical data. The horizontal lines show timestamps with a significant decrease compared to the mean values of the first 10 minutes, using the right-tailed two-sample t-test, with a p-value threshold of 0.025 dividing the total number of timestamps (for Bonferroni correction). The red vertical dashed line shows the sleep onset. **b**. The magnitude changes of s (Δs) as a function of time. **c**. The kinetics energy potential ($E = \Delta s^2/2$) as a function of s. The blue line in this subplot is for visual illustration only but not from the model.

is a bifurcation dynamic, I examined whether the data can be described with a bifurcation model (Equation (4.1)) with an assumption that the control parameter c increases linearly (the variable s was used without smoothing etc.). I set the initial condition of the system state x to be the starting value of s (i.e., s at -60 min point) and the initial value of the control variable c to 1. The other model parameters were then optimized to maximize the fit, resulting in the following optimal parameters: $r = 5.73$, $K = 4.73$, $h = 0.58$, and $m = 0.11$. See Figure 4.3 for the theoretical diagrams of these parameters. Here, the m parameter is important since it controls the rate by which the control parameter increases over time. The resulting bifurcation model is shown in blue in Figure 4.5a. I found that the bifurcation model accurately describes the dynamic of the sleep state variable with an R-squared of 0.94 and a root-mean-squared error of 0.1.

Figure 4.5b shows the change in the first-order difference Δs of the sleep state variable, i.e., the change in the distance to the sleep onset centroid per time step, corresponding to the velocity towards sleep onset. It shows a sharp acceleration of the sleep state variable velocity just before sleep onset.

Figure 4.5c plots $E = \Delta s^2/2$ related to the kinetic energy of the sleep state variable as a function of the sleep state variable. It demonstrates that the transition from wakefulness ($s > 3$) to sleep ($s = 0$) requires overcoming an energy barrier with a peak at a Euclidian distance at $s \approx 1.5$.

These results demonstrate that the sleep state variable s (i.e., the distance to the sleep onset centroid) does not change much till the bifurcation transition. Does the state variable x follow the same dynamic? Figure 4.6a shows the change in the time derivative $\Delta\vec{x}$ of the state variable, i.e., the change in the feature space per time step, corresponding to the velocity at an arbitrary direction. It reveals that the state variables decreased monotonically from about 20 minutes

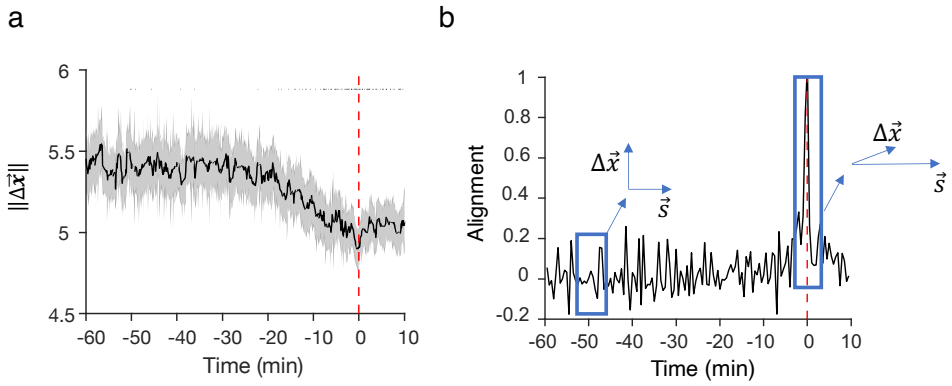


Figure 4.6: The dynamics of the feature state space variable. **a**. The temporal dynamics of the absolute feature state space variable velocity $\|\Delta\vec{x}\|$. Shown are mean ± s.e.m. across 200 samples. The horizontal lines on top show the significant changes compared to the mean of the first 10 minutes, using a two-sided two-sample t-test with a p-value threshold of 0.05 dividing the total number of timestamps. The red vertical dashed line shows the sleep onset. **b**. The temporal changes of the alignment between the feature state space velocity vector $\Delta\vec{x}$ and the vector towards sleep onset $\vec{s}$. The insets show that when they are not aligned (orthogonal), the alignment value will be around 0, while large values imply aligned directions.

before the sleep onset, independent of the sleep onset latency (right-tailed two-sample t-test to test with p-value Bonferroni corrected). This result suggests that the first stage of falling asleep transition is governed by fast movements in the feature state space (i.e., fast state variable) that is not aligned with the trajectory to the sleep onset centroid (i.e., the orientation of the sleep state vector). Before the sleep bifurcation, the movements in the feature state space slow down (slowing of the sleep state variable) and gradually align with the trajectory to the sleep onset centroid (i.e., the orientation of the sleep state vector), resulting in acceleration of the sleep state variable. To directly test whether the sleep bifurcation is associated with an alignment of the state variable velocity with the trajectory to the sleep onset centroid, I computed the dot product between feature state space velocity vector $\Delta\vec{x}$ and the vector towards sleep onset $\vec{s}$:

$$Alignment = \frac{\Delta\vec{x} \cdot \vec{s}}{\|\vec{s}\|} = \|\Delta\vec{x}\| \cdot \cos\theta, \tag{4.5}$$

where θ is the angle between these two vectors. The change in the vector alignment is shown in Figure 4.6b, confirming a sharp alignment at the sleep bifurcation. The result clearly shows that, although the brain dynamics were still changing, they were not directed towards the sleep onset until the very last moments of the transition, and the general dynamics of the brain were slowing down.

Finally, to validate that the above results are not confounded by setting the time zero at sleep onset, I repeated the analysis for half the period, where the s time series and $\|\Delta\vec{x}\|$ time series was aligned to half of the falling asleep period for each participant and then averaged. The results for the confounding checks can be seen in Appendix B.13. It can be seen clearly that no nonlinear dynamics could be observed in both plots. Even though the statistics showed some

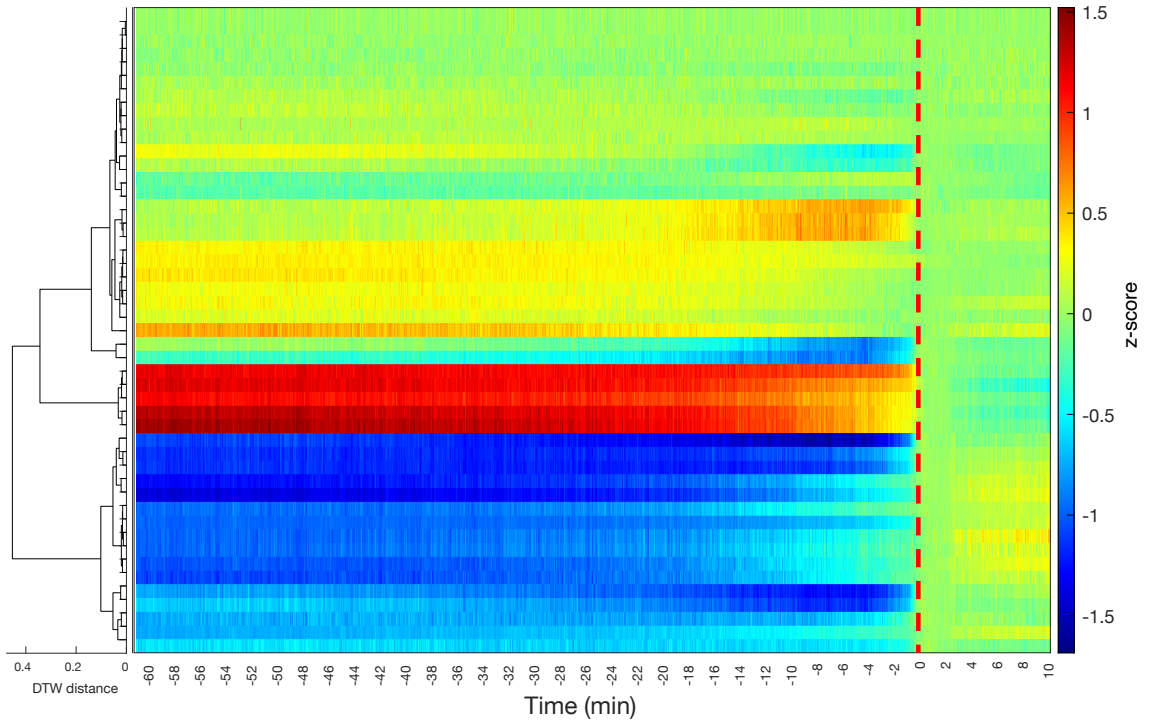


Figure 4.7: The feature dynamic time series during the falling asleep transition from 60 minutes before sleep onset to 10 minutes after. The feature values are globally z-score normalised and then normalised to the sleep onset (red line, $t = 0$), where colours show the z-score values. The 47 features are hierarchically arranged based on their DTW distance with the dendrogram on the left side. Refer to Appendix B.15 for the list of feature names in the hierarchical order. The vertical dashed red line shows the sleep onset.

significant changes compared to the first 10 minutes for s, the data was very noisy, and no clear patterns could be observed. I further separated the participants into a fast group (less than 40 minutes sleep onset latency) and a slow group (more than 60 minutes latency), and showed their sleep state variable dynamics. The group level sleep state variable s time series and $\|\Delta\vec{x}\|$ time series were evaluated in the same bootstrapping way with a minimum sample number of 150 for these two groups and plotted together (see in Appendix B.14). It can be found that the dynamics of the fast group followed the same path as that of the slow sleepers (although slightly noisier), with no qualitative differences. This implies that the bifurcation dynamics during the actual 'transition' are independent of the latency.

4.3.2. The key dynamical EEG features during falling asleep

After discovering that the changes in the brain EEG during the transition from wakefulness to sleep involve a bifurcation dynamic in the feature state space, I aimed to identify the EEG features that govern the bifurcation dynamic.

Figure 4.7 shows the time series of all the features (i.e., 47 features) with the feature values

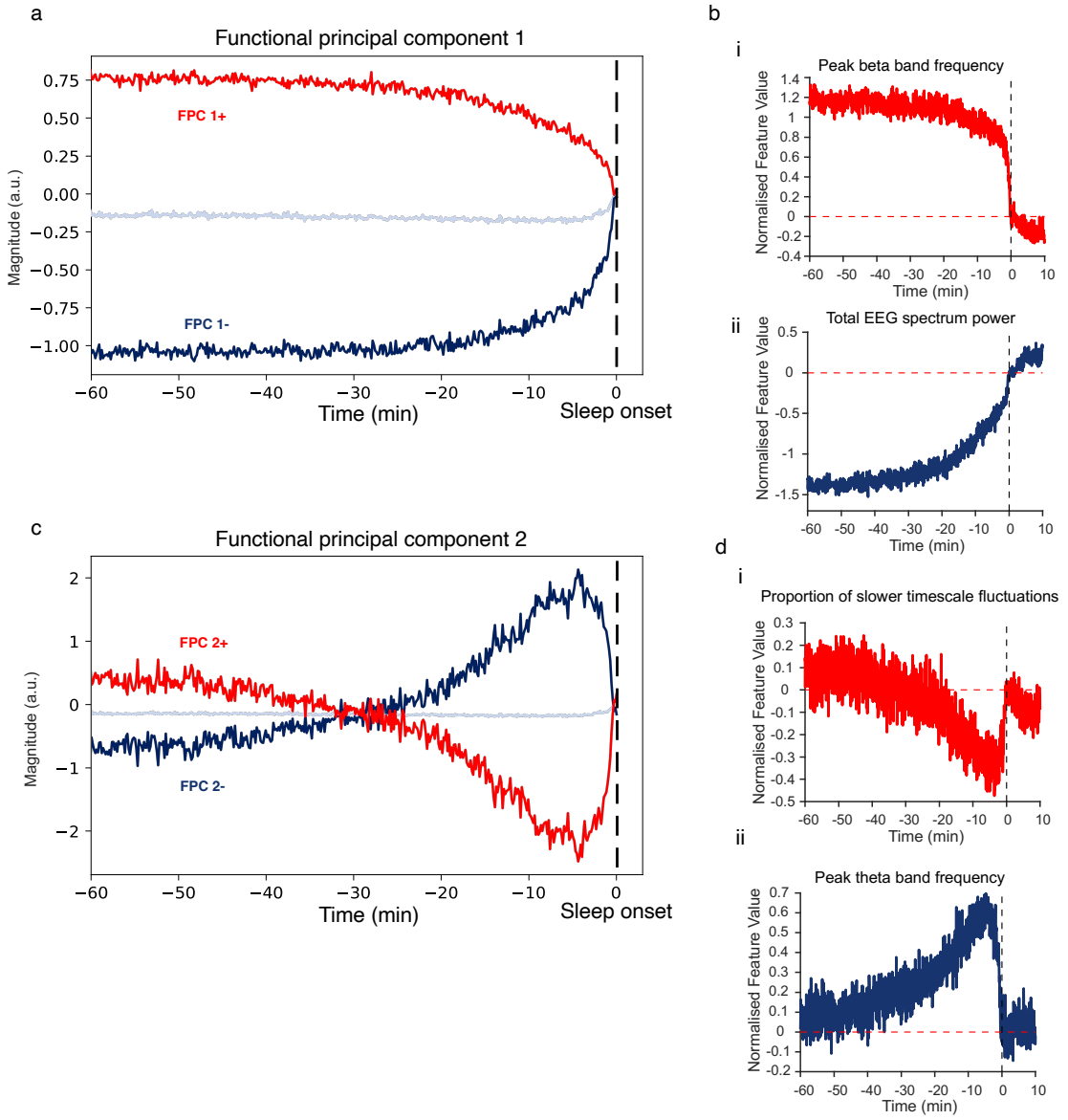


Figure 4.8: FPCA analysis in identifying key feature dynamic patterns and representative features. FPCA analysis is done on feature time series from 60 minutes before sleep onset up to sleep onset. Two functional principal components (FPC) are identified from the FPCA analysis, with FPC1 in (**a**) taking 96.03% of the total variance, and FPC2 in (**b**) taking 3.32%. Each FPC is shown in both positive (FPC+, red line) and negative (FPC-, dark blue line) directions relative to the mean of all feature time series (light blue line), with the y-axis having arbitrary units. The 'positive' and 'negative' means 'addition' or 'subtraction' of the actual FPC for the mean. For each FPC, the most representative features (both positive and negative) are selected based on the similarity score (maximum and minimum), and shown on the right. For FPC1, the most representative features are the peak beta band frequency (**bi**) for positive and the total EEG spectrum power (**bii**) for negative. For FPC2, the most representative features are the proportion of slower timescale fluctuations (**di**) for positive and the peak theta band frequency (**dii**) for negative. The horizontal dashed red lines indicate a 0 feature value, and the vertical dashed black lines indicate the sleep onset. See Appendix B.17 for similarity scores between the two FPCs and all features, as well as detailed explanations of the selected features.

normalised to sleep onset as before. The features were arranged in a hierarchical order based on their dynamical-time-warping distance (see methods), and neighbouring features with sim-

ilar patterns. The feature names list arranged in the same hierarchical order can be found in Appendix B.15. It shows that most of the features had constantly larger (lower) values but monotonically decreased (increased) shortly before the sleep onset, i.e., at the sleep bifurcation. In addition, a small subset of features had similar values as sleep onset apart from a local peak shortly before sleep onset, at the sleep bifurcation. Linking back to the results and discussions in Chapter 3, the features that had the burst type dynamics did not have good classification accuracy because the changes happen transiently before sleep onset. To validate that the above results are not confounded by setting time zero at sleep onset, I repeated the analysis for half the period (i.e., from -60 to -30 minutes, setting -30 minutes as time zero) as before. The result, which is reported in Appendix B.16, did not show clear patterns in the features' time series data.

To identify the most informative features, I deployed FPCA with the Ridge regularization to the features' time series prior to sleep onset, Figure 4.8. I found that the first functional principal component (FPC1), explaining 96.03% of the total variance, captured the monotonically decreased (increased), i.e., a dynamic similar to the sleep state space variable. To find the EEG features that best represent FPC1, I evaluated the similarity score of all the features to FPC1 and extracted the features with the highest and lowest scores corresponding to positive + and negative - directions (see methods). I found that FPC1 was best represented by the beta band peak frequency feature and the total EEG spectrum power feature in the + and –directions respectively, both with a similarity score of 1 (Figure 4.8b). These two features were found in Chapter 3 to best distinguish sleep onset state from bedtime, together with alpha-to-beta band ratio and the low-to-high band power ratio which also had high similarity scores to FPC1. In fact, many features in Figure 4.7 with monotonically decreased (increased) shortly before the sleep onset have a high similarity score to FPC1 (e.g., the power band features, entropy, etc.).

The second FPC (FPC2), explaining 3.33% of the total variance, captured the features with a local peak shortly before sleep onset (Figure 4.8c). I found that FPC2 was best represented by the theta-band peak frequency feature (-) (similarity score -0.77) and the proportion of slower timescale fluctuations (+) (similarity score 1) (Figure 4.8d, see Appendix B.17 for feature interpretation and al similarity scores). Specifically, before the bifurcation, there was a short increase in the peak theta band frequency and a decrease in the proportion of slower timescale fluctuations –implying a reduced predictability of the EEG signal.

The four features representing the FPCs, i.e., the beta band peak frequency, the total EEG spectrum power, the theta-band peak frequency feature, and the proportion of slower timescale fluctuations are later referred to as the 'key dynamical features'.

I next examined the relationship between the key dynamical features. Figure 4.9 shows the values of one key dynamical feature relative to another key dynamical feature (the four key representative features corresponding to FPC1± and FPC2±) over time. The nonlinear dynamic

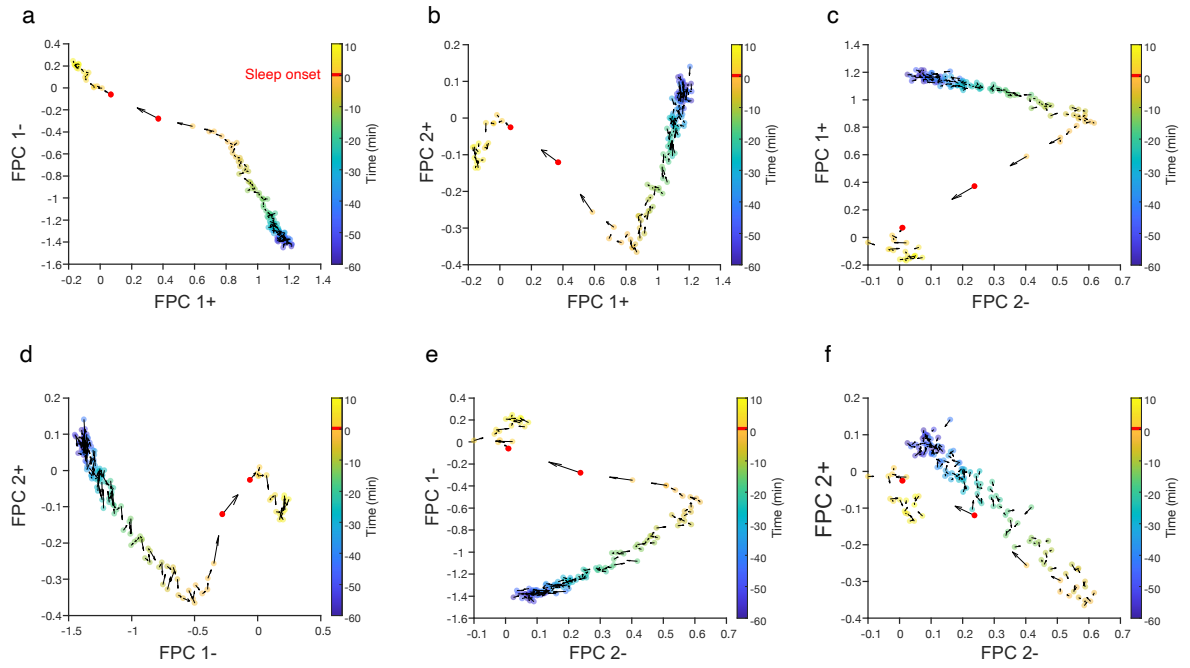


Figure 4.9: Transition dynamics in the 2D feature state space between each pair of the four key dynamical features. For the four key dynamical features, there are six combinations in total (shown from (**a-f**)), and the feature names are replaced by the FPC they represent. FPC1+: the peak beta band frequency; FPC1-: the total EEG spectrum power; FPC2+: the proportion of slower timescale dynamics; FPC2-: the peak theta band frequency. The colour shows the time, from 60 minutes before sleep onset to 10 minutes after. Each point in the figure represents the median of the feature values in every 30s segment. The arrows show the median direction vector from all data points of the current 30s segment to those in the next (with arrow length indicating the magnitude of the vector), with the 1-minute sleep onset marked in red. The x and y-axis units are all z-score normalised feature values showing relative changes to the sleep onset.

is evident in the accelerated change (small number of data points) at the sleep bifurcation. The panels in Figure 4.9b-e show a classical catastrophic bifurcation (folded system state variables). This further demonstrates that bifurcation dynamics observed in the sleep state variable were well underpinned by FPC1 dynamics (the dominant pattern), while although FPC2 took only a small variance, it is likely to represent the important control parameter of the system.

4.3.3. The key dynamical features govern the falling asleep bifurcation

I next aspire to identify the features that control the bifurcation dynamic of the sleep state variable by examining the relationship between the key dynamical features and the control parameter time series c of the bifurcation model (i.e., $c(t) = mt + 1$ with $m = 0.11$, Figure 4.5a). Figure 4.10 shows the values of the key dynamical features relative to the values of the bifurcation control variable. To quantify the level of similarity to the bifurcation control variable, I computed the absolute similarity score and the dynamical-time-warping (DTW) distance (see methods) between each key dynamic feature and the bifurcation control variable. The results

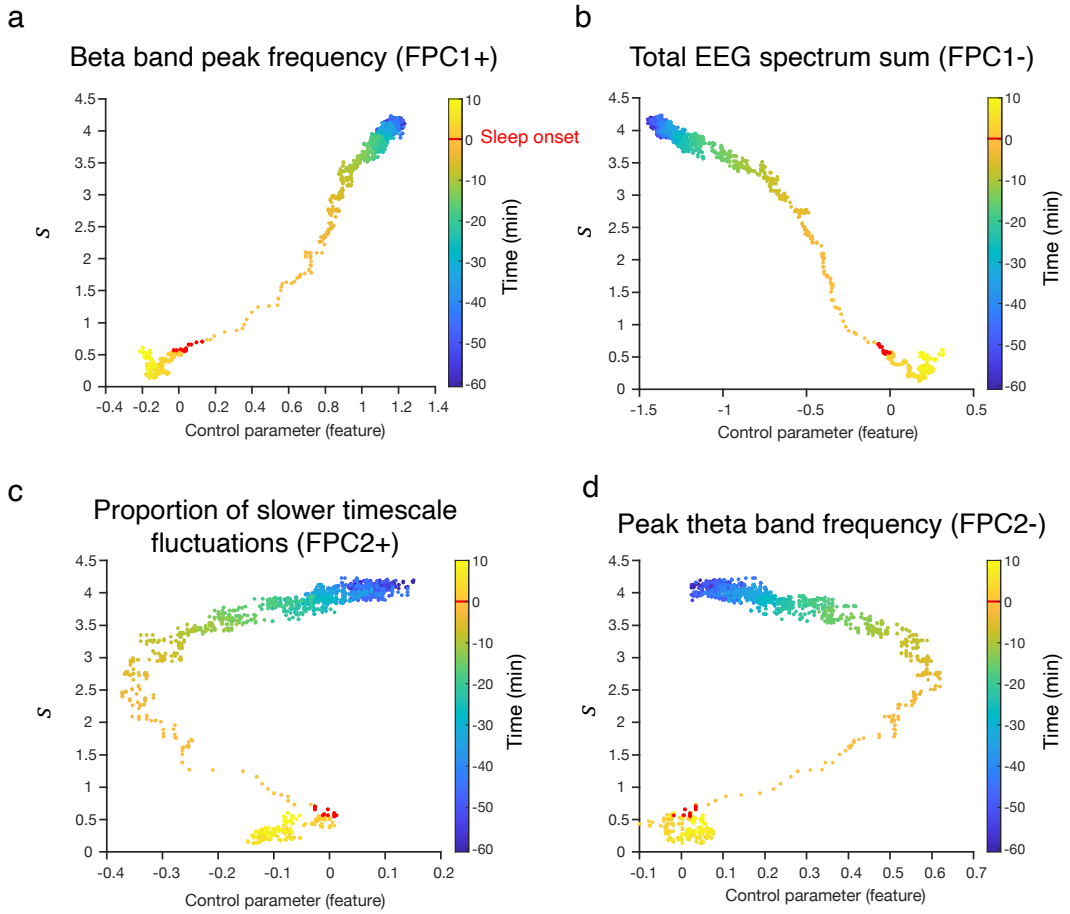


Figure 4.10: The sleep state variable as a function of the four key dynamical features as control parameters. The four key dynamical features are plotted against the s variable. (**a**), the peak beta band frequency (FPC1+) as the control parameter; (**b**), the total EEG spectrum power (FPC1-); (**c**), the proportion of slower timescale fluctuations (FPC2+); (**d**), the peak theta band frequency (FPC2-). The colour bar shows the time from 60 minutes before sleep onset to 10 minutes after, with 1-minute sleep onset marked in red. The feature values are z-score normalised to the sleep onset.

are summarized in Appendix B.18. I found that across the two assessment metrics, the peak theta band frequency (FPC2-) had the strongest similarity to the control parameter time series, which had the minimum DTW distance (7.03), and the second-highest similarity score (0.94). The theta-to-alpha band power ratio had the highest absolute similarity score (0.96) but a very large DTW distance (63.87). To further compare, the other three key dynamical features: peak beta band frequency (DTW distance 14.21 and similarity score 0.86), total EEG spectrum power (DTW distance 65.62 and similarity score 0.79), and the proportion of slower timescale fluctuations (DTW distance 44 and similarity score 0.62). Across both metrics, these features are inferior compared to the peak theta band frequency.

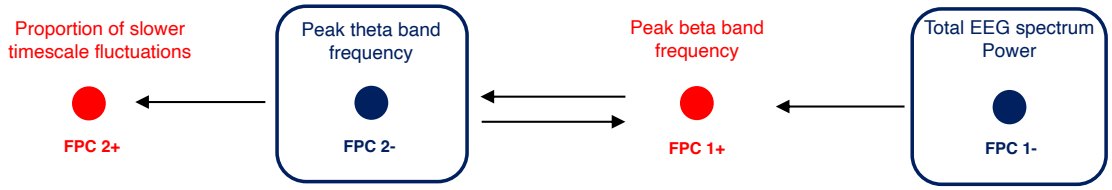


Figure 4.11: Granger causal interactions between the four key feature dynamical features. The four key dynamical features: the peak beta band frequency (FPC1+), the total EEG spectrum power (FPC1-), the proportion of slower timescale fluctuations (FPC2+) and the peak theta band frequency (FPC2-). The arrows indicate a significant Granger causality from one feature to another, using the Toda-Yamamoto approach with a p-value threshold of 0.025 (corrected for two directions). The p-value table for all directions can be found in Appendix B.19. The final VAR model has maximum integration 1, $p_0 = 5$, and extra lags $p_1 = 4$.

4.3.4. Causal interactions between the key dynamical features

After establishing the relationship between the key dynamical features and their relationship to the bifurcation control parameter, I next aimed to explore whether there are causal interactions between these features since causal interactions could help guide the development of modulation strategies. I applied the Toda-Yamamoto approach of Granger causality to test pairwise causality among the key dynamical features (see methods). The result is shown in Figure 4.11, with a significant causality indicated as an arrow. It reveals a causal interaction from the total spectrum power (FPC 1-) to the peak beta-band frequency (FPC 1+), a bi-directional causal interaction between the peak beta-band frequency (FPC 1+) and the peak theta band frequency (FPC 2-), and causal interaction from the peak theta band frequency (FPC 2-) to the proportion of slower timescale fluctuations (FPC 2+). By looking at the p-values, it seems that the most significant causal interaction was from the peak theta band frequency to the proportion of slower timescale dynamics (within FPC2). In general, it shows that FPC1 causally affects FPC2, with a feedback input as well from the peak theta band frequency.

4.3.5. Critical slowing down during falling asleep

For a system that is gradually pushed towards the bifurcation, critical slowing down can be observed in its system state dynamics (see method section for detailed theory explanation). The previous results built the bifurcation model under the assumption that the control parameter of the sleeping brain linearly increases until the bifurcation happens and the brain goes into the sleep state. Here, I aim to test whether the bifurcation dynamic of the brain EEG during the falling asleep process shows a critical slowing down characteristic.

The last 20 minutes of the sleep state variable s was extracted, and the two early warning signals (EWS), the autocorrelation and variance (standard deviation here, which is essentially the same), were evaluated in a 5-minute running window on the s time series of each participant. The group-level results (for $N = 210$ participants specifically selected; see methods section) are

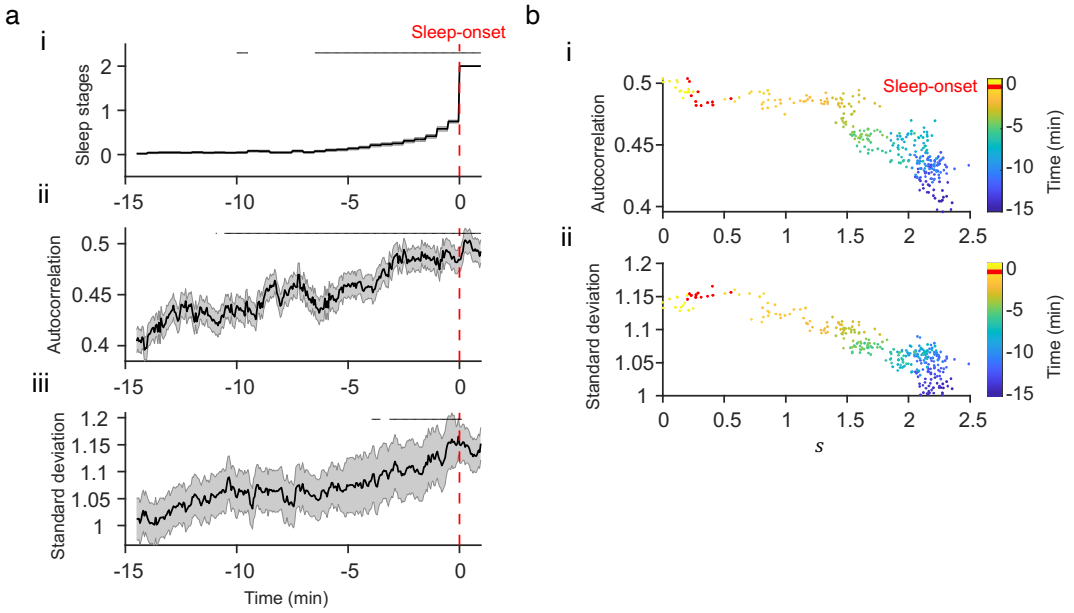


Figure 4.12: The critical slowing down dynamics in the sleep state variable during the falling asleep transition. a. The group-averaged (across 210 participants selected) autocorrelation (ii) and standard deviation (iii) evaluated on the last 20 minutes of the s time series for each participant. The shaded area shows the standard error of the mean (s.e.m.) around the mean value (black line). The actual time starts from 15 minutes before sleep onset due to the running window size (5 minutes). The average sleep stages (i) for these participants are also plotted for comparison. Horizontal lines show significant increasing changes compared to the mean of the first 3 minutes, using a left-tailed two-sample t-test to detect significant increases (for all three), with a p-value threshold of 0.025. The vertical red dashed line shows the sleep onset. **b**. The autocorrelation (i) and standard deviation (ii) are plotted against the averaged sleep state variable s time series, with colour showing time (from 15 minutes before sleep onset to 1 minute after sleep onset), and with the 1-minute sleep onset marked in red.

shown in Figure 4.12aii and aiii. Visually, there was a very clear and almost linear increase in both autocorrelation and standard deviation throughout the last 15 minutes prior to the transition. The LME models were fitted to test the linear relationship between time and these two EWSs, with a random effect as the participant ID. For autocorrelation, the model had an R-squared of 0.65. The linear coefficient between autocorrelation and time was 0.005 with a p-value smaller than 0.0005 (F-stat 31.97). For standard deviation, the R-squared was 0.74, with a linear coefficient of 0.01 and with p-value of 0.02 (F-stat 5.26), which was just below the significance line. The statistics showed that both EWSs significantly increased with time during the falling asleep transition. When compared to the sleep stages (Figure 4.12ai), it was found that the autocorrelation started to increase significantly around 4 minutes earlier than the sleep stages (see the horizontal lines in the figure). However, the standard deviation did not increase significantly until the last 3 minutes before transition, which was also later than the sleep stages. From the plots, it seems that the standard deviation had a much larger variability as well. The autocorrelation and standard deviation are also plotted against the sleep state variable s as shown in Figure 4.12b. The direct visualisation makes it obvious that even before obvious system state shifts started in s, the increase of autocorrelation and variance had

already started. The changes in the autocorrelation, but not the standard deviation, preceded the changes in the sleep staging, highlighting the utility of this measure to monitor and predict the falling asleep process.

To further validate that such an 'increase' is related to sleep onset transition specifically rather than an artefact itself, a non-falling asleep transition period (still 20 minutes) was taken to compare against the final transition period). Similarly, the participants were selected using the same artifactual criterion, but now the participants were required to have a falling asleep period longer than 50 minutes. This resulted in $N = 118$ participants left, and for each one, the s variable time series between 50 to 30 minutes before sleep onset was taken for evaluation of the EWSs. The same methods were used for evaluations, and the group-averaged results are shown in Appendix B.20. The results illustrate that for this control period where no falling asleep transition happened (the averaged sleep stages remain 0), there were no signs of increases in autocorrelation and standard deviation at all. The LME model also did not show significance in the time coefficients ($p = 0.19$, F-stat 1.66 for autocorrelation and $p = 0.31$, F-stat 1.05 for standard deviation).

To explore the feasibility of using the EWS to predict the sleep bifurcation of individual subjects, I examined the autocorrelation and standard deviation of the sleep state variable of a subset of participants. The results of three exemplary participants are shown in Appendix B.21, with the relevant description of the specific statistical methods used. The participant in Figure B.8a showed a clear critical slowing down with a significant linear increase in both the autocorrelation and variance (tested against surrogates). The participant in Figure B.8b showed a significant increase in the variance but not the autocorrelation. The sleep state variable resembled a flickering-like bifurcation dynamic (i.e., the system state fluctuates between two system states until finally attracted by another state), which could explain the early plateau in the EWS. implying that the transition could have started far before the sleep stage N2. The participant in Figure B.8c did not have a significant increase in either of the parameters. However, there was a local peak. Overall, the sleep state variable of individual recordings is noisy. The lack of ground truth in the actual transition point made it hard to quantify.

4.4. Discussion

In this chapter, I constructed a dynamical systems framework based on a novel feature-based EEG state-space analysis and demonstrated that the process of falling asleep can be modelled and described as a catastrophic bifurcation. I showed that the sleep state variable (distance to sleep onset) followed the bifurcation dynamics, while in the feature state space, the dynamic changes were slowing down. I further showed the key feature dynamics underpinning this transition and identified specific bursting-type dynamics as the control parameter of the system.

Using a Granger-causality-based analysis, I showed that there were causal interactions between the different feature dimensions. Finally, I showed that critical slowing down dynamics can be found during falling asleep transition, and showed its potential utility for tracking sleep onset.

4.4.1. The falling asleep transition can be described as a catastrophic bifurcation

To the best of my knowledge, this is the first empirical illustration of the bifurcation dynamics of human EEG during the falling asleep transition, with all previous studies relying solely on neurophysiological modelling. In general, it was shown that the global brain dynamics relative to the sleep onset state exhibit a catastrophic bifurcation process at a group level. The actual brain dynamics actively changed during the falling asleep process in the feature state space, with a minor slowing down, but the dynamics did not move towards the sleep onset state until a 'tipping point', where a fast nonlinear 'catastrophic' state shift happens. The actual transition process (catastrophic bifurcation dynamics) was independent of the latency from bedtime to sleep onset.

This nonlinear process was also never reported in any quantitative EEG analysis previously. The lack of empirical demonstration to date may be due to the difference in sample size, where most studies to date analyzed cohorts that were smaller than 100 participants. A second reason could be the shorter period of sleep onset that has been studied to date. Most studies either only analyzed the very last few minutes before sleep onset (normally defined as the start of N2 or even N1) (Gennaro et al., 2005) or averaged across small segments of the sleep onset period (Marzano et al., 2013). By focusing on the last few minutes these earlier studies might captured the quasi-linear changes after the bifurcation tipping point. A third reason could be related to the limited EEG feature bases (mostly power spectrum features) that were used in these studies. By examining sufficiently long periods of the transition to sleep using a bootstrapping method to account for differences in sleep latency and a comprehensive EEG feature base, this chapter could reveal the true dynamic of the falling asleep process.

The 'sleep onset period' deserves a more detailed discussion here, as most studies define a starting point for this period (for instance, the start of alpha trains (Ogilvie, 2001; Hori, 1985; Guerrero and Achermann, 2019; Siclari et al., 2014)). Here, the sleep onset period is broken out of this framework, which captures the entire period from bedtime to sleep onset. Looking retrospectively, the participants in the MESA dataset had indeed a much longer latency from bedtime to sleep onset than normal. When observing the actual dynamics, the participants' brain dynamics strolled around for a long period before the actual transition started in around the last 15 minutes. Previously in Chapter 3, there was an attempt to predict sleep onset latency using the features. In hindsight with the results of this chapter, before the final transition process (fast bifurcation), the brain dynamics showed very little changes for the features to be capable of distinguishing well.

By studying the specific feature dynamics underlying this process using FPCA, it was found that there are two types of transition dynamics, with the dominant pattern showing exactly the bifurcation dynamics, with a monotonic but nonlinear increase or decrease shift. The features that best represented this dynamic pattern were the total EEG spectrum power and the peak beta band frequency. As mentioned, these two features were also identified as key features that distinguish between bedtime and sleep onset in Chapter 3. The other two key features, the alpha-to-beta band ratio and the low-to-high band power ratio, also showed this dynamical pattern and had a very high similarity score (-0.99 and -0.94, respectively). Previously, we only knew that there was a key statistical difference, the results now can help better understand the actual temporal dynamics of these features. The second dynamic pattern, with a burst of change prior to sleep onset, was quite unexpected, but it represented the most important part of this bifurcation process: the control parameter. Specifically, the two key dynamical features of this pattern show that the key control parameter of this transition is an augmentation of theta oscillations and faster-changing EEG dynamics. Recall that the peak theta band frequency was one of the key distinguishing features identified in the Montreal dataset, but not the MESA. One important potential factor that distinguishes these two datasets is the huge difference in the mean actual sleep onset latency from bedtime to sleep onset (see Chapter 3 Discussion as well). Given the results here, the reason could be that the bedtime states of participants from the Montreal dataset were closer to the peak of the burst of changes (see Figure 4.8dii), and thus a decreasing change of the peak theta band frequency could be observed compared to sleep onset. For the MESA dataset, the bedtime states of the participants were much further away, and compared to the sleep onset state, there was barely any difference in the peak theta band frequency (as well as the other features showing the burst type dynamics). Looking independently, the burst type changing pattern is counter-intuitive, but in the framework of the dynamical systems and catastrophic bifurcation models, these features can be well described as the underpinning control parameters of the system, where their gradual increase/decrease serves as a pre-requisite for the transition to happen.

Linking to the previous point about the 'falling asleep period', this also means that if one cuts short the period of observation by defining a starting point, this 'burst' pattern might never be seen. This reversed change can also be thought of as an initial transition into a transient 'pre-sleep' state, where there is a continuous process of deactivation of the thalamus-cortical connection (de-arousal) for the real transition to be prepared (Kinreich et al., 2014; Merica and Fortune, 2004). From another perspective, it could be that the participants started from a very heterogeneous bedtime state, and their brain dynamics started to align prior to the 'ready transition' pre-sleep stage for the bifurcation to happen. The augmentation of the theta oscillations could also be related to hypnagogic mentation (a visual hallucination like a dream) before sleep onset (Gorgoni et al., 2019).

Using Granger causality analysis, I showed that there were causal interactions among the EEG

features that govern the bifurcation dynamics. It suggests that the sleep bifurcation dynamic may start with an increase in the total spectrum power that causes a decrease in the peak beta-band frequency that then causes an increase in the peak theta band frequency. The increase in the peak theta band frequency causes a decrease in the slower timescale fluctuations but also feedback to enhance the decrease in the beta-band frequency. The increase of the total EEG spectrum power and the decrease of beta band frequency implies a slowing down of the brain oscillations and the disappearance of high-level cognitive activities. Again, linking to previous findings, the sleep initiation started from a gradual de-arousal process before the transition into sleep (Merica and Fortune, 2004; Kinreich et al., 2014; Sulaman et al., 2023), consistent with the direction of the causal interactions shown here. A gradual de-arousal can drive the underpinning sleep control parameter (or sleep drives), which then leads to a bifurcation of the sleep state transition. Overall, these results imply that the increase in the total EEG spectrum power and the bi-directional interaction between the slowing of the beta-band frequency and the acceleration of the theta-band frequency (i.e., between FPC1 and FPC2) present tangible markers to monitor and manipulate the falling asleep process.

4.4.2. Using early warning signal (EWS) to predict sleep onset

At the end of the results, by quantifying the change in the autoregression and variance of the sleep state variable (aka, early warning signals), I showed that the falling asleep bifurcation involves critical slowing down. Yet, at the subject level, the sleep state variable was too noisy to show convincing early warning signals. One reason is the high stochasticity in the individual brain dynamics that is mitigated when averaging across participants. As shown previously, the individual-level dynamics might be described by a different model or scenario (for instance, the flickering pattern in Figure B.8bi). Another reason could be the fact that the sleep state variable s combines all the features, with some of them showing no changes and adding noise to it. Other reasons pertain to the complexity of the brain as a dynamical system (Stam, 2005) and the limitations of the EEG that only partially captures cortical activities. There could be better and more sensitive ways or features at the individual level to capture the falling asleep bifurcation dynamic. In addition, the high stochasticity of the brain could result in an unstable evaluation of the autocorrelation or variance, resulting in false positives (wrong warnings) (Dakos et al., 2012).

There are also challenges in deploying EWS to predict sleep onset in real time. The first challenge is the statistical difficulty in determining at which point the critical transition might happen without the knowledge of the actual governing laws of the system. For instance, the detection of the gradual increase of the autocorrelation might take a long time to observe, and how to statistically demonstrate that the duration of this gradual increase is not a random event (or significant enough) to imply upcoming bifurcation. The statistical method introduced

in Appendix B.21 is not feasible for real-time computation due to its heavy computational demands. A second challenge is the fact that computation of the sleep state variable requires estimation of the sleep onset feature values. In a real-time scenario, this could only be achieved by using group-level sleep onset feature data or using multiple night recordings from the same individual with the assumption that such sleep onset is consistent within participants across nights (which needs to be tested). Aside from these, there are also challenges in computing the EWS, such as the pre-processing of the time series data (detrending), and the choice of hyperparameters, such as the running window sizes for evaluating the EWSs and for statistical testing, etc. (Dakos et al., 2012). There could be alternative approaches to detect upcoming tipping points (Dakos et al., 2012), but deploying them to predict sleep onset requires much more engineering effort.

The lack of a ground truth about whether the participant is transitioning to sleep or not renders the development of detection strategies difficult. Similarly, it was reported that the success of applying EWS in seizure detection is dependent on clear visualisation of the seizure onsets (Maturana et al., 2020). In the case of sleep, the commonly used sleep staging has too poor temporal resolution, as well as insufficient matching of subjective experience (Yang et al., 2010; Rezaie et al., 2018). For example, in the participant in Figure B.8b, the transition occurred much earlier than the detection in sleep staging (a potential false negative). This uncertainty could also contribute to potential false positives, as sometimes spurious perturbations or stochasticity could lead to surges in autocorrelation and variance.

The next step is to use a multi-night recording dataset and try to understand individual-level brain dynamics during falling asleep. If there is consistency in the sleep onset state in the feature state space, one can then build a personalised falling asleep dynamical systems model. With an individualised model, one could potentially identify individual control parameters of the system for tracking the tipping point. There could also be a personalised way to best capture the sleep onset bifurcation dynamics (individualised s variable) and apply EWS analysis to detect sleep onset in future unseen nights. In the meantime, there would be more explorations into the engineering challenges of applying EWS in a real-time scenario, with deep learning models as a new potential to detect upcoming tipping points (Bury et al., 2021). Although, one has to be aware of the challenges in general. The most successful reference for us in neuroscience, the prediction of epilepsy, is still in high controversy as to whether critical slowing down dynamics do exist prior the the transition (Wilkat et al., 2019). The fundamental challenge would still be the lack of in-depth understanding of the dynamical system in the brain, and the governing laws behind it.

4.4.3. General limitations and future directions

Similar to previous limitations, the feature base used could be further extended, to capture other dimensions of brain activity and provide extra sensitivity to the dynamical changes. Although additional features most likely will not change the qualitative patterns of the bifurcation dynamics, with only minor quantitative differences. In addition, the bifurcation dynamics were observed on the global brain dynamics (averaged across all EEG channels), and there could be distinct regional dynamics not captured here. More importantly, there could be inter-regional interaction dynamics (how the interactions across the brain change) during the falling asleep process (De Gennaro et al., 2001b). This will be further explored in Chapter 5.

The theoretical catastrophic bifurcation model used here is descriptive only, which provides an intuitive understanding and framework to explore. Intuitively, the gradual accumulation of sleep pressure (and the gradual inhibition of wake-promoting neuronal populations underlying) leading to the global sleep state transition could be comparably thought of as the critical transition arising from over-grazing in this model. Still, the fitted constant values in the model are hard to explain and interpret in this falling asleep transition context. In the future, other simpler descriptive models could be explored to describe the dynamics and could provide probably more interpretation value. In addition, the assumption that the control parameter linearly increases with time might not be true. One could reasonably doubt that the control parameter actually changes in a nonlinear way as well (for instance, 2nd-order polynomials for the peak theta band frequency dynamics), which could be potentially tested.

The dynamical properties identified in this chapter came from a feature time series evaluated in 6 s EEG mini-epochs with 3 s intervals. These parameters were chosen assuming that EEG signals are relatively stationary within a 6 s time range, and the feature dynamics were relatively stationary within 3 s intervals. Changing it within a small acceptable range most likely will not incur qualitative changes in the bifurcation dynamics observed. However, this could potentially be important in EWS evaluations and capturing the critical slowing down dynamics (Dakos et al., 2012; Thompson and Sieber, 2011), although it might be very difficult to systematically test it.

From the insights of Chapter 3, one important future direction is to study how this temporal bifurcation dynamics change with age, which could provide significant insights into understanding the sleep regulations in the elderly population. It could also aid in phenotyping and designing a more targeted neuromodulation system to address sleep initiation problems in elderly individuals. Similarly, other demographic factors could also be studied, for instance, cognitive impairment. These factors are closely related to the wake-sleep regulations, and thus a detailed exploration into them would be an important step to do next. In the meantime, a replication of the results in another independent dataset would also be valuable to further consolidate the

results. A quick exploration into the Montreal dataset can be found in Appendix B.22, where there were no differences in the sleep state variable s between the healthy and MCI cohorts, although the dynamics also clearly show the bifurcation pattern.

4.4.4. Implications

The bifurcation model of the brain EEG during the falling asleep process provides a uniform governing law which generalises the changes of different EEG features (including those normally used power band features), with mechanistic understanding insights. The provided dynamical systems framework could also be impactful in changing the future quantification of the sleep onset process. For instance, the identification of the system control parameters or the existence of critical slowing down dynamics all provide immense opportunities for tracking the sleep onset and modulating the sleep onset in a closed-loop system. This will be further explored in the subject-level dynamics, and potentially being tested in a closed-loop neuromodulation paradigm.

Putting under a bigger scope, the falling asleep transition is just one of the many phase transitions in the brain. The direct reversed process of falling asleep is the waking-up transition. Theoretically, and in most neurophysiological modelling, the transition to wakefulness is also a bifurcation process, which follows a different transition path compared to falling asleep (the hysteresis and sleep inertia discussed before). Studying the dynamics in empirical EEG data could be extremely hard since the actual waking-up process is highly heterogeneous in its types and functionalities (Halász et al., 2004). The waking-up transition dynamics have been explored in Ilina's Master's thesis (Ilina, 2022) using the same feature-based approach, and it was shown to be more complicated to study than the falling-asleep transition. For instance, waking up from REM is most likely to be a natural process, but waking up from deep NREM sleep could be an abrupt perturbation of a stable state, indicating external disruptions or potential diseases. This dynamical system framework could be extended towards the waking-up transition, which could provide more mechanistic insights, as well as build closed-loop intervention systems for sleep fragmentation in elderly and dementia patients (Qian and Wong, 2022). In addition to sleep, many other phase transitions occur in the brain, such as the well-studied seizure onset, and high-level cognitive processing (attention) during cognitive tasks (Stam, 2005). These transitions could also be studied using the same conceptualisation, which has great potential to quantify and mechanistically understand dynamic changes in the brain.

5. The regional and inter-regional brain dynamics during falling asleep

5.1. Introduction

The previous two chapters extensively explored the EEG brain dynamics from bedtime to sleep onset, but on a global level, with all brain areas (EEG electrodes) considered together. It is a good and valid starting point, still, there has been a substantial amount of literature suggesting that the falling asleep process is not a spatially uniform process across the brain (Guerrero and Achermann, 2019; Gorgoni et al., 2019). Different brain areas might display a different temporal EEG dynamic pattern during sleep onset transition (Gorgoni et al., 2021; Marzano et al., 2013), due to the local neuronal networks and use-dependent sleep regulation patterns (Krueger et al., 2008). For instance, it has been established that the slow wave activities (delta band oscillations) start from frontal-medial regions and propagate to the entire cortex after (Tanaka et al., 2000; Nir et al., 2011; Siclari et al., 2014; Gennaro et al., 2005). On the other hand, spindle activities are more predominately a local event around the central-parietal cortex (Nir et al., 2011; Siclari et al., 2014). The regional differences make the sleep initiation process (falling asleep process) a very interesting transient state, with the mixing of both wakefulness and sleep state across the cortex (an interesting name called fluid boundaries Sarasso et al. (2014)). See Chapter 1 for a review of the 'local sleep' concept and the current findings of the regional EEG changes during falling asleep.

Aside from the regional differences, the other important aspect during the falling asleep transition is the inter-regional interactions. For instance, one widely used method is the EEG spectrum coherence analysis, to see the synchronisation between cortical regions (Morikawa et al., 1997; Tanaka et al., 2000). Aside from the coherence analysis, Gennaro et al. used a causality-inference method called directed transfer function (DTF) to detect causal couplings and information flows, and found that there is an important role in the frontal region to synchronise the global activity via information transfer (Gennaro et al., 2005). With the advances in mathematical algorithms, there is a plethora of different linear or nonlinear methods for capturing causal or non-directional EEG interactions (Sakkalis, 2011). As another example, Ferri et

al. used a functional connectivity EEG metric called synchronisation likelihood and found that the sleeping brain is closer to small-world network design (Ferri et al., 2007). Similarly, most of the analyses are done as a statistical inference only constrained to the sleep stages, and there are lots of other metrics not tested and analysed for the falling asleep brain dynamics. It would be beneficial to put them together to compare and understand the most prominent inter-regional interactional changes during the falling asleep process, which would provide further mechanistic insights, and potential novel biomarkers with extra sensitivity outside the normal regional dynamic features.

In this chapter, following the analysis of the previous two chapters, the regional cortical dynamics are explored. Firstly, I test the classification between bedtime and sleep onset on each EEG channel, to identify the best informative channel for the brain dynamic differences. The continuous regional brain dynamics during the falling asleep transition are explored as well, as it would be interesting to see whether the same bifurcation dynamics pattern would still exist in all locations. To study the inter-regional interactions, a comprehensive EEG relationship or connectivity measurement feature base using different types of linear and nonlinear metrics is first developed. I then use a feature category-wise classification to capture the most prominent EEG interactional changes at sleep onset compared to bedtime and see if these features would provide extra sensitivity. The continuous feature dynamics are then computed to see how different inter-regional features change during the falling asleep transition, with the same FPCA analysis done to extract the key dynamical patterns and features. Finally, the key representative regional and inter-regional feature dynamics are put together to study the causal interactions between them, which will provide additional insights into the initiation process of sleep. The analysis and methodologies in this chapter are explorative following the previous lines of research, with most work done on the MESA dataset and some replications in the Montreal one for the inter-regional features.

5.2. Methods

5.2.1. Analysis of regional differences

The regional EEG dynamics analysis followed the same approaches as in Chapter 3 for classification between bedtime and sleep onset. Basically, the same features were computed in the same way, but now the features of the EEG mini-epochs from different channels were studied independently. Recall that in the MESA dataset, EEG channels with a quality index smaller than 4 were excluded, thus not all participants would have 3 channels complete. Therefore, in the current analysis, only those participants with all 3 channels available were included in the analysis, with a total of $N = 408$ participants included. In all later analyses, the three

channels would be referred to as the cortical region they are measuring: Frontal (Fz-Cz), Posterior (Cz-Oz) and Right-Central (C4-M1). This is an approximation to facilitate discussion and interpretations. The same linear SVM classifier and hierarchical clustering were applied to the data of each channel and identified the key features that distinguish between bedtime and sleep onset state. The accuracy was also compared to the global accuracy as in Chapter 3. Note that the feature values were still globally (all 3 channels together, bedtime and sleep onset combined) Sigmoid normalised to [0,1] before running the analysis.

Similarly, the continuous feature dynamics followed the same approaches as in Chapter 4. Previously, the features from the 3 channels were averaged to get a 'global' representative feature value. Here, the three channels were studied separately. For each channel, the group-averaged sleep state variable s time series was extracted using the same bootstrapping technique with 200 samples per timestamp. The same bifurcation model as in Equation (4.1) was fitted to the s time series traces of each channel, and the differences were compared. Note that the model would start with the same assumption of a linear increase of the control parameter over time. FPCA was applied to the feature dynamics of each channel to see the extracted key dynamical patterns in each channel separately. The main aim was to learn the key dynamic patterns in each channel, thus further analyses were not replicated.

5.2.2. Development of the feature base for quantifying inter-regional interactions and its implementations

To study the inter-regional interactions between different cortical regions during the falling asleep transition, I first set out to develop a representative interactional feature base. There have been several studies done using different linkage metrics trying to understand the interactional changes (see Introduction). Still, their interpretations are constrained within the linkage method they used specifically, as different algorithms would estimate different aspects of the relationship between the cortical regions. A more comprehensive feature base has been developed to capture different types or aspects of EEG interactions during the process of falling asleep. It is important to note that not all features in the feature base are 'connectivity' metrics, as connectivity is a specific concept related to statistical dependency (functional connectivity) or causal links (effective connectivity) between EEG time series (Sakkalis, 2011). There is one more type of feature base called structural connectivity, which is only applicable in neuroimaging data with great spatial resolution (e.g., fMRI), as EEG data does not provide information outside the local cortical areas. Here, specifically, the features that assess relative band power ratios between channels are not connectivity measures but are still essential in studying the transition dynamics to capture relative EEG band oscillation changes. Therefore, the feature base developed is called inter-regional interactional features in general. Of course, on the interpretation side, the connectivity features themselves are 'actions' which quantify the linkages between regions,

but those non-connectivity features (cross-channel band power ratios) are most likely driven by underlying regional power-band changes.

The following section will introduce the complete list of inter-regional interactional features, grouped into different categories, along with their computational details. Please refer to Appendix A.3 for the full list of feature categories. It is important to note that not all types of features (especially connectivity features) can be included here, and only the most relevant and significant features for sleep analysis have been selected and implemented. There are several other toolboxes that contain comprehensive connectivity estimation methods, such as HERMES (Niso et al., 2013), that serve as good references. However, direct use of these toolboxes may result in poor transparency of computational details, and thus affect the interpretation of the results. These toolboxes will only be used to assist with some of the algorithm implementations later on. The mathematics behind the features will be briefly introduced for intuitive understanding purposes, where rigorous theories and proofs can be found in relevant references.

Correlation features

The most basic and straightforward metrics are the pair-wised correlation metrics. Given two time series, the similarity between them can be easily measured via a Pearson linear correlation. The correlation coefficient estimates a linear or monotonous dependence relationship between the two signals. The correlation metric was evaluated using MATLAB function `corr`. The input EEG data of two different channels was subdivided into several 1s segments, where the correlation was evaluated between each pair of 1s segments and then averaged.

As a discrete time series, it would also be important to evaluate the cross-correlation or cross-covariance between the two signals. The cross-correlation measures the similarity between signal X and the time-shifted (lagged) version of Y. The output is the similarity or correlation strength at different time lags. For cross-covariance, the only difference is that the covariance metrics will first centre the mean of both signals to 0 (subtract the mean values respectively). The output cross-correlation or covariance sequence will have a maximum value at lag 0 and decreases with lags increasing. The most commonly used metric is the width (or the number of lags), where the correlation or covariance value drops to half of the maximum (called half-width), which generally provides a measurement of the time range where the signals are similar to each other. The cross-correlation and covariance were evaluated using MATLAB function `xcorr` and `xcov`. Similarly, the input EEG was divided into several 1s segments for evaluation, and then the half-width values were averaged across all pairs of segments. These correlation features are non-directed, and there is one feature for each pair of EEG channels.

Cross-channel power band features

When it comes to the frequency domain, multiple measurements can be referenced from single-channel analysis. Previously the power band ratios were evaluated on one channel EEG itself, but now, one can evaluate the cross-channel band power ratios to see how different frequency band activities change relative to each other. The same methodology for estimating the power spectrum density was used as in 2.3.2. The input signals were first differentiated-filtered to remove 1/f noise, and then Thompson's multi-taper approach was used for power spectrum estimation. The estimated spectrum was then normalised to the total power for each channel separately. The normalised power ratio between the different frequency bands of the two EEG channels was then computed and log-transformed. Since the spectrum was normalised to itself prior to computing the ratios, this ratio should be interpreted as a 'relative dominance' of the specific frequency band activities. For instance, a ratio larger than 1 between the Frontal delta and Posterior delta implies the Frontal EEG has a larger delta activity compared to the total spectrum activity, but the absolute delta power in the Frontal region might not be larger than that of the Posterior. The same four frequency band definitions were used here (delta, theta, alpha and beta). The cross-channel band power ratios are bi-directional features, where for each pair of channels and frequency bands pair there are two independent features (ratio from X to Y and vice versa).

Similarly, for the EEG signals from two different channels, the magnitude-squared coherence was estimated between them to see the coherent oscillatory activity between them. The coherence was calculated on 1s segments of the input EEG signals between channels and then averaged. To get more sensitivity in the coherence in different beta frequency ranges, the beta band here was further divided into beta1 (13-18Hz), beta2 (18-24Hz) and beta3 (24-30Hz). The coherence is a non-directional measurement, where for each pair of channels and frequency bands pair there is one feature.

Granger-causality-based features

Estimating causal interactions is an important component in studying EEG connectivities. There are two types of most commonly used metrics coming from the theoretical basis of Granger causality: the partial directed coherence (PDC) and the directed transfer function (DTF). They are especially common in neuroscience to investigate the flow of information between different brain regions. Based on the formation of a multi-variate vector autoregression (MVAR) model (which is a multi-variate extension of the VAR model):

$$\boldsymbol{x}(t) = \sum_{r=0}^{p} \boldsymbol{A_r}\boldsymbol{x}(t-r) + \epsilon(t), \tag{5.1}$$

where $\boldsymbol{x}(t)$ represents an N-dimensional data (e.g., multi-channel EEG signals), p denotes the model order, $\boldsymbol{A_r}$ denotes the coefficient matrix (N by N) for each time lag r, and $\epsilon(t)$ is the error term. Each component in $\boldsymbol{A_r}$ (denoted as $a_{ij}(r)$) is the specific coefficient that indicates the causal influence from $x_j(t-r)$ to $x_i(t)$. As the equation suggests, the MVAR model captures how the past of the multi-variate data can predict the future (depending on the coefficients), which is the basic definition of Granger causality. By taking the Fourier transform of $\boldsymbol{A_r}$ into $\boldsymbol{A}(f)$, the PDC can then be defined as:

$$PDC_{j\to i}(f) = \frac{a_{ij}(f)}{\sqrt{a_j^*(f)a_j(f)}}, \tag{5.2}$$

where $a_{ij}(f)$ denotes the components of $\boldsymbol{A}(f)$, $a_j(f)$ denotes the entire $j_$th column, and $*$ denotes the complex conjugate. The $PDC_{j\to i}(f)$ is then the Granger causality from jth channel to ith in the frequency domain, but normalised to the entire outflow from jth channel. As one can see, this normalised value ranges from 0 to 1. Still, in this multi-variate scenario, this normalisation means that the strength of causal linkage between two channels can be affected by adding more channels into the analysis (because they need to sum to 1), even though the underlying dynamics do not change (Schelter et al., 2009).

The DTF, as its name, comes from the transfer function of the system by doing z-transform to Equation (5.1) on the unit cycle: $X(f) = H(f)E(f)$, where $H(f) = (I - A(f))^{-1}$ with I an identity matrix of the same size as $A(f)$, and $E(f)$ is the Fourier transform of the noise term. The normalised DTF is then defined as:

$$DTF_{j\to i}(f) = \frac{H_{ij}(f)}{\sqrt{\sum_{j=1}^{N} H_{ij}^*(f)H_{ij}(f)}}, \tag{5.3}$$

where $H_{ij}(f)$ is the elements of $H(f)$. $DTF_{j\to i}(f)$ is the normalised DTF from jth to ith channel at different frequencies, whose values range from 0 to 1 as well. Specifically, it is important to understand the fundamental differences between DTF and PDC, that DTF captures the causal influence from jth to ith channel divided by the total inflow into ith channel, so it emphasizes the sources that how the other channel is causally influencing this one. The mathematical descriptions provided above are just for an intuitive understanding, rather than a rigorous proof of theories behind them. The two metrics are complementary to each other. Gennaro et al. used DTF and showed the anterior-to-posterior information transfer and synchronisation process (Gennaro et al., 2004, 2005), which is an important reason that this feature is included in the study here as a comparison.

The PDC and DTF are multi-variate approaches, where all EEG channels can be fitted into the same model and quantified together. Both metrics were implemented using the Time-varying EEG connectivity toolbox (Omidvarnia et al., 2011), which evaluated the two metrics in a time-

varying approach using the Kalman filter. EEG signals from all channels were input to fit the model, with DTF and PDC output as a 4-D matrix (N by N by N_f by T, where N is the total number of channels, N_f is the number of frequency bins and T as the time resolution for time-varying estimation). The final feature values took the sum of the desired frequency bands averaged across time (for each pair of channels). Features were evaluated for the six frequency bands: delta, theta, alpha, beta1, beta2 and beta3 (see previous definitions) separately, with both directions considered for each pair of channels.

Cross channel frequency couplings

The same phase-amplitude coupling (PAC) method was applied here, but to cross-channel EEG data, to see if there would be such couplings across different cortical areas. Specifically, the PAC can capture whether the phase of a lower frequency in one channel is coupled to the amplitude envelope of a higher frequency in another channel. The same evaluation method for PAC was applied here (see Section 2.3.2), with the only difference that here the phases and amplitudes came from different EEG channels. The PAC feature is bi-directional, as for every channel pair, each channel will provide either phase or amplitude.

PAC is a very basic method which quantifies one type of coupling. The cross-bispectrum is a more comprehensive way to detect different types of cross-frequency couplings (specifically, the quadratic phase-coupling) between EEG channels (Jirsa and Müller, 2013; Kovach et al., 2018). The EEG bispectrum features have been widely applied, and proven to be valuable biomarkers for seizure prediction (Mahmoodian et al., 2019) or Alzheimer's disease (Klepl et al., 2022). Bispectrum analysis is a higher-order spectrum analysis (Swami et al., 1998), which computes a 2D spectrum from the EEG signals of two channels:

$$S_{XY}(f_1, f_2) = |\langle X(f_1) \cdot Y(f_2) \cdot Y^*(f_1, f_2)\rangle|, \tag{5.4}$$

where $X(f1)$ and $Y(f1)$ denotes the Fourier transform of two EEG signals x and y, $\langle\cdot\rangle$ denotes average. Essentially, the output bispectrum S_{XY} has a directionality, indicating the coupling from f_1 in signal x to f_2 in signal y. The existence of a quadratic phase coupling actually implies a nonlinear underlying phenomenon (Swami et al., 1998). The cross-bispectrum analysis was done using the Higher Order Spectral Analysis Toolbox (Swami et al., 1998) in MATLAB. The spectrum was evaluated with a Rao-Gabr optimal window of 6 and a window size of 256 samples. The coupling relationship (feature value) was then taken as the average spectrum values within the relevant spectrum windows in the 2D bispectrum for the required frequency bands. Since this feature is directional, therefore, for each pair of channels and frequency bands, there was a cross-bispectrum estimated in both directions and, thus, two distinct feature outputs. For both coupling features, there were four frequency bands considered: delta, theta, alpha and beta as a whole.

General synchronisation

In this category, I considered one feature called synchronisation likelihood (Stam and van Dijk, 2002). Synchronisation likelihood is based on the nonlinear dynamical system theory, which measures the level of similarity for the two systems' state space attractors (reconstructed from the EEG time series using time-delayed embedding). As an easy interpretation outside maths terms, it measures both linear or nonlinear dependences between the two signals, or more specifically, the correlated structure of the underlying dynamical system governing their dynamics. The core mathematics behind can be referred to in Stam and van Dijk (2002). Ferri et al. deployed the synchronisation likelihood in EEG during sleep and showed that the sleeping brain shows a small-world network property for better information processing (Ferri et al., 2007, 2008).

The synchronisation likelihood was evaluated using the HERMES toolbox (Niso et al., 2013). The optimal time-delayed embedding parameters (embedding delay and dimension) were estimated using MATLAB function `phaseSpaceReconstruction`. The input EEG signals were divided into several 1s segments for state-space reconstructions and then the estimation of the synchronisation likelihood. This is a non-directed feature, and for each pair of EEG channels, there is one feature value ranging from 0 to 1.

Information based metrics

With the development of information theory in recent years, a lot of novel algorithms have been developed and widely applied in EEG time series analysis. One most commonly used metrics is mutual information, which captures the statistical dependence between two EEG time series. Briefly, by treating two different EEG time series as two different random variables, the mutual information (MI) is evaluated as the shared entropy between them:

$$MI = I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X), \tag{5.5}$$

where $H(X)$ and $H(Y)$ are the entropy estimated from EEG time series x and y, $H(X|Y)$ and $H(Y|X)$ denote the conditional entropy (the entropy of one variable given the knowledge of the other variable). An easy illustration is given in Figure 5.1a. As entropy is a measure of the level of uncertainty in the data (estimated from the data distribution), the MI quantifies the shared uncertainty between these two variables. When X and Y share the same data statistics, the MI will be equivalent to the entropy of either one, implying that by knowing X, one can infer all the statistical properties of Y. Notice that there are many different algorithms to estimate entropy H from time series data. The simplest way is to estimate the probability distribution of the time series samples directly from the data. There are other approaches, such as approximate entropy (Pincus et al., 1991), which will not be further introduced here. MI has been applied to the EEG

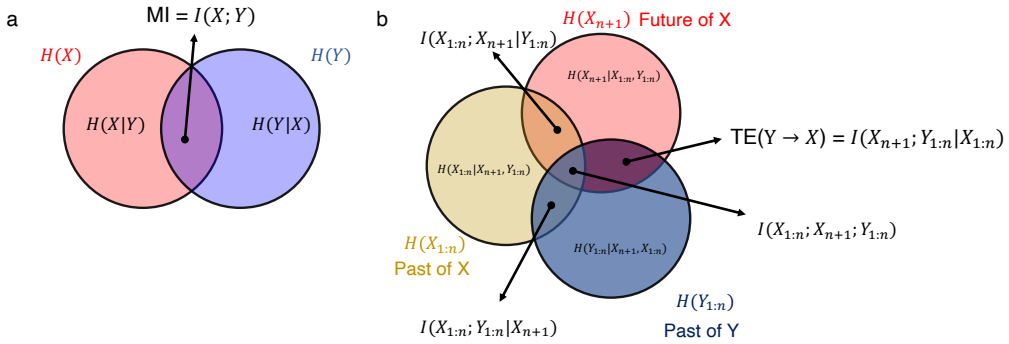


Figure 5.1: Visual illustration of the computation of mutual information (a) and transfer entropy (b). H generally denotes entropy of the variable; $I(X;Y)$ denotes mutual information (MI) between X and Y; $H(X|Y)$ denotes entropy of X conditioned on Y (vice versa). In (**b**), $X_{1:n}$ and $Y_{1:n}$ denotes the past of X and Y from time 1 to n, X_{n+1} denotes the one step future of X. $I(X_{n+1};Y_{1:n}|X_{1:n})$ denotes the mutual information between X_{n+1} and $Y_{1:n}$ but conditioned on the past of X as $X_{1:n}$, and this is the definition of transfer entropy (TE) from Y to X. The meanings of the other denotations can be deduced accordingly.

time series to quantify the impaired connectivity functionality in Alzheimer's patients (Jeong et al., 2001) or quantifying sleepiness level (Melia et al., 2015). Ramanand et al. showed that the MI between cortical regions increased from wakefulness to sleep (Ramanand et al., 2010).

By approximating the time series as a Markov process, the MI is further extended into another metric called transfer entropy (TE) (Schreiber, 2000). TE is a directed metric indicating the amount of information transferred from one channel to another. Figure 5.1b illustrates how TE is evaluated:

$$TE_{Y\to X} = I(X_{n+1};Y_{1:n}|X_{1:n}), \tag{5.6}$$

where I denotes the mutual information. As shown in the figure, TE is the shared information between the future of X (as X_{n+1}) and the past of Y (as $Y_{1:n}$), but conditioned on the past of X itself (as $X_{1:n}$). In non-mathematical terms, it means how well the information of the past of Y can predict the future of another channel X. One can easily find out that this is a very similar definition to Granger causality, and actually, it has been shown that for Gaussian variables, TE is equivalent to Granger causality (Barnett et al., 2009). TE has been demonstrated to be a very powerful brain connectivity feature, which can well distinguish anaesthetic depth (Sanjari et al., 2021) or detecting drowsiness levels (Huang et al., 2015).

Both MI and TE were implemented using the HERMES toolbox. The MI was evaluated by dividing the input EEG signals into several 0.5s segments and then averaging the output values, and the MI value was estimated using k-nearest neighbour distances (Kraskov et al., 2004) with a minimum neighbour size as 4. To estimate TE, the EEG time series was firstly time delay embedded (same as previous) to get the 'past' and 'future' sets and the same nearest neighbour approach was used for evaluation. Note that TE is a directed feature, which means that for each pair of channels, both directions were estimated.

Limitations and future extensions

This feature base is sufficient to start with, but of course, there could be many other features that can be added. For instance, in the correlation part, only linear correlation is considered, but nonlinear metrics such as distance correlation could also be added. All of the features are 'paired-wised' metrics, and even though PDC and DTF consider all channels in the same model, the returned metrics are between pairs only. One potential important addition in the future is the higher-order interactions, which use one single metric to quantify the general network's synergy or redundancy level (called O-information and S-information) (Rosas et al., 2019). This one value metric could be useful in providing novel insights into the brain functioning (Varley et al., 2023; Gatica et al., 2021). In Ilina's thesis that I supervised, these two information metrics have been applied to the MESA dataset for an initial exploration, but essentially the result could be constrained by the very limited number of channels available (only three) (Ilina, 2023). As in the regional feature base, most parameters for the evaluation of features could only be fixed, where actually, for a lot of these connectivity metrics, hyperparameter tuning could be important for the sensitivity of the features. The specific algorithms chosen to evaluate the features could also impact final results and feature interpretations (for example, entropy estimation). The other limitation lies in the very limited EEG channels available in the MESA dataset itself. With more EEG channels, one can get more spatial resolution and more complicated spatial-temporal analytic approaches can be taken, for example, source locational analysis, although the number of pair-wise connectivity features will become unscalable (which can also be resolved by graph theory analysis).

5.2.3. The analysis of the inter-regional interaction dynamics

For the MESA dataset, as previously introduced, after the exclusion of participants with incomplete EEG channels, there were in total $N = 408$ participants left for analysis. Similarly, the channels will be referred to as Frontal (Fz-Cz), Posterior (Cz-Oz) and Right-Central (C4-M1). The acronym 'R-Central' will be used in most figures to denote 'Right-Central'. All feature extractions were done on 6s EEG mini-epochs same as before. As described, the features were evaluated between each pair of channels (and frequency pairs if applicable). Therefore, there were in total 240 features: 3 features each for Pearson correlation, cross-correlation and cross-covariance; 30 features for band power ratios and 21 features for coherence; 36 features for DTF and PDC respectively; 36 for PAC and 60 for cross-bispectrum; 3 for synchronisation likelihood; 3 for MI and 6 for TE.

For the classification between bedtime and sleep onset, the same feature evaluation method was applied as in Chapter 3 for bedtime and sleep onset state separately, but here the features were not globally normalised after evaluation. The same linear SVM classifier was applied

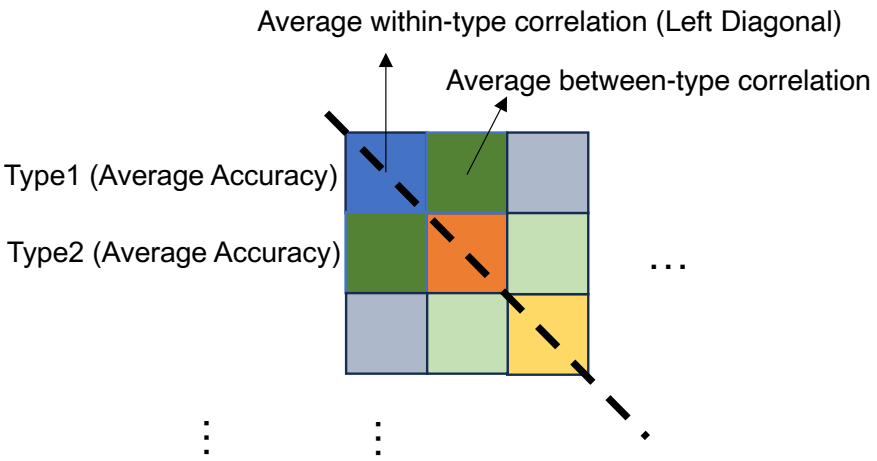


Figure 5.2: Illustration of feature type-wise hierarchical clustering approach. The average within-type accuracy will be shown along with the feature type name. The correlation values on the left diagonal will be the average within-type correlations between each other. The others will be average between-type correlations, which will also be the distance criterion for hierarchical clustering between them.

with leave-one-out cross-validation. Given the main aim was to identify the most prominent interactional changes, and for easier interpretation (as the feature space is much larger), a type-wise classification and hierarchical clustering approach was taken (see Figure 5.2). For a given feature type, the within-type correlation was evaluated as the average Pearson correlation value between all possible pairs of features inside this type. Likewise, the between-type correlation was the average Pearson correlation value between all possible pairs of features from each type. This between-type correlation value was the distance criterion for hierarchical clustering (similar to before, 1 minus the correlation value is the distance). The performance for each feature type was the average classification accuracy of all the features in that type. There would be some constraints, for instance, the features of the same type could be completely uncorrelated to each other and have huge performance gaps. But as one will see, as an initial exploration, it could already be informative in identifying key interactional changes. As this is a novel feature base, it would be meaningful to replicate it in the Montreal dataset. Only healthy participants were considered (in total $N = 65$ without any MCI patients), with three EEG channels taken for each one: Fz, Cz and Pz (linked-ear reference). To facilitate discussion, in later analysis, they are denoted as Frontal (Fz), Central (Cz) and Pariental (Pz). It is important to note that the big differences in the EEG montages between these two datasets might make these two datasets not really comparable. This is even more the case in connectivity features, and this will be further discussed in the results.

As an initial exploration, the continuous feature state space analysis was done on a subset of $N = 197$ participants who had longer than 60 minutes period from bedtime to sleep onset. The same computational approach was applied as in Chapter 4, with feature time series computed on 6s EEG mini-epochs and 3s intervals. Each feature was globally z-score normalised similarly

(to the mean and standard deviation of all values from all timestamps of all participants). The same bootstrapping technique was applied to get the group-averaged dynamics, but due to the smaller N number here, the sample number for each timestamp was restricted to 150. The sleep state variable s was evaluated in this 240-dimensional inter-regional feature space. There were some features with extremely small standard deviation at the group level, which would bias the results as a whole, therefore those features with a standard deviation smaller than 0.001 were excluded when calculating s. The same theoretical bifurcation model was used to fit the interactional feature state space dynamics to see whether the dynamical pattern can be described as the same model. FPCA analysis was applied to all the features' time series to extract key dynamic patterns and features.

5.2.4. Other statistical methods

For most between-group analyses, a double-sided two-sample t-test was applied (as the sample number is sufficiently large), where the p-value threshold for significance will be specified when the test is applied. The causality between the time series was done using the same Toda-Yamamoto approach (see Chapter 4 Methods). There are some post hoc analyses which will be introduced where applicable.

5.3. Results

5.3.1. The regional specific EEG dynamic features at sleep onset compared to bedtime

Previously in Chapter 3, the features from the 3 channels were taken as a whole to classify between bedtime and sleep onset state. However, it could be that different channels provide different information and have distinct sensitivity. To start with, the classification was done again on each channel (region) separately for each feature, and compared to the global accuracy (put them together to classify). The performance comparison for all 47 features can be seen in Appendix B.23, with the four key governing features identified previously marked in red, and for each feature, the best-performing region is marked in bold. It is obvious that, for all features, there was one or more regions outperforming the global accuracy. For the four key features specifically, it can be found that the best performance was either in the Posterior (for the low-to-high frequency band power ratio and the total EEG spectrum power), or the Right-Central (for the alpha-to-beta band power ratio and the peak beta band frequency). The Posterior region had the highest average accuracy across all features (62.66%). The Right-Central had an average of 61.53%, similar to the average global accuracy of 61.12%. The Frontal region had the worst performance (average performance of 59.18%).

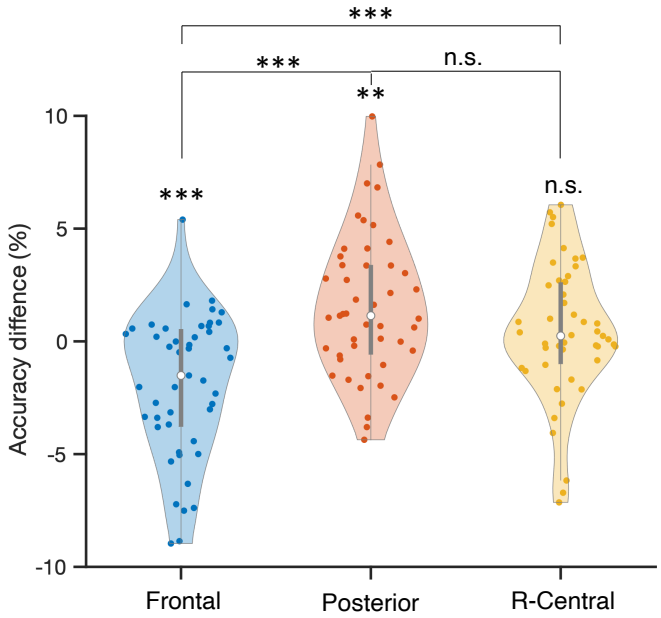


Figure 5.3: Distribution of the accuracy differences for all features between the global and each EEG region. Three regions are shown: Frontal (Fz-Cz channel), Posterior (Cz-Oz channel) and R-Central (an acronym for Right-Central, C4-M1 channel). The one-sample t-test is done for the accuracy differences in each channel separately (compared to 0). A double-sided two-sample t-test is done for comparisons between channels. ** indicates $p < 0.005$ for the test; *** for $p < 0.0005$; n.s. for non significance. Violins show the kernel density plot (external shape), median values (white non-filled circle) and upper and lower 25% quantiles (grey bar).

The differences between the global accuracy and that of each region for all features are also shown in Figure 5.3. The result further shows that the Frontal region had significantly worse performance than the global (mean ± st.d. -1.94 ± 3.15), and was worse than the other two as well. The Posterior had the best performance (1.54 ± 3.13), and Right-Central had a similar performance to the global (0.41 ± 3.0). The results illustrate that the Posterior region has the best sensitivity in showing the key EEG dynamics differences between bedtime and sleep onset.

The four key governing features identified previously were done on the global level. The performance difference between the regions implies that there could be region-specific key features that provide better sensitivity, and other potential mechanistic insights. Therefore, the same hierarchical clustering pipeline was applied to each region separately to extract the key features. The result for the Posterior region is shown as an exemplary result in Figure 5.4. The features listed in the same hierarchical order as in the figure can be found in the same Appendix. The key features were selected based on the same criterion: the cluster centroids with a mean cluster accuracy above 65%. It can be seen that the Posterior had four key features: a decrease in the entropy, an increase of the alpha-to-beta band power ratio, the delta band peak frequency and the low-to-high frequency band power ratio. Reflecting on the four key features identified globally, the total EEG spectrum power was captured in the entropy's cluster, while the peak beta band frequency had only 61.88% accuracy (still as an independent cluster, though).

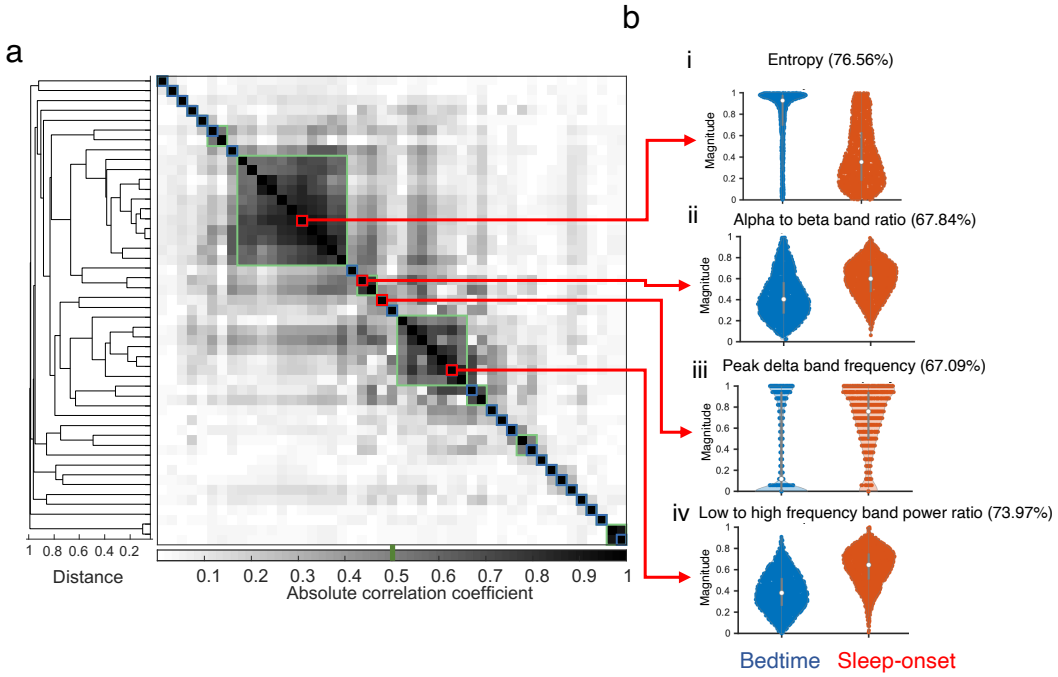


Figure 5.4: Hierarchical clustering of the features that distinguish bedtime and sleep-onset in the Posterior region (Cz-Oz channel). a. The top 40 features that best classify wakefulness (bedtime) and sleep-onset are chosen and hierarchically clustered based on the absolute correlation coefficient with a clustering threshold of 0.5. The features are re-arranged according to the clustering, with the dendrogram shown on the left; Green box, outlines of a feature cluster; Blue square, central feature of a cluster; Red square, central feature with mean classification accuracy in that cluster above 65% and chosen as a key governing feature. The Pearson correlation coefficients between pairs of features are shown as the grey-scale colour bar. **b**. The distribution of the four key features identified from (**a**): the entropy (**i**), the alpha-to-beta band ratio (**ii**), the peak delta band frequency (**iii**), and the low-to-high frequency band power ratio (**iv**). Blue indicates bedtime, and red indicates sleep onset.

The results for the other two regions can be referred to in Appendix B.24, along with all the features interpretation. For the Frontal region, only 2 features satisfied the selection criterion, the total EEG spectrum power and the peak beta band frequency (both increasing at sleep onset). The low-to-high frequency band power ratio and the alpha-to-beta band power ratio were also cluster centres, but their accuracies were bad. Interestingly, the entropy feature was not the cluster centroid but the total EEG spectrum sum. For the Right-Central, there were 6 features: the low-to-high frequency band power ratio (increasing at sleep onset), the theta band power (increasing), the peak delta band frequency (increasing), the auto mutual information (decreasing), the total EEG spectrum power (increasing), and the proportion of slower timescale dynamics (increasing). Interestingly, the theta band power was captured by the low-to-high frequency band power in the other two regions as well as in the global case, but here it became an independent cluster centre in the Right-Central. The auto mutual information feature did not have good accuracy in the other two regions. The proportion of slower timescale dynamics was previously captured inside the cluster of total EEG spectrum power for the Posterior and

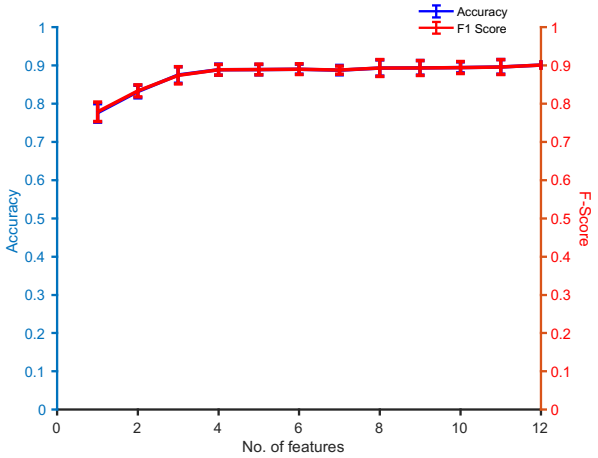


Figure 5.5: Classification accuracy between bedtime and sleep onset using a combination of the key features from the 3 different regions in the MESA dataset. The features were input in a sequential order based on their accuracies from high to low (regardless of region). The Red line shows the F-score, and the blue line shows the accuracy. The error bar shows the mean ± st.d. for 10-fold leave-one-out cross-validation. The feature added at each step, and the stats for each step can be seen in Appendix B.25.

global. It became an independent cluster here but did not have good accuracy in the Frontal. The key features from all these regions are listed along with their accuracies in Table 5.1, and compared to the global accuracies. In general, there were some region-specific features that have extra sensitivity, with the four key features identified previously showing up in different regions.

Table 5.1: List of all the key features selected from each region separately with their accuracies. The accuracies for all 3 regions and global are shown for each feature, with 'R-Central' standing for Right-Central (C4-M1 channel). The key features chosen specifically from each separate region are marked in bold and red. The values shown are all in units percentage (%).

Feature name	***Frontal***	***Posterior***	***R-Central***	***Global***
Theta band power	67.84	77.16	**77.78**	74.10
Alpha to beta band ratio	64.44	**67.84**	72.03	69.39
Low to high frequency band power ratio	62.62	**73.97**	**73.34**	69.85
Peak delta band frequency	57.25	**67.09**	**70.84**	64.79
Peak beta band frequency	**66.00**	61.88	68.94	66.24
Auto mutual information	62.69	65.00	**70.62**	66.49
Total EEG spectrum power	**69.84**	73.16	**72.47**	72.54
Entropy	71.88	**76.56**	68.72	72.15
Proportion of slower timescale dynamics	58.72	65.91	**68.97**	63.76

The key features' performance was then compared for all these regions (channels). For the Frontal, combining the two key features, the model could achieve 70.8% accuracy (F-score 0.70). For the Posterior, the model combining the four key features could achieve 86.5% accuracy (F-score 0.86), and for Right-Central, the model combining all six features could achieve 83.6% (F-score 0.83). Recall from Chapter 3 that the four key features globally only achieved 80% accuracy (F-score 0.8). The results clearly show as well that the Posterior had the best sensitivity in telling the differences between bedtime and onset (with even 2 features less than Right-Central). If one combines all the features from all the channels (in a sequential order of performance from high to low), the model achieved 90.1% accuracy with F-score 0.90 (Figure 5.5, which is an impressive performance already. It is also interesting to see that, the model performance already reached a plateau after around the first 4 features input: the theta band power from the Right-Central, the entropy from the Posterior, and the low-to-high frequency band power ratio from the Posterior and Right-Central. The key information for this exploration is clear here, of course, the same analysis can be replicated in the Montreal dataset, and there needs to be even more channels to systematically disentangle which exact region could provide maybe even better sensitivity.

5.3.2. The region-specific bifurcation dynamics during falling asleep

Following the analysis, the region-specific continuous EEG dynamics during falling asleep were explored. The same methods were applied to get the group-averaged sleep state variable s for each channel separately (200 samples per timestamp). The results of all 3 channels are shown in Figure 5.6a, with each channel having a right-tailed two-sample t-test at each timestamp compared to the first 10 minutes of its own. The s time series was smoothed using a moving-median filter of window size 20 (same as before). It can be seen that the Frontal region had the smallest dynamic range (from 3 to 0), with the Posterior in the middle (3.5 to 0), and the Right-Central area having the largest range (from 4 to 0). From the statistical tests, it can be seen that Right-Central and Frontal had earlier decreasing changes compared to the Posterior. When comparing each other, it can be found that the s dynamics in the Frontal region were significantly different from that of Right-Central, but there were not many differences between the other two pairs (Figure 5.6b). The kinetics energy plot for each region (Figure 5.6c) shows that the dynamics in the Frontal region required the least energy from one state to another. This is also clear in the s dynamics showing distinct dynamic ranges. The same bifurcation model could be fitted for the s dynamics of each region separately, which can be referred to in Appendix B.26. In general, for each region, the s dynamics could all be well captured using the catastrophic bifurcation model, with the R-squared for each region similar (0.928 for Frontal, 0.933 for Posterior, and 0.927 for Right-Central).

The feature dynamics for each region could be similarly evaluated, as shown in Appendix B.27. As a brief overview of the results, the key dynamical patterns during falling asleep did not differ

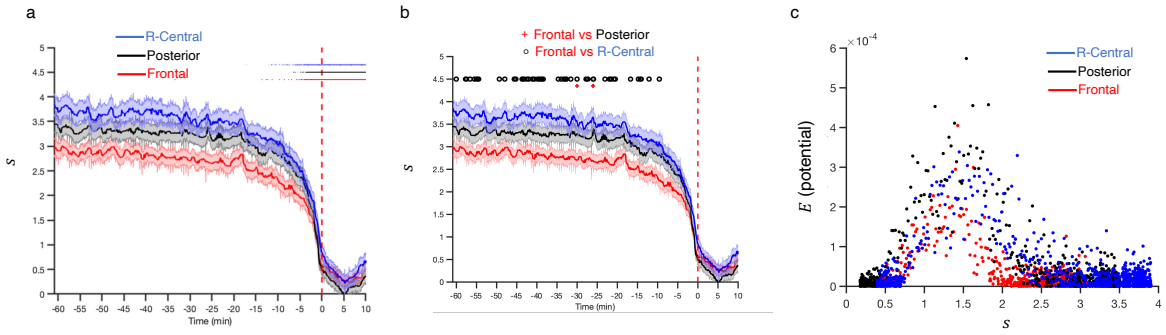


Figure 5.6: The sleep state variable dynamics in three regions separately, with the potential analysis. **a**. The sleep state variable s dynamics from 60 minutes before sleep onset to 10 minutes after in 3 channels separately: Frontal in red, Posterior in black, and Right-Central in blue (denoted as R-central in plots). Shown are the mean ± standard error of the mean (s.e.m.) across 200 samples. The horizontal lines of each respective colour show significant decreasing changes compared to the first 10 minutes of the dynamics of each respective channel, using a right-tailed two-sample t-test at each timestamp, with the p-value threshold corrected to be 0.05 dividing the total number of timestamps. The vertical red dashed line shows the sleep onset. **b**. The same plot as (**a**), but now a double-sided two-sample t-test is done at each timestamp between every pair of the three channels, with the p-value threshold corrected to be 0.05 dividing the total number of timestamps. '+' shows significance between Frontal and Posterior; '∘' shows significance between Frontal and Right-Central; No significance is found between Posterior and Right-Central. **c**. The kinetics energy potential ($E = \Delta s^2/2$) as a function of s for each channel separately (in respective colour).

from each other significantly. In the continuous feature dynamics plot, one obvious difference would be that the Right-Central region had the most dramatic changes, especially in those decreasing pattern features in the middle of the plots, as well as the first few increasing pattern features just below them. This was also illustrated in the s dynamics previously, that the Right-Central region had the largest dynamic changes (range). When comparing all of them to the global dynamics in Figure 4.7, it seems that the global dynamics were most similar to the Right-Central area, although here, the z-score ranges were slightly larger.

The FPCA analysis was applied to the feature dynamics of each region separately (Figure 5.7), with the FPC similarity scores of each feature shown in Appendix B.27. For the Frontal region, FPC1 explained 91.16% of the total variance, and FPC2 was 7.35%. For the Posterior region, FPC1 explained 93.17% of the total variance and FPC2 5.79%; For the Right-Central region, FPC1 explained 97.36% of the total variance and FPC2 1.77%. The FPCA results for the three regions showed no qualitative difference, with simply trivial differences in the amount of variance explained by the FPC1 or FPC2 patterns. In sum, all FPC1 patterns showed bifurcation-type dynamics, with FPC2 showing a burst type of change as previously. The key representative dynamic features were slightly different in each region. For the Frontal region, the FPC1 was best represented by the total EEG spectrum sum and the entropy; FPC2 was best represented by the delta-to-alpha band ratio and peak theta band frequency. Note that although peak beta band frequency was not the most representative one for FPC1, it still had a similarity score of -0.99. For the Posterior region, FPC1 was best represented by theta band power and peak beta band frequency, while for FPC2, they were the delta-to-theta band ratio and the

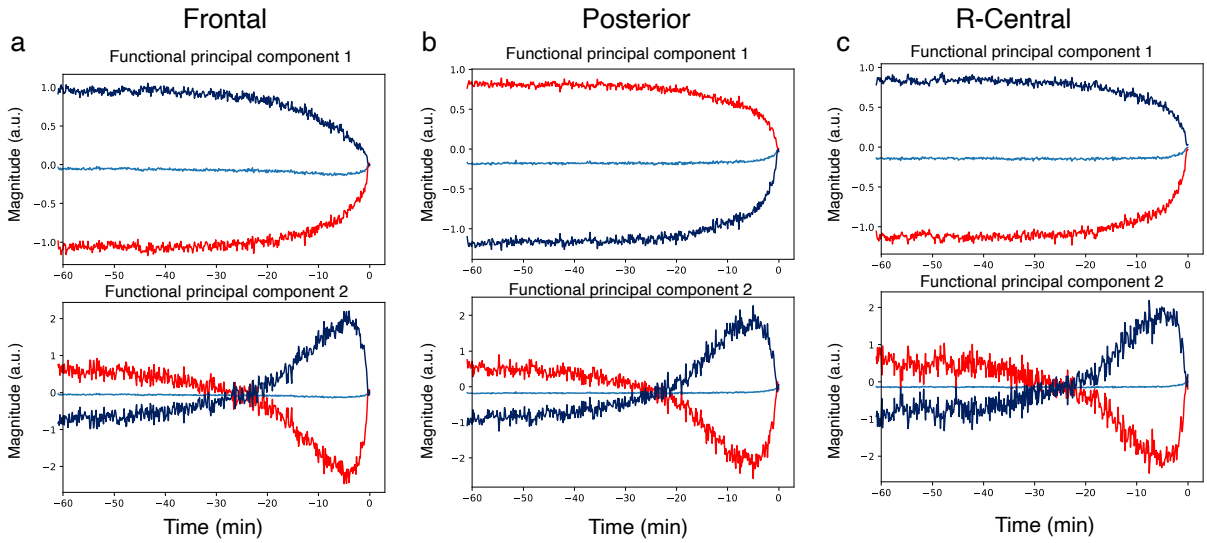


Figure 5.7: The FPCA analysis for the feature dynamics of the three different regions in the MESA dataset. The three different regions: Frontal (**a**), Posterior (**b**) and Right-Central (**c**). The FPCA analysis is done on the feature time series from 60 minutes before sleep onset to sleep onset (without the post-sleep-onset period). For each functional principal component (FPC), shown are the negative (dark blue line) and positive (red line) directions around the mean time series of all the features (light blue line). The y-axis has arbitrary units. The 'positive' and 'negative' means 'addition' or 'subtraction' of the actual FPC to the mean.

temporal coherence in the theta band. Still, if one observes the FPC1 similarity scores to the total EEG spectrum sum and the entropy feature (representative FPC1 features in the Frontal region), they were also very high (-0.99 and 0.99 respectively). For the Right-Central region, the FPC1 was best represented by the same two features as those of the Frontal, while for FPC2, they were the delta-to-alpha band ratio (same as the Frontal) and the temporal coherence in the theta band (same as the Posterior). It might seem quite confusing that so many different features came out, but if one refers back to the similarity score results of the global dynamics (Appendix Table B.14), one can directly see that all the FPC1 features from these 3 regions (the total EEG spectrum power, the theta band power, the entropy and the peak beta band frequency) were either already key representing features, or had a similarity score of either 1 or -1 in the global dynamics already, which means that they are generally inter-changeable as the key representative dynamic features. Similarly, for the FPC2, the temporal coherence in the theta band had a score of -0.76 (just below that of -0.77 for peak theta band frequency), the delta-to-theta (or alpha) band power ratio had a score of 0.94 (or 0.97), both very high as well.

5.3.3. The inter-regional interactions dynamics at sleep onset compared to bedtime

To study the inter-regional interaction differences between bedtime and sleep onset, the inter-actional feature base developed previously was applied here (on the same participants with all 3 channels complete), and the feature type-wise hierarchical clustering was applied (see previ-

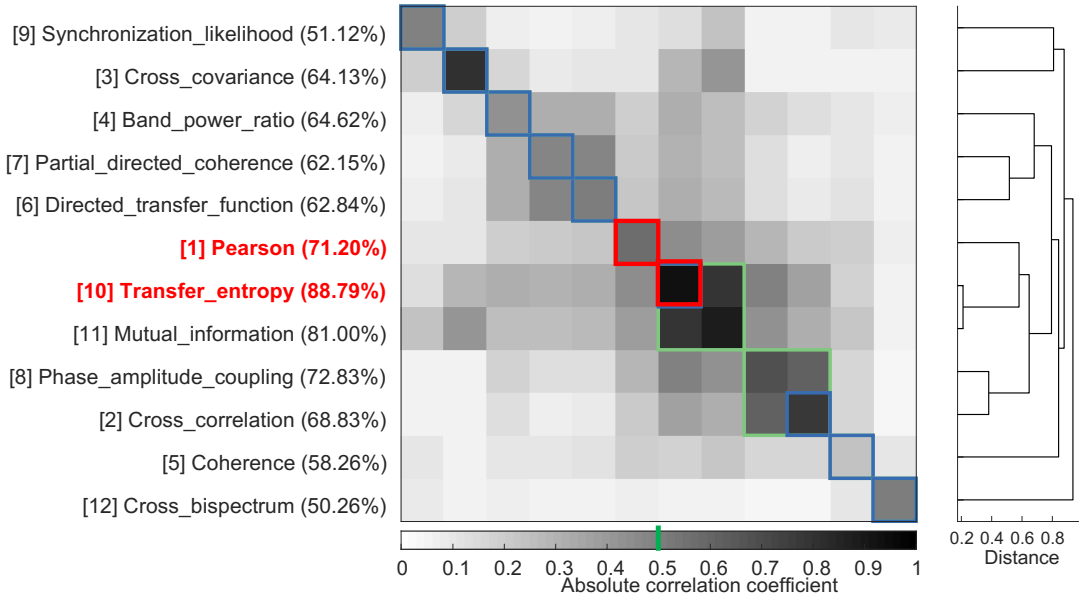


Figure 5.8: Hierarchical clustering of the inter-regional interaction features (arranged in each category) that distinguish bedtime and sleep-onset in the MESA dataset. The feature type name, along with the average accuracy, is shown on the left side. The average within-type correlation is shown on the left diagonal, with the between-type correlation shown between each pair, and the correlation strength is shown as the grey scale (colour bar at the bottom). The dendrogram is shown on the right side, with a clustering threshold of 0.5. The green boxes show the feature types clustered together, and the blue boxes show cluster centres. The selected key feature types are marked with red names and in red squares.

ous methods) to identify the most prominent interactional changes at sleep onset compared to bedtime. For the 3 EEG channels, there were 240 interactional features in total, which were then grouped into each type for analysis. The result can be seen in Figure 5.8. The key feature types were selected as the cluster centroids with above 70% average within-type classification accuracy. Obviously, in the figure, two feature types satisfied this criterion: the Pearson correlation (71.2% average accuracy) and the transfer entropy (TE) (88.79% average accuracy). It is not surprising to see that mutual information (MI) was highly correlated with TE, and had a similarly good accuracy. It is interesting to see that the cross-channel phase-amplitude coupling (PAC) and cross-correlation were clustered together, where both of them had a decent accuracy as well. The other feature types did not show a good performance, but of course, some of the feature types had very low within-category correlation (for instance, the coherence and band power ratio features), and they could potentially be studied from within further for more specific features inside that type.

For the specific two feature types (Pearson and TE), the distributions of the features within them between all pairs of regions (channels) and directions are shown in Figure 5.9 and compared between bedtime and sleep onset. It can be seen that the Frontal region had a decreased Pearson correlation with either Posterior or Right-Central at sleep onset, while the correlation between Right-central and Posterior increased significantly. For TE, it is interesting to see that all

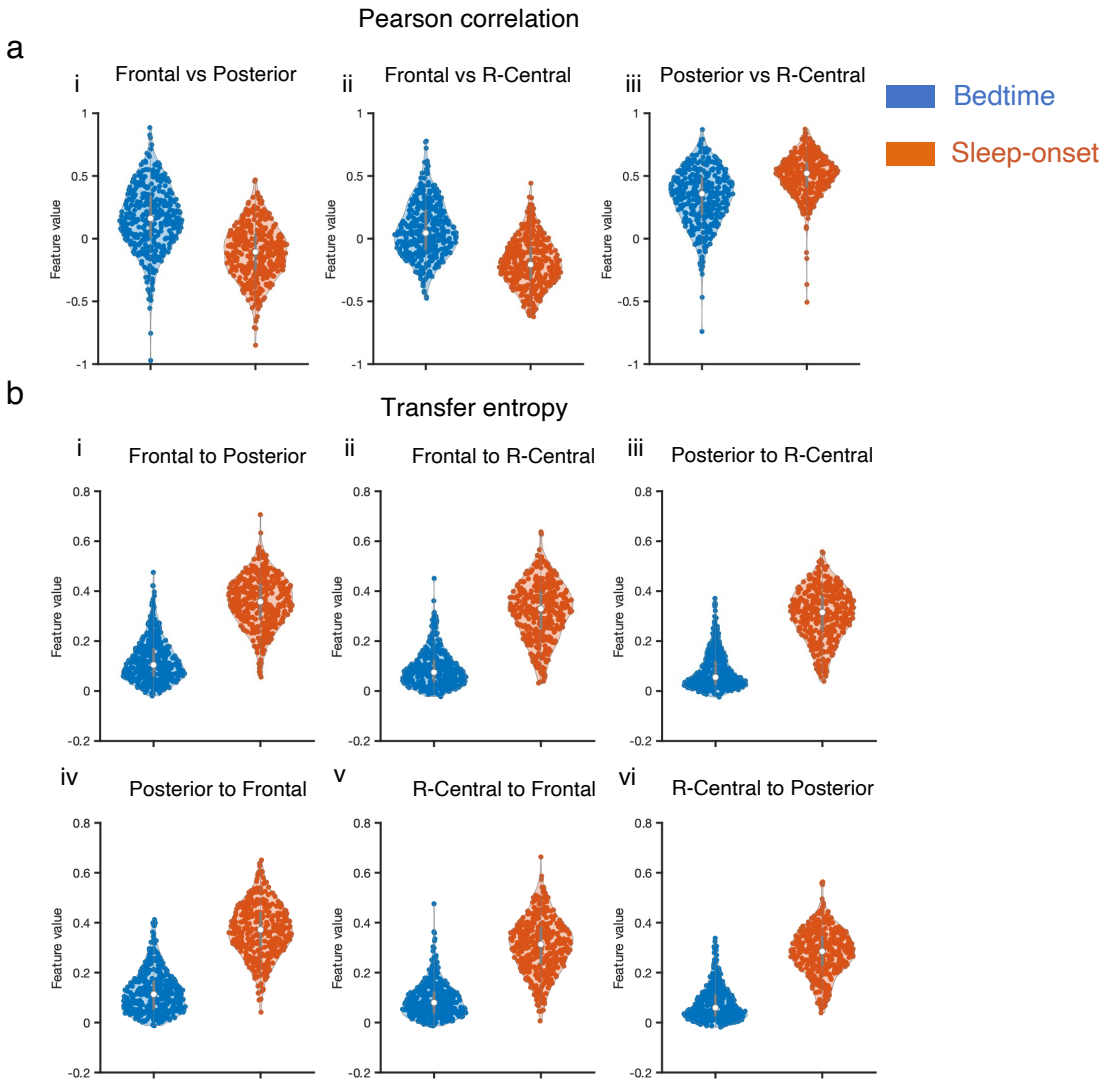


Figure 5.9: The distribution of the features in the two key inter-regional feature types at bedtime and sleep-onset in the MESA dataset. All pairs of regions are shown in the figure. **a**. Pearson correlation distribution at bedtime and sleep onset between different pairs of regions: (i) Frontal vs Posterior; (ii) Frontal vs R-Central (stands for Right-Central); (iii) Posterior vs R-Central. (**b**). Transfer entropy distribution at bedtime and sleep onset for all directions of regions: (i) Frontal to Posterior; (ii) Frontal to R-Central; (iii) Posterior to R-Central. The reversed directions are shown in (iv-vi) respectively. The detailed statistics of all distributions and two-sided two-sample t-test results between them can be found in Appendix B.28. Essentially, all p-values are smaller than 0.0005 and thus not marked in the figure. Violins show the kernel density plot (external shape), median values (white non-filled circle) and upper and lower 25% quantiles (grey bar).

directions of information flow increased dramatically. Combining all the 9 features in these two types, the linear SVM model could achieve an accuracy of 91.7% (F-score 0.92).

Outside the feature types, normal hierarchical clustering was also applied to the top 60 features selected from the 240 features (regardless of their types) to choose the specific governing features. The results, along with the specific features chosen, can be found in Appendix B.24. In general, there were four key governing interactional features that best capture the changes at sleep onset

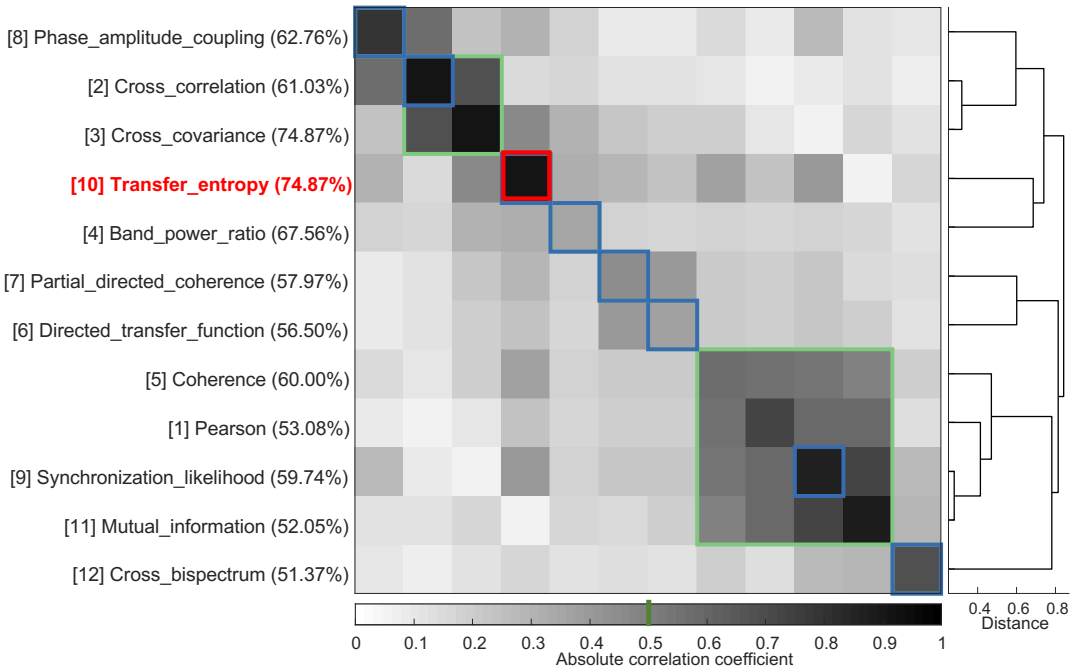


Figure 5.10: Hierarchical clustering of the inter-regional interaction features (arranged in each category) that distinguish bedtime and sleep-onset in the Montreal dataset. The feature type name, along with the average accuracy, is shown on the left side. The average within-type correlation is shown on the left diagonal, with the between-type correlation shown between each pair, and the correlation strength is shown as the grey scale (colour bar at the bottom). The dendrogram is shown on the right side, with a clustering threshold of 0.5. The green boxes show the feature types clustered together, and the blue boxes show cluster centres. The selected key feature types are marked in red names and red squares.

compared to bedtime: a decrease in the coupling between theta band phase of Posterior and beta band amplitude of Right-Central, a decrease of the DTF at theta band from Right-Central to Frontal, an increase of TE from Posterior to Frontal, and a decrease in the Pearson correlation between Frontal and Right-Central. When observing the hierarchical clustering plot, it can be found that all TE and MI features were still clustered together with high correlations between each other, and as the distribution plots show, there was an increase of TE regardless of the direction. Individual TE features all reached an accuracy of around 88%, which is impressive. Similarly, a lot of the PAC features (or DTF and PDC features) were clustered together, and represented by the specific ones described before. The Pearson correlation features became an independent cluster. Combining all four features from this analysis, the linear SVM model achieved an accuracy of 91.9% (F-score 0.92). The results also confirm that the type-wise feature selection can well extract the most important interactional changes at sleep onset compared to bedtime.

Given this is a new feature base, a replication in the Montreal dataset was done (only for healthy participants). As specified before, there were 3 channels chosen for the Montreal dataset: Fz, Cz and Pz, all with linked-ear reference, and they are denoted here as Frontal, Central and Parietal, respectively. The feature type-wise hierarchical clustering results can be seen in Figure 5.10.

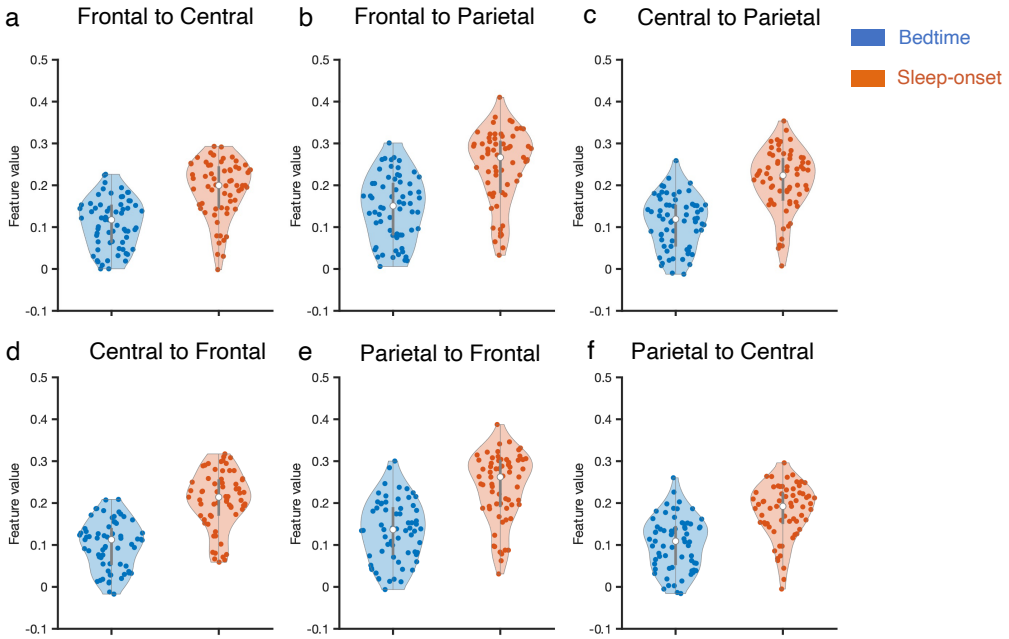


Figure 5.11: The transfer entropy (TE) feature values distribution differences between bedtime and sleep-onset in the Montreal dataset with 3 regions (Frontal Fz, Central Cz and Parietal Pz). **a**. TE from Frontal to Central; **b**. TE from Frontal to Parietal; **c**. TE from Central to Parietal. The reversed direction can be found in (**d-f**) respectively. Blue indicates bedtime, and red indicates sleep onset. A two-sided two-sample t-test is applied between bedtime and sleep onset for each one of them, and all significant with $p < 0.0005$. The full stats can be referred to in Appendix B.30. Violins show the kernel density plot (external shape), median values (white non-filled circle) and upper and lower 25% quantiles (grey bar).

Based on the same criterion, only TE became the key feature type. Looking at the clustering structures, it can be found that MI was uncorrelated with TE and had a very bad accuracy (only 52.05%) on average in this dataset. Similarly, the Pearson correlation feature type only had an average of 53.08% accuracy. The distributions of TE between all pairs of regions and directions can be found in Figure 5.11. The same findings can be found that for all region pairs and directions, there was a significant increase in information transfer (all p-values smaller than 0.0005).

The accuracy gap, as well as the big differences in the clustering structures between these two datasets, is interesting. The most likely reason was the EEG referencing differences. Therefore, an exploration test was then done by re-referencing the electrodes of the Montreal dataset similar to that of the MESA dataset: Fz-Cz (Frontal region), Cz-Pz (Posterior region), and C4 (Right-Central, linked ear referenced, however). The re-referencing was simply done by subtraction of the two relevant channels, for example, Fz-Cz was derived by subtracting the Cz from the Fz EEG signal. The same montage could not be achieved since there was no Oz channel in the Montreal dataset, as well as that the original electrodes were linked-ear referenced. Therefore, this was simply an explorative step trying to approximate that of the MESA to see whether the differences could be somewhat explained by the montage. The results for the re-referenced

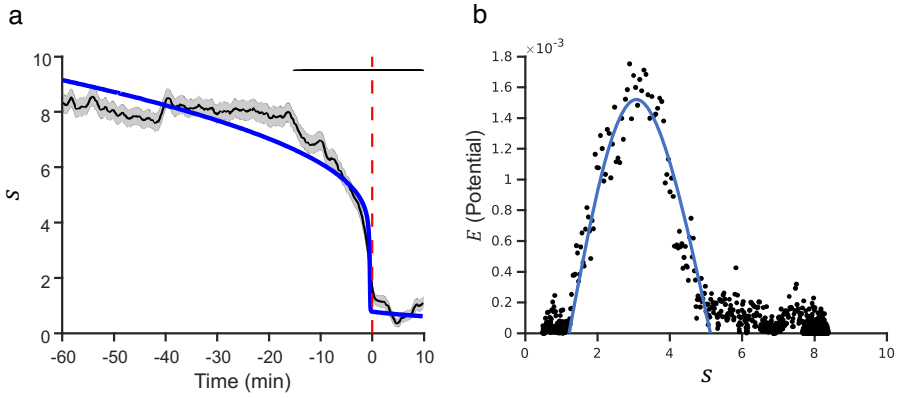


Figure 5.12: The sleep state variable dynamics in the interactional feature state space. a. The group-averaged temporal dynamics of sleep state variable s from 60 minutes before sleep onset to 10 minutes after, shown are mean ± standard error of the mean (s.e.m.) across 150 samples. The blue line is the theoretical model from Equation (4.1) with parameters optimally fitted to the empirical data. The horizontal lines show timestamps with a significant decrease compared to the mean values of the first 10 minutes, using the right-tailed two-sample t-test, with a p-value threshold of 0.025 dividing the total number of timestamps (for Bonferroni correction). The red vertical dashed line shows the sleep onset. **b**. The kinetics energy potential ($E = \Delta s^2/2$) as a function of s. The blue line in this subplot is for visual illustration only but not from the model.

Montreal data and comparisons to the previous results can be found in Appendix B.31. The most straightforward observation is that now the accuracy for MI increased drastically from 52.05% to 77.69%, and it was clustered together with TE now. The re-referencing also had a positive impact on the Pearson correlation measurements (around 8% increase in accuracy). Some feature types would have negative impacts, for instance, the synchronisation likelihood dropped by 8.20% in accuracy. The other features, including TE, could generally be concluded to be robust to EEG montages, especially linked referencing.

5.3.4. The continuous inter-regional interaction dynamics during falling asleep transition

The same analysis approach was applied here to study the continuous inter-regional dynamic changes during falling asleep. Firstly, the sleep state variable s time series was computed in this interactional feature state space (240 features) using the same bootstrapping technique, with 150 samples at each timestamp. There were some features excluded due to very small standard deviations (smaller than 0.001), which would bias the entire results. The s time series was then smoothed using a moving-median filter with a window size of 20. The result is shown in Figure 5.12, along with the potential analysis. The same bifurcation model was also fitted to the interactional feature state space s dynamics (Figure 5.12a blue line), with the R-squared of fit 0.940. The optimal model parameters: $r = 54.69$, $K = 9.20$, $h = -1.24$ and $m = 2.27$. It is obvious that m, the increasing rate of the control parameter, was much larger than those fitted to the regional dynamics. The larger dynamical range (from around 8 to 0) is also illustrated in the much higher potential wall (energy required) from one state to another, where the energy

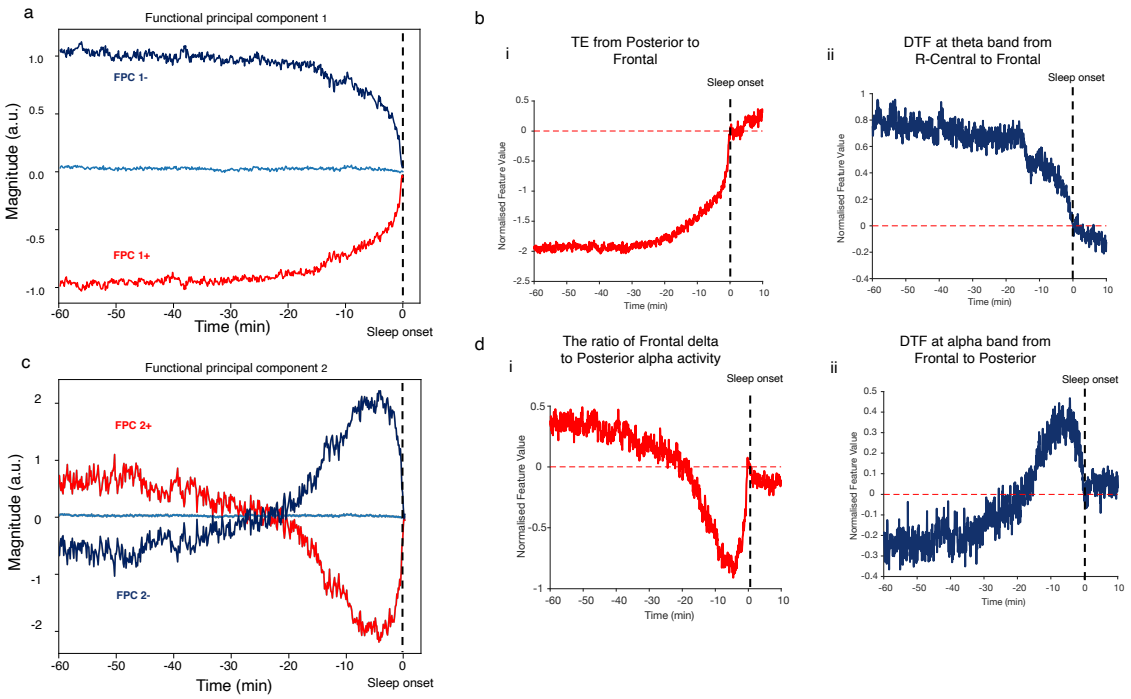


Figure 5.13: Functional Principal component analysis (FPCA) for extracting key feature dynamical patterns for inter-regional features. FPCA analysis is done on the inter-regional feature dynamics from 60 minutes before sleep onset up to sleep onset. Two functional principal components (FPC) are identified, with FPC1 in (**a**) taking 95.55% of the total variance, and FPC2 in (**b**) taking 3.38%. Each FPC is shown in both positive (FPC+, red line) and negative (FPC-, dark blue line) directions relative to the mean of all feature time series (light blue line), with the y-axis having arbitrary units. The 'positive' and 'negative' means 'addition' or 'subtraction' of the actual FPC to the mean. For each FPC, the most representative features (both positive and negative) are selected based on the similarity score (maximum and minimum), and shown on the right. For FPC1, the most representative features are the transfer entropy from Posterior to Frontal (**bi**) for positive and the DTF at theta band from Right-Central to Frontal (**bii**) for negative. For FPC2, the most representative features are the ratio of delta power in Frontal to alpha power in Posterior (**di**) for positive and the DTF at alpha band from Frontal to Posterior (**dii**) for negative. The horizontal dashed red lines indicate a 0 feature value, and the vertical dashed black lines indicate the sleep onset.

magnitude is 10^{-3} rather than 10^{-4} previously (Figure 5.12b). This result illustrates that the inter-regional interactional feature dynamics also follow the bifurcation dynamics during the falling asleep transition.

Coming to the specific feature dynamics, the continuous feature time series for all the 240 features (normalised to the 1-minute sleep onset) can be seen in Appendix B.32, arranged in a sequential order. The feature space was very huge, therefore FPCA analysis was important here as a dimensionality reduction to extract the key dynamical features. The FPCA results are shown in Figure 5.13, along with the key representative dynamical features chosen based on the similarity scores. Firstly, the FPC1 and FPC2 dynamics were still similar to all the other regional dynamics observed before, where FPC1 (explains 95.55% of total variance) showed similar bifurcation type dynamics and FPC2 (explains 3.38% of total variance) showed a burst type of change. For FPC1, the representative features were the TE from Posterior to Frontal (similarity score

1, positive), and the DTF at theta band from Right-Central to Frontal (similarity score -1, negative). For FPC2, the representative features were the ratio of delta in Frontal to alpha in Posterior (similarity score 1, positive), and the DTF at alpha band from Frontal to Posterior (similarity score -0.92, negative). The two representative features for FPC1 were exactly the same as the two identified previously from hierarchical clustering (outside the feature types, see Appendix B.29). Many other features had a similarity score of 1 to FPC1 (see below), but for FPC2, those two were the only representative ones.

It was previously observed that TE has significantly increased between all region pairs and directions at sleep onset. The continuous feature dynamics of those TE features (see Appendix Figure B.20) also all showed a monotonic nonlinear increasing pattern, and the similarity scores between them and FPC1 were all 1 (see Appendix B.33), which means that they are interchangeable for the representative dynamics of FPC1. The DTF features were also interesting, as both FPC1 and FPC2 could be represented by some specific ones in it, although most DTF features had a dynamical pattern similar to FPC1 (see the similarity scores table in the same Appendix). When looking at the continuous dynamics of the DTF features as a whole (see Appendix Figure B.21), it can be observed that there was a general increase of information outflow from Frontal to Posterior and Right-Central during the falling asleep transition at especially low-frequency bands (delta and theta), and while the reversed directions decreased. Interestingly, DTF had both increasing and decreasing causal information flows while TE showed global information transfer increasing (sharing or dependence). This is most likely because this normalised DTF is a relative metric (to all the incoming information flows) when interpreting its changes. The continuous dynamics are also shown for the cross-channel power band ratio features (Appendix Figure B.22). As an overview of the results, it can be seen that the Frontal region had a much larger delta and theta dominance compared to the Posterior and Right-Central regions before sleep onset. The specific feature, the ratio of Frontal delta to Posterior alpha, was the best representative feature of FPC2. From the plot, it seems that the ratios between the Frontal delta to the alpha or beta activities in either Posterior or Right-Central showed the same pattern (a burst of decreasing changes prior to sleep onset). As previously discussed, this cross-channel band power ratio is not a connectivity feature, and its dynamics should reflect underlying power-band dynamics in different regions. By looking at the specific dynamics of the Frontal delta and the alpha in Posterior or Right-Central (see Appendix B.34), it seems that this burst of change in these cross-channel band power ratios was governed by the Frontal delta activity (which showed the same burst of reduction and then went up), while the alpha activity in both Posterior and Right-central showed a monotonic increasing pattern. Note again that when interpreting these features, it is important to know that the 'band power' here is a ratio itself normalised to the total spectrum powers.

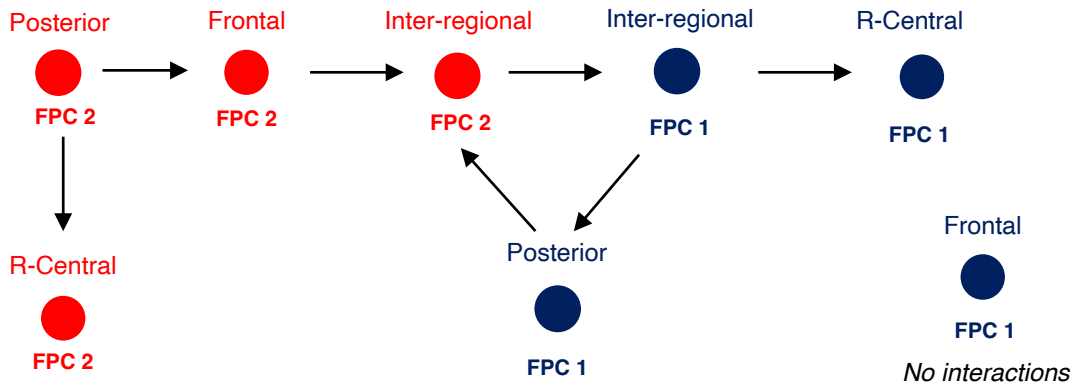


Figure 5.14: Granger causal interactions between all pairs of regional and inter-regional FPC dynamics. The FPC time series used are the FPC1 and FPC2 from the Frontal, Posterior, Right-Central regional dynamics, and the Inter-regional interactional dynamics. The arrows indicate a significant Granger causality from one feature to another, using the Toda-Yamamoto approach with a p-value threshold of 0.025 (corrected for two directions). The p-value table for all directions can be found in Appendix B.35. The final VAR model has maximum integration 1, optimal model order $p_0 = 3$, and extra lags $p_1 = 6$.

5.3.5. A Granger-causality based framework between regional dynamics and inter-regional interactions during falling asleep

Now given all these regional and inter-regional dynamics, I here aim to test how are these different dynamic 'movements' linked to each other during the sleep initiation process. Therefore, all the regional and inter-regional FPC time series were taken (to represent their key feature dynamics), and the same Toda-Yamamoto approach (introduced in Chapter 4) was used to test the Granger causality between the different pairs of FPC dynamics. The results are shown in Figure 5.14, with significant causal effects shown in arrows. The results in sum show that, the FPC2 Posterior was the source that caused the dynamic changes in all other regional and inter-regional FPC2 dynamics, as well as the Right-central, Posterior and inter-regional FPC1 dynamics. The Frontal FPC1 dynamics had no causal interactions with any others. There was an interesting loop inside, where the inter-regional FPC2 caused its own FPC1 changes, and this inter-regional FPC1 caused the Posterior FPC1 dynamics, which in turn caused back to the inter-regional FPC2 dynamics. Bringing the specific features into consideration, the results imply that a burst of increase in the information transfer (DTF) in the alpha band from the Frontal to the Posterior region, or the decrease burst in the Frontal delta power normalised to the entire spectrum (the unique representative FPC2 dynamics for inter-regional features) would cause a general increase in the brain's information sharing (an increase in TE in almost all regions directions). This increase in the general information sharing (for future prediction) between regions caused the increase in total EEG spectrum power or the decrease in the peak beta band frequency in the Posterior or Right-Central region, as a continuous de-arousal process. This gradual de-arousal in the Posterior region was then fed back to enhance the Frontal-to-Posterior information transfer and the dynamical changes in this enhancing loop led to the

sleep onset. If one changes the representative features, there could be other different interesting interpretations, but essentially, this self-enhancing loop of dynamics for synchronisation and de-arousals, as well as Frontal-Posterior connections, shed light on novel insights for sleep initiation.

5.4. Discussion

5.4.1. The regional EEG brain dynamics during falling asleep change quantitatively but not qualitatively

The regional EEG brain dynamics during the falling asleep transition were studied following the routes of the previous two chapters. Firstly, it was found that different cortical regions do have differences in the sensitivity to detect the EEG dynamics changes at sleep onset compared to bedtime, with the Posterior region having the best performance. The Frontal region was dragging down the global performance, and using only four features from the Posterior (the entropy and the low-to-high frequency band power ratio) and Right-Central regions (the theta band power and the low-to-high frequency band power ratio), one could already achieve around 90% accuracy which is much higher than the global performance using all features. Different regions also had some different key features (although mostly captured on the global level), which makes sense that originally, these different derivations were to capture different cortical changes at different regions for sleep scoring purposes (Silber et al., 2007). Still, one has to be careful when interpreting these results, as the number of EEG channels in the MESA dataset fundamentally limited the cortical regions that could be studied. It could be further replicated in other datasets with more EEG channels, to see which cortical region (as well as EEG derivations) could have the most sensitivity in telling the differences at sleep onset compared to bedtime, and potentially gain more insights.

When it comes to the continuous dynamic changes during the falling asleep transition, it can be concluded that the global dynamics (combining all channels) can well represent the regional dynamics. All three different regions showed bifurcation dynamics when observing the sleep state variable s, with slight parameter differences. The Frontal region showed the smallest dynamic changes and seemed to be much closer to the sleep onset state in the feature state space prior to the transition. This is also illustrated in the specific regional feature dynamics, that features in the Frontal region showed smaller changes compared to Posterior and especially Right-Central. The results here could potentially explain the reasons that the Frontal region had the worst classification accuracy between bedtime and sleep onset. This result also corresponds well to the previous literature findings that Frontal regions have naturally more homeostasis forces (and thus sleep pressure), and show the strongest slow-wave activities (implying that it is closer to the sleep state). Despite these slight differences, the FPCA analysis showed great

consistency, that almost the same feature set as from the global analysis can well represent the key governing feature dynamic patterns in all three regions, as most of them have similarly high similarity scores and were interchangeable for representing the dynamics. This implies that, despite all the small quantitative changes in different regions, there is no qualitative change in the underlying governing laws, and that the falling asleep transition is a bifurcation process with the same driving parameters. This finding is consistent with the proposal from Merica and Fortune (Merica and Fortune, 2003), that the shape of the time course during the falling asleep transition is determined by underpinning brain stem controls, with site-dependent quantitative differences controlled by the cortical-thalamic interactions.

In summary, the findings suggest that although different brain regions display quantitative variations, their dynamics are driven by global state transitions and can be described uniformly. These quantitative differences might still be useful, though, for instance, the extra sensitivity in the Posterior or Right-Central regions could better help sleep-onset tracking. The specific regional dynamics could provide better value in detecting early warning signals or tipping points prior to the transition, which could be a very interesting next step. The region-specific dynamics could also be important in quantifying the paradoxical insomnia problem, as proposed by Ramjford et al. (Stålesen Ramfjord et al., 2020), patients with insomnia potentially have more specific regions with wakefulness patterns that cannot be detected using standard polysomnography and sleep stages. This means that by applying the specific regional features or using bifurcation models, one could identify regional transition patterns, for a better marker of sleep states (instead of standard sleep staging or global patterns) or clinical diagnosis of different sleep issues.

5.4.2. Sleep onset is hallmarked by a global increase of information transfer between regions

Using the developed interactional feature base, I studied inter-regional EEG dynamics that differ the sleep onset states to bedtime, and it was found that transfer entropy (TE) was the most important feature. Specifically, the TE increased at sleep onset in all directions, showing a global increase in information transfer between regions.

Using a feature type-wise hierarchical clustering, there were two most prominent interactional changes were identified: the Pearson correlation and the transfer entropy (TE). The Frontal region had a decreased correlation to the other two regions at sleep onset, while the other two regions became correlated with each other. Still, the Pearson correlation only quantifies a linear relationship between the EEG data samples, which is very hard to interpret with regard to how this correlation is linked to actual dynamics aside from a 'signal similarity'. Therefore, it could be used as a biomarker but without much mechanistic value. TE showed a significant increase between all regions and all directions. This indicates a global-level information-sharing structure emerging at sleep onset. Specifically, TE features all had a very good classification performance,

with individual accuracies all above 88%, which is almost equivalent to the best performance that can be achieved using regional EEG features. By combining TE and Pearson features, the accuracies can reach 91.7%. The analyses here show the specific values of using the inter-regional features, as they provide extra sensitivity to distinguish these two states even using just one measurement from the brain, which could be a very powerful biomarker for detecting state transitions.

The replication analysis in the Montreal dataset further showed that TE was the most prominent interactional change at sleep onset compared to bedtime, with the same increasing pattern. However, TE had a lower accuracy here, and the clustering structure was also significantly changed, with the performance of some other features (MI and Pearson specifically) significantly degraded. Aside from the other potential facts of the dataset differences (no MCI patients were included), the most potential cause is the EEG montages. How the EEG electrodes are referenced could indeed be an important factor affecting the connectivity evaluations (Chella et al., 2016; Huang et al., 2017). For instance, it has been shown that the synchronisation likelihood works best for linked-ear references (Stam and de Bruin, 2004). On the other hand, DTF was shown to be agnostic to montage (Gennaro et al., 2005). The specific challenge could be that the bipolar EEG channels in the MESA datasets were already linked (Fz-Cz and Cz-Oz), which could result in spurious statistical dependencies or correlations. By re-referencing the EEG data in the Montreal dataset, it was indeed shown that MI had a drastic increase in performance (more than 25%), which could potentially be due to these linked EEG measurements. The synchronisation likelihood had a lower accuracy (8%), which matched the previous findings that it works best for linked-ear references (Stam and de Bruin, 2004; Ferri et al., 2007). The DTF was robust to this re-referencing, which also matches the previous conclusion. It is interesting to see that TE was also not affected. This makes sense in considering that, TE is a conditional mutual information, where the information of the past of either X or Y will be excluded. This means that even if X and Y have spurious statistical dependency (illustrated by the MI feature), it will not affect the TE, as the shared information will be removed from evaluation anyway. Of course, as a future step, the effects of EEG montages should be studied more rigorously, and taken into consideration in the analysis. For instance, from hindsight, MI should have a different interpretation or even be excluded in the bipolar montages.

There are limitations to this analysis. The feature type-wise hierarchical clustering was useful in identifying the most prominent 'type' of interactional changes, but the averaged within-type accuracy could not represent every specific feature in that type. By doing the hierarchical clustering outside the feature types, it was shown that the general clustering structures are preserved, where most of the features from the same type were still clustered together. There were four features selected here, with one feature from each type (TE, Pearson, DTF and PAC) representing most of the other features of the same type and with similar performances. TE and Pearson's features have already been identified before, with good within-type consistencies

and performances. The other two types (DTF and PAC) did not have a good average accuracy, but this analysis showed that there were subsets of them which could be useful as well in distinguishing the two states. Specifically, the decreased coupling between the Posterior theta phase and Right-Central beta amplitude captured a lot of other PAC features in its cluster, and decreased DTF at the theta band from Right-Central to Frontal captured a lot of other DTF features and some power band features. The general performance of the four features selected here was not different from the previous. Therefore, as an initial exploration, this feature type-wise hierarchical clustering was generally sufficient for identifying the most prominent type of interactional changes that were also easy to interpret. But as a future step, one could delve into each type specifically to identify the specific one within each type that could be useful for distinguishing the two states, and gain more mechanistic insights potentially.

5.4.3. A Granger-causality based framework eluding to a self-enhancing loop between regional and inter-regional dynamics during sleep initiation

It was found that the continuous inter-regional interaction feature state space dynamics could also be described as a catastrophic bifurcation but with a much larger dynamical range than regional dynamics (around doubled). This extra dynamical range was also consistent with the binary classification results, that these inter-regional features showed extra sensitivity. The FPCA analysis showed the same two types of governing dynamics, with FPC1 showing a bifurcation pattern, and FPC2 showing a burst of change prior to sleep onset and then returning. For the representative features of FPC1, a lot of features were interchangeable, as the similarity scores showed. All six TE features had similar bifurcation-type increasing patterns and similarity scores of 1 to FPC1. Most of the DTF features also had bifurcation dynamics with a maximum (or minimum) similarity score to FPC1, with most of them showing a relative increase of Frontal to Posterior information flow in the low-frequency bands (or decrease in the reversed direction). This result also matches the previous literature; for instance, Gennaro et al. concluded that the frontal region leads the propagation of global synchronisation, especially in low-frequency bands (Gennaro et al., 2004). Summarising them all, this means that the inter-regional FPC1 dynamics captured a global information transfer (dependence), with the Frontal to Posterior information flow growing larger than its reversed direction (recall again that the normalised DTF is relative). The representative features for FPC2 were unique, with one very specific increasing burst of alpha band information flow (DTF) from Frontal to Posterior, and another feature, the ratio between Frontal delta and Posterior alpha activity (a non-connectivity feature), which was demonstrated to be governed by the dynamics of the Frontal delta (the decreasing burst shape). The decreasing burst of delta in the Frontal did not necessarily mean a disappearance of delta rhythms but was more likely due to an increase in theta or alpha activities in the Frontal region, as this was a normalised measurement of the entire spectrum. Of course, how these features are related to the underpinning neurophysiological mechanisms behind the falling asleep transition

is not clear here.

In summary, the initial exploration revealed significant changes in EEG interactions between regions during the transition to sleep. These changes were most effectively captured through Transfer Entropy (TE) and Directed Transfer Function (DTF) analyses, which examined the connectivity dynamics between EEG signals in different brain regions. Notably, these methods demonstrate greater sensitivity than traditional regional features. While these findings offer valuable insights, there is still much to uncover for a whole picture to emerge. Specifically, there needs to be a deeper understanding of how regional dynamics relate to each other and how the interactional actions are linking the different cortical regions for sleep initiation. To address this, a Granger causality analysis framework was proposed to examine the relationships between different levels of EEG dynamics. These findings suggested an intriguing loop, with the DTF in the alpha band from the Frontal to the Posterior (or the Frontal delta activity) acting as a driving force (or the control parameter) to generate a global cortical dependence, ultimately leading to gradual Posterior de-arousal. This key interactional dynamic then appears to reinforce itself, culminating in a global state transition indicative of sleep initiation. The importance of Frontal regions for sleep initiation is clear, for instance, an examination of cross-channel power band ratios reveals a Frontal low-frequency (delta and theta) dominance at bedtime, gradually diminishing during the process of falling asleep. This diminishing relative dominance is related to the increased information flow at the low-frequency bands from the Frontal to the Posterior region, which represents the spread of sleep pressure (Gennaro et al., 2005). The Frontal dominance of low-frequency activities (relative to other regions) also matches the existing literature, indicating that the Frontal region experiences a distinct activation input with more homeostasis pressure (D' Ambrosio et al., 2019), and has a larger sleep need (Ferrara et al., 2002). The burst increase in the information transfer from the Frontal to the Posterior (the representative feature of inter-regional FPC2 dynamics) corresponds to the shift from posterior to frontal alpha dominance (see Section 1.3 for details). This framework further consolidates this pattern as a self-enhancing loop with key initiation features. However, it is essential to acknowledge the primitive nature of this framework, as several other complex causal components remain challenging to interpret, such as the key driving sources of all dynamics from Posterior FPC2 in the causal diagram. The limitations of the Granger causality methodology and the specific type of 'causal' relationship (as predictability of future) it is capturing might also limit the general interpretations. There could be many of other information inside the other detailed interactional feature dynamics as well. The most challenging part could be the link from any of these regional dynamics and inter-regional interaction changes to underlying neuronal mechanisms related to wake-sleep regulations. Specifically, the functional meanings of these interactional features (such as TE and DTF) are also unclear. How is this information transfer or interactions in different frequency bands observed in the EEG dynamics happening, and how are they related to underlying deeper brain structure control inputs or cortical-cortical linkages such as frequency mixing? One future

step, which is currently still in exploration, is frequency mixing as the fundamental mechanism of such cortical synchronisation and 'spread' of information (Luff et al., 2023; Haufler and Paré, 2019). Relevant questions can also be asked for a more detailed causality test between different interacting dynamics (such as power changes), or a more in-depth algorithm to evaluate TE in different frequency bands or even oscillation phases.

5.4.4. General limitations and future directions

Aside from the interactional feature base, as mentioned, the very limited EEG channels available in the MESA dataset did not provide many options. The EEG regions mentioned in the text are also an approximation of the cortical region that is measured, due to the specific bi-polar montage in the dataset. This could also be a constraint in actually interpreting these connectivity features, as they could be a very rough estimation of the cortical connections. Of course, the ultimate constraint lies in EEG itself, as it measures the underlying surface cortical activity as a whole, which makes it difficult to pinpoint the exact sources of the signals, let alone the knowledge of deeper brain interactions. There are methods such as dipole analysis or source locational analysis to locate the EEG sources, but that normally requires high-density EEG data, with complex factors and parameters to consider in the analysis to get sensible results (Eom, 2022). Still, future work would involve datasets with more EEG channels and uniform references (thus better spatial resolutions), and trying to incorporate other spatial analysis approaches such as the graph theory (on the connectivity maps built on these interactional features), or spatial-temporal analysis approaches such as dynamic mode decomposition (Brunton et al., 2016a). Compared to previous work doing graph theoretical analysis on the falling asleep brain dynamics (Koenis et al., 2013; Ferri et al., 2007; Vecchio et al., 2017), the comparative advantage here is the more comprehensive feature base to construct the graph (network) based on different interactional features.

There is much additional information that could be delved into for the analyses in this chapter, for instance, the specific informative features that can distinguish bedtime and sleep onset states or the other detailed feature time series dynamics that are not looked into due to the large feature number. For the specific informative features, such as TE or DTF, there should be a more in-depth study for the exact functional meaning of the critical information flow.

Many potential future actions have already been mentioned previously. One other important step could be incorporating this comprehensive inter-regional feature base into the local features framework, and studying them together. This novel feature base would then have both local and spatial information captured and could be applied for more generalised biomarker research. Similar to all previous chapters, the analysis could be expanded to ask specific questions, such as demographic effects, or how different neurological diseases could impact the brain's regional and inter-regional dynamics.

6. Testing sleep onset augmentation via phase-locked auditory stimulation

6.1. Introduction

In the previous chapters, I developed a strategy to track the transition from wakefulness to sleep by unveiling the governing EEG dynamics and biomarkers to support the development of closed-loop neuromodulation therapies. In this chapter, I aimed to use this knowledge to test the feasibility of augmenting the falling asleep process using phase-locked auditory stimulation.

Earlier studies deploying auditory stimulation targeted at the up-phase of the slow-wave activity reported improvement in the slow-wave activities and declarative memory performance compared to sham (Ngo et al., 2013, 2015; Cox et al., 2014). Studies using open-loop audio stimulations were reported (Forquer and Johnson, 2007; Messineo et al., 2017; Kawada and Suzuki, 1993; Montgomery-Downs et al., 2010; Nasari et al., 2018). These studies used subjective sleep questionnaires or even just immobility periods to assess sleep, without EEG recording (see Chapter 1 for a review). Conceptually, the auditory stimulation phase locked to brain rhythms might be able to promote sleep by inducing low-frequency oscillations, as has been demonstrated in other sensory stimulations (Bayer et al., 2011) and slow-wave sleep examples. Earlier feasibility studies in our group found that auditory stimulation (pink noise sound) phase-locked to the trough of the theta-band (4-8 Hz) recorded from Fz (mid-frontal region) in the eyes-closed increase the low-to-high power ratio as well as the theta-to-alpha band power ratio, while peak stimulations had the reversed effect (unpublished results), suggesting reduced vigilance.

The Frontal region plays an important role in sleep initiation, and the Frontal theta band activity has also been shown as an important cognitive control of the global cognitive process (Nigbur et al., 2011; Cavanagh and Frank, 2014). Theta band oscillations were also proven to be important during sleep initiation, and as low-frequency activity, it is relatively easy to track for real-time stimulation and more relevant to augmenting sleep onset (where the brain slows down). A second study conducted by the lab systematically tested the effect of different phase conditions (pre and post-peak 0°, pre- and post-trough 180°) of the theta band oscillation in the

Frontal region (Fz channel) on the brain EEG. The experiment was conducted in an eyes-closed resting state scenario, with one block for each phase condition, and each block containing 120 stimulation pulses with 1s inter-stimulation interval (in total 4 minutes for each block including baselines). Using the four EEG biomarkers identified in Chapter 3 as the outcome measurements (specifically, the alpha-to-beta band power ratio, the low-to-high frequency band power ratio, peak beta band frequency and the total EEG spectrum power), it was found that at the group level, post-peak stimulation was optimal in driving these features in the same direction as falling asleep at the group level (unpublished).

These preliminary studies provided the rationales to test whether phase-locked auditory stimulation can augment the falling asleep process. The need to do phase-locking and the closed-loop stimulation paradigm was discussed in Section 2.6. The experiment used the auditory stimulation phase-locked to the post-peak (72°) of the theta (5 Hz) oscillations (recorded from the midline Frontal region). The experiment included a 15-minute day nap with either stimulation or a sham. Specifically, the experiment will include cognitive tests before and after a 15-minute nap. Short naps have been previously related to potential cognitive benefits such as verbal memory performances (Tucker et al., 2006; Schmidt et al., 2006; Mednick et al., 2013), and this positive value was correlated to the intensity of slow wave activities during the nap. Thus, the experiment also tested the effect of cognitive tests, to see whether the auditory-augmented nap could also improve cognitive performance. Further experimental designs are explained in detail below. I declare here again that the experiment was co-designed by me and Dr. Edward Rhodes in the lab. The data collection was done with the help of other Master's students in the lab. The data analyses presented belong to my authentic work.

6.2. Methods

6.2.1. Participants recruitment

The participant recruitment was done at Imperial College London (an opportunity sampling), where most of the participants were current students or staff from the college. The participants were screened to ensure that no presence of personal or family history of neurological disorders or psychiatric illnesses, no current brain-related issues (such as migraine), and no implants in the head. As this study will use auditory stimulation, the participants were ensured that they had normal hearing. For the specific purpose of studying sleep onset, the participants were ensured that they had no use of sleep-related drugs (especially to treat insomnia). All participants were required to fill out a 7-day sleep diary before their test day, and participants with abnormal sleep patterns or suspicion of sleep disorders were excluded. The study was approved by the Imperial College Research Ethics Committee. All participants provided their written informed consent

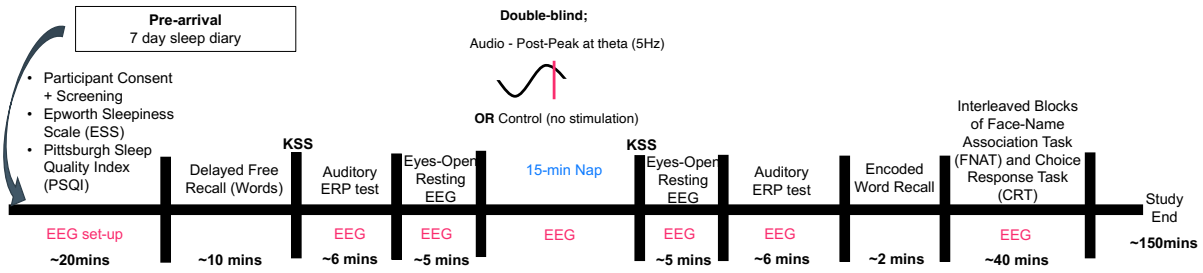


Figure 6.1: The entire experimental setup and procedure. Before the experimental visit, each participant must have a 7-day sleep diary. On the day of the visit, the participant will first sign the consent form and undergo a screening. Two sleep questionnaires will follow: the Epworth Sleepiness Scale (ESS) and the Pittsburg Sleep Quality Index (PSQI), during which EEG set-up will be performed (approximately 20 minutes). The participant will then undergo a series of tests, including the Delayed Free-Recall task (around 10 minutes pre-nap and 2 minutes post-nap), auditory event-related potential (ERP) test (around 6 minutes both pre and post-nap), resting state EEG recordings (around 5 minutes both pre and post-nap), a 15-minute nap, and the Face-Name Association Task (FNAT) interleaved with Choice-Response Task (CRT) (around 40 minutes), in sequential order as in the figure. All times shown are estimated to include preparation. The nap will be either control (no audio) or audio condition (stimulation phase-locked to post-peak of the 5Hz theta band). The tests with concurrent 32-channel EEG recording are marked with 'EEG' in red text below. There will be two Karolinska Sleepiness Scale (KSS) tests before and after nap. The entire visit would take approximately 150 minutes, including all setup times, in an ideal scenario.

prior to participation and received monetary compensation for their participation. There are in total $N = 36$ with complete data (17 for audio conditions and 19 for sham), with ages ranging from 18 to 34 (mean $\pm$ st.d. 23.03 ± 2.97). Further demographic details and comparisons will be shown in the results.

6.2.2. Experimental design

The experiment generally took a double-blind between-subjects design. Participants were randomly streamed into two parallel experimental condition groups: audio (with auditory stimulation phase-locked to the Fz 5 Hz theta band) or control (sham with no stimulation). Each participant was randomly assigned to one of the two sites in the same building ('Floor 7' or 'Floor 9') and chose their arrival time as 11 a.m. or 2:30 p.m. There were many differences between the two sites, such as ambient sound volume and temperature, therefore, the random assignments equally delivered participants to each location. For each visit, the experimental setup and procedures can be seen in Figure 6.1. Each visit took approximately 150 minutes, as shown in the figure, with a series of tests including subjective sleep metrics, behavioural measurements, and the 15-minute nap with auditory stimulation or sham. From arrival, participants gave their consent and underwent additional screening to ensure that inclusion criteria were met. Then, the tests were done sequentially, as shown in the figure. The details of different tests and data recordings (including EEG) will be introduced in the following.

The subjective sleep metrics

The experiment had three subjective metrics: the Epworth Sleepiness Scale (ESS), the Pittsburg Sleep Quality Index (PSQI), and the Karolinska Sleepiness Scale (KSS). ESS and PSQI were performed at the beginning of the experiment after the participants had given their content. These two tests have been introduced previously in Section 2.1 for the Montreal dataset. ESS measures subjective sleepiness before testing, and PSQI measures subjective sleep quality in the last month. The EEG set-up was done during this period (more details to follow). The KSS (Shahid et al., 2012) measures the subjective sleepiness at the time of the test, with a scale from 1 to 10 and 10 indicating the most sleepiness. The KSS was performed twice, once before the nap (just prior to the first auditory event-related potential task) and once just after the nap, so that the differences before and after the nap could be recorded and compared. The questionnaires were done digitally on the laptop and then saved onto the local laptop after they were finished.

The Delayed Free-Recall task

The design of the Delayed Free-Recall task can be found in Figure 6.2. The task was adapted from Jacobs et al. (Jacobs et al., 2016), designed to probe verbal episodic memory. The task was implemented in MATLAB using the Psychtoolbox (Kleiner et al., 2007). Basically, the task was divided into two parts, one before the nap and one after. The pre-nap recall period contained 5 blocks and a full-list recall. For each block, the participants were presented with a list of 12 words (nouns only) to remember. Each word was on-screen for 1.6 seconds with a randomised inter-stimulus interval of 800-1200 milliseconds (ms). The participants were allowed to speak out the word when presented. After the entire word list had been presented, an 18-second distractor task was performed, during which the participants were asked to determine whether the sum on the screen was correct or incorrect. The distractor task was designed to prevent the participants from constantly reciting the words in their minds. Finally, there was a 45-second free recall period during which the participants verbally recited as many of the words they could remember as possible. The five blocks resulted in a total of 60 words encoded. After all blocks were completed, the participants had 90 seconds to speak out as many of those words they could remember as possible as the pre-nap full-list recall. After the nap, the participants performed a second free recall, where they tried to recite as many of the 60 encoded words as possible. As with the pre-nap, they had 90 seconds to do so. The audio tracks for the recall periods were recorded and saved on local laptops for researchers to mark their performance later. Prior to the task, the participants were ensured to be seated at a 60cm distance from the monitor. The Delayed Free-Recall task was performed without concurrent EEG recordings.

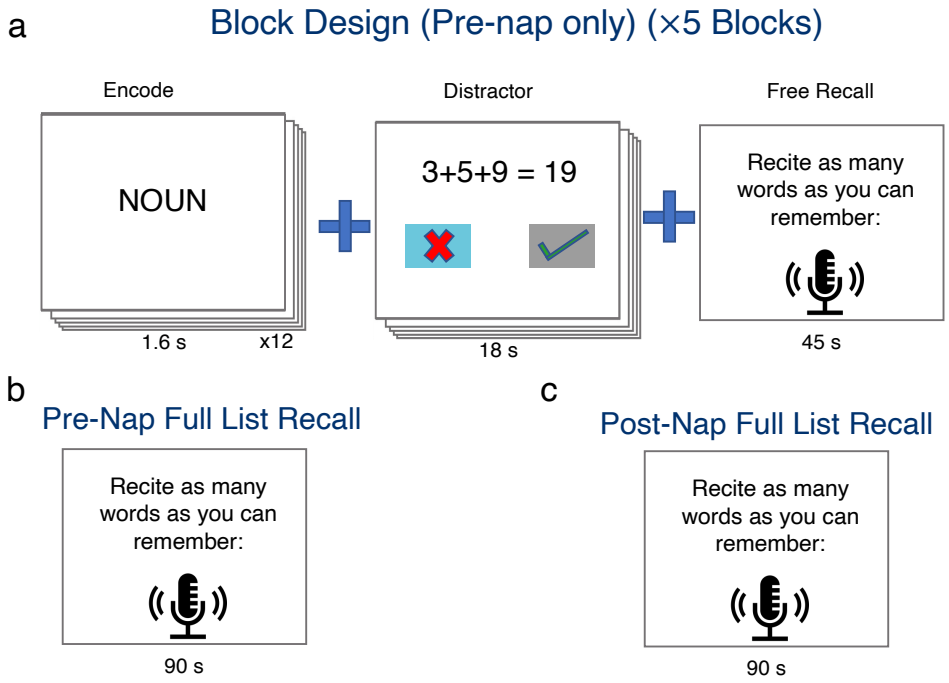


Figure 6.2: The design of the Delayed Free-Recall task. a. The block design for the pre-nap delayed free-recall. In each block (in total 5 blocks), there is a list of 12 nouns to encode. A simple 18-second mathematical task is then presented as a distractor. Afterwards, the participant will be asked to recite as many words as they can remember from the 12 words encoded previously in 45 seconds. **b**. After all five blocks, the participants are asked to do a full-list recall (of the 60 words) in 90 seconds. **c**. The full-list recall is repeated after the nap.

The auditory event-related potential test and the resting-state EEG recordings

The auditory event-related potential (ERP) test was done to probe the hearing functions, during which the participants were asked to count the number of audio clicks they heard as a measurement of attention. The 3-minute resting-state EEG was recorded for additional purposes, such as phenotyping of participants to predict auditory responses (more details will follow in discussions). Both tests were done twice, before and after nap. The data from them will not be analysed in this thesis. Therefore, they will not be further introduced in detail.

The nap period

The nap period lasted for 15 minutes, with an additional 20 seconds as baseline (no stimulation). Prior to the nap, the participants were moved to a reclined chair and allowed to adjust their position to their comfort. They were offered a blanket and an eye mask to make them feel comfortable in a sleeping environment. Then, the lights are turned off and all researchers will leave the room, leaving the participant alone quietly to sleep for this 15-minute period. The participants were instructed to try to relax and fall asleep, but they were also assured that it did

not matter even if they could not. The situations inside the room during the nap were monitored by researchers outside, and the participants could call out if there were any emergencies. The EEG cap was set up on the participants' heads, with in-ear earphones worn before the nap started.

The participant was either in the 'control' condition (no stimulation) or the 'audio' condition (phase-locked auditory stimulation). This was blind to both researchers and participants, and participants were instructed as 'you may or may not hear a sound. To ensure blindness, the codes for controlling the ecHT device for all participants were generated pseudo-randomly in MATLAB and pre-saved on the hard drive for use during the experiment. The codes control the output sound volume (0 decibels for control and 40 decibels for audio condition), and the participants were instructed not to tell the researchers whether they heard a sound after the nap. Prior to the nap, the setup was tested using a test code which delivered periodic audio to the participants, to ensure that the audio output was normal, and the volume was the same in both ears. Auditory stimulation was delivered using the ecHT device (see Methods Section 2.6). Specifically, the ecHT device would record the Fz-channel EEG signals, and estimate the phases of the 5 Hz theta band oscillation in real time to deliver audio outputs. The stimulation was delivered in 1-second trains, with each train containing five 6 ms pink noise pulses, and each pulse was delivered in phase to the target post-peak (72°) of 5 Hz theta band oscillation. The pulses were separated from each other by at least 130 ms, and each stimulation train had 1-second intervals between each other. The stimulation was delivered using audiometric in-ear earphones with insulation from electrical noise. The sound volume was set at 40 decibels (dB), which was not too loud to arouse the participants. Concurrent 32-channel EEG data were also recorded during the nap. The detailed EEG setup will be introduced later. For the 20 s baseline, there was no stimulation for both conditions.

The Face-Name Association Task (FNAT)

The Face-Name Association Task (FNAT) has been demonstrated to be a very sensitive measurement of associative memory, which is largely dependent on regional hippocampal activity (Zeineh et al., 2003) and relates to the amyloid-β burden in the elderly population (Rentz et al., 2011). The task was adjusted from that of Violante et al. (Violante et al., 2023). Basically, the task was designed in MATLAB using the Psychotoolbox as well. All faces were collected from the Chicago Face Database v.2.0.3 (Ma et al., 2015), with an equal number of males and females and an equal amount of black and white faces. The names were collected from the Office for National Statistics (Baby Names, England and Wales, 1996). The year 1996 was chosen as it was the closest survey year to the mean age (28 years old) of the faces selected, and all names had between 5 and 7 letters. Unisex names were removed from either the male or female list, based on name frequency, to prevent duplication.

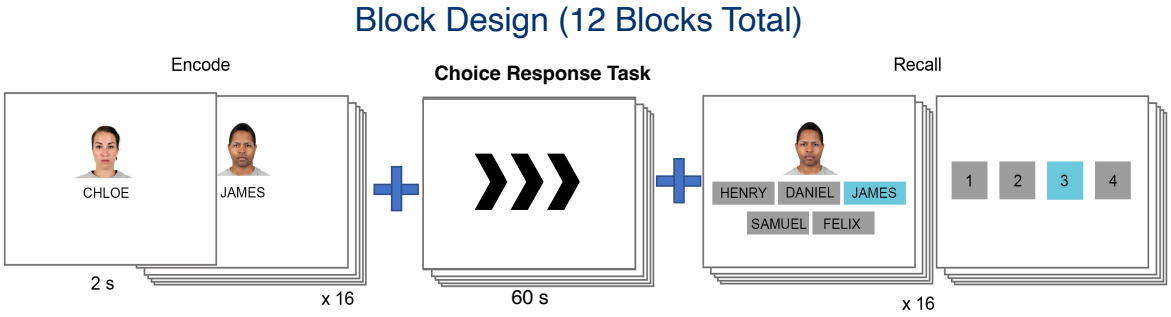


Figure 6.3: Block design of the Face-Name Association Task (FNAT) interleaved with Choice Response Task as a distractor between encode and recall period. During encoding, 16 combinations of face and name will be shown to the participant, with each one presented for 2 seconds. The participants will then do a 60-second Choice Response Task, during which they are asked to press left or right corresponding to the directions shown on the screen. During recall, the encoded faces will be shown with five name choices, and the participants need to make a decision within 8 seconds. The five names consist of the correct name, two distractors (novel names not encoded) and two foils (names encoded to different faces). After making the choice, they will be asked about their confidence level from 1 to 4 (most confident). This is repeated 12 times (12 blocks) in total, with a total of 192 trials. Figure adapted from Violante et al. (2023) under Creative Commons Attribution 4.0 International License.

The task was designed to be interleaved with the Choice Response task, as shown in Figure 6.3. Within each block of the task, there was an encoding and recall period, with a Choice Response task in the middle. During the encoding period, a single face and name were presented to participants for a duration of 2 seconds (16 combinations in total). The 16 unique faces consisted of different ethnicities: 4 black females, 4 black males, 4 white females and 4 white males per block; all with neutral facial expressions. After that, they were asked to do a 60-second Choice Response Task (respond left or right in the direction of the arrow on the screen), as a distractor to prevent them from constantly reciting in mind. The accuracy and response time of the Choice Response Task could also be used as a probe of the attention level of the participant with time passing. In the recall, the participant will be presented with the 16 faces encoded, with five name choices for each face. The five names consist of one correct answer, two distractors (novel names not encoded), and two foil names (names seen during encoding but correspond to different faces). The choice must be made within 8 seconds, and after that, they were asked to make a confidence-level judgement about their decision, on a scale of 1 to 4 (4 indicates the most confidence). They had 3 seconds to choose their confidence level. There were in total 12 blocks (192 trials), and the order of the blocks was kept constant across participants, but the order of the face-name pairs was pseudo-randomised across participants, such that no more than three faces of the same type appeared in a row. Concurrent EEG data were recorded throughout the period of this task, with trigger connections between the task laptop and the recording laptop, where there were trigger markers in the EEG simultaneously with the faces shown to the participants. During the experiment, the 12 blocks were further divided into three sessions with a short break between sessions, during which the EEG cap was checked and re-moisturised (see below). Similarly, prior to the task, the participants were ensured to be seated at a 60cm

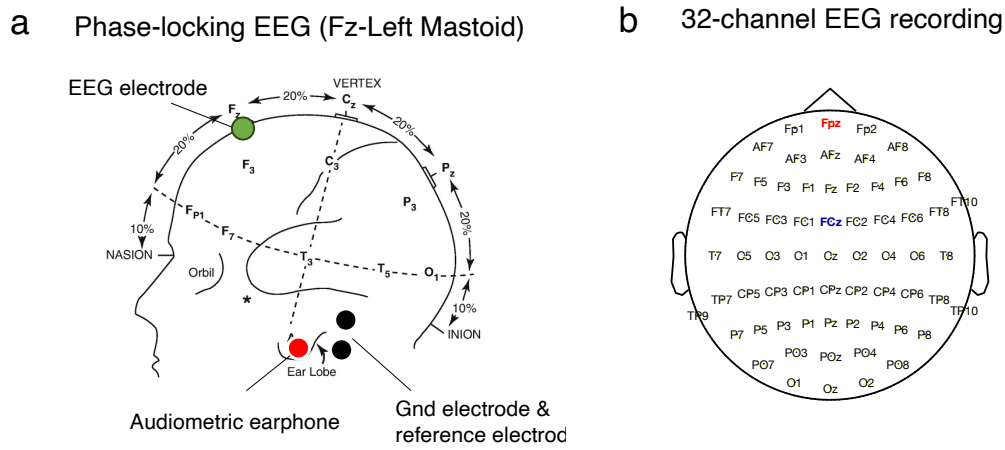


Figure 6.4: The setup of the two EEG systems. (**a**) for phase-locking EEG system, and (**b**) for 32-channel EEG recordings. The phase-locking EEG system records EEG at the Fz channel (green dot) to track phases, with reference and ground both at the left mastoid (black). The red shows the earphones. The 32-channel EEG system has ground at Fpz (red) and is referenced to FCz (blue). (**a**) was adapted from Sazgar and Young (2019) with kind permission from the publisher Springer Nature.

distance from the monitor, where they could access a keyboard by hand to do the task and make choices with the arrow keys and the space button. Normally, this task can be completed in 40 minutes, including all preparation time and EEG setup times.

EEG setup and recordings

There were two EEG systems for this experiment, one for the phase-locking system, and another for 32-channel EEG recordings (Figure 6.4). The setup for the phase-locking EEG system is shown in Figure 6.4a, where Fz channel EEG was recorded with ground and reference both at the left mastoid. The ecHT phase-locking system computed the phases of the recorded EEG in real time and delivered audio outputs to the participant. The electrodes were 1.5 millimetre (mm) Ag/AgCl ring electrodes, glued to the clean scalp of the participant using a conductive gel. The participants' heads were cleansed with a cotton bulb using an alcohol solution and conductive scrub prior to attaching all electrodes. The electrodes were further covered with gauze to fix them, as well as to separate them from the cap above (see below). The ecHT EEG system has a sampling rate of 1000 Hz.

The 32-channel EEG system used saltwater sponge-based R-Net caps (Brain Products GmBH, Germany). The sponge electrodes of the cap have Ag/AgCl sensors inside. There are four different sizes of caps to use according to the head circumference of the participant (55 cm, 57 cm, 59 cm, and 61 cm, cm stands for centimetre). To set up the cap, the jug provided was filled with 2 litres of distilled water at normal room temperature, and 2 spoons of granular KCL (around 11 g) was added to the water to create the saline solution, with a little bit of

baby shampoo (Johnson & Johnson). The cap was then soaked in the solution for 15 minutes before application to the participants (or 30 minutes if the cap was never used before). These caps are extremely quick to apply compared to other EEG caps, but there are limitations. For example, they typically function at a much higher impedance range (30-150 kOhm). Also, the conductivity of the electrodes is mostly dependent on the saline solution rather than the gels, and when dried out, they will have a much higher impedance level and lower signal quality. Therefore, the cap required constant impedance checks (after every task), and they were re-moisturised using the saline solution, to ensure that the impedances were smaller than 50 kOhm prior to any task recordings. After use, the cap was cleaned and washed with a baby shampoo solution, and then naturally dried. There are 32 EEG channels recorded using the cap (listed from Frontal to Posterior): Fp1, Fp2, Fz, F3, F4, F7, F8, F9, F10, FC1, FC2, FC5, FC6, Cz, C3, C4, T7, T8, CP1, CP2, CP5, CP6, Pz, P3, P4, P7, P8, P9, P10, Oz, O1 and O2 (all 10-20 system based; see also the figure). In general, 'F' denotes the Frontal region, 'C' denotes the Central, capitalised 'P' denotes the Posterior, 'T' denotes the Temporal, and 'O' denotes the Occipital. The common ground is Fpz (the very front one), and the reference electrode is FCz. The EEG signals are amplified using the shielded Brain Amp MR+ amplifier (Brain Products GmBH, Germany), with a digitalisation of 1000 Hz (sampling frequency), and a 16-bit A/D converter. Notice that the EEG system is also connected to task laptops to get triggers of events (such as the presence of a stimulus).

The EEG preparation and setup were done prior to all tasks (see Figure 6.1). Tasks with EEG recordings are also shown in the figure. Since the main aim of the phase-locking system was to deliver audio outputs, during the FNAT task, only the 32-channel cap was actively recording data, with the phase-locking system disconnected. The two systems overlapped at the Fz location. To ensure the phase-locking, the Fz electrode of the cap was actually 'floating' above the phase-locking system one, not recording anything. The Fz signal of the cap was later interpolated (see Data analysis).

6.2.3. Data analysis

Data cleaning and pre-processing

First, the data for all participants were checked through, and participants with missing data were excluded. The demographic data were then digitalised and loaded into MATLAB. The ESS scores were the sum of all the question scores in the questionnaire (maximum score 24, indicating highest sleepiness), and the PSQI was scored according to its normal scoring standard (Buysse et al., 1989), with the final score as the sum of the scores for all seven components (maximum score 21, indicating the worst sleep quality). After checking the demographical and sleep questionnaires, there was a participant with an extreme value of the ESS score, therefore

excluded from further analysis. Two participants were further removed due to technical issues in the phase-locking system (auditory stimulation was not phase-locked). Two left-handed participants were excluded as outliers. There are in total $N = 36$ participants with complete data for later analyses. The demographic details of the included participants will be shown in the results, and the statistical methods used will be introduced where applied.

The audio tracks for the recall periods of the Delayed Free-Recall task were listened to by three researchers independently. The words heard for each recall period were transcribed and listed down, and the majority were taken if there were conflicts. By matching the lists to the actual word list encoded, one can get the many different metrics straight away, for instance, the number of words or the percentage of words correctly recalled. The pre-nap and post-nap recall can also be compared using the data.

The FNAT task results were recorded by MATLAB after the task and saved on the local laptop. The results for all participants were combined into one 'CSV' file, with all trials from all participants concatenated altogether. For each trial, there would be a record of response time (RT), the answer chosen (Target or accurate name, foil, or distractor), the confidence level for that trial, and the corresponding block ID and subject ID of that trial. The result analysis was done using R Version 3.6.3 via RStudio, further analysis methods and details will follow below.

The one-channel EEG data for the ecHT phase-locking system was saved as CSV files, loaded into MATLAB, aligned to the auditory markers (audio stimulation trigger markers), and then saved for further analysis. Most of the analyses later are done on the 32-channel EEG data. The data was loaded into MATLAB and pre-processed using the FieldTrip toolbox (Oostenveld et al., 2010). Firstly, the missing Fz channel in the data was interpolated using the `ft_channelrepair` function with spline methods. The P9 and P10 channels were also re-interpolated as both channels are on the very edge of the cap, which had very bad contact with the scalp in general. The P9 channel was also situated very close to the ground and reference electrodes of the phase-locking ecHT system. After that, signals from all channels were bandpass filtered to [0.5,100] Hz to remove the DC component as well as high-frequency noises, using a 2nd-order Butterworth zero-phase filter (MATLAB function `filtfilt`). A 50 Hz notch filter was applied to suppress the line noise. The EEG data was then average referenced to all channels. To further remove EEG artefacts, an adaptive mixture of independent component analysis (Palmer et al., 2012) was applied to reject artifactual components from the signals, which was effective in excluding blink and muscle artefacts from the EEG signals. The 32-channel EEG signals were also aligned to the auditory markers prior to further analysis. The EEG signal for every task (nap period, auditory ERP, resting state and FNAT task) was loaded and pre-processed in the same way, and saved separately. Note again that for both EEG systems, the sampling rate was 1000 Hz.

Analysis on phase-locking accuracy

Phase-locking accuracy is important to ensure that the right phase is targeted for the desired stimulation outcomes. The phases of the EEG signal at the desired frequency band were computed using the ecHT algorithm to ensure the same causal phase estimations. It was also shown that non-causal computation of the phases could cause biases and information leakage (Vinao-Carl et al., 2023). The target frequency was 5 Hz, therefore, the ecHT algorithm causally bandpass filtered the EEG signal between 3.75 Hz and 6.25 Hz, using a 2nd-order Butterworth filter. The filtering was done in the frequency domain, where the signal was firstly Fourier transformed, and then multiplied by the coefficients of the filter's impulse response, and an inverse Fourier transform was applied to get the filtered signal. The phases of the bandpass-filtered signal were then estimated using the Hilbert transform (the angles of the analytic signal). Using the audio markers in the signal, the phases at which the audio signal was delivered can then be extracted. The phase-locking accuracy and statistics of the phase distribution were then analysed using the CircStat toolbox in MATLAB (Berens, 2009). Basically, the phase resultant was computed to show how uniform the phase distribution was. The phase resultant ranges from 0 to 1, where 1 indicates that all phases are perfectly aligned (in a circular plot), indicating a good phase-locking. Rayleigh's test was also applied to test whether there is significant uniformity of the phase distribution. The same analysis can be done for every EEG channel to see the distribution of phases at which the stimulation was delivered. The phase-locking analysis was only done for nap-period EEG and for participants with audio stimulation.

Sleep staging for the nap period

Given the time limit, the sleep staging was currently done using an automatic Artificial Intelligence-based model called U-Sleep (Perslev et al., 2021) (`https://sleep.ai.ku.dk/home`). The model was benchmarked against the AASM criterion, although the model had a powerful capacity to do a high-frequency sleep staging up to 128 Hz, where the model will output a confidence score that the subject is in one of the 5 sleep stages (Wake, Stage N1, N2, N3 and REM) at a given time. Here, the model was used to make predictions on 30 s epochs (1/30 Hz) same as the AASM criterion, although internally, the model simply aggregates the confidence score in the 30 s epoch and chooses the best representative one as the staging outcome. The original model (in the published paper) required at least one EEG and one electrooculography (EOG) input, now there are several updated and re-trained models where only EEG signal inputs. As specified in the paper, the model made no assumptions about any data pre-processing, and internally it will resample the EEG to 128 Hz anyway.

The pre-processed 32-channel 15-minute nap EEG data (see previous, 20 s baseline removed) was firstly converted to European Data Format (EDF) using the FieldTrip toolbox. The model

used was 'U-Sleep-EEG v2.0'. A maximum of 24 channels are allowed to input to the model, thus all peripheral channels (F7, F8, T7, T8, P7, P8, P9 and P10) were removed. The output was selected to be 1/30 Hz (one sleep stage every 30 s). The model ran predictions on every input EEG channel (one hypnogram per channel), and the final output hypnogram took the sleep stage of the maximum probability (mostly predicted from all channels). For the 15-minute nap period, the hypnogram had 30 sleep stages predicted. The Application Programming Interface was requested and approved by the developer, and the code implementation was done in Python.

To analyse the sleep staging data, the hypnogram results were first checked manually to ensure validity. Two participants were further excluded from this analysis, as their hypnograms started from stage N2 sleep, which was impossible to happen. The average sleep stage trajectories were then plotted for the audio and control groups for comparison. The sleep onset was defined the same as previous (continuous 1-minute stage N2 sleep), and the sleep onset latency was computed and compared between the two groups. The Wilcoxon rank-sum test was applied to test whether the latencies were significantly different between the two conditions. Not all subjects were asleep based on this sleep onset definition, therefore, a Fisher's exact test was done with the null hypothesis that the phase-locked auditory stimulation did not have a positive effect leading to sleep onset. Survival analysis was also applied, to see the probability of still awake (no presence of sleep onset as defined). The empirical probability density of 'awake' was estimated using MATLAB function `ecdf`, with the option to plot 'survival function'. One important concept in survival analysis is censoring. For instance, the experiment stopped before the total survival times of all individuals could be observed, or the participants dropped out of the study. In the current case, it means that participants who did not have sleep onset (thus were not asleep at all) should be right-censored because the exact survival time (asleep time) is unknown. The KSS data from pre and post-nap were also analysed, as a subjective measurement of potential effects of nap with or without audio stimulation. The changes in KSS score (post minus pre) were calculated for each participant, and compared between the two groups. The statistical testing was done using the Wilcoxon rank-sum test, a non-parametric test suitable for the limited sample size and discrete score values. Some post hoc analysis will be introduced along with the results.

Feature-based EEG dynamics analysis during nap period

The same feature-based analysis approach was applied here to study the EEG brain dynamics during the nap period and make comparisons between the two conditions. The same feature base was used, and most of the feature evaluation methods were kept the same. One important difference in this experimental data is the sampling rate of the EEG, where all previous datasets had 256 Hz. To make sure that the feature interpretations would not be affected (especially CATCH-22 features such as the total EEG spectrum sum, which was computed as the first 1/5

of the spectrum power), the EEG signals were downsampled to 250 Hz prior to any feature evaluation. The downsampling was done using MATLAB function `decimate` with a decimation factor of 4, and the decimated signal was lowpass filtered to ensure spectral properties (no aliasing).

The same classification pipeline with hierarchical clustering was applied to probe whether there would be any significant differences in EEG brain dynamics at the end of the nap due to the phase-locked stimulation. The last 3 minutes of the nap EEG data was taken, and three channels were taken into analysis: F3, Cz and Oz, which measures the regional cortical activities of interest. The Fz channel was not used since it was interpolated in the 32-channel system, therefore F3 was taken to replace in the analysis. For the 3-minute EEG period, ten non-overlapping 6-second EEG mini-epochs were extracted for feature extraction (after RMS-based artefact rejections with threshold 3 times the median), and then linear SVM was used to classify between audio and control condition groups. Hierarchical clustering was then applied to extract key features, with all methodology details the same as before.

Based on the foundation of previous analysis, the four key features that best distinguish bedtime wakefulness and sleep onset states were also studied here: the alpha-to-beta band ratio, the low-to-high frequency band power ratio, the peak beta band frequency, and the total EEG spectrum power. To explore the whole-brain changes of the four features throughout the 15-minute nap, the features were firstly computed continuously using the same approach as in Chapter 4. For each participant and EEG channel, the 15-minute period was firstly divided into 6-second mini-epochs with 3-second overlapping. The features were then evaluated for all mini-epochs. The same RMS-based artefact rejection method was applied, where mini-epochs with RMS values higher than four times the median of all epochs will be excluded as artefacts, where the corresponding feature values were marked as NaN. To ensure the continuity of the feature time series, a running-window median filtering with a window size of 20 data points was done to interpolate the NaN values with the median of nearby window values. The feature values were then globally z-score normalised: the mean and standard deviation of all timestamps from all channels and all participants were computed to normalise on. The feature time series are then normalised to the first 30 seconds (as a baseline to compare the changes) by subtracting the median value of the feature data points in the first 30 seconds (10 in total). After all the steps, there was one feature time series for every EEG channel of every participant.

To compare the global changes in these features between conditions over time, a cluster-based permutation test was used. The whole 15-minute period was first divided into the first 3-minute, the middle 3-minute (from 6 to 9 minutes), and the last 3-minute periods, or the whole 15-minute was used as a whole. The cluster-based permutation test was then done for each period and for each feature separately. For each feature and each period, the average feature values of all mini-epochs inside this period were computed for each channel and participant separately.

By concatenating them together, it became a 2-D feature matrix: number of participants by 32 channels for each condition. The feature matrices for the two conditions were then used to compute the two-sample t-statistic for each channel:

$$t\text{-}stat = \frac{\overline{x} - \overline{y}}{s_p(\sqrt{1/N_1 + 1/N_2})}, \tag{6.1}$$

where $s_p = \sqrt{\frac{(N_1-1)s_1^2+(N_2-1)s_2^2}{N_1+N_2-2}}$, x and y denote the feature value vector of one channel for each condition, and the horizontal bar denotes the population mean (across all participants), N_1 and N_2 denote the number of participants in each condition, and s_1 and s_2 denote the population standard deviation. Basically, this was equivalent to doing a two-sample t-test for each channel separately. The returned *t-stat* for all the channels were then used for the cluster-based permutation tests. Briefly, the cluster-based permutation test first identifies a bunch of significant clusters of channels (based on the previous two-sample t-test, with a p-value threshold of 0.05). The representative t-value for that cluster is the sum of all the *t-stat* of the channels in that cluster. Then, permutations will be done to randomly shuffle the channel indices, and re-evaluate clusters and cluster t-values (if available). This was done 10000 times. This is similar to the concept of surrogate testing but is now done on a spatial basis. The cluster t-values from all these permutations would be put together, and if taking a p-value threshold of 0.05, this implies that the final significant clusters would have a t-value higher than 95% of all the other cluster t-values from the 10000 permutations (based on the Monte-Carlo method). The explanation here is for intuitive understanding purposes only; fundamentally, this is a strong non-parametric method to detect significant regional changes in the features compared between audio and control conditions. The actual implementation of the cluster-based permutation test was using the FieldTrip toolbox function `ft_timelockstatistics`, using a within-subject design, and the t-statistic was computed by subtracting control from audio (so that positive t-values indicating a larger value in audio condition). The p-value threshold for both channel significance and cluster significance was 0.05.

The cluster-based permutation test was done on four periods without observing the feature time series. To see the actual feature dynamic trajectory over time, the average feature time series was averaged across the three channels (F3, Cz, and Oz) to see the dynamics globally during the 15-minute nap and compare between audio and control conditions. The comparison between the two conditions was done at each timestamp using a two-sample t-test on the time series. A generalised linear mixed effect model (GLME) was further applied to test whether auditory stimulation significantly changed the linear coefficient of the feature changes. The basic theory of the regression models has been introduced in Section 3.2. The GLME model took the feature values as the response variable, with time, stimulation conditions, and their multiplication as predictors (fixed effects), and included a random effect intercept for participant ID (as the grouping variable which quantifies the inter-dependence of the data samples). The GLME

equation generally takes: 'Feature = Time × Condition + (1|Participant-ID)'. The model fitting was done in MATLAB using function `fitglme`, with identity linkage function. The statistics of interest from the model is how 'Time × Condition' as a predictor changes the feature dynamics, where significance implies that the stimulation significantly changed the feature dynamical changes over time.

Delayed Free-Recall and FNAT analysis

As described previously, the Delayed Free-Recall task data was transcribed and matched to the actual list presented. There are many specific questions that can be asked in this task, but given the scope of this thesis, I only focus on the effect of auditory-intervened naps on performance. Firstly, the 90-second full-list recall performance difference between post and pre-nap will be compared between audio and control conditions. The performance was quantified via the percentage of words correctly recalled. Specifically, there could be new words recalled or words forgotten after the nap, which were also compared between the two conditions. Wilcoxon rank-sum test was used for the percentage changes, and for the number of word differences, the Kruskal-Wallis test was applied to test the distribution differences.

The statistical analyses for the FNAT data were done in R via RStudio. Again, there could be numerous questions to ask for this task, given the richness of the measurements. Within the scope of this thesis, the analyses focused on how auditory-intervened nap could change the number of different response types: Target or correct response, Foil and Distractor. Similarly, the effects on the response time (RT) and the confidence level could also be tested. The analysis generally followed the same approaches as in Violante et al. (2023). Prior to the analysis, the data was structured in the CSV file as specified previously, in a trial-by-trial way.

Firstly, to investigate whether the auditory-intervened nap significantly changed the accuracy of FNAT performances (correct: Target response; incorrect: Foil and Distractor responses), a binomial logistic regression model was fitted. The response variable is a binary variable (correct or not), while the predictor is the stimulation condition, with random intercepts of the participant ID and block ID. The model statistics were evaluated using the Type II Wald Chi-squared tests with Kenward-Roger correction for degrees of freedom, and the p-values would tell whether the stimulation significantly changed the accuracy level. Post hoc analysis was also done by estimating the marginal mean values (a.k.a., least-squares mean, evaluated in a certain subgroup such as the audio or control subgroup) in each condition, to see in which direction (positive or negative effects) the stimulation was taking effect. In addition to this, to investigate whether the auditory-intervened nap significantly changed the probability of different response types (Target, Foil and Distractor), a multi-nomial logistic regression model was used, with the 'Target' as the level of intercept. The response variable is the probability of different response types, while the predictor is the stimulation condition. Basically, the multi-nomial

logistic regression solves the regression problem by estimating the coefficients for the equation:

$$ln(\frac{P(Response = \texttt{'}Foil\texttt{'})}{P(Response = \texttt{'}Target\texttt{'})}) = b_0 + b_1(Condition = \texttt{'}Audio\texttt{'}), \tag{6.2}$$

$$ln(\frac{P(Response = \texttt{'}Distractor\texttt{'})}{P(Response = \texttt{'}Target\texttt{'})}) = b_2 + b_3(Condition = \texttt{'}Audio\texttt{'}), \tag{6.3}$$

where $P(.)$ denotes the probability of the specific response type, b_0 and b_2 denote the intercept of the linear regression results, and b_1 and b_3 are the coefficients, implying the effects of the stimulation on the probability of different responses. Positive coefficients mean that the auditory-intervened nap is making the performance generally worse, where the probability of Foil or Distractor response becomes higher. The model statistics were evaluated using Type II Wald Chi-squared tests with Kenward-Roger correction for degrees of freedom, which basically did the analysis of variance (ANOVA) across all response types. The Type II sum of squares generally considers the main effect conditioned on all other main effects (main predictors). To see the effects of the stimulation in each response type, for each response type, the marginal mean values were estimated and contrasted for each condition (audio or control). The post hoc contrasts could tell whether the stimulation significantly changed the probability of the specific response type. The implementation of the multi-nomial logistic regression model was done using the `multinom` function in the 'nnet' package in R. The statistical testing (ANOVA) of the model was done using `Anova` function, with the estimation of the marginal mean was done using the function `emmeans`.

To explore the effects of stimulation on the response time (RT) of the FNAT trials in different response types, a GLME model was used, with the stimulation condition, the response type, and their multiplication (as the interaction) as predictors to predict the RT. Given the specific distribution of RT (positive only), an inverse-Gaussian distribution was assumed in the GLME model, with an identity linkage function. The random effects (grouping variables) of the participant ID and the block ID (from 1 to 12) were considered in the model. Similarly, the model statistics were evaluated using the Type II Wald chi-squared test. Post hoc contracts were done by evaluating the marginal mean of the RT for each condition and each response type, to see whether the stimulation significantly changed the RT in each response type. The GLME model was fitted using the `glmer` function in the 'lme4' package, with the same functions to do the statistical testing and evaluating marginal means.

To test whether stimulation changed the confidence level in each response type, the cumulative linked mixed model (CLMM) with logit linkage function was used. CLMM model is actually the same regression model as GLME, but specifically designed for response variables as an ordinal variable, which has natural, ordered categories and the distances between the categories are not known. The CLMM model was fitted to use the stimulation condition, the response type, and their multiplication (as the interaction) as predictors to predict the confidence level, with

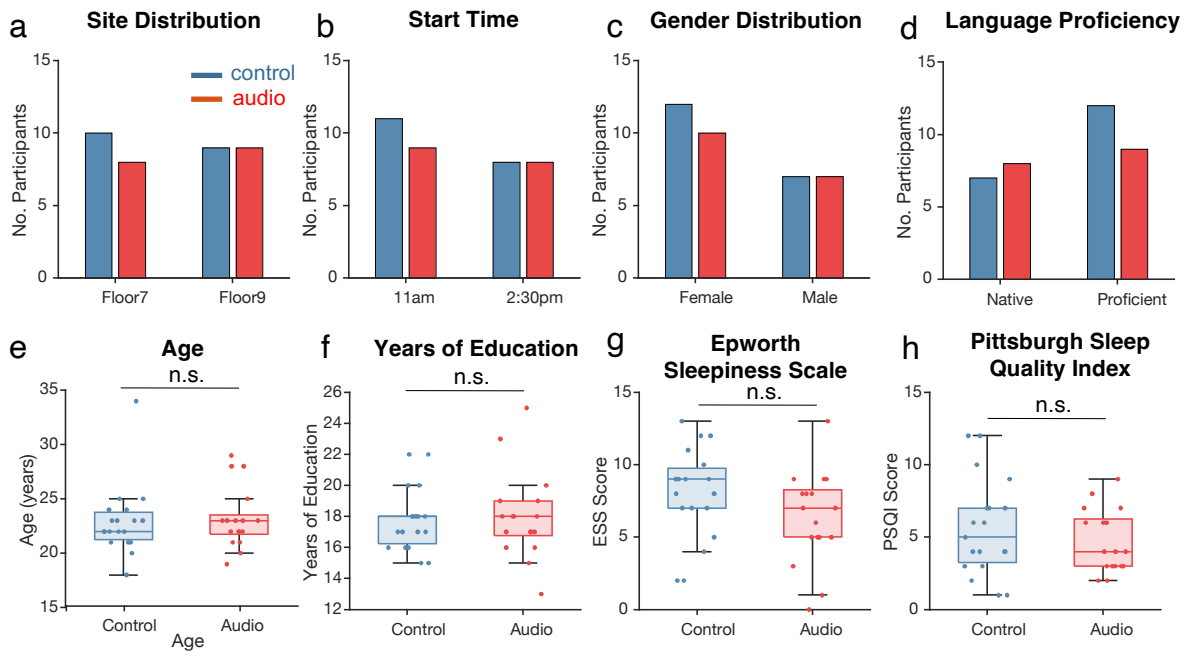


Figure 6.5: Demographics details of the participants contrasted between two conditions. Blue denotes control, and red denotes audio condition (same across all plots). **a**. The number of participants of each stimulation condition at each site; the distribution difference was tested using Fisher's exact test, which was not significant ($p = 1$). **b**. The number of participants of each stimulation condition for the two different experimental starting times; the distribution difference was tested using Fisher's exact test, which was not significant ($p = 1$). **c**. The number of male and female participants in each stimulation condition; the distribution difference was tested using Fisher's exact test, which was not significant ($p = 1$). **d**. The language proficiency distribution of the two stimulation conditions; the distribution difference was tested using Fisher's exact test, which was not significant ($p = 0.74$). **e**. The Age distribution comparison between audio and control conditions. Significance was tested using the Wilcoxon rank-sum test, and $p = 0.79$. n.s., non-significance (same across all later plots). **f**. The years of education distribution comparison between audio and control conditions. Significance was tested using the Wilcoxon rank-sum test, and $p = 0.85$. **g**. The Epworth sleepiness scale (ESS) distribution comparison between audio and control conditions. Significance was tested using the Wilcoxon rank-sum test, and $p = 0.09$. **h**. The Pittsburg Sleep Quality Index (PSQI) distribution comparison between audio and control conditions. Significance was tested using the Wilcoxon rank-sum test, and $p = 0.43$.

random intercepts as the participant ID and the block ID. The statistics were evaluated in the same way as the previous two models. The CLMM model was implemented using the function `clmm` in the 'ordinal' package in R.

6.3. Results

6.3.1. Demographics

The initial exclusion of participants has been illustrated before. After all exclusion, there were in total $N = 36$ participants for analysis. The demographic details are shown in Figure 6.5, along with all the sleep questionnaire results, contrasted between the two conditions. As mentioned

previously, there were two sites and two starting times for the experiment. The participant distribution is shown in Figure 6.5a and b. There were 18 participants on Floor 7 (10 control and 8 audio conditions), and 18 participants on Floor 9 (9 control and 9 audio conditions). 20 participants started at 11 am (11 control and 9 audio conditions), and 16 participants at 2:30 pm (8 control and 8 audio conditions). For the gender, there were 12 females and 7 males in the control condition, and 10 females and 7 males in the audio condition (Figure 6.5c). Fisher's exact test for all of them showed no significant difference between the two conditions ($p = 1$). All the participants had a language proficiency above the proficient level, with 15 native speakers (7 controls and 8 audio), and 21 proficient speakers (12 control and 9 audio conditions), and there were no significant differences (Fisher's exact test $p = 0.74$, Figure 6.5d). The control participants had an age distribution of 22.89 $\pm$ 3.18 years old (mean $\pm$ st.d.), while audio participants were 23.18 $\pm$ 2.81 years old (Figure 6.5e). The Wilcoxon rank-sum test showed no significant difference in the age distribution ($p = 0.78$). The control participants had 17.79 $\pm$ 2.02 years of education, while audio participants had 18.00 $\pm$ 2.81 years (Figure 6.5f, Wilcoxon rank-sum test $p = 0.85$). The control participants had an ESS score of 8.05 $\pm$ 3.12, while audio participants had 6.35 $\pm$ 3.18 (Figure 6.5g, Wilcoxon rank-sum test $p = 0.09$). The control participants had a PSQI score of 5.63 $\pm$ 3.32, while audio participants had 4.71 $\pm$ 2.17 (Figure 6.5h, Wilcoxon rank-sum test $p = 0.43$). In sum, there are no significant demographic differences between the audio and control groups that could confound later analysis results.

6.3.2. Phase-locking accuracy

Delivering the audio stimulation targeted at the correct phase is important to achieve the desired effects. The same ecHT algorithm was applied in computing the phases of the EEG signal at the 5 Hz theta band (target frequency in stimulation, see Methods for computation details), and the phase-locking accuracy results are shown in Figure 6.6. The whole head phase resultants are shown in Figure 6.6a, computed for all 32 channels (with Fz channel interpolated). The phase resultants are high around the phase-locking area (Fz and the left mastoid area), while for other locations, the resultants are low. The phase histograms (at which the audio stimulation was delivered) for all channels, and relevant statistics (mean phase, standard deviation of phase distributions, and Rayleigh's test for uniformity of the phases) can be found in Appendix B.36.

Specifically, the exact phases that were targeted are also important. The phase histogram at stimulation onset for the Fz channel (in the 32-channel system) is shown in Figure 6.6b. The phase resultant at Fz is 0.97 (Rayleigh's test $p = 6.36 \times 10^{-11}$), with a mean phase of 84.93° $\pm$ 13.55°. Recall that the original phase target was 72° (post-peak), and the phase-locking in the Fz channel in the 32-channel system has a 12° difference with it. The phase-histogram for the ecHT system can be seen in Appendix B.37, where the mean phase was 74.99° $\pm$ 1.39°, with phase resultant 0.99 (Rayleigh's test $p = 2.66 \times 10^{-12}$). The results clearly show that the

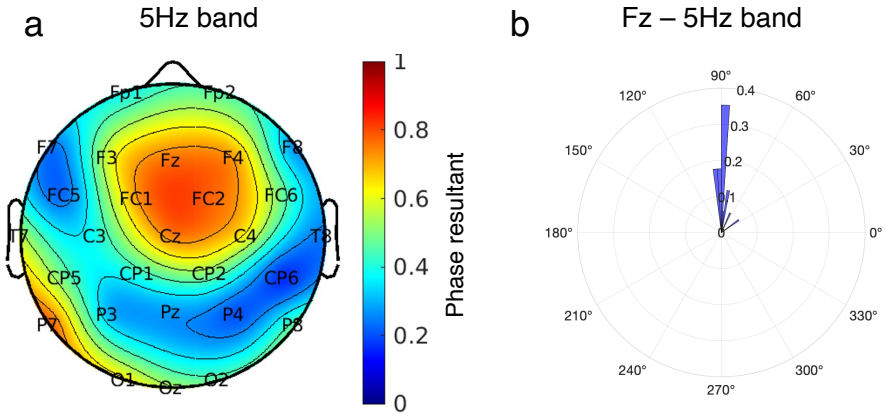


Figure 6.6: Phase-locking accuracy across the brain in the 32-channel EEG system. a. The phase resultant at 32 EEG channels for the phase distribution at which the stimulation was delivered. The channel names are marked on the figure, with the colour bar showing the phase-resultant value (from 0 to 1). **b**. The phase histogram at which the audio stimulation was delivered, plotted with a bin size of 10°. All phases were computed at the 5 Hz theta band, and only 32-channel EEG data was used with Fz-channel interpolated.

phase-locking in the original ecHT system was pretty accurate and consistent. The differences in the phase-locking results in the 32-channel system and ecHT system are expected, since the ecHT system had a very specific reference (to left mastoid), while the 32-channel system was referenced to FCz. The most important reason could be that the Fz EEG signal used in the analysis was actually interpolated using post hoc algorithms, thus leading to potential phase distortions.

6.3.3. The effect of auditory stimulation on nap

Sleep staging and KSS results

As previously mentioned, the sleep staging was done using a pre-trained AI model. There were two participants further excluded from the sleep staging results analysis, due to the fact that their sleep stages started from stage N2 ($N = 34$ for analyses). To see whether the phase-locked auditory stimulation impacted the sleep stage results, the average sleep stage trajectories for the two conditions are plotted and compared in Figure 6.7a. There are no statistical differences in the sleep stage trajectories, but when observing the figure, at around the last 3 minutes of the nap, the participants with audio stimulation seem to be in a deeper average sleep stage. Therefore, the sleep stages in the last 3 minutes were taken to compare between audio and control groups, and the left-tailed Wilcoxon rank-sum test was applied (to see if the values in the audio group were significantly higher than those of the control group). Still, the p-value was 0.30, indicating that statistically, there was no difference even in the last 3 minutes of the nap.

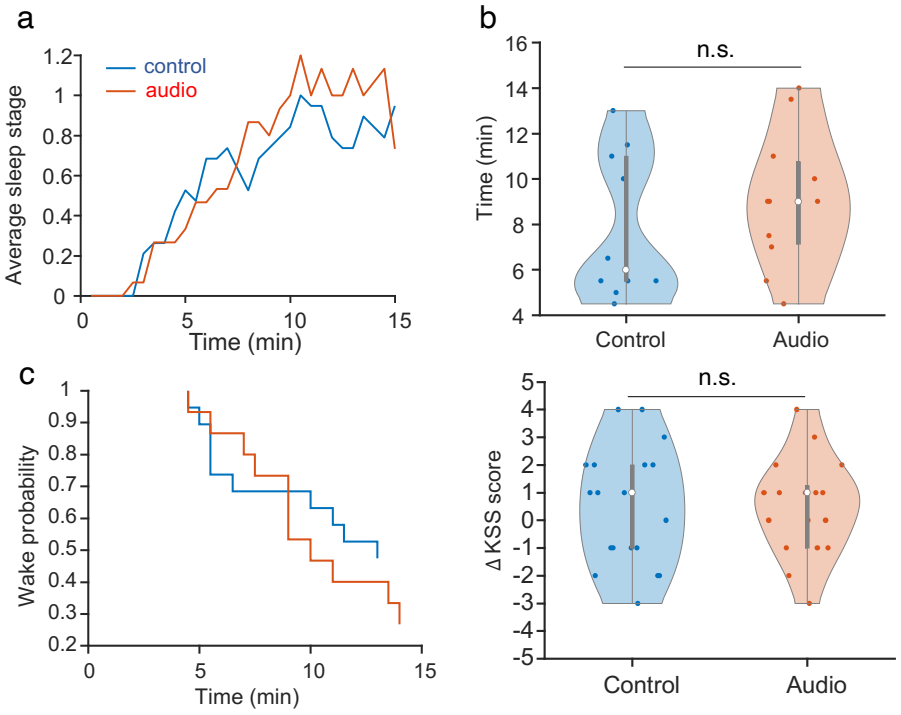


Figure 6.7: Differences between audio and control in sleep stages and KSS tests. a. The average sleep stages in the 15-minute nap compared between the control (blue) and audio (red) conditions. At each timestamp, the Wilcoxon rank-sum test was applied to see whether the sleep stages differed significantly between conditions, and no significance was found. **b**. The time to sleep onset (defined as the first 1-minute stage N2 sleep) comparison between the control (blue) and audio (red) conditions. The Wilcoxon rank-sum test showed no significance ($p = 0.38$). **c**. Survival analysis showing the probability of wake with time increasing in both conditions. Fisher's exact test showed no significant differences in the number of participants who fell asleep or not between the two conditions ($p = 0.30$). **d**. The changes in the KSS score (post minus pre-nap) compared between the two conditions. The Wilcoxon rank-sum test showed no significant difference between the two conditions ($p = 1$). n.s., non-significance in all relevant plots. Violins in (**b**) and (**d**) show the kernel density plot (external shape), median values (white non-filled circle) and upper and lower 25% quantiles (grey bar).

To get deeper into the sleep staging outcomes, the sleep onset was defined the same as previously: the continuous 1-minute stage N2 (2 consecutive N2 scores). The sleep onset latency was then defined as the time from the start of the nap up to the sleep onset. The latencies are compared between the control and audio groups (Figure 6.7b). Since this was a very short nap, not all participants could have reached stage N2 sleep. For the control participants, there were 10 who fell asleep based on this sleep onset definition and had mean latencies of 7.80 minutes ($\pm$ 3.19, st.d.). For the audio participants, there were 11 who fell asleep based on this sleep onset definition and had mean latencies of 9.09 minutes ($\pm$ 2.98, st.d.). Still, the Wilcoxon rank-sum test between them showed no significant differences. The survival analysis was also applied, to see the probability of falling asleep in each condition over time (Figure 6.7c). No obvious difference can be seen visually, and Fisher's exact test also showed no significance in the number of participants who fell asleep in each condition ($p = 0.30$). From the perspective of subjective sleepiness metrics (the KSS test), the changes in the KSS scores were also not significantly

different between conditions (Wilcoxon rank-sum test $p = 1$): control participants had a change of 0.47 ± 2.12, while audio participants had 0.47 ± 1.77. Interestingly, both groups showed a positive change in the KSS score, implying more sleepiness after the short nap, although neither group itself was significant (One-sample t-test, $p = 0.34$ for control and $p = 0.29$ for audio). The potential reason for this could be the sleep inertia effect (see later).

Feature-based classification between audio and control in the last 3 minutes of the nap

Although not significant, the previous results showed that the potential effect could be in the last period of nap. Therefore, the feature-based statistical learning pipeline was applied to classify between audio and control conditions in the last 3 minutes of the nap. The EEG was downsampled to 250 Hz prior to feature evaluation, and the same evaluation methods and hierarchical clustering as before were used (see Methods). Three EEG channels were taken for the analysis (F3, Cz and Oz). The results can be seen in Appendix B.38. There were no features with accuracy above 60%. Although this analysis did not have good sensitivity here, it was interesting to observe that the clustering structure was similar to that of the classification between bedtime and sleep onset (Chapter 3), with some of the previously identified features still at the centroid of clusters (especially the total EEG spectrum power and the peak beta band frequency).

Key governing sleep features differences between audio and control

The post-peak of theta band stimulation condition was identified as the optimal condition based on the four key governing features identified in Chapter 3 for the MESA dataset: the alpha-to-beta band ratio, the low-to-high frequency band power ratio, the peak beta band frequency, and the total EEG spectrum sum. The stimulation time in the previous experiment was much shorter, and not in a 'falling asleep' scenario, however. It is thus important to study the impacts of auditory stimulation during the 15-minute nap. Firstly, the changes in the four features normalised to the 30-second baseline throughout the nap period were contrasted between the audio and control condition on the whole brain (all 32 channels), using the cluster-based permutation test (see Methods). The results are shown in Figure 6.8, with the t-statistics shown in colours and significant EEG channel clusters marked. Only the total EEG spectrum power was significant in the comparison, with the audio condition having significantly higher feature values in almost all channels, and throughout the entire nap period. Recall the previous results that the total EEG spectrum power was higher at sleep onset compared to bedtime, which implies that the auditory stimulation is pushing the brain closer to sleep onset based on this feature. For the alpha-to-beta band power ratio, there seems to be a slight increase across all brain areas, which matches the changes at sleep onset compared to bedtime (increasing change), although

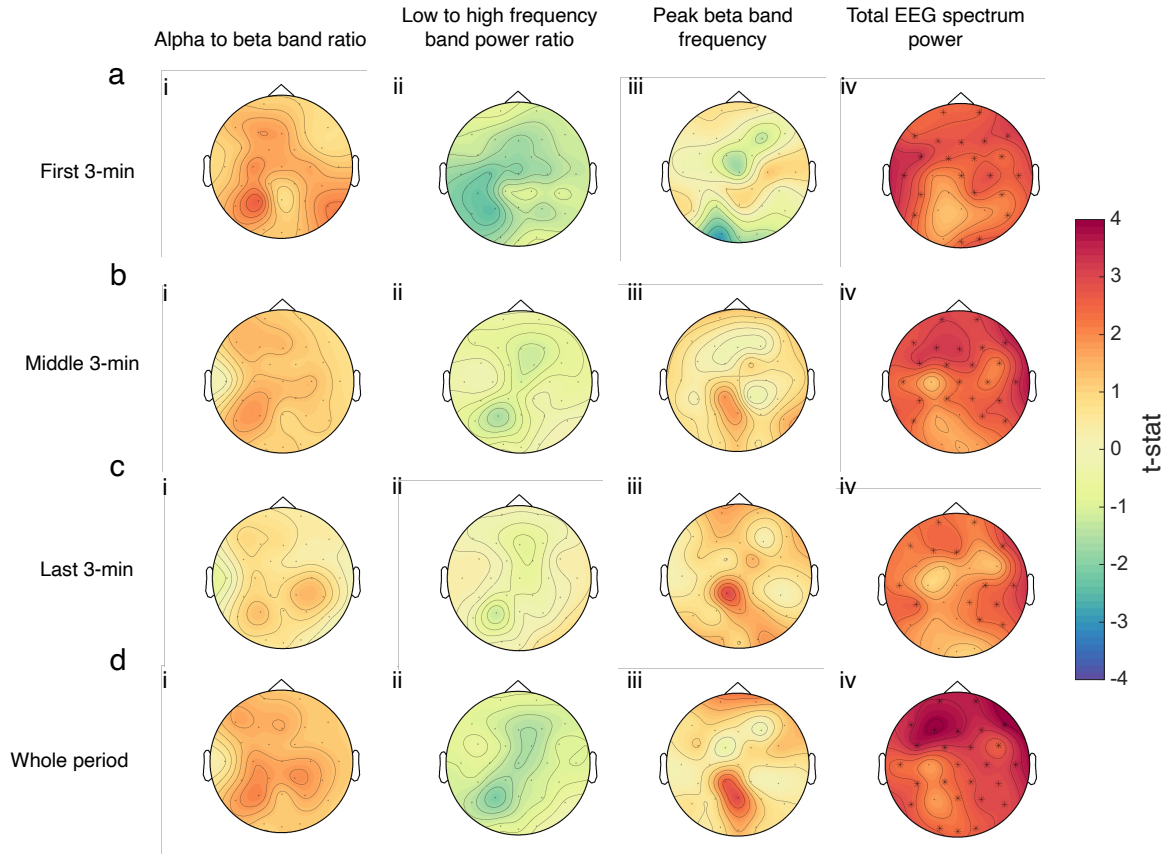


Figure 6.8: Comparison of the four key feature changes on the whole brain between audio and control across the nap period. The entire nap was divided into four different periods to study: (**a**) for the first 3 minutes; (**b**) for the middle 3 minutes (6-9 minute period); (**c**) for the last 3 minutes; and (**d**) for the entire nap period. Four key features are shown: (i) Alpha to beta band ratio; (ii) Low to high frequency band power ratio; (iii) Peak beta band frequency; and (iv) Total EEG spectrum power. The changes in the feature values are contrasted between the audio and control groups (audio minus control), and the two-sample t-statistics (t-stat) are shown in the plots as the colour scales (from -4 to 4), where positive values imply larger positive changes in the audio condition compared to the baseline. Cluster-based permutation tests were done, with significant EEG channel clusters shown in '*' markers in the plot. The curves in the plots are isolines (contour lines) on which the t-stat values are the same, where the density of the contour lines shows the gradient (denser lines imply a larger gradient).

this change was not significant. The low-to-high frequency band power ratio showed barely any changes. The peak beta band frequency showed no changes at first, but with stimulation going on, it seems that the Parietal region (around Pz) had an increased feature value. Recall that peak beta band frequency should decrease at sleep onset. No significant clusters were identified, which means that this increase was very restricted in that small cortical area, implying a more aroused state in that region only. Interestingly, the same regions around the left of Pz also showed the least positive change in the total EEG spectrum power as well.

The cluster-based permutation test allowed a whole-brain view of the changes in the features but lost some level of temporal information. To see how the features change continuously in each condition, and compare them across conditions, the four feature time series are plotted to

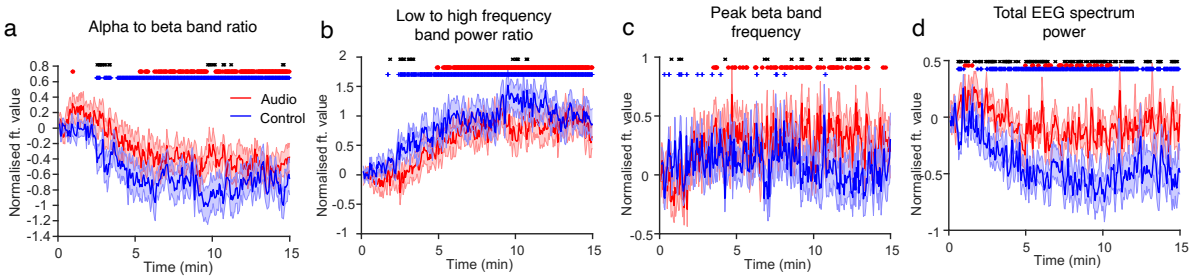


Figure 6.9: The key feature time series during nap normalised to baseline and compared between audio and control. The key feature time series are averaged from F3, Cz and Oz channels during 15-minute nap, z-score normalised to 30 s baseline and compared between audio (red lines) and control (blue lines). **a**. Alpha to beta band ratio; **b**. Low to high frequency band power ratio; **c**. The peak beta band frequency; and **d**. Total EEG spectrum power. Shown are the mean (solid line) ± s.e.m. (shaded area) across all participants in each condition group. One-sample t-test was done at each timestamp for the normalised time series of each group, to see significant changes compared to 30 s baseline; The Red asterisk '*' denotes significance for the audio-group time series, blue addition '+' denotes significance for the control-group time series ($p < 0.05$). The time series are also compared across groups, using a two-sided two-sample t-test, with the black cross '×' indicating significance ($p < 0.05$).

see their temporal dynamics (Figure 6.9). The feature time series were averaged from the F3, Cz and Oz channels to get one representative time series per participant. All feature values were z-score normalised to the 30 s baseline (see Methods), and the statistics at the group level were shown and compared in the plot. Aside from the low-to-high frequency band power ratio which had similar significant increases in both conditions, the other features did not change in the same direction as one would expect during falling asleep (according to previous results) for both control and audio condition participants. For the alpha-to-beta band power ratio, both control and audio participants had significantly decreased feature value compared to baseline, although the change for audio participants started later. There was not much difference between the two conditions for this feature except the first few minutes. For the peak beta band frequency, participants in both conditions did not change much, and the feature values seemed to be noisy. For the total EEG spectrum power, the control participants had significantly decreased feature values compared to the baseline throughout the nap, while for audio participants, they remained constant with a slight decrease in the middle. It is now clear that the differences that were previously seen in Figure 6.8 were not due to the fact that the total EEG spectrum power was increasing more in the audio participants, but rather it was decreasing in the control participants. The total EEG spectrum power also had the most significant differences between audio and control conditions throughout the nap period (Figure 6.9d black crosses). In sum, compared to the control condition, the phase-locked audio stimulation did not seem to improve sleep onset but also did not make it worse.

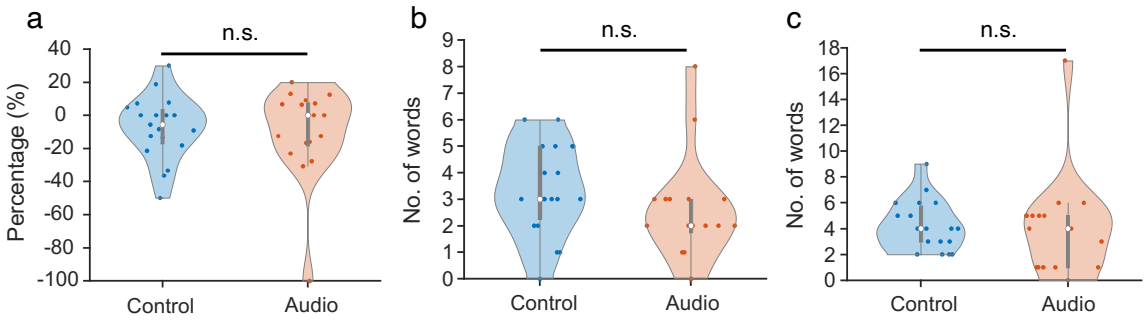


Figure 6.10: Results comparison in the Delayed Free-Recall task performance between audio (red) and control (blue) groups. a. The changes in the percentage of words recalled (post minus pre-nap) compared between audio and control. n.s., non-significance using two-sided two-sample t-test ($p = 0.77$). **b**. The number of new words recalled after nap compared to pre-nap and contrasted between audio and control groups. n.s., non-significance using Kruskal-Wallis test ($p = 0.07$). **c**. The number of new words forgotten after nap compared to pre-nap and contrasted between audio and control groups. n.s., non-significance using Kruskal-Wallis test ($p = 0.48$). Violins show the kernel density plot (external shape), median values (white non-filled circle) and upper and lower 25% quantiles (grey bar).

6.3.4. Effect of auditory-intervened nap on memory performances

Finally, I analysed the Delayed Free-Recall task and the FNAT results to probe the effects of the auditory-intervened nap on memory performances. For the Delayed Free-Recall task, the performance differences between the pre and post-nap full-list recall are compared between the two conditions (Figure 6.10), to see whether auditory-intervened nap could enhance the verbal episodic memory encoding more than control. Firstly, the changes in the percentage of words recalled (from post minus pre-nap) are contrasted between the two conditions (Figure 6.10a). Both groups had a decrease in the percentage of words recalled after the nap, with the audio group -9.69% $\pm$ 27.89% (mean $\pm$ st.d.) and the control group -7.36% $\pm$ 19.14%. Using a two-sample t-test, the performance changes were not significantly different between them ($p = 0.77$). It is interesting that, in the audio group, there was one participant who did not recall any word that was encoded pre-nap (-100%). Aside from the percentage, there could be new words recalled, or previously recalled words forgotten after the nap. Therefore, they are also contrasted between audio and control groups (Figure 6.10b and c). Audio-group participants had, on average, 2.53 $\pm$ 1.97 (st.d.) new words recalled, while control participants had 3.37 $\pm$ 1.71. Since the number of words is discrete values, the Kruskal-Wallis distribution test was applied, and it was not significantly different between the two conditions ($p = 0.07$). Similarly, audio-group (or control group) participants forgot on average 4.12 $\pm$ 3.90 (or 4.26 $\pm$ 1.91) words and the Kruskal-Wallis test was not significant ($p = 0.48$). In general, the Delayed Free-Recall performances were not different after the auditory-intervened nap compared to the control.

The FNAT analysis provided much richer cognitive metrics to analyse the potential effects of the auditory-intervened nap. Prior to conducting the analysis, the accuracies for all participants

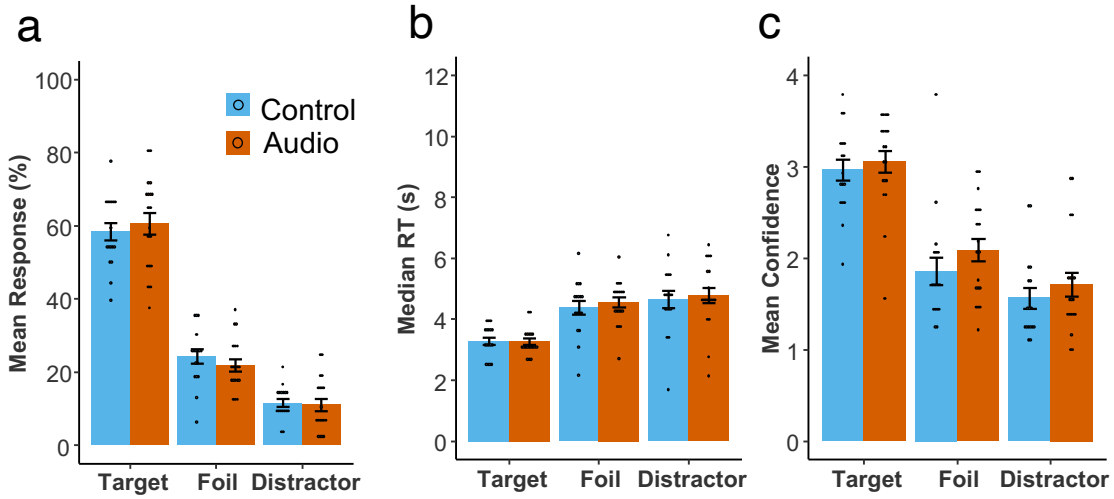


Figure 6.11: Results comparison in FNAT performance between the control (blue) and audio (red) groups. **a**. The percentage of Target, Foil and Distractor response compared between the control and audio groups. Post hoc contrasts between audio and control in each response type showed no significance (Target $p = 0.08$, Foil $p = 0.08$, Distractor $p = 0.65$). **b**. The response time (RT) distribution of Target, Foil and Distractor response compared between the control and audio groups. Post hoc contrasts between audio and control in each response type showed no significance (Target $p = 0.98$, Foil $p = 0.21$, Distractor $p = 0.14$). **c**. The confidence distribution of Target, Foil and Distractor response compared between the control and audio groups. Post hoc contrasts between audio and control in each response type showed no significance (Target $p = 0.63$, Foil $p = 0.20$, Distractor $p = 0.15$). The distributions were shown in bars with mean ± s.e.m. (the error bar), grouped in each response type and stimulation condition. Black dots show individual participant's data (mean or median across all trials). The statistics were done using all trial data rather than averaged data within participants. Refer to Appendix B.39 for all model statistics (post hoc contrasts) and distribution statistics.

were checked to make sure that all participants had accuracies (percentage of Target response) above 30%, with a mean 60.47% ± 11.77% (st.d.) accuracy across all participants. The trials without a response (the participant failed to respond within 8 seconds) were removed (4.19%), and in total, there were 60.47% Target responses, 23.54% Foil responses, and 11.79% Distractor responses. Firstly, a binomial logistic regression model was fitted to predict accuracy (Target response) or not (Foil and Distractor combined) with the same predictor (stimulation condition) and random intercepts. The model was not significant, with $\chi^2 = 0.57$ and $p = 0.45$, with an estimated mean accuracy of 62.9% ± 3.3% for the control cohort, and 66.0% ± 3.1% for the stimulation cohort. Further, the analysis was expanded to answer whether the stimulation significantly changed the response types, the multi-nomial logistic regression model was fitted to the data, with the 'Target' response (correct response) as the intercept for comparison (see Methods for details). The model showed a significant effect of the stimulation: $\chi^2 = 6.33$, $p = 0.04$. Specifically, the model coefficients showed a significantly lower probability of Foil response compared to the Target under the stimulation condition ($P(Foil)/P(Target) = e^{b_1} = 0.86$, $p = 0.013$). The probability of Distractor compared to Target was lower, but not significant ($P(Distractor)/P(Target) = e^{b_1} = 0.92$, $p = 0.31$). Still, when comparing and contrasting the percentage of each response type between the two conditions, no significance was found (Figure

6.11a), although the p-value for the Target response and Foil response was very close to 0.05. These results imply that although the general accuracy was not affected by the stimulation, there was a weak effect in augmenting the correct bindings. For the response time (RT), a GLME model was fitted to predict it using the stimulation condition and response type, with random intercept as the participant and blocks. The model statistics were analysed using the ANOVA (see Appendix B.39 for full statistics), and it was found that the stimulation type and the response type had a significant interactional effect on the RT ($\chi^2 = 13.40$, $p = 0.001$). This implies that the stimulation significantly changed the RT distributions across all response types. However, the post hoc contrasts in each response type showed no significant differences (Figure 6.11b), and barely any differences can be visually seen in the RT distribution in each response type as well. For the confidence, a CLMM model was fitted to predict it using the stimulation condition and response type with random intercept as the participant and blocks. Similarly, the model statistics showed a significant interactional effect on the confidence by stimulation across all the response types ($\chi^2 = 7.03$, $p = 0.03$). The post hoc contrasts, again, showed no significant differences in the confidence between conditions in each response type. All these results imply that the stimulation would only have significant impacts on the FNAT performance dependent on the response type, but the impact cannot be seen directly in a specific response type.

6.4. Discussion

6.4.1. Reflections on experimental results

In this chapter, I presented a nap experiment trying to use brain oscillation synchronised audio stimulation to augment sleep onset and potentially improve memory performances. I showed that the audio was delivered accurately to the target phases. There was no effect found in the augmentation of sleep onset from the perspective of sleep staging and key sleep onset features. However, I finally showed that there was a weak improvement in associative memory in the FNAT task.

I first demonstrated that the phase-locking was accurate in targetting the desired phase of the ongoing neuronal oscillations. However, it can be found that, due to the specific reference that was taken (Fz to left mastoid), the audio outputs were actually phase-locked to the relative theta activities from frontal minus temporal activity. This might need to be taken into consideration for future analysis, given the specific neuronal oscillations that are targeted.

The effect of auditory intervention in the 15-minute nap was analysed from several different aspects. Firstly, based on the hypnograms (sleep stages), there was no effect in augmenting sleep using the phase-locked auditory intervention. However, it has to be acknowledged that the AI-based staging should in no case be taken as the ground truth. As a future plan, there

will be human experts to manually score this data, and further results will be revised according to the stages. Still, sleep staging itself is not a very good outcome measurement to quantify the auditory effects, since it has a very low temporal resolution (30 seconds), and as discussed previously, stage N1 has been the the most controversial sleep stage, with the lowest inter-rater consistency (Rosenberg and Van Hout, 2013). This also holds for the model used, as specified in the original paper, the F-score for stage N1 was the lowest (around 0.5) (Perslev et al., 2021). From the perspective of subjective sleepiness metrics (KSS scores), it was found that all participants seemed to be drowsier after the nap (although not significant). This could potentially be related to the sleep inertia effects (Lovato and Lack, 2010; Ficca et al., 2010), caused by the abrupt disruption of sleep (recall the concept of hysteresis in state transitions mentioned in Chapter 4 as well). Since the nap period was only 15 minutes, at the end of the nap, most participants might just reach the sleep onset, or they were even woken up while their brains were just prepared for the transition. Sleep inertia would cause a short period of cognitive impairment, and the longer the nap, the longer time the person would need to recover (Lovato and Lack, 2010). Here in the present study, the next cognitive task was at least 12 minutes away from the nap, therefore the effect of inertia should have gone by then, leaving only the benefits from the short nap (Lovato and Lack, 2010).

From the perspective of the EEG dynamic features, a classification analysis was done between control and audio conditions, where no key features could be found. Still, the similar structure in the hierarchical clustering to previous results implies that the effects of this phase-locked audio might still be related to sleep features somehow. As a future step, the previously developed inter-regional features could be deployed here and explore further the potential dynamic differences. When studying the four key sleep-related features, combining the results from the cluster-based permutation test and the feature time series dynamics, although no significant effect of 'augmentation of sleep onset' was found, at least, it did not make it worse by making the participants more aroused (except the left Parietal area, which is interesting to study further on its own). From the perspective of the total EEG spectrum power, it could even be concluded that the auditory stimulation induced slower EEG dynamics in general, although this effect could be due to the entrainment of theta rhythms, as it was not obvious in the other features that the stimulation is augmenting sleep onset more than control. It was interesting to observe that even in the control condition, the feature dynamics did not change in the expected direction while falling asleep. There could be many specific points to notice and improve in this analysis. Firstly, the difference in the analysis methods should be emphasized. The previous trajectories in the falling asleep dynamics were observed by aligning all participants to the sleep onset, where here the feature time series are normalised to the first 30 seconds of the nap period. Specifically, for the audio condition participants, there were stimulations in the first 30 seconds as well, which could bias the normalised results, as the novel exposure to the stimuli could cause arousal or other effects. Secondly, the restriction of the nap length might not represent the complete

falling asleep transition for many participants. Further analysis could divide the participants into successful asleep ones (after manual sleep staging is done) and non-successful ones, and redo the analysis again to see potential effects. Thirdly, the specific cohort this study recruited is completely different to previous datasets, as the two datasets used in previous analyses have an elderly population. Lastly, the nap was done during the day instead of an overnight study, which could affect the falling asleep transition. Further, there were no 'adaptations' normally present in other sleep experiments. Similar to the 'first-night effect' (Tamaki et al., 2005), there could also be a 'first-day effect' that is changing the brain dynamics seen in the nap.

The cognitive benefits of nap have been shown previously (Lovato and Lack, 2010; Mednick et al., 2013; Tucker et al., 2006), and the experiment incorporated the cognitive tasks to see whether by boosting the nap (facilitating the sleep onset), one can achieve augmented cognitive benefits as well. In comparing the cognitive performances between the two stimulation conditions, no positive effects of the stimulation were found in the Delayed Free-Recall performance. In the FNAT analysis, it was found that although the general accuracy was not changed by the stimulation, the stimulation could significantly reduce the probability of a misbinding (Foil response) compared to the correct response, implying a potential benefit in associative memory. Of course, this benefit seems to be very restricted in this comparison only, as there was no difference in the actual accuracy (correct response percentages) and that compared to the Distractor response. Similarly, for the response time and confidence, the stimulation only had an effect dependent on the response type, but not within any specific response. It also needs to be acknowledged that due to the limitation of sample size, and a between-subject design, there could be many other confounding factors making it difficult to link the auditory intervention effects to cognitive performance changes.

Still, all these analyses presented were very preliminary, where the 'weak' effects found here could be further augmented with refined methods. Firstly, the 'optimal' stimulation condition (the theta band post-peak) was identified at the group level. In our lab's preliminary experiment, 12 out of 18 participants responded best to this condition, however, the remaining subjects responded best to other stimulation conditions. Therefore, the very next step would be to further group the participants into potential 'responding' and 'non-responding' groups, and redo the analyses again to see potential effects. How to do this re-grouping might also be important, as this study was not sham-controlled within the subject. Starting from the preliminary experiment, one could apply the feature-based statistical learning pipeline to classify the baselines of the 12 responding participants and the non-responding ones. Then, one can use the trained classifier or potential key features that can 'predict' whether the participant can respond or not, to distinguish the $N = 17$ participants into responding and non-responding cohorts. Aside from predicting whether the participant would respond to theta post-peak auditory stimulation, we could also potentially phenotype the participants, and predict the exact stimulation condition that they would respond to. Again, one could run the feature-based statistical learning pipeline

and identify the features that could tell the personalised 'optimal' response condition. Another important difference between the previous experiment and the current one is the stimulation time and the analysis method for determining the optimal condition. The previous experiment had in total of 4-minute resting state EEG recordings, while here the stimulation lasted for 15 minutes, and in a falling-asleep scenario. Secondly, the knowledge of dynamical systems and the bifurcation theory during the falling asleep transition was not considered in the analysis. The bifurcation-based biomarkers (such as tipping point prediction, or control variables) could be potentially used in the future to probe the effects of this phase-locked auditory stimulation.

On the other hand, the potential mechanism behind how phase-locked auditory stimulation could augment sleep onset remains unclear. In deep sleep, it has been proposed that the auditory signals activate the nonlemniscal pathways, which then leads to widespread depolarisation across the cortex and, thus, synchronisation of the oscillations (Bellesi et al., 2014). However, in triggering the sleep state transitions, the mechanistic pathways could be different. Linking to the previous results and discussions, the need to sleep depends on homeostasis pressures, and the 'spring theory' (Yamagata et al., 2021) also indicates the same thing, that the brain needs to be active with enough intensity for homeostasis to build up, and thus a smoother sleep state transition. Potentially, the auditory stimulation could work by actually arousing the brain activity, facilitating the homeostasis pressure and then augmenting the sleep onset. If that is the case, the auditory stimulation might work better for those insomnia patients with excessive daytime sleepiness, rather than healthy participants (in this experiment, for example). This is only a hypothesis, without any evidence. In future research, there could potentially be more mechanistic studies into how the phase-locked audio inputs change the brain dynamics (such as EEG features) and gain more mechanistic insights.

6.4.2. Potential personalised stimulation design in augmenting sleep onset

Coming back to the topic of developing a closed-loop neuromodulation system, the ultimate reason for those functional biomarkers and models in quantifying the falling asleep process was to inform the stimulation system in optimising the stimulation in real time. Given the complexity and non-stationarity of the brain dynamics, it is conceivable that the 'optimal' phase condition could be changing over time as well. This implies that the stimulation should be adjusted to the actual brain states in a real-time scenario to achieve optimal benefits (Fang and Yang, 2022). There can be many potential approaches to do this. For instance, by tracking the four key governing sleep-related features in real-time, one can define a threshold for improvement (features moving in the direction of falling asleep), thus keeping the current stimulation condition; Or if the features show no changes, or undesired changes (moving in the opposite directions), the target phases of the stimulation would be changed. In more complicated approaches, if the personalised bifurcation model was successful (see Chapter 4), then tracking the tipping point

or control variables of the system could be a more sensitive way to inform the system to change the conditions. These approaches are very basic, a more advanced approach could take into consideration the control theory, and develop "self-adaptive" models such as Bayesian optimisation and reinforcement learning models to adjust the stimulation parameters with the quantifiable metrics described before as the reward function, where the stimulation will be optimised to "push" the brain state to sleep-onset.

Outside the target phases, there are many other parameters in the auditory stimulation paradigm that can be optimised, for example, the audio intensity (volume). Similar to slow-wave sleep augmentation, in augmenting sleep onset, the volumes cannot be too large to cause arousal but also need to be strong enough for the stimulation to be effective. Moreover, during sleep, auditory stimulation functions depend on the activation of the nonlemniscal pathways, which overlap anatomically with the ascending arousal systems (Bellesi et al., 2014). Therefore, the benefit range of the audio volume could be small, and difficult to optimise. The current experiment used a 40 dB stimulation, and the results did not show an arousal effect in general, but this volume could be too small for some participants to gain benefits, or too high for other participants and disrupt their falling asleep attempt. There are many other parameters, such as the target neuronal oscillation frequency or the output sound frequencies. As Bellesi et al. proposed, the output sound frequencies should vary in general to avoid habituation in the nonlemniscal pathways (Bellesi et al., 2014). Of course, achieving a closed-loop system in fine-tuning all these parameters could be extremely difficult. Essentially, how different auditory stimulation actually changes the EEG brain dynamics has not been well studied. There is a need to develop these perturbation-based biomarkers (outside the sleep features) as well, as to answer the question of how these audio outputs are changing the EEG brain dynamics (from a feature-based perspective, for instance). Similar to the previous comment, these features could also provide important mechanistic insights about the auditory stimulation itself, and also better help develop the system.

6.4.3. Implications and other potential work

There are many additional analyses to be done for the experimental data, and if successful, this is going to have an important clinical impact, as a novel treatment approach for insomnia patients, and is especially beneficial for those cognitively impaired and dementia patients who suffer from sleep-onset insomnia. As an additional value, the augmentation of sleep onset could help the patients benefit more from actual deep sleep, and prevent further cognitive declines. Deeper analyses of the data could also provide important mechanistic insights on how the phase-synchronised auditory stimulation could augment falling asleep, thus informing future experiments and designing better stimulation paradigms. Outside the augmentation of sleep onset, similar mechanisms could also work to prevent sleep fragmentation, which is also one of

the sleep disturbances experienced by cognitively impaired and dementia patients.

The richness of this experimental data also allows for many other explorations, for instance, one can use the concurrent EEG recordings during the FNAT task, and identify potential functional biomarkers in predicting or correlating to the task performance (Gutierrez, 2023). The 3-minute resting state EEG recording prior to and after the nap could also be analysed to see the brain dynamics differences, so as to better understand the stimulation effects and to phenotype the participants' responses or predict the stimulation conditions as well.

7. Conclusions and discussion

7.1. Summary of the results

In Chapter 3, 4 and 5, I presented a series of computational studies which investigated the falling asleep brain dynamics measured via EEG from different aspects based on meta sleep EEG datasets. Starting from Chapter 3, I developed the feature-based statistical learning pipeline and tested its efficacy in identifying a set of orthogonal EEG biomarkers for distinguishing the bedtime wakefulness and sleep onset states, based on two large different sleep datasets. Four orthogonal key governing features were found to distinguish the two states best: the alpha-to-beta band power ratio, the low-to-high frequency power band ratio, the peak beta band frequency, and the total EEG spectrum power. The importance of these features was validated in two datasets, with slight differences that could be explained by the differences in the datasets. The classification model trained on selected key features could reach a comparable accuracy to all features, which further confirmed the validity of this dimensionality-reduction approach for selecting the key features in best describing the brain dynamic differences between the two states. I further tested how different demographic factors could affect the EEG brain dynamics at bedtime wakefulness and sleep onset state. Ageing brains have slower brain rhythms and more noise-like EEG dynamics with less nonlinear structures during normal wakefulness, and I further found that this also exists at the sleep onset state. Lastly, I systematically explored the features that were correlated to and could predict the sleep onset latency. Contrary to our current understanding, some features suggested a spring-like phenomenon for sleep regulation: the further to the sleep-onset state, the more the homeostasis forces, and the smoother the sleep onset transition. Similar findings were also found in mild cognitive impairment (MCI) patients, who had excessive daytime sleepiness (features closer to sleep onset state at bedtime), took longer to fall asleep and reached a shallower sleep state. I also found that age and gender could significantly impact the sleep onset latencies. Using all these factors (age, gender, and MCI) combined with the six most informative features (derived from classification and linear regression analyses), I built a Ridge regression model and showed that sleep-onset latency could be predicted. Of course, further improvements and replications can be done to further consolidate the findings and improve performances.

In Chapter 4, inspired by the underpinning neurophysiology of wake-sleep regulatory brain circuits, I provided the empirical evidence in human sleep EEG that the falling asleep transition can be modelled as a catastrophic bifurcation in the feature-state-space. By transforming the EEG data into the feature state space, one can observe the transition dynamics in the transformed space. The sleep state variable (distance to sleep onset) was found to follow the bifurcation dynamics, while the feature state space speed slowed down a bit during the transition. By studying the specific feature dynamics behind this transition, I found two types of transition dynamics using functional principal component analysis (FPCA). The first type, which is also the dominant one, resembled the bifurcation dynamics. The second type showed a burst of change before the transition and then went back, and although taking only a trivial amount of variance, it was found to be the crucial control parameter of the entire bifurcation transition process. For each type, there were two governing features identified, with important causal interactions between them. Specifically, it was found that the total EEG spectrum power causally drove the dynamics in the other feature dimensions, implying a potential neuromodulation target or trackable biomarker for this bifurcation process. Further, it provided potential insights about sleep initiation, that a gradual de-arousal (slowing down) could be the driving force before bifurcation. Finally, I showed the utility of such functional models in describing this transition process, in answering a 'so what' question. Before bifurcation, there would be a critical slowing down in the system dynamics, which contains early warning signals (increased autocorrelation and variance) that could help predict the tipping point.

Chapter 5 extended from the previous two chapters in resolving one of the limitations of the previous chapters, that is the lack of consideration of regional differences. It was found that different cortical regions had slight quantitative differences but qualitatively consistent dynamics. Different brain regions had specific features with different sensitivities to distinguish the two states. These regional gaps imply a potential functionality for detecting paradoxical insomnia, or better sensitivity in marking sleep onset states. The same bifurcation-type dynamics further confirm the solidity of the descriptive functional model developed in Chapter 4. In addition, I further developed an inter-regional interaction feature base to quantify different connectivity and non-connectivity relationships (e.g., power-band ratios) across channels. It was then found that the sleep onset state was hallmarked by an increase of transfer entropy between different brain regions in all directions, implying global information sharing between different regions at the sleep onset state. In this inter-regional feature state space, bifurcation dynamics could also be observed, with key dynamical features governed mostly by transfer entropy and directed transfer function. Lastly, I tested the causal interactions between the interactional features and regional feature dynamics and found a significant self-enhancing loop shedding light on more insights about sleep initiation: the burst increase in the information transfer at the alpha band from the Frontal to Posterior region (or the Frontal delta activity) drives a global information dependence, which further leads to the posterior de-arousal. Still, I acknowledge the primitiveness of this

framework and the limitations of the datasets, and the results require further polishing.

In sum, the first three results chapters presented the key representative EEG dynamics changes at sleep onset compared to bedtime wakefulness (key governing features), how they are changing (descriptive functional model as bifurcation), whether and how they are different across brain regions (regional dynamics), and how the regional dynamics interact and connect for sleep initiation. Of course, I acknowledge that the fundamental limit of all these results is the database. MESA was recorded at home around 13 years ago from the time of this thesis, therefore, the data quality could be poor compared to other laboratory-based newer datasets (such as the Montreal). The specific EEG montages and the specific participant cohort also determine the background of all the findings in this thesis. However, the use of meta-datasets is an important aspect of this work, and the sufficient participant number at least ensured the quality of findings on the group statistical level, with most replications done on the Montreal dataset confirming the findings as well. Aside from that, as previously discussed, EEG is a constrained method for measuring brain dynamics, and all current findings are founded on EEG data.

In Chapter 6, I presented the nap experiment to test the feasibility of using phase-locked (theta band post-peak) auditory stimulation to augment sleep onset and achieve cognitive (memory) benefits, as one of the key aims of this thesis. The choice of the stimulation condition (target phases) and the outcome measurements were anchored on the feature-based analysis presented before. I first showed that the audio was delivered to the desired phases. I then explored the effects of this phase-locked auditory stimulation on augmenting sleep. From the perspective of sleep staging (although done by artificial intelligence), there was no effect of the stimulation on the nap: no differences in sleep stages, sleep onset latencies, and drowsiness changes. From the perspective of key governing features, it was found that the auditory-intervened participants had slower brain rhythms (higher total EEG spectrum power) compared to the control. However, the changes were due to a larger decrease in this feature in the control participants, rather than augmentations in the stimulation participants. Finally, I showed that there was a weak benefit in associative memory after the auditory-intervened nap, based on the FNAT results. Of course, as discussed in the chapter, the analyses presented were preliminary with many potential confounding factors, and further analyses with refined methodologies should be done to further consolidate the stimulation effects, as well as to probe the potential mechanisms of auditory stimulation in sleep-onset augmentation. The results also shed light on potential personalised stimulation designs to augment sleep onset.

7.2. Novel contributions and implications

Specific findings and their implications were mostly discussed in each chapter. A list of key highlights of this thesis:

- The development and implementation of a spatial-temporal feature-based sleep EEG analysis pipeline into a meta-sleep dataset analysis; (Methodological advances)
- The first empirical illustration and modelling of the human brain falling asleep dynamics as a catastrophic bifurcation process, with the previous literature foundations set on theoretical neurophysiological models only (Knowledge advances).
- A Granger-causality-based framework to better understand the self-enhancing interactions in sleep initiation, which combines the regional brain dynamics, and their interactions (Knowledge advances).
- The feasibility test of auditory stimulation in the augmentation of sleep onset, with previous literature focusing on deep sleep (Translational and clinical advances).

Broader implications of the work:

- **Methodology**. The spatial-temporal feature-based statistical learning analysis pipeline could be used in other sleep neuroscience scenarios, for mechanistic biomarkers or functional neuromodolatory targets. The dynamical systems framework for quantifying other phase-transitions in the brain, especially in the context of arousal transitions (such as anaesthesia).
- **Sleep Neuroscience**. The bifurcation model of the falling asleep transition could be a potential novel continuous standard of a description of the falling asleep process outside sleep staging. Similarly, the model and the framework provide significant value in promoting the development of disease diagnosis and therapy treatments, as well as neuromodulation targets
- **Neuromodulation**. The preliminary intervention experiment supports further mechanistic understandings of phase-locked auditory interventions on brain dynamics; it also provides important insights for future personalized closed-loop neuromodulation strategy development;

7.3. Outlook

Taking a broader view of this thesis, there are some action points and interesting open questions to explore further. The first straightforward action is to combine the two feature bases (one local and one inter-regional) into a general feature base and build a systematic pipeline around it for the identification of the most informative local or connectivity feature. This would provide a more generalised biomarker research and a better sensitivity. As discussed in the main text, the feature base itself could also be refined to include more novel and question-specific features.

Concerning the field of sleep and falling asleep brain dynamics, it could be further studied how different demographic factors and neurological diseases can impact the transition dynamics (e.g., the bifurcation process), as well as the inter-regional interactions during sleep initiation.

Specifically, this could further help the design of a patient-specific closed-loop neuromodulation system that is better suited for their brain dynamics. Another aspect which is very common in the falling asleep dynamics literature is to study the effect of sleep deprivation on the dynamics (Gorgoni et al., 2021). Studying the impacts of demographic factors, brain diseases, and sleep deprivation is also an important pathway for a deeper mechanistic understanding of sleep regulations and sleep functionalities in the brain. It would also be interesting to further link the bifurcation transition model (and its control parameters) to underpinning neuronal mechanisms. The results shown took a data-driven approach in constructing the descriptive model, while the background literature used neurophysiological modelling. How are the cortical-level bifurcation dynamics driven by underlying neuronal bifurcation switches demonstrated in the model? And how could the neuronal control parameters (excitatory or inhibitory level) match the system-level control parameters found in the results? Specifically, the transition of neuronal firing rates happens at the timescale of a few seconds (Saper et al., 2010), but the transition at the whole-brain cortical level takes several minutes. A further in-depth study could probably answer these questions, particularly with the aid of other neuroimaging (fMRI for instance) with better spatial resolution and information on deeper brain activities. Similarly, for the results in Chapter 5, there needs a further functional meaning in the 'information transfer' between cortical areas from the underpinning neuronal dynamics (see details in that discussion).

Concerning the field of computational neuroscience and system dynamics, there has been a wide repertoire of data-driven approaches to studying the brain system dynamics, state transitions, and self-organised criticality (Plenz et al., 2021), revealing insights into brain functional mechanisms, and building models for neuromodulation aims (Deco et al., 2019). One relevant field would be reverse engineering: using a data-driven approach to discover underpinning governing equations (or laws) of the dynamical system (Schmidt and Lipson, 2009; Brunton et al., 2016b; Rudy et al., 2017). In the current thesis, the bifurcation model is purely descriptive, however, using those approaches, one might get the actual governing equation of the brain's dynamic (as a dynamical system) during the falling asleep transition, which would surely have huge interpretative values. Of course, this is just a naive thought, the complexity level of the brain (high-dimensional and multi-levels) could make the task impossible at all.

Concerning the field of neuromodulation for sleep interventions, aside from the phase-synchronised auditory stimulation used in this thesis, electrical stimulation could also be a potential tool. Transcranial current stimulation has achieved success in augmenting deep NREM sleep as well (Marshall et al., 2006), but again, yet to be explored in the augmentation of sleep onset. Recently developed non-invasive deep brain stimulation techniques such as temporal interference (Grossman et al., 2017) targeting deep brain structures make this path more promising. Of course, auditory stimulation should remain the priority as it is safer and easier to apply compared to electrical stimulations.

Bibliography

Acharya, U. R., Bhat, S., Faust, O., Adeli, H., Chua, E. C.-P., Lim, W. J. E., and Koh, J. E. W. (2015). Nonlinear Dynamics Measures for Automated EEG-Based Sleep Stage Detection. *European Neurology*, 74(5-6):268–287.

Acharya U., R., Faust, O., Kannathal, N., Chua, T., and Laxminarayan, S. (2005). Non-linear analysis of EEG signals at various sleep stages. *Computer Methods and Programs in Biomedicine*, 80(1):37–45.

Achermann, P., Rusterholz, T., Stucky, B., and Olbrich, E. (2019). Oscillatory patterns in the electroencephalogram at sleep onset. *Sleep*, 42(8):1–12.

Aeschbach, D. and Borbély, A. A. (1993). All-night dynamics of the human sleep EEG. *Journal of Sleep Research*, 2(2):70–81. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-2869.1993.tb00065.x.

Allada, R. and Bass, J. (2021). Circadian Mechanisms in Medicine. *New England Journal of Medicine*, 384(6):550–561. Publisher: Massachusetts Medical Society _eprint: https://doi.org/10.1056/NEJMra1802337.

Alvaro, P. K., Roberts, R. M., and Harris, J. K. (2013). A Systematic Review Assessing Bidirectionality between Sleep Disturbances, Anxiety, and Depression. *Sleep*, 36(7):1059–1068.

Anderson, K. N. and Bradley, A. J. (2013). Sleep disturbance in mental health problems and neurodegenerative disease. *Nature and Science of Sleep*, 5:61–75. Publisher: Dove Medical Press _eprint: https://www.tandfonline.com/doi/pdf/10.2147/NSS.S34842.

Aserinsky, E. and Kleitman, N. (1955). Two types of ocular motility occurring in sleep. *Journal of applied physiology*, 8(1):1–10.

Bandarabadi, M., Vassalli, A., and Tafti, M. (2020). Sleep as a default state of cortical and subcortical networks. *Current Opinion in Physiology*, 15:60–67.

Barnett, L., Barrett, A. B., and Seth, A. K. (2009). Granger Causality and Transfer Entropy Are Equivalent for Gaussian Variables. *Physical Review Letters*, 103(23):238701. Publisher: American Physical Society.

Barnett, L. and Seth, A. K. (2014). The MVGC multivariate Granger causality toolbox: A new approach to Granger-causal inference. *Journal of Neuroscience Methods*, 223:50–68. Publisher: Elsevier.

Bayer, L., Constantinescu, I., Perrig, S., Vienne, J., Vidal, P.-P., Mühlethaler, M., and Schwartz, S. (2011). Rocking synchronizes brain waves during a short nap. *Current Biology*, 21(12):R461–R462.

Bear, M., Connors, B., and Paradiso, M. A. (2020). *Neuroscience: Exploring the Brain, Enhanced Edition: Exploring the Brain*. Jones & Bartlett Learning.

Bellesi, M., Riedner, B. A., Garcia-Molina, G. N., Cirelli, C., and Tononi, G. (2014). Enhancement of sleep slow waves: underlying mechanisms and practical consequences. *Frontiers in Systems Neuroscience*, 8.

Berens, P. (2009). CircStat: a MATLAB toolbox for circular statistics. *Journal of statistical software*, 31:1–21.

Berger, H. (1929). Über das elektroenkephalogramm des menschen. *Archiv für psychiatrie und nervenkrankheiten*, 87(1):527–570.

Berndt, D. J. and Clifford, J. (1994). Using dynamic time warping to find patterns in time series. In *Proceedings of the 3rd International Conference on Knowledge Discovery and Data Mining*, AAAIWS'94, pages 359–370, Seattle, WA. AAAI Press.

Bliwise, D. L. (1993). Sleep in normal aging and dementia. *Sleep*, 16(1):40–81. Publisher: Oxford University Press.

Borb, A. A. and Achermann, P. (1999). Sleep homeostasis and models of sleep regulation. *Journal of biological rhythms*, 14(6):559–570. Publisher: Sage Publications Sage CA: Thousand Oaks, CA.

Borbély, A. A., Daan, S., Wirz-Justice, A., and Deboer, T. (2016). The two-process model of sleep regulation: a reappraisal. *Journal of Sleep Research*, 25(2):131–143. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/jsr.12371.

Born, J., Rasch, B., and Gais, S. (2006). Sleep to remember. *The Neuroscientist*, 12(5):410–424. Publisher: Sage Publications Sage CA: Thousand Oaks, CA.

Boser, B. E., Guyon, I. M., and Vapnik, V. N. (1992). A training algorithm for optimal margin classifiers. In *Proceedings of the fifth annual workshop on Computational learning theory*, COLT '92, pages 144–152, New York, NY, USA. Association for Computing Machinery.

Brunton, B. W., Johnson, L. A., Ojemann, J. G., and Kutz, J. N. (2016a). Extracting spatial-temporal coherent patterns in large-scale neural recordings using dynamic mode decomposition. *Journal of Neuroscience Methods*, 258:1–15. arXiv: 1409.5496 Publisher: Elsevier B.V.

Brunton, S. L., Proctor, J. L., Kutz, J. N., and Bialek, W. (2016b). Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences of the United States of America*, 113(15):3932–3937. arXiv: 1509.03580.

Bury, T. M., Sujith, R. I., Pavithran, I., Scheffer, M., Lenton, T. M., Anand, M., and Bauch, C. T. (2021). Deep learning for early warning signals of tipping points. *Proceedings of the National Academy of Sciences*, 118(39):e2106140118. Company: National Academy of Sciences Distributor: National Academy of Sciences ISBN: 9782106140118 Institution: National Academy of Sciences Label: National Academy of Sciences Publisher: Proceedings of the National Academy of Sciences.

Buysse, D. J., Reynolds, C. F., Monk, T. H., Berman, S. R., and Kupfer, D. J. (1989). The Pittsburgh sleep quality index: A new instrument for psychiatric practice and research. *Psychiatry Research*, 28(2):193–213.

Campos-Beltrán, D. and Marshall, L. (2021). Changes in sleep EEG with aging in humans and rodents. *Pflügers Archiv - European Journal of Physiology*, 473(5):841–851.

Canolty, R. T. and Knight, R. T. (2010). The functional role of cross-frequency coupling. *Trends in Cognitive Sciences*, 14(11):506–515. Publisher: Elsevier Ltd.

Capezuti, E., Pain, K., Alamag, E., Chen, X., Philibert, V., and Krieger, A. C. (2022). Systematic review: auditory stimulation and sleep. *Journal of Clinical Sleep Medicine*, 18(6):1697–1709. Publisher: American Academy of Sleep Medicine.

Carrier, J., Land, S., Buysse, D. J., Kupfer, D. J., and Monk, T. H. (2001). The effects of age and gender on sleep EEG power spectral density in the middle years of life (ages 20–60 years old). *Psychophysiology*, 38(2):232–242. Publisher: Cambridge University Press.

Cassone, V. M., Chesworth, M. J., and Armstrong, S. M. (1986). Entrainment of rat circadian rhythms by daily injection of melatonin depends upon the hypothalamic suprachiasmatic nuclei. *Physiology & Behavior*, 36(6):1111–1121.

Cavanagh, J. F. and Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends in Cognitive Sciences*, 18(8):414–421. Publisher: Elsevier.

Chella, F., Pizzella, V., Zappasodi, F., and Marzetti, L. (2016). Impact of the reference choice on scalp EEG connectivity estimation. *Journal of Neural Engineering*, 13(3):036016. Publisher: IOP Publishing.

Chen, X., Wang, R., Zee, P., Lutsey, P. L., Javaheri, S., Alcántara, C., Jackson, C. L., Williams, M. A., and Redline, S. (2015). Racial/Ethnic Differences in Sleep Disturbances: The Multi-Ethnic Study of Atherosclerosis (MESA). *Sleep*, 38(6):877–888.

Colombo, M. A., Napolitani, M., Boly, M., Gosseries, O., Casarotto, S., Rosanova, M., Brichant, J.-F., Boveroux, P., Rex, S., Laureys, S., Massimini, M., Chieregato, A., and Sarasso, S. (2019). The spectral exponent of the resting EEG indexes the presence of consciousness during unresponsiveness induced by propofol, xenon, and ketamine. *NeuroImage*, 189:631–644.

Comsa, I. M., Bekinschtein, T. A., and Chennu, S. (2019). Transient Topographical Dynamics of the Electroencephalogram Predict Brain Connectivity and Behavioural Responsiveness During Drowsiness. *Brain Topography*, 32(2):315–331. Publisher: Springer US ISBN: 0123456789.

Cormier, R. E. (1990). Sleep Disturbances. In Walker, H. K., Hall, W. D., and Hurst, J. W., editors, *Clinical Methods: The History, Physical, and Laboratory Examinations*. Butterworths, Boston, 3rd edition.

Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3):273–297.

Cox, R., Korjoukov, I., de Boer, M., and Talamini, L. M. (2014). Sound asleep: processing and retention of slow oscillation phase-targeted stimuli. *PLoS One*, 9(7). Publisher: Public Library of Science.

Dakos, V., Carpenter, S. R., Brock, W. A., Ellison, A. M., Guttal, V., Ives, A. R., Kéfi, S., Livina, V., Seekell, D. A., van Nes, E. H., and Scheffer, M. (2012). Methods for detecting early warnings of critical transitions in time series illustrated using simulated ecological data. *PLoS ONE*, 7(7).

Dakos, V., van Nes, E. H., and Scheffer, M. (2013). Flickering as an early warning signal. *Theoretical Ecology*, 6(3):309–317.

Daum, I., Leonard, J. P., and Hehl, F. J. (1988). Development of sleep during monotonous stimulation as related to individual differences. *The Pavlovian Journal of Biological Science*, 23(3):118–124.

Dauvilliers, Y. (2007). Insomnia in patients with neurodegenerative conditions. *Sleep Medicine*, 8:S27–S34.

Dauwels, J., Srinivasan, K., Ramasubba Reddy, M., Musha, T., Vialatte, F.-B., Latchoumane, C., Jeong, J., and Cichocki, A. (2011). Slowing and Loss of Complexity in Alzheimer's EEG: Two Sides of the Same Coin? *International Journal of Alzheimer' s Disease*, 2011:e539621. Publisher: Hindawi.

Davis, H., Davis, P. A., Loomis, A. L., Harvey, E. N., and Hobart, G. (1937). Changes in human brain potentials during the onset of sleep. *Science*, 86(2237):448–450. Publisher: American Association for the Advancement of Science.

Davis, H., Davis, P. A., Loomis, A. L., Harvey, E. N., and Hobart, G. (1938). Human brain potentials during the onset of sleep. *Journal of Neurophysiology*, 1(1):24–38.

De Gennaro, L., Devoto, A., Lucidi, F., and Violani, C. (2005). Oculomotor changes are associated to daytime sleepiness in the multiple sleep latency test. *Journal of Sleep Research*, 14(2):107–112.

De Gennaro, L., Ferrara, M., and Bertini, M. (2001a). The boundary between wakefulness and sleep: Quantitative electroencephalographic changes during the sleep onset period. *Neuroscience*, 107(1):1–11.

De Gennaro, L., Ferrara, M., Curcio, G., and Cristiani, R. (2001b). Antero-posterior EEG changes during the wakefulness-sleep transition. *Clinical Neurophysiology*, 112(10):1901–1911.

De Gennaro, L., Ferrara, M., Ferlazzo, F., and Bertini, M. (2000). Slow eye movements and EEG power spectra during wake-sleep transition. *Clinical Neurophysiology*, 111(12):2107–2115.

De Gooijer, J. G. and Hyndman, R. J. (2006). 25 years of time series forecasting. *International Journal of Forecasting*, 22(3):443–473.

de Mooij, S. M., Blanken, T. F., Grasman, R. P., Ramautar, J. R., Van Someren, E. J., and van der Maas, H. L. (2020). Dynamics of sleep: Exploring critical transitions and early warning signals. *Computer Methods and Programs in Biomedicine*, 193:105448. Publisher: Elsevier B.V.

de Zambotti, M., Goldstone, A., Forouzanfar, M., Javitz, H., Claudatos, S., Colrain, I. M., and Baker, F. C. (2020). The falling asleep process in adolescents. *Sleep*, 43(6):zsz312.

Debellemanière, E., Pinaud, C., Schneider, J., Arnal, P. J., Casson, A. J., Chennaoui, M., Galtier, M., Navarrete, M., and Lewis, P. A. (2022). Optimising sounds for the driving of sleep oscillations by closed-loop auditory stimulation. *Journal of Sleep Research*, 31(6):e13676. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/jsr.13676.

Deco, G., Cruzat, J., Cabral, J., Tagliazucchi, E., Laufs, H., Logothetis, N. K., and Kringelbach, M. L. (2019). Awakening: Predicting external stimulation to force transitions between different brain states. *Proceedings of the National Academy of Sciences of the United States of America*, 116(36):18088–18097. ISBN: 1905534116.

Delbeuck, X., Van der Linden, M., and Collette, F. (2003). Alzheimer' Disease as a Disconnection Syndrome? *Neuropsychology Review*, 13(2):79–92.

Demanuele, C., James, C. J., and Sonuga-Barke, E. J. (2007). Distinguishing low frequency oscillations within the 1/f spectral behaviour of electromagnetic brain signals. *Behavioral and Brain Functions*, 3:1–14.

Dement, W. and Kleitman, N. (1957a). Cyclic variations in EEG during sleep and their relation to eye movements, body motility, and dreaming. *Electroencephalography and Clinical Neurophysiology*, 9(4):673–690.

Dement, W. and Kleitman, N. (1957b). The relation of eye movements during sleep to dream activity: An objective method for the study of dreaming. *Journal of Experimental Psychology*, 53(5):339–346.

Diaz, B. A., Hardstone, R., Mansvelder, H. D., Van Someren, E. J., and Linkenkaer-Hansen, K. (2016). Resting-state subjective experience and EEG biomarkers are associated with sleep-onset latency. *Frontiers in Psychology*, 7(APR):1–11.

Dijk, D.-J. and Landolt, H.-P. (2019). Sleep Physiology, Circadian Rhythms, Waking Performance and the Development of Sleep-Wake Therapeutics. In Landolt, H.-P. and Dijk, D.-J., editors, *Sleep-Wake Neurobiology and Pharmacology*, Handbook of Experimental Pharmacology, pages 441–481. Springer International Publishing, Cham.

Dimitriadis, S. I., Salis, C. I., and Liparas, D. (2021). An automatic sleep disorder detection based on EEG cross-frequency coupling and random forest model. *Journal of Neural Engineering*, 18(4):046064. Publisher: IOP Publishing.

Diniz Behn, C. G., Klerman, E. B., Mochizuki, T., Lin, S. C., and Scammell, T. E. (2010). Abnormal sleep/wake dynamics in orexin knockout mice. *Sleep*, 33(3):297–306.

Doneh, B. (2015). Epworth Sleepiness Scale. *Occupational Medicine*, 65(6):508.

Donoghue, T., Haller, M., Peterson, E. J., Varma, P., Sebastian, P., Gao, R., Noto, T., Lara, A. H., Wallis, J. D., Knight, R. T., Shestyuk, A., and Voytek, B. (2020). Parameterizing neural power spectra into periodic and aperiodic components. *Nature Neuroscience*, 23(12):1655–1665. Publisher: Springer US.

Duan, K.-B. and Keerthi, S. S. (2005). Which Is the Best Multiclass SVM Method? An Empirical Study. In Oza, N. C., Polikar, R., Kittler, J., and Roli, F., editors, *Multiple Classifier Systems*, Lecture Notes in Computer Science, pages 278–285, Berlin, Heidelberg. Springer.

D' Ambrosio, S., Castelnovo, A., Guglielmi, O., Nobili, L., Sarasso, S., and Garbarino, S. (2019). Sleepiness as a Local Phenomenon. *Frontiers in Neuroscience*, 13.

Ebisuzaki, W. (1997). A Method to Estimate the Statistical Significance of a Correlation When the Data Are Serially Correlated. *Journal of Climate*, 10(9):2147–2153. Publisher: American Meteorological Society Section: Journal of Climate.

Eggert, T., Dorn, H., Sauter, C., Nitsche, M. A., Bajbouj, M., and Danker-Hopfe, H. (2013). No Effects of Slow Oscillatory Transcranial Direct Current Stimulation (tDCS) on Sleep-Dependent Memory Consolidation in Healthy Elderly Subjects. *Brain Stimulation*, 6(6):938–945.

Eom, T.-H. (2022). Electroencephalography source localization. *Clinical and Experimental Pediatrics*, 66(5):201–209.

Escudero, J., Hornero, R., and Abásolo, D. (2009). Interpretation of the auto-mutual information rate of decrease in the context of biomedical signal analysis. Application to electroencephalogram recordings. *Physiological Measurement*, 30(2):187.

Esfahani, M. J., Farboud, S., Ngo, H.-V. V., Schneider, J., Weber, F. D., Talamini, L. M., and Dresler, M. (2023). Closed-loop auditory stimulation of sleep slow oscillations: Basic principles and best practices. *Neuroscience & Biobehavioral Reviews*, 153:105379.

Fang, H. and Yang, Y. (2022). Designing and validating a robust adaptive neuromodulation algorithm for closed-loop control of brain states. *Journal of Neural Engineering*, 19(3).

Fattinger, S., de Beukelaar, T. T., Ruddy, K. L., Volk, C., Heyse, N. C., Herbst, J. A., Hahnloser, R. H. R., Wenderoth, N., and Huber, R. (2017). Deep sleep maintains learning efficiency of the human brain. *Nature Communications*, 8(1):15405. Number: 1 Publisher: Nature Publishing Group.

Fava, G. A. (1996). Amnestic syndrome induced by zopiclone. *European Journal of Clinical Pharmacology*, 50(6):509–509.

Feldman, D. P. (2019). *Chaos and Dynamical Systems*. Princeton University Press. Google-Books-ID: zweaDwAAQBAJ.

Fernandez Guerrero, A. and Achermann, P. (2018). Intracortical Causal Information Flow of Oscillatory Activity (Effective Connectivity) at the Sleep Onset Transition. *Frontiers in Neuroscience*, 12(December):1–15.

Ferrara, M. and De Gennaro, L. (2000). The sleep inertia phenomenon during the sleep-wake transition: Theoretical and operational issues. *Aviation Space and Environmental Medicine*, 71(8):843–848.

Ferrara, M., De Gennaro, L., Curcio, G., Cristiani, R., Corvasce, C., and Bertini, M. (2002). Regional differences of the human sleep electroencephalogram in response to selective slow-wave sleep deprivation. *Cerebral Cortex*, 12(7):737–748.

Ferri, R., Rundo, F., Bruni, O., Terzano, M. G., and Stam, C. J. (2007). Small-world network organization of functional connectivity of EEG slow-wave activity during sleep. *Clinical Neurophysiology*, 118(2):449–456.

Ferri, R., Rundo, F., Bruni, O., Terzano, M. G., and Stam, C. J. (2008). The functional connectivity of different EEG bands moves towards small-world network organization during sleep. *Clinical Neurophysiology*, 119(9):2026–2036.

Ficca, G., Axelsson, J., Mollicone, D. J., Muto, V., and Vitiello, M. V. (2010). Naps, cognition and performance. *Sleep Medicine Reviews*, 14(4):249–258.

Flanigan, W. F. (1972). Behavioral states and electroencephalograms of reptiles. *The sleeping brain. Perspectives in the brain sciences, Los Angeles: Brain Information Service/Brain Research Institute, UCLA*, pages 14–18.

Fletcher, A., Van Den Heuvel, C., and Dawson, D. (1999). Sleeping with an electric blanket: Effects on core temperature, sleep, and melatonin in young adults. *Sleep*, 22(3):313–318.

Fong, S., Ho, C., and Wing, Y. (2005). Comparing MSLT and ESS in the measurement of excessive daytime sleepiness in obstructive sleep apnoea syndrome. *Journal of Psychosomatic Research*, 58(1):55–60.

Forquer, L. and Johnson, C. M. (2007). Continuous white noise to reduce sleep latency and night wakings in college students. *Sleep and Hypnosis*, 9(2):60. Publisher: KURE ILETISIM GRUBU AS.

Fulcher, B. D. and Jones, N. S. (2017). hctsa: A Computational Framework for Automated Time-Series Phenotyping Using Massive Feature Extraction. *Cell Systems*, 5(5):527–531.e3. Publisher: Elsevier.

Fulcher, B. D., Little, M. A., and Jones, N. S. (2013). Highly comparative time-series analysis: the empirical structure of time series and their methods. *Journal of The Royal Society Interface*, 10(83):20130048. Publisher: Royal Society.

Gagnon, K., Baril, A. A., Montplaisir, J., Carrier, J., Chami, S., Gauthier, S., Lafond, C., Gagnon, J. F., and Gosselin, N. (2018). Detection of mild cognitive impairment in middle-aged and older adults with obstructive sleep apnoea. *European Respiratory Journal*, 52(5).

Gatica, M., Cofré, R., Mediano, P. A., Rosas, F. E., Orio, P., Diez, I., Swinnen, S. P., and Cortes, J. M. (2021). High-Order Interdependencies in the Aging Brain. *Brain Connectivity*, 11(9):734–744. Publisher: Mary Ann Liebert, Inc., publishers.

Gauthier, S., Reisberg, B., Zaudig, M., Petersen, R. C., Ritchie, K., Broich, K., Belleville, S., Brodaty, H., Bennett, D., Chertkow, H., Cummings, J. L., de Leon, M., Feldman, H.,

Ganguli, M., Hampel, H., Scheltens, P., Tierney, M. C., Whitehouse, P., and Winblad, B. (2006). Mild cognitive impairment. *The Lancet*, 367(9518):1262–1270. Publisher: Elsevier.

Gennaro, L. D., Vecchio, F., Ferrara, M., Curcio, G., Rossini, P. M., and Babiloni, C. (2004). Changes in fronto-posterior functional coupling at sleep onset in humans. *Journal of sleep research*, 13(3):209–217. Publisher: Wiley Online Library.

Gennaro, L. D., Vecchio, F., Ferrara, M., Curcio, G., Rossini, P. M., and Babiloni, C. (2005). Antero-posterior functional coupling at sleep onset: changes as a function of increased sleep pressure. *Brain Research Bulletin*, 65(2):133–140.

Gersch, W. and Yonemoto, J. (1977). Parametric time series models for multivariate EEG analysis. *Computers and Biomedical Research*, 10(2):113–125.

Geweke, J. (1984). Chapter 19 Inference and causality in economic time series models. In *Handbook of Econometrics*, volume 2, pages 1101–1144. Elsevier.

Giles, D. (2011). Econometrics Beat: Dave Giles' Blog: Testing for Granger Causality.

Goldstein, A. N. and Walker, M. P. (2014). The Role of Sleep in Emotional Brain Function. *Annual Review of Clinical Psychology*, 10(1):679–708. _eprint: https://doi.org/10.1146/annurev-clinpsy-032813-153716.

Gorgoni, M., D'Atri, A., Scarpelli, S., Ferrara, M., and De Gennaro, L. (2020). The electroencephalographic features of the sleep onset process and their experimental manipulation with sleep deprivation and transcranial electrical stimulation protocols. *Neuroscience and biobehavioral reviews*, 114(March):25–37. Publisher: Elsevier.

Gorgoni, M., D' Atri, A., Scarpelli, S., Ferrara, M., and De Gennaro, L. (2019). Chapter 2 - Timing and Topography of Sleep Onset: Asynchronies and Regional Changes of Brain Activity. In Dringenberg, H. C., editor, *Handbook of Behavioral Neuroscience*, volume 30 of *Handbook of Sleep Research*, pages 19–31. Elsevier.

Gorgoni, M., Scarpelli, S., Annarumma, L., D' Atri, A., Alfonsi, V., Ferrara, M., and De Gennaro, L. (2021). The Regional EEG Pattern of the Sleep Onset Process in Older Adults. *Brain Sciences*, 11(10):1261. Number: 10 Publisher: Multidisciplinary Digital Publishing Institute.

Grandner, M. A. and Fernandez, F.-X. (2021). The translational neuroscience of sleep: a contextual framework. *Science*, 374(6567):568–573. Publisher: American Association for the Advancement of Science.

Grossman, N., Bono, D., Dedic, N., Kodandaramaiah, S. B., Rudenko, A., Suk, H. J., Cassara, A. M., Neufeld, E., Kuster, N., Tsai, L. H., Pascual-Leone, A., and Boyden, E. S.

(2017). Noninvasive Deep Brain Stimulation via Temporally Interfering Electric Fields. *Cell*, 169(6):1029–1041.e16. Publisher: Elsevier.

Guerrero, A. F. and Achermann, P. (2019). Brain dynamics during the sleep onset transition: an EEG source localization study. *Neurobiology of sleep and circadian rhythms*, 6:24–34. Publisher: Elsevier.

Gutierrez, J. (2023). Neuromodulation of the Resting-State Correlates of Associative Memory Using Phase-Locked Auditory Stimulation During a Short Daytime Nap. Master's thesis, Imperial College London, London, UK.

Gyurkovics, M., Clements, G. M., Low, K. A., Fabiani, M., and Gratton, G. (2022). Stimulus-Induced Changes in 1/f-like Background Activity in EEG. *Journal of Neuroscience*, 42(37):7144–7151. Publisher: Society for Neuroscience Section: Research Articles.

Hall-Porter, J. M., Schweitzer, P. K., Eisenstein, R. D., Ahmed, H. A. H., and Walsh, J. K. (2014). The effect of two benzodiazepine receptor agonist hypnotics on sleep-dependent memory consolidation. *Journal of Clinical Sleep Medicine*, 10(01):27–34. Publisher: American Academy of Sleep Medicine.

Halász, P. (2016). The K-complex as a special reactive sleep slow wave –A theoretical update. *Sleep Medicine Reviews*, 29:34–40.

Halász, P., Terzano, M., Parrino, L., and Bódizs, R. (2004). The nature of arousal in sleep. *Journal of Sleep Research*, 13(1):1–23.

Hamilton, J. D. (2020). *Time Series Analysis*. Princeton University Press. Google-Books-ID: BeryDwAAQBAJ.

Hancock, F., Rosas, F. E., Mediano, P. A. M., Luppi, A. I., Cabral, J., Dipasquale, O., and Turkheimer, F. E. (2022). May the 4C's be with you: an overview of complexity-inspired frameworks for analysing resting-state neuroimaging data. *Journal of The Royal Society Interface*, 19(191):20220214. Publisher: Royal Society.

Hara, J., Beuckmann, C. T., Nambu, T., Willie, J. T., Chemelli, R. M., Sinton, C. M., Sugiyama, F., Yagami, K.-i., Goto, K., and Yanagisawa, M. (2001). Genetic ablation of orexin neurons in mice results in narcolepsy, hypophagia, and obesity. *Neuron*, 30(2):345–354. Publisher: Elsevier.

Harsh, J. R., Voss, U., Hull, J., Schrepfer, S., and Badia, P. (1994). ERP and behavioral changes during the wake/sleep transition. *Psychophysiology*, 31 3:244–252.

Hastie, T. (2020). Ridge Regularization: An Essential Concept in Data Science. *Technometrics : a journal of statistics for the physical, chemical, and engineering sciences*, 62(4):426–433.

Haufler, D. and Paré, D. (2019). Detection of Multiway Gamma Coordination Reveals How Frequency Mixing Shapes Neural Dynamics. *Neuron*, 101(4):603–614.e6. Publisher: Elsevier Inc.

Hauri, P. J. and Wisbey, J. (1992). Wrist actigraphy in insomnia. *Sleep*, 15(4):293–301.

Hebron, H., Lugli, B., Dimitrova, R., Jaramillo, V., Rhodes, E., Grossman, N., Dijk, D.-J., and Violante, I. R. (2023). Alpha closed-loop auditory stimulation modulates waking alpha oscillations and sleep onset dynamics in a phase-dependent manner in humans. Pages: 2022.11.15.512722 Section: New Results.

Helfrich, R. F., Mander, B. A., Jagust, W. J., Knight, R. T., and Walker, M. P. (2018). Old Brains Come Uncoupled in Sleep: Slow Wave-Spindle Synchrony, Brain Atrophy, and Forgetting. *Neuron*, 97(1):221–230.e4. Publisher: Elsevier Inc.

Hermans, L. W., Huijben, I. A., van Gorp, H., Leufkens, T. R., Fonseca, P., Overeem, S., and van Gilst, M. M. (2022). Representations of temporal sleep dynamics: Review and synthesis of the literature. *Sleep Medicine Reviews*, 63:101611.

Hermans, L. W., Leufkens, T. R., van Gilst, M. M., Weysen, T., Ross, M., Anderer, P., Overeem, S., and Vermeeren, A. (2019). Sleep EEG characteristics associated with sleep onset misperception. *Sleep Medicine*, 57:70–79. Publisher: Elsevier Ltd.

Hewitt, E. and Hewitt, R. E. (1979). The Gibbs-Wilbraham Phenomenon: An Episode in Fourier Analysis. *Archive for History of Exact Sciences*, 21(2):129–160. Publisher: Springer.

Hoerl, A. E. and Kennard, R. W. (1981). Ridge Regression —1980: Advances, Algorithms, and Applications. *American Journal of Mathematical and Management Sciences*, 1(1):5–83. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/01966324.1981.10737061.

Holth, J. K., Patel, T. K., and Holtzman, D. M. (2017). Sleep in Alzheimer's Disease–Beyond Amyloid. *Neurobiology of Sleep and Circadian Rhythms*, 2:4–14.

Homan, R. W. (1988). The 10-20 Electrode System and Cerebral Location. *American Journal of EEG Technology*, 28(4):269–279. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/00029238.1988.11080272.

Hori, T. (1985). Spatiotemporal changes of EEG activity during waking-sleeping transition period. *International Journal of Neuroscience*, 27(1-2):101–114. Publisher: Taylor & Francis.

Hori, T., Hayashi, M., and Morikawa, T. (1994). Topographical EEG changes and the hypnagogic experience. Publisher: American Psychological Association.

Horikawa, T., Tamaki, M., Miyawaki, Y., and Kamitani, Y. (2013). Neural Decoding of Visual Imagery During Sleep. *Science*, 340(6132):639–642. Publisher: American Association for the Advancement of Science.

Huang, C.-S., Pal, N. R., Chuang, C.-H., and Lin, C.-T. (2015). Identifying changes in EEG information transfer during drowsy driving by transfer entropy. *Frontiers in Human Neuroscience*, 9.

Huang, Y., Zhang, J., Cui, Y., Yang, G., He, L., Liu, Q., and Yin, G. (2017). How Different EEG References Influence Sensor Level Functional Connectivity Graphs. *Frontiers in Neuroscience*, 11.

Huang, Z.-L., Qu, W.-M., Eguchi, N., Chen, J.-F., Schwarzschild, M. A., Fredholm, B. B., Urade, Y., and Hayaishi, O. (2005). Adenosine A2A, but not A1, receptors mediate the arousal effect of caffeine. *Nature Neuroscience*, 8(7):858–859. Number: 7 Publisher: Nature Publishing Group.

Huber, R., Felice Ghilardi, M., Massimini, M., and Tononi, G. (2004). Local sleep and learning. *Nature*, 430(6995):78–81. Number: 6995 Publisher: Nature Publishing Group.

Hwang, E., Kim, S., Han, K., and Choi, J. H. (2012). Characterization of Phase Transition in the Thalamocortical System during Anesthesia-Induced Loss of Consciousness. *PLOS ONE*, 7(12):e50580. Publisher: Public Library of Science.

Iber, C. (2007). The AASM manual for the scoring of sleep and associated events: Rules. *Terminology and Technical Specification*. Publisher: American Academy of Sleep Medicine.

Ilina, A. (2022). Uncovering brain dynamics underlying night-time awakenings using feature-based statistical learning of sleep EEG. Master's thesis, Imperial College London, London, UK.

Ilina, A. (2023). Forecasting Sleep Onset from Brain Dynamics: Advanced Ensemble and Deep Learning Approaches Applied to Feature-Based EEG Data. Master's thesis, Imperial College London, London, UK.

Ives, A. R. and Dakos, V. (2012). Detecting dynamical changes in nonlinear time series using locally linear state-space models. *Ecosphere*, 3(6):art58. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1890/ES11-00347.1.

Jacobs, J., Miller, J., Lee, S. A., Coffey, T., Watrous, A. J., Sperling, M. R., Sharan, A., Worrell, G., Berry, B., Lega, B., Jobst, B. C., Davis, K., Gross, R. E., Sheth, S. A., Ezzyat, Y., Das, S. R., Stein, J., Gorniak, R., Kahana, M. J., and Rizzuto, D. S. (2016). Direct Electrical Stimulation of the Human Entorhinal Region and Hippocampus Impairs Memory. *Neuron*, 92(5):983–990. Publisher: Elsevier.

James, G., Witten, D., Hastie, T., and Tibshirani, R. (2021a). Statistical Learning. In James, G., Witten, D., Hastie, T., and Tibshirani, R., editors, *An Introduction to Statistical Learning: with Applications in R*, Springer Texts in Statistics, pages 15–57. Springer US, New York, NY.

James, G., Witten, D., Hastie, T., and Tibshirani, R. (2021b). Support Vector Machines. In James, G., Witten, D., Hastie, T., and Tibshirani, R., editors, *An Introduction to Statistical Learning: with Applications in R*, Springer Texts in Statistics, pages 367–402. Springer US, New York, NY.

Janjarasjitt, S., Scher, M. S., and Loparo, K. A. (2008). Nonlinear dynamical analysis of the neonatal EEG time series: The relationship between sleep state and complexity. *Clinical Neurophysiology*, 119(8):1812–1823.

Jap, B. T., Lal, S., Fischer, P., and Bekiaris, E. (2009). Using EEG spectral components to assess algorithms for detecting fatigue. *Expert Systems with Applications*, 36(2, Part 1):2352–2359.

Jaussent, I., Bouyer, J., Ancelin, M.-L., Berr, C., Foubert-Samier, A., Ritchie, K., Ohayon, M. M., Besset, A., and Dauvilliers, Y. (2012). Excessive sleepiness is predictive of cognitive decline in the elderly. *Sleep*, 35(9):1201–1207. Publisher: Oxford University Press.

Jensen, O., Kaiser, J., and Lachaux, J.-P. (2007). Human gamma-frequency oscillations associated with attention and memory. *Trends in Neurosciences*, 30(7):317–324.

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

Jeong, J., Gore, J. C., and Peterson, B. S. (2001). Mutual information analysis of the EEG in patients with Alzheimer's disease. *Clinical Neurophysiology*, 112(5):827–835.

Jiang, X., Bian, G. B., and Tian, Z. (2019). Removal of artifacts from EEG signals: A review. *Sensors (Switzerland)*, 19(5):1–18.

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

Johns, M. W. (1991). A New Method for Measuring Daytime Sleepiness: The Epworth Sleepiness Scale. *Sleep*, 14(6):540–545.

Ju, Y.-E. S., Lucey, B. P., and Holtzman, D. M. (2014). Sleep and Alzheimer disease pathology—a bidirectional relationship. *Nature Reviews Neurology*, 10(2):115–119. Number: 2 Publisher: Nature Publishing Group.

Kang, J.-E., Lim, M. M., Bateman, R. J., Lee, J. J., Smyth, L. P., Cirrito, J. R., Fujiki, N., Nishino, S., and Holtzman, D. M. (2009). Amyloid-beta Dynamics Are Regulated by Orexin and the Sleep-Wake Cycle. *Science*, 326(November):1005–1008.

Kaplan, A. Y., Fingelkurts, A. A., Fingelkurts, A. A., Borisov, S. V., and Darkhovsky, B. S. (2005). Nonstationary nature of the brain activity as revealed by EEG/MEG: Methodological, practical and conceptual challenges. *Signal Processing*, 85(11):2190–2212.

Kawada, T. and Suzuki, S. (1993). Sleep Induction Effects of Steady 60 dB (A) Pink Noise. *Industrial health*, 31(1):35–38.

Kay, A., Trinder, J., Bowes, G., and Kim, Y. (1994). Changes in airway resistance during sleep onset. *Journal of Applied Physiology*, 76(4):1600–1607.

King, A. P. and Eckersley, R. J. (2019). Chapter 5 - Inferential Statistics II: Parametric Hypothesis Testing. In King, A. P. and Eckersley, R. J., editors, *Statistics for Biomedical Engineers and Scientists*, pages 91–117. Academic Press.

Kinreich, S., Podlipsky, I., Jamshy, S., Intrator, N., and Hendler, T. (2014). Neural dynamics necessary and sufficient for transition into pre-sleep induced by EEG NeuroFeedback. *NeuroImage*, 97:19–28.

Kjaer, T. W., Law, I., Wiltschiøtz, G., Paulson, O. B., and Madsen, P. L. (2002). Regional cerebral blood flow during light sleep –a H215O-PET study. *Journal of Sleep Research*, 11(3):201–207. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1046/j.1365-2869.2002.00303.x.

Kleiner, M., Brainard, D., and Pelli, D. (2007). What's new in Psychtoolbox-3? Publisher: Pion Ltd.

Klepl, D., He, F., Wu, M., Blackburn, D. J., and Sarrigiannis, P. G. (2022). Cross-Frequency Multilayer Network Analysis with Bispectrum-based Functional Connectivity: A Study of Alzheimer' s Disease. Pages: 2021.08.07.455499 Section: New Results.

Klimesch, W., Sauseng, P., and Hanslmayr, S. (2007). EEG alpha oscillations: The inhibition–timing hypothesis. *Brain Research Reviews*, 53(1):63–88.

Koenis, M. M., Romeijn, N., Piantoni, G., Verweij, I., Van der Werf, Y. D., Van Someren, E. J., and Stam, C. J. (2013). Does sleep restore the topology of functional brain networks? *Human Brain Mapping*, 34(2):487–500. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/hbm.21455.

Konjarski, M., Murray, G., Lee, V. V., and Jackson, M. L. (2018). Reciprocal relationships between daily sleep and mood: A systematic review of naturalistic prospective studies. *Sleep Medicine Reviews*, 42:47–58.

Kovach, C. K., Oya, H., and Kawasaki, H. (2018). The bispectrum and its relationship to phase-amplitude coupling. *NeuroImage*, 173:518–539.

Kovrov, G. V., Merkulova, T. B., Posokhov, S. I., and Garakh, Z. V. (2018). Dynamics of Alpha Rhythm Peak Frequency during Falling Asleep. *Bulletin of Experimental Biology and Medicine*, 165(2):177–180.

Kraskov, A., Stögbauer, H., and Grassberger, P. (2004). Estimating mutual information. *Physical Review E*, 69(6):066138. Publisher: American Physical Society.

Krueger, J. M., Nguyen, J. T., Dykstra-Aiello, C. J., and Taishi, P. (2019). Local sleep. *Sleep Medicine Reviews*, 43:14–21.

Krueger, J. M., Rector, D. M., Roy, S., Van Dongen, H. P. A., Belenky, G., and Panksepp, J. (2008). Sleep as a fundamental property of neuronal assemblies. *Nature Reviews Neuroscience*, 9(12):910–919. Number: 12 Publisher: Nature Publishing Group.

Kräuchi, K., Cajochen, C., Werth, E., and Wirz-Justice, A. (1999). Warm feet promote the rapid onset of sleep. *Nature*, 401(6748):36–37. Publisher: Nature Publishing Group.

Kräuchi, K., Cajochen, C., Werth, E., and Wirz-Justice, A. (2000). Functional link between distal vasodilation and sleep-onset latency? *American Journal of Physiology - Regulatory Integrative and Comparative Physiology*, 278(3 47-3):741–748.

Lal, C., Weaver, T. E., Bae, C. J., and Strohl, K. P. (2021). Excessive Daytime Sleepiness in Obstructive Sleep Apnea. Mechanisms and Clinical Management. *Annals of the American Thoracic Society*, 18(5):757–768. Publisher: American Thoracic Society - AJRCCM.

Lal, S. K. L. and Craig, A. (2001). A critical review of the psychophysiology of driver fatigue. *Biological Psychology*, 55(3):173–194.

Landolt, H.-P. and Borbély, A. A. (2001). Age-dependent changes in sleep EEG topography. *Clinical Neurophysiology*, 112(2):369–377.

Landolt, H.-P., Dijk, D.-J., Achermann, P., and Borbély, A. A. (1996). Effect of age on the sleep EEG: slow-wave activity and spindle frequency activity in young and middle-aged men. *Brain Research*, 738(2):205–212.

Lavie, P. (2007). Insomnia and sleep-disordered breathing. *Sleep Medicine*, 8:S21–S25.

Lehnertz, K. (1999). Non-linear time series analysis of intracranial EEG recordings in patients with epilepsy —an overview. *International Journal of Psychophysiology*, 34(1):45–52.

Lehnertz, K. (2008). Epilepsy and Nonlinear Dynamics. *Journal of Biological Physics*, 34(3):253–266.

Lendner, J. D., Helfrich, R. F., Mander, B. A., Romundstad, L., Lin, J. J., Walker, M. P., Larsson, P. G., and Knight, R. T. (2020). An electrophysiological marker of arousal level in humans. *eLife*, 9:e55092.

Leng, Y., McEvoy, C. T., Allen, I. E., and Yaffe, K. (2017). Association of Sleep-Disordered Breathing With Cognitive Function and Risk of Cognitive Impairment: A Systematic Review and Meta-analysis. *JAMA Neurology*, 74(10):1237–1245.

Leocani, L. and Comi, G. (1999). EEG Coherence in Pathological Conditions. *Journal of Clinical Neurophysiology*, 16(6):548.

Lewis, L. D. (2021). The interconnected causes and consequences of sleep in the brain. *Science*, 374(6567):564–568. Publisher: American Association for the Advancement of Science.

Li, J., Vitiello, M. V., and Gooneratne, N. S. (2022). Sleep in Normal Aging. *Sleep Medicine Clinics*, 17(2):161–171.

Liberson, W. T. and Liberson, C. W. (1966). EEG records, reaction times, eye movements, respiration, and mental content during drowsiness. In *Recent advances in biological psychiatry*, pages 295–302. Springer.

Lim, A. S. P., Kowgier, M., Yu, L., Buchman, A. S., and Bennett, D. A. (2013). Sleep Fragmentation and the Risk of Incident Alzheimer's Disease and Cognitive Decline in Older Persons. *Sleep*, 36(7):1027–1032.

Lin, Y.-Q. and Chen, S.-D. (2018). RBD: a red flag for cognitive impairment in Parkinson's disease? *Sleep Medicine*, 44:38–44.

Liu, H., Wu, X., and Zhang, S. (2011). Feature selection using hierarchical feature clustering. In *Proceedings of the 20th ACM international conference on Information and knowledge management*, CIKM '11, pages 979–984, New York, NY, USA. Association for Computing Machinery.

Loomis, A. L., Harvey, E. N., and Hobart, G. A. (1937). Cerebral states during sleep, as studied by human brain potentials. *Journal of experimental psychology*, 21(2):127. Publisher: Psychological Review Company.

Lopes da Silva, F., Blanes, W., Kalitzin, S. N., Parra, J., Suffczynski, P., and Velis, D. N. (2003). Epilepsies as dynamical diseases of brain systems: basic models of the transition between normal and epileptic activity. *Epilepsia*, 44 Suppl 12:72–83.

Lovato, N. and Lack, L. (2010). The effects of napping on cognitive functioning. In Kerkhof, G. A. and Dongen, H. P. A. v., editors, *Progress in Brain Research*, volume 185, pages 155–166. Elsevier.

Lubba, C. H., Sethi, S. S., Knaute, P., Schultz, S. R., Fulcher, B. D., and Jones, N. S. (2019). catch22: CAnonical Time-series CHaracteristics. *Data Mining and Knowledge Discovery*, 33(6):1821–1852.

Luff, C. E., Peach, R., Mallas, E.-J., Rhodes, E., Laumann, F., Boyden, E. S., Sharp, D. J., Barahona, M., and Grossman, N. (2023). The neuron mixer and its impact on human brain dynamics. Pages: 2023.01.05.522833 Section: New Results.

Ma, D. S., Correll, J., and Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior research methods*, 47:1122–1135. Publisher: Springer.

Ma, Y., Shi, W., Peng, C.-K., and Yang, A. C. (2018). Nonlinear dynamical analysis of sleep electroencephalography using fractal and entropy approaches. *Sleep Medicine Reviews*, 37:85–93.

Maccone, C. (2009). A simple introduction to the KLT (Karhunen—Loève Transform). In *Deep Space Flight and Communications: Exploiting the Sun as a Gravitational Lens*, Springer Praxis Books, pages 151–179. Springer, Berlin, Heidelberg.

Magnin, M., Rey, M., Bastuji, H., Guillemant, P., Mauguière, F., and Garcia-Larrea, L. (2010). Thalamic deactivation at sleep onset precedes that of the cerebral cortex in humans. *Proceedings of the National Academy of Sciences of the United States of America*, 107(8):3829–3833.

Mahmoodian, N., Boese, A., Friebe, M., and Haddadnia, J. (2019). Epileptic seizure detection using cross-bispectrum of electroencephalogram signal. *Seizure*, 66:4–11.

Manconi, M., Ferri, R., Sagrada, C., Punjabi, N. M., Tettamanzi, E., Zucconi, M., Oldani, A., Castronovo, V., and Ferini-Strambi, L. (2010). Measuring the error in sleep estimation in normal subjects and in patients with insomnia. *Journal of Sleep Research*, 19(3):478–486. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-2869.2009.00801.x.

Mander, B. A., Marks, S. M., Vogel, J. W., Rao, V., Lu, B., Saletin, J. M., Ancoli-Israel, S., Jagust, W. J., and Walker, M. P. (2015). -amyloid disrupts human NREM slow waves and related hippocampus-dependent memory consolidation. *Nature Neuroscience*, 18(7):1051–1057. Number: 7 Publisher: Nature Publishing Group.

Mander, B. A., Rao, V., Lu, B., Saletin, J. M., Ancoli-Israel, S., Jagust, W. J., and Walker, M. P. (2014). Impaired Prefrontal Sleep Spindle Regulation of Hippocampal-Dependent Learning in Older Adults. *Cerebral Cortex*, 24(12):3301–3309.

Mander, B. A., Rao, V., Lu, B., Saletin, J. M., Lindquist, J. R., Ancoli-Israel, S., Jagust, W., and Walker, M. P. (2013). Prefrontal atrophy, disrupted NREM slow waves and impaired hippocampal-dependent memory in aging. *Nature Neuroscience*, 16(3):357–364. Number: 3 Publisher: Nature Publishing Group.

Mander, B. A., Winer, J. R., Jagust, W. J., and Walker, M. P. (2016). Sleep: A Novel Mechanistic Pathway, Biomarker, and Treatment Target in the Pathology of Alzheimer's Disease? *Trends in Neurosciences*, 39(8):552–566. Publisher: Elsevier Ltd.

Mander, B. A., Winer, J. R., and Walker, M. P. (2017). Sleep and Human Aging. *Neuron*, 94(1):19–36. Publisher: Elsevier.

Maquet, P. (2000). Functional neuroimaging using positron emission tomography has recently yielded original data on the functional neuroanatomy of human sleep. This paper attempts to describe the possibilities and limitations of the technique and clarify its usefulness in s. *J Sleep Res.*, 9:207 – 231.

Marshall, L., Helgadóttir, H., Mölle, M., and Born, J. (2006). Boosting slow oscillations during sleep potentiates memory. *Nature*, 444(7119):610–613.

Martin, J. L. and Hakim, A. D. (2011). Wrist actigraphy. *Chest*, 139(6):1514–1527. Publisher: The American College of Chest Physicians.

Marzano, C., Ferrara, M., Sforza, E., and De Gennaro, L. (2008). Quantitative electroencephalogram (EEG) in insomnia: a new window on pathophysiological mechanisms. *Current pharmaceutical design*, 14(32):3446–3455. Publisher: Bentham Science Publishers.

Marzano, C., Fratello, F., Moroni, F., Pellicciari, M. C., Curcio, G., Ferrara, M., Ferlazzo, F., and De Gennaro, L. (2007). Slow eye movements and subjective estimates of sleepiness predict EEG power changes during sleep deprivation. *Sleep*, 30(5):610–616.

Marzano, C., Moroni, F., Gorgoni, M., Nobili, L., Ferrara, M., and De Gennaro, L. (2013). How we fall asleep: regional and temporal differences in electroencephalographic synchronization at sleep onset. *Sleep medicine*, 14(11):1112–1122. Publisher: Elsevier.

Maturana, M. I., Meisel, C., Dell, K., Karoly, P. J., D' Souza, W., Grayden, D. B., Burkitt, A. N., Jiruska, P., Kudlacek, J., Hlinka, J., Cook, M. J., Kuhlmann, L., and Freestone, D. R. (2020). Critical slowing down as a biomarker for seizure susceptibility. *Nature Communications*, 11(1). Publisher: Nature Research.

May, R. M. (1977). Thresholds and breakpoints in ecosystems with a multiplicity of stable states. *Nature*, 269(5628):471–477. Number: 5628 Publisher: Nature Publishing Group.

May, R. M., Levin, S. A., and Sugihara, G. (2008). Ecology for bankers. *Nature*, 451(7181):893–894. Number: 7181 Publisher: Nature Publishing Group.

McCormick, D. A. (1989). Cholinergic and noradrenergic modulation of thalamocortical processing. *Trends in Neurosciences*, 12(6):215–221. Publisher: Elsevier.

McSharry, P. E., Smith, L. A., Tarassenko, L., Martinerie, J., Le Van Quyen, M., Baulac, M., and Renault, B. (2003). Prediction of epileptic seizures: Are nonlinear methods relevant? [1] (multiple letters). *Nature Medicine*, 9(3):241–242.

Mediano, P. A. M., Rosas, F. E., Luppi, A. I., Noreika, V., Seth, A. K., Carhart-Harris, R. L., Barnett, L., and Bor, D. (2023). Spectrally and temporally resolved estimation of neural signal diversity. Pages: 2023.03.30.534922 Section: New Results.

Mednick, S. C., McDevitt, E. A., Walsh, J. K., Wamsley, E., Paulus, M., Kanady, J. C., and Drummond, S. P. A. (2013). The Critical Role of Sleep Spindles in Hippocampal-Dependent Memory: A Pharmacology Study. *The Journal of Neuroscience*, 33(10):4494 LP–4504.

Melia, U., Guaita, M., Vallverdú, M., Embid, C., Vilaseca, I., Salamero, M., and Santamaria, J. (2015). Mutual information measures applied to EEG signals for sleepiness characterization. *Medical Engineering & Physics*, 37(3):297–308.

Merica, H. and Fortune, R. D. (2003). A unique pattern of sleep structure is found to be identical at all cortical sites: A neurobiological interpretation. *Cerebral Cortex*, 13(10):1044–1050.

Merica, H. and Fortune, R. D. (2004). State transitions between wake and sleep, and within the ultradian cycle, with focus on the link to neuronal activity. *Sleep Medicine Reviews*, 8(6):473–485.

Messineo, L., Taranto-Montemurro, L., Sands, S. A., Oliveira Marques, M. D., Azabarzin, A., and Wellman, D. A. (2017). Broadband Sound Administration Improves Sleep Onset Latency in Healthy Subjects in a Model of Transient Insomnia. *Frontiers in Neurology*, 8.

Micheloyannis, S., Flitzanis, N., Papanikolaou, E., Bourkas, M., Terzakis, D., Arvanitis, S., and Stam, C. J. (1998). Usefulness of non-linear EEG analysis. *Acta Neurologica Scandinavica*, 97(1):13–19. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1600-0404.1998.tb00603.x.

Mollayeva, T., Thurairajah, P., Burton, K., Mollayeva, S., Shapiro, C. M., and Colantonio, A. (2016). The Pittsburgh sleep quality index as a screening tool for sleep dysfunction in clinical and non-clinical samples: A systematic review and meta-analysis. *Sleep Medicine Reviews*, 25:52–73.

Montgomery-Downs, H. E., Insana, S. P., and Miller, E. A. (2010). Effects of Two Types of Ambient Sound During Sleep. *Behavioral Sleep Medicine*, 8(1):40–47. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/15402000903425686.

Morikawa, T., Hayashi, M., and Hori, T. (1997). Auto power and coherence analysis of delta-theta band EEG during the waking-sleeping transition period. *Electroencephalography and Clinical Neurophysiology*, 103(6):633–641.

Morikawa, T., Hayashi, M., and Hori, T. (2002). Spatiotemporal Variations of Alpha and Sigma Band EEG in the Waking-Sleeping Transition Period. *Perceptual and Motor Skills*, 95(1):131–154. Publisher: SAGE Publications Inc.

Murphy, M., Huber, R., Esser, S., A. Riedner, B., Massimini, M., Ferrarelli, F., Felice Ghilardi, M., and Tononi, G. (2011). The Cortical Topography of Local Sleep. *Current Topics in Medicinal Chemistry*, 11(19):2438–2446.

Mölle, M., Eschenko, O., Gais, S., Sara, S. J., and Born, J. (2009). The influence of learning on sleep slow oscillations and associated spindles and ripples in humans and rats. *European Journal of Neuroscience*, 29(5):1071–1081.

Naifeh, K. H. and Kamiya, J. (1981). The nature of respiratory changes associated with sleep onset. *Sleep*, 4(1):49–59.

Naismith, S. L., Lewis, S. J. G., and Rogers, N. L. (2011). Chapter 2 - Sleep–wake changes and cognition in neurodegenerative disease. In Van Dongen, H. P. A. and Kerkhof, G. A., editors, *Progress in Brain Research*, volume 190 of *Human Sleep and Cognition Part II*, pages 21–52. Elsevier.

Naismith, S. L. and Mowszowski, L. (2018). Sleep disturbance in mild cognitive impairment: a systematic review of recent findings. *Current Opinion in Psychiatry*, 31(2):153.

Nasari, M., Ghezeljeh, T. N., and Haghani, H. (2018). Effects of Nature Sounds on Sleep Quality among Patients Hospitalized in Coronary Care Units: A Randomized Controlled Clinical Trial. *Nursing and Midwifery Studies*, 7(1):18.

Navarrete, M., Schneider, J., Ngo, H.-V. V., Valderrama, M., Casson, A. J., and Lewis, P. A. (2020). Examining the optimal timing for closed-loop auditory stimulation of slow-wave sleep in young and older adults. *Sleep*, 43(6):zsz315.

Ngo, H. V. V., Martinetz, T., Born, J., and Mölle, M. (2013). Auditory closed-loop stimulation of the sleep slow oscillation enhances memory. *Neuron*, 78(3):545–553.

Ngo, H. V. V., Miedema, A., Faude, I., Martinetz, T., Mölle, M., and Born, J. (2015). Driving sleep slow oscillations by auditory closed-loop stimulation—A self-limiting process. *Journal of Neuroscience*, 35(17):6630–6638.

Nigbur, R., Ivanova, G., and Stürmer, B. (2011). Theta power as a marker for cognitive interference. *Clinical Neurophysiology*, 122(11):2185–2194.

Nir, Y., Staba, R. J., Andrillon, T., Vyazovskiy, V. V., Cirelli, C., Fried, I., and Tononi, G. (2011). Regional Slow Waves and Spindles in Human Sleep. *Neuron*, 70(1):153–169. Publisher: Elsevier Inc.

Niso, G., Bruña, R., Pereda, E., Gutiérrez, R., Bajo, R., Maestú, F., and del Pozo, F. (2013). HERMES: Towards an Integrated Toolbox to Characterize Functional and Effective Brain Connectivity. *Neuroinformatics*, 11(4):405–434.

Nittono, H., Momose, D., and Hori, T. (2001). The vanishing point of the mismatch negativity at sleep onset. *Clinical Neurophysiology*, 112(5):732–739. ISBN: 8182424658.

Nobili, L., Ferrara, M., Moroni, F., De Gennaro, L., Russo, G. L., Campus, C., Cardinale, F., and De Carli, F. (2011). Dissociated wake-like and sleep-like electro-cortical activity during sleep. *NeuroImage*, 58(2):612–619.

Noy-Meir, I. (1975). Stability of Grazing Systems: An Application of Predator-Prey Graphs. *Journal of Ecology*, 63(2):459–481. Publisher: [Wiley, British Ecological Society].

Ogilvie, R. D. (2001). The process of falling asleep. *Sleep Medicine Reviews*, 5(3):247–270.

Ogilvie, R. D., Simons, I. A., Kuderian, R. H., MacDonald, T., and Rustenburg, J. (1991). Behavioral, Event-Related Potential, and EEG/FFT Changes at Sleep Onset. *Psychophysiology*, 28(1):54–64. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1469-8986.1991.tb03386.x.

Ogilvie, R. D. and Wilkinson, R. T. (1984). The Detection of Sleep Onset: Behavioral and Physiological Convergence. *Psychophysiology*, 21(5):510–520. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1469-8986.1984.tb00234.x.

Ogilvie, R. D., Wilkinson, R. T., and Allison, S. (1989). The Detection of Sleep Onset:Behavioral, Physiological, and Subjective Convergence. *Sleep*, 12(5):458–474.

Okamoto-mizuno, K., Yamashiro, Y., Tanaka, H., Komada, Y., Mizuno, K., Tamaki, M., Kitado, M., Inoue, Y., and Shirakawa, S. (2008). Heart rate variability and body temperature during the sleep onset period. *Sleep and Biological Rhythms*, 6(1):42–49.

Olejniczak, P. (2006). Neurophysiologic Basis of EEG. *Journal of Clinical Neurophysiology*, 23(3):186.

Oleksenko, A., Mukhametov, L., Polyakova, I., Supin, A. Y., and Kovalzon, V. (1992). Unihemispheric sleep deprivation in bottlenose dolphins. *Journal of Sleep Research*, 1(1):40–44. Publisher: Wiley Online Library.

Omidvarnia, A., Mesbah, M., O'Toole, J. M., Colditz, P., and Boashash, B. (2011). Analysis of the time-varying cortical neural connectivity in the newborn EEG: A time-frequency approach. In *International Workshop on Systems, Signal Processing and their Applications, WOSSPA*, pages 179–182.

Ong, J. L., Lo, J. C., Chee, N. I. Y. N., Santostasi, G., Paller, K. A., Zee, P. C., and Chee, M. W. L. (2016). Effects of phase-locked acoustic stimulation during a nap on EEG spectra and declarative memory consolidation. *Sleep medicine*, 20:88–97. Publisher: Elsevier.

Oostenveld, R., Fries, P., Maris, E., and Schoffelen, J.-M. (2010). FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Computational Intelligence and Neuroscience*, 2011:e156869. Publisher: Hindawi.

Osorio, R. S., Gumb, T., Pirraglia, E., Varga, A. W., Lu, S.-e., Lim, J., Wohlleber, M. E., Ducca, E. L., Koushyk, V., Glodzik, L., Mosconi, L., Ayappa, I., Rapoport, D. M., Leon, M. J. d., and Initiative, F. t. A. D. N. (2015). Sleep-disordered breathing advances cognitive decline in the elderly. *Neurology*, 84(19):1964–1971. Publisher: Wolters Kluwer Health, Inc. on behalf of the American Academy of Neurology Section: Article.

Ott, E. (2002). Chaos in Dynamical Systems. *Chaos in Dynamical Systems*. Publisher: Cambridge University Press ISBN: 9780521811965.

Palmer, J. A., Kreutz-Delgado, K., and Makeig, S. (2012). AMICA: An adaptive mixture of independent component analyzers with shared components. *Swartz Center for Computatonal Neursoscience, University of California San Diego, Tech. Rep.*

Pan, J. and Tompkins, W. J. (1985). A Real-Time QRS Detection Algorithm. *IEEE Transactions on Biomedical Engineering*, BME-32(3):230–236. Conference Name: IEEE Transactions on Biomedical Engineering.

Park, C. H. (2013). A Feature Selection Method Using Hierarchical Clustering. In Prasath, R. and Kathirvalavakumar, T., editors, *Mining Intelligence and Knowledge Exploration*, Lecture Notes in Computer Science, pages 1–6, Cham. Springer International Publishing.

Parlitz, U. (1998). Nonlinear Time-Series Analysis. In Suykens, J. A. K. and Vandewalle, J., editors, *Nonlinear Modeling: Advanced Black-Box Techniques*, pages 209–239. Springer US, Boston, MA.

Peng, C. K., Havlin, S., Stanley, H. E., and Goldberger, A. L. (1995). Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *Chaos*, 5(1):82–87.

Perslev, M., Darkner, S., Kempfner, L., Nikolic, M., Jennum, P. J., and Igel, C. (2021). U-Sleep: resilient high-frequency sleep staging. *npj Digital Medicine*, 4(1):1–12. Publisher: Springer US.

Petersen, R. C. and Morris, J. C. (2005). Mild Cognitive Impairment as a Clinical Entity and Treatment Target. *Archives of Neurology*, 62(7):1160–1163.

Phillips, A. J. and Robinson, P. A. (2007). A quantitative model of sleep-wake dynamics based on the physiology of the brainstem ascending arousal system. *Journal of Biological Rhythms*, 22(2):167–179.

Pincus, S. M., Gladstone, I. M., and Ehrenkranz, R. A. (1991). A regularity statistic for medical data analysis. *Journal of Clinical Monitoring*, 7(4):335–345.

Pivik, R. T. and Harman, K. (1995). A reconceptualization of EEG alpha activity as an index of arousal during sleep: all alpha activity is not equal. *Journal of Sleep Research*, 4(3):131–137. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-2869.1995.tb00161.x.

Plenz, D., Ribeiro, T. L., Miller, S. R., Kells, P. A., Vakili, A., and Capek, E. L. (2021). Self-Organized Criticality in the Brain. *Frontiers in Physics*, 9(July):1–23. arXiv: 2102.09124.

Posner, J. B. (2007). *Plum and Posner's diagnosis of stupor and coma*, volume 71. OUP USA.

Prerau, M. J., Hartnack, K. E., Obregon-Henao, G., Sampson, A., Merlino, M., Gannon, K., Bianchi, M. T., Ellenbogen, J. M., and Purdon, P. L. (2014). Tracking the sleep onset process: an empirical model of behavioral and physiological dynamics. *PLoS computational biology*, 10(10). Publisher: Public Library of Science.

Pritchard, W. S. and Duke, D. W. (1992). Measuring chaos in the brain: A tutorial review of nonlinear dynamical eeg analysis. *International Journal of Neuroscience*, 67(1-4):31–80.

Pritchard, W. S. and Duke, D. W. (1995). Measuring Chaos in the Brain - A Tutorial Review of EEG Dimension Estimation. *Brain and Cognition*, 27(3):353–397.

Pritchard, W. S., Duke, D. W., and Coburn, K. L. (1991). Altered EEG Dynamical Responsivity Associated with Normal Aging and Probable Alzheimer's Disease. *Dementia*, 2(2):102–105.

Purcell, S. M., Manoach, D. S., Demanuele, C., Cade, B. E., Mariani, S., Cox, R., Panagiotaropoulou, G., Saxena, R., Pan, J. Q., Smoller, J. W., Redline, S., and Stickgold, R. (2017). Characterizing sleep spindles in 11,630 individuals from the National Sleep Research Resource. *Nature Communications*, 8(May):1–16. Publisher: Nature Publishing Group.

Qian, D. and Wong, J. (2022). Sleep disturbances, sleep-disordered breathing, and cognitive impairment in older adults: perioperative implications. *International Anesthesiology Clinics*, 60(2):20–26.

Raggi, A. and Ferri, R. (2010). Sleep disorders in neurodegenerative diseases. *European Journal of Neurology*, 17(11):1326–1338. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1468-1331.2010.03034.x.

Ramanand, P., Bruce, M. C., and Bruce, E. N. (2010). Mutual Information Analysis of EEG Signals Indicates Age-Related Changes in Cortical Interdependence during Sleep in Middle-aged vs. Elderly Women. *Journal of clinical neurophysiology : official publication of the American Electroencephalographic Society*, 27(4):274–284.

Rasch, B. and Born, J. (2013). About sleep's role in memory. *Physiological reviews*. Publisher: American Physiological Society Bethesda, MD.

Raymann, R. J., Swaab, D. F., and Van Someren, E. J. (2005). Cutaneous warming promotes sleep onset. *American Journal of Physiology - Regulatory Integrative and Comparative Physiology*, 288(6 57-6):1589–1597.

Raymann, R. J., Swaab, D. F., and Van Someren, E. J. (2008). Skin deep: Enhanced sleep depth by cutaneous temperature manipulation. *Brain*, 131(2):500–513.

Rechtschaffen, A. and Kales, A. (1968). A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects. 1968. Washington, DC. *Public Health Service, US Government, Printing Office*.

Rentz, D. M., Amariglio, R. E., Becker, J. A., Frey, M., Olson, L. E., Frishe, K., Carmasin, J., Maye, J. E., Johnson, K. A., and Sperling, R. A. (2011). Face-name associative memory performance is related to amyloid burden in normal elderly. *Neuropsychologia*, 49(9):2776–2783.

Rezaie, L., Fobian, A. D., McCall, W. V., and Khazaie, H. (2018). Paradoxical insomnia and subjective–objective sleep discrepancy: A review. *Sleep Medicine Reviews*, 40:196–202. Publisher: W.B. Saunders Ltd.

Robinson, P. A., Phillips, A. J., Fulcher, B. D., Puckeridge, M., and Roberts, J. A. (2011). Quantitative modelling of sleep dynamics. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 369(1952):3840–3854.

Roh, J. H., Huang, Y., Bero, A. W., Kasten, T., Stewart, F. R., Bateman, R. J., and Holtzman, D. M. (2012). Disruption of the Sleep-Wake Cycle and Diurnal Fluctuation of -Amyloid in Mice with Alzheimer' s Disease Pathology. *Science Translational Medicine*, 4(150):150ra122–150ra122. Publisher: American Association for the Advancement of Science.

Rosas, F. E., Mediano, P. A. M., Gastpar, M., and Jensen, H. J. (2019). Quantifying high-order interdependencies via multivariate extensions of the mutual information. *Physical Review E*, 100(3):032305. Publisher: American Physical Society.

Rosenberg, R. S. and Van Hout, S. (2013). The American Academy of Sleep Medicine interscorer reliability program: sleep stage scoring. *Journal of clinical sleep medicine*, 9(1):81–87. Publisher: American Academy of Sleep Medicine.

Roth, T. (2007). Insomnia: Definition, Prevalence, Etiology, and Consequences. *Journal of Clinical Sleep Medicine*, 3(5 suppl):S7–S10. Publisher: American Academy of Sleep Medicine.

Roth, T. and Roehrs, T. (2003). Insomnia: Epidemiology, characteristics, and consequences. *Clinical Cornerstone*, 5(3):5–15.

Rudy, S. H., Brunton, S. L., Proctor, J. L., and Kutz, J. N. (2017). Data-driven discovery of partial differential equations. *Science Advances*, 3(4):1–7. arXiv: 1609.06401.

Rué-Queralt, J., Stevner, A., Tagliazucchi, E., Laufs, H., Kringelbach, M. L., Deco, G., and Atasoy, S. (2021). Decoding brain states on the intrinsic manifold of human brain dynamics across wakefulness and sleep. *Communications Biology*, 4(1):1–11. Publisher: Springer US.

Sakkalis, V. (2011). Review of advanced techniques for the estimation of brain connectivity measured with EEG/MEG. *Computers in Biology and Medicine*, 41(12):1110–1117.

Samiee, S. and Baillet, S. (2017). Time-resolved phase-amplitude coupling in neural oscillations. *NeuroImage*, 159(July):270–279. Publisher: Elsevier Ltd.

Sanders, T. H., McCurry, M., and Clements, M. A. (2014). Sleep stage classification with cross frequency coupling. In *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 4579–4582. ISSN: 1558-4615.

Sanjari, N., Shalbaf, A., Shalbaf, R., and Sleigh, J. (2021). Assessment of Anesthesia Depth Using Effective Brain Connectivity Based on Transfer Entropy on EEG Signal. *Basic and Clinical Neuroscience*, 12(2):269–280.

Saper, C. B., Chou, T. C., and Scammell, T. E. (2001). The sleep switch: hypothalamic control of sleep and wakefulness. *Trends in Neurosciences*, 24(12):726–731. Publisher: Elsevier.

Saper, C. B., Fuller, P. M., Pedersen, N. P., Lu, J., and Scammell, T. E. (2010). Sleep State Switching. *Neuron*, 68(6):1023–1042. Publisher: Elsevier Inc.

Saper, C. B., Scammell, T. E., and Lu, J. (2005). Hypothalamic regulation of sleep and circadian rhythms. *Nature*, 437(7063):1257–1263.

Sarasso, S., Pigorini, A., Proserpio, P., Gibbs, S. A., Massimini, M., and Nobili, L. (2014). Fluid boundaries between wake and sleep: experimental evidence from Stereo-EEG recordings. *Archives italiennes de biologie*, 152(2-3):169–177.

Sazgar, M. and Young, M. G. (2019). Overview of EEG, Electrode Placement, and Montages. In Sazgar, M. and Young, M. G., editors, *Absolute Epilepsy and EEG Rotation Review: Essentials for Trainees*, pages 117–125. Springer International Publishing, Cham.

Scammell, T. E., Gerashchenko, D. Y., Mochizuki, T., McCarthy, M. T., Estabrooke, I. V., Sears, C. A., Saper, C. B., Urade, Y., and Hayaishi, O. (2001). An adenosine A2a agonist increases sleep and induces Fos in ventrolateral preoptic neurons. *Neuroscience*, 107(4):653–663.

Scheffer, M., Bascompte, J., Brock, W. A., Brovkin, V., Carpenter, S. R., Dakos, V., Held, H., Van Nes, E. H., Rietkerk, M., and Sugihara, G. (2009). Early-warning signals for critical transitions. *Nature*, 461(7260):53–59. Publisher: Nature Publishing Group.

Scheffer, M., Carpenter, S. R., Lenton, T. M., Bascompte, J., Brock, W., Dakos, V., Van De Koppel, J., Van De Leemput, I. A., Levin, S. A., Van Nes, E. H., Pascual, M., and Vandermeer, J. (2012). Anticipating critical transitions. *Science*, 338(6105):344–348.

Schelter, B., Timmer, J., and Eichler, M. (2009). Assessing the strength of directed influences among neural signals using renormalized partial directed coherence. *Journal of Neuroscience Methods*, 179(1):121–130.

Schmidt, C., Peigneux, P., Muto, V., Schenkel, M., Knoblauch, V., Münch, M., Dominique, J.-F., Wirz-Justice, A., and Cajochen, C. (2006). Encoding difficulty promotes postlearning changes in sleep spindle activity during napping. *Journal of Neuroscience*, 26(35):8976–8982. Publisher: Soc Neuroscience.

Schmidt, M. and Lipson, H. (2009). Distilling Natural Laws. *Science*, 324(April):81–85.

Schneider, J., Lewis, P. A., Koester, D., Born, J., and Ngo, H.-V. V. (2020). Susceptibility to auditory closed-loop stimulation of sleep slow oscillations changes with age. *Sleep*, 43(12):zsaa111.

Schoch, S. F., Werth, E., Poryazova, R., Scammell, T. E., Baumann, C. R., and Imbach, L. L. (2017). Dysregulation of Sleep Behavioral States in Narcolepsy. *Sleep*, 40(12):zsx170.

Schomer, D. L. and Lopes da Silva, F. H., editors (2017). *Niedermeyer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Oxford University Press.

Schreglmann, S. R., Wang, D., Peach, R. L., Li, J., Zhang, X., Latorre, A., Rhodes, E., Panella, E., Cassara, A. M., Boyden, E. S., Barahona, M., Santaniello, S., Rothwell, J., Bhatia, K. P., and Grossman, N. (2021). Non-invasive suppression of essential tremor via phase-locked disruption of its temporal coherence. *Nature Communications*, 12(1). Publisher: Springer US ISBN: 4146702020581.

Schreiber, T. (2000). Measuring Information Transfer. *Physical Review Letters*, 85(2):461–464. Publisher: American Physical Society.

Sedgwick, P. (2012). Multiple significance tests: the Bonferroni correction. *BMJ*, 344:e509. Publisher: British Medical Journal Publishing Group Section: Endgames.

Seneviratne, U. and Puvanendran, K. (2004). Excessive daytime sleepiness in obstructive sleep apnea: prevalence, severity, and predictors. *Sleep Medicine*, 5(4):339–343.

Shahid, A., Wilkinson, K., Marcu, S., and Shapiro, C. M. (2012). Karolinska Sleepiness Scale (KSS). In Shahid, A., Wilkinson, K., Marcu, S., and Shapiro, C. M., editors, *STOP, THAT and One Hundred Other Sleep Scales*, pages 209–210. Springer, New York, NY.

Shen, Y., Olbrich, E., Achermann, P., and Meier, P. F. (2003). Dimensional complexity and spectral properties of the human sleep EEG. *Clinical Neurophysiology*, 114(2):199–209. Publisher: Elsevier.

Sherin, J. E., Elmquist, J. K., Torrealba, F., and Saper, C. B. (1998). Innervation of Histaminergic Tuberomammillary Neurons by GABAergic and Galaninergic Neurons in the Ventrolateral Preoptic Nucleus of the Rat. *Journal of Neuroscience*, 18(12):4705–4721. Publisher: Society for Neuroscience Section: ARTICLE.

Shi, L., Chen, S.-J., Ma, M.-Y., Bao, Y.-P., Han, Y., Wang, Y.-M., Shi, J., Vitiello, M. V., and Lu, L. (2018). Sleep disturbances increase the risk of dementia: a systematic review and meta-analysis. *Sleep medicine reviews*, 40:4–16. Publisher: Elsevier.

Shih, H.-I., Lin, C.-C., Tu, Y.-F., Chang, C.-M., Hsu, H.-C., Chi, C.-H., and Kao, C.-H. (2015). An increased risk of reversible dementia may occur after zolpidem derivative use in the elderly population a population-based case-control study. *Medicine (United States)*, 94(17):e809.

Shin, D., Sakai, H., and Uchiyama, Y. (2011). Slow eye movement detection can prevent sleep-related accidents effectively in a simulated driving task. *Journal of Sleep Research*, 20(3):416–424.

Shinar, Z., Akselrod, S., Dagan, Y., and Baharav, A. (2006). Autonomic changes during wake-sleep transition: A heart rate variability based approach. *Autonomic Neuroscience: Basic and Clinical*, 130(1-2):17–27.

Siclari, F., Bernardi, G., Riedner, B. A., LaRocque, J. J., Benca, R. M., and Tononi, G. (2014). Two distinct synchronization processes in the transition to sleep: a high-density electroencephalographic study. *Sleep*, 37(10):1621–1637. Publisher: Oxford University Press.

Silber, M. H., Ancoli-Israel, S., Bonnet, M. H., Chokroverty, S., Grigg-Damberger, M. M., Hirshkowitz, M., Kapen, S., Keenan, S. A., Kryger, M. H., Penzel, T., Pressman, M. R., and Iber, C. (2007). The visual scoring of sleep in adults. *Journal of Clinical Sleep Medicine*, 3(2):121–131.

Simor, P., van der Wijk, G., Nobili, L., and Peigneux, P. (2020). The microstructure of REM sleep: Why phasic and tonic? *Sleep Medicine Reviews*, 52.

Slipher, S., Williams, S., Hopp, J., Kaplan, P., and Gordon, B. (2014). Alpha plus Beta/Delta plus Theta Ratio is a better measure of consciousness than Alpha/Delta Ratio in the acute care setting: Preliminary results. volume 76, Baltimore, Maryland.

Somaskandhan, P., Terrill, P., Korkalainen, H., Kainulainen, S., Leppänen, T., Islind, A., Grétarsdóttir, H., and Nikkonen, S. (2022). P113 A detailed analysis of multicentric sleep staging inter-rater variabilities. *Sleep Advances: A Journal of the Sleep Research Society*, 3(Suppl 1):A66.

Sousouri, G., Krugliakova, E., Skorucak, J., Leach, S., Snipes, S., Ferster, M. L., Da Poian, G., Karlen, W., and Huber, R. (2022). Neuromodulation by means of phase-locked auditory stimulation affects key marker of excitability and connectivity during sleep. *Sleep*, 45(1):zsab204.

Spiess, M., Bernardi, G., Kurth, S., Ringli, M., Wehrle, F. M., Jenni, O. G., Huber, R., and Siclari, F. (2018). How do children fall asleep? A high-density EEG study of slow waves in the transition from wake to sleep. *NeuroImage*, 178(May):23–35. Publisher: Elsevier Inc.

Srinivasan, R., Winter, W. R., and Nunez, P. L. (2006). Source analysis of EEG oscillations using high-resolution EEG and MEG. In Neuper, C. and Klimesch, W., editors, *Progress in Brain Research*, volume 159 of *Event-Related Dynamics of Brain Oscillations*, pages 29–42. Elsevier.

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

Stam, C. J. (2014). Modern network science of neurological disorders. *Nature Reviews Neuroscience*, 15(10):683–695. Publisher: Nature Publishing Group.

Stam, C. J. and de Bruin, E. A. (2004). Scale-free dynamics of global functional connectivity in the human brain. *Human Brain Mapping*, 22(2):97–109.

Stam, C. J., Jelles, B., Achtereekte, H. A. M., van Birgelen, J. H., and Slaets, J. P. J. (1996). Diagnostic Usefulness of Linear and Nonlinear Quantitative EEG Analysis in Alzheimer's Disease. *Clinical Electroencephalography*, 27(2):69–77. Publisher: SAGE Publications.

Stam, C. J. and van Dijk, B. W. (2002). Synchronization likelihood: an unbiased measure of generalized synchronization in multivariate data sets. *Physica D: Nonlinear Phenomena*, 163(3):236–251.

Staresina, B. P., Bergmann, T. O., Bonnefond, M., Van Der Meij, R., Jensen, O., Deuker, L., Elger, C. E., Axmacher, N., and Fell, J. (2015). Hierarchical nesting of slow oscillations, spindles and ripples in the human hippocampus during sleep. *Nature Neuroscience*, 18(11):1679–1686. Publisher: Nature Publishing Group.

Stevner, A. B., Vidaurre, D., Cabral, J., Rapuano, K., Nielsen, S. F., Tagliazucchi, E., Laufs, H., Vuust, P., Deco, G., Woolrich, M. W., Van Someren, E., and Kringelbach, M. L. (2019). Discovery of key whole-brain transitions and dynamics during human wakefulness and non-REM sleep. *Nature Communications*, 10(1). Publisher: Springer US ISBN: 4146701908934.

Stewart, E. M., Landry, S., Edwards, B. A., and Drummond, S. P. A. (2020). The Bidirectional Relationship Between Sleep and Health. In *The Wiley Encyclopedia of Health Psychology*, pages 165–188. John Wiley & Sons, Ltd. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/9781119057840.ch200.

Strauss, M., Sitt, J. D., Naccache, L., and Raimondo, F. (2022). Predicting the loss of responsiveness when falling asleep in humans. *NeuroImage*, 251. Publisher: Academic Press Inc.

Stålesen Ramfjord, L., Hertenstein, E., Fehér, K., Mikutta, C., Schneider, C. L., Nissen, C., and Maier, J. G. (2020). Local sleep and wakefulness—the concept and its potential for the understanding and treatment of insomnia disorder. *Somnologie*, 24(2):116–120.

Sulaman, B. A., Wang, S., Tyan, J., and Eban-Rothschild, A. (2023). Neuro-orchestration of sleep and wakefulness. *Nature Neuroscience*, 26(2):196–212. Number: 2 Publisher: Nature Publishing Group.

Swami, A., Mendel, J. M., and Nikias, C. L. (1998). Higher-order spectral analysis toolbox. *The Mathworks Inc*, 3:22–26.

Takahashi, K., Kayama, Y., Lin, J. S., and Sakai, K. (2010). Locus coeruleus neuronal activity during the sleep-waking cycle in mice. *Neuroscience*, 169(3):1115–1126. Publisher: Elsevier Inc.

Tamaki, M., Nittono, H., Hayashi, M., and Hori, T. (2005). Examination of the first-night effect during the sleep-onset period. *Sleep*, 28(2):195–202.

Tanaka, H., Hayashi, M., and Hori, T. (1997). Topographical characteristics and principal component structure of the hypnagogic EEG. *Sleep*, 20(7):523–534.

Tanaka, H., Hayashi, M., and Hori, T. (1999). Topographic mapping of EEG spectral power and coherence in delta activity during the transition from wakefulness to sleep. *Psychiatry and Clinical Neurosciences*, 53(2):155–157. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1046/j.1440-1819.1999.00509.x.

Tanaka, H., Hayashi, M., and Hori, T. (2000). Topographical characteristics of slow wave activities during the transition from wakefulness to sleep. *Clinical Neurophysiology*, 111(3):417–427. ISBN: 8182424658.

Tassi, P., Bonnefond, A., Engasser, O., Hoeft, A., Eschenlauer, R., and Muzet, A. (2006). EEG spectral power and cognitive performance during sleep inertia: The effect of normal sleep duration and partial sleep deprivation. *Physiology & Behavior*, 87(1):177–184.

Thompson, J. M. T. and Sieber, J. (2011). Predicting climate tipping as a noisy bifurcation: A review. *International Journal of Bifurcation and Chaos*, 21(2):399–423.

Toda, H. Y. and Yamamoto, T. (1995). Statistical inference in vector autoregressions with possibly integrated processes. *Journal of Econometrics*, 66(1):225–250.

Torsvall, L. and Åkerstedt, T. (1988). Extreme sleepiness: Quantification of EOG and spectral EEG parameters. *International Journal of Neuroscience*, 38(3-4):435–441.

Trinder, J., Kleiman, J., Carrington, M., Smith, S., Breen, S., Tan, N., and Kim, Y. (2001). Autonomic activity during human sleep as a function of time and sleep stage. *Journal of Sleep Research*, 10(4):253–264.

Tucker, M. A., Hirota, Y., Wamsley, E. J., Lau, H., Chaklader, A., and Fishbein, W. (2006). A daytime nap containing solely non-REM sleep enhances declarative but not procedural memory. *Neurobiology of learning and memory*, 86(2):241–247. Publisher: Elsevier.

Urigüen, J. A. and Garcia-Zapirain, B. (2015). EEG artifact removal - State-of-the-art and guidelines. *Journal of Neural Engineering*, 12(3). Publisher: IOP Publishing.

Valderrama, M., Crépon, B., Botella-Soler, V., Martinerie, J., Hasboun, D., Alvarado-Rojas, C., Baulac, M., Adam, C., Navarro, V., and Le Van Quyen, M. (2012). Human gamma oscillations during slow wave sleep. *PLoS ONE*, 7(4):1–14.

van Nes, E. H. and Scheffer, M. (2005). Implications of Spatial Heterogeneity for Catastrophic Regime Shifts in Ecosystems. *Ecology*, 86(7):1797–1807. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1890/04-0550.

Van Someren, E. J. (2000). More than a marker: Interaction between the circadian regulation of temperature and sleep, age-related changes, and treatment possibilities. *Chronobiology International*, 17(3):313–354.

Van Someren, E. J. (2004). Sleep propensity is modulated by circadian and behavior-induced changes in cutaneous temperature. *Journal of Thermal Biology*, 29(7-8 SPEC. ISS.):437–444.

Van Someren, E. J. W. (2021). Brain mechanisms of insomnia: new perspectives on causes and consequences. *Physiological Reviews*, 101(3):995–1046. Publisher: American Physiological Society.

Varley, T. F., Pope, M., Faskowitz, J., and Sporns, O. (2023). Multivariate information theory uncovers synergistic subsystems of the human cerebral cortex. *Communications Biology*, 6(1):1–12. Number: 1 Publisher: Nature Publishing Group.

Vecchio, F., Miraglia, F., Gorgoni, M., Ferrara, M., Iberite, F., Bramanti, P., De Gennaro, L., and Rossini, P. M. (2017). Cortical connectivity modulation during sleep onset: a study via graph theory on EEG data. *Human brain mapping*, 38(11):5456–5464. Publisher: Wiley Online Library.

Vinao-Carl, M. L., Gal-Shohet, Y. R., Rhodes, E., Li, J., Hampshire, A., Sharp, D. J., and Grossman, N. (2023). Just a phase? Causal probing reveals spurious phasic dependence of sensory perception. *bioRxiv*, pages 2023–08. Publisher: Cold Spring Harbor Laboratory.

Violante, I. R., Alania, K., Cassarà, A. M., Neufeld, E., Acerbo, E., Carron, R., Williamson, A., Kurtin, D. L., Rhodes, E., Hampshire, A., Kuster, N., Boyden, E. S., Pascual-Leone, A., and Grossman, N. (2023). Non-invasive temporal interference electrical stimulation of the human hippocampus. *Nature Neuroscience*, 26(11):1994–2004. Number: 11 Publisher: Nature Publishing Group.

Voytek, B., Kramer, M. A., Case, J., Lepage, K. Q., Tempesta, Z. R., Knight, R. T., and Gazzaley, A. (2015). Age-related changes in 1/f neural electrophysiological noise. *Journal of Neuroscience*, 35(38):13257–13265.

Vyazovskiy, V. V., Olcese, U., Hanlon, E. C., Nir, Y., Cirelli, C., and Tononi, G. (2011). Local sleep in awake rats. *Nature*, 472(7344):443–447. Number: 7344 Publisher: Nature Publishing Group.

Wang, J., Chen, L., and Fei, X. (2007a). Analysis and control of the bifurcation of Hodgkin–Huxley model. *Chaos, Solitons & Fractals*, 31(1):247–256.

Wang, J. L., Chiou, J. M., and Müller, H. G. (2016). Functional Data Analysis. *Annual Review of Statistics and Its Application*, 3:257–295.

Wang, X., Wirth, A., and Wang, L. (2007b). Structure-Based Statistical Features and Multivariate Time Series Clustering. In *Seventh IEEE International Conference on Data Mining (ICDM 2007)*, pages 351–360. ISSN: 2374-8486.

Wei, Y., Colombo, M. A., Ramautar, J. R., Blanken, T. F., van der Werf, Y. D., Spiegelhalder, K., Feige, B., Riemann, D., and Van Someren, E. J. W. (2017). Sleep Stage Transition Dynamics Reveal Specific Stage 2 Vulnerability in Insomnia. *Sleep*, 40(9):zsx117.

Weigenand, A., Mölle, M., Werner, F., Martinetz, T., and Marshall, L. (2016). Timing matters: open-loop stimulation does not improve overnight consolidation of word pairs in humans. *European Journal of Neuroscience*, 44(6):2357–2368. Publisher: Wiley Online Library.

Westerberg, C. E., Florczak, S. M., Weintraub, S., Mesulam, M.-M., Marshall, L., Zee, P. C., and Paller, K. A. (2015). Memory improvement via slow-oscillatory stimulation during sleep in older adults. *Neurobiology of aging*, 36(9):2577–2586. Publisher: Elsevier.

Wilkat, T., Rings, T., and Lehnertz, K. (2019). No evidence for critical slowing down prior to human epileptic seizures. *Chaos*, 29(9). arXiv: 1908.08973 Publisher: AIP Publishing LLC.

Wisden, W., Yu, X., and Franks, N. P. (2019). GABA Receptors and the Pharmacology of Sleep. In Landolt, H.-P. and Dijk, D.-J., editors, *Sleep-Wake Neurobiology and Pharmacology*, Handbook of Experimental Pharmacology, pages 279–304. Springer International Publishing, Cham.

Worsnop, C., Kay, A., Pierce, R., Kim, Y., and Trinder, J. (1998). Activity of respiratory pump and upper airway muscles during sleep onset. *Journal of Applied Physiology*, 85(3):908–920.

Wright, K. P. (2009). EEG in Sleep States. In Binder, M. D., Hirokawa, N., and Windhorst, U., editors, *Encyclopedia of Neuroscience*, pages 1031–1031. Springer, Berlin, Heidelberg.

Wright, K. P., Badia, P., and Wauquier, A. (1995). Topographical and temporal patterns of brain activity during the transition from wakefulness to sleep. *Sleep*, 18(10):880–889.

Xie, L., Kang, H., Xu, Q., Chen, M. J., Liao, Y., Thiyagarajan, M., O' Donnell, J., Christensen, D. J., Nicholson, C., Iliff, J. J., Takano, T., Deane, R., and Nedergaard, M. (2013). Sleep Drives Metabolite Clearance from the Adult Brain. *Science*, 342(6156):373–377. Publisher: American Association for the Advancement of Science.

Yamagata, T., Kahn, M. C., Prius-Mengual, J., Meijer, E., Šabanović, M., Guillaumin, M. C. C., van der Vinne, V., Huang, Y.-G., McKillop, L. E., Jagannath, A., Peirson, S. N., Mann, E. O., Foster, R. G., and Vyazovskiy, V. V. (2021). The hypothalamic link between arousal and sleep homeostasis in mice. *Proceedings of the National Academy of Sciences*, 118(51):e2101580118. Publisher: Proceedings of the National Academy of Sciences.

Yang, C. M., Han, H. Y., Yang, M. H., Su, W. C., and Lane, T. (2010). What subjective experiences determine the perception of falling asleep during sleep onset period? *Consciousness and Cognition*, 19(4):1084–1092. Publisher: Academic Press.

Yang, D. P., McKenzie-Sell, L., Karanjai, A., and Robinson, P. A. (2016). Wake-sleep transition as a noisy bifurcation. *Physical Review E*, 94(2):1–10.

Yankova, G., Bogomyakova, O., and Tulupov, A. (2021). The glymphatic system and meningeal lymphatics of the brain: new understanding of brain clearance. *Reviews in the Neurosciences*, 32(7):693–705. Publisher: De Gruyter.

Yoon, H. and Baek, H. J. (2022). External Auditory Stimulation as a Non-Pharmacological Sleep Aid. *Sensors*, 22(3):1264. Number: 3 Publisher: Multidisciplinary Digital Publishing Institute.

Yuan, Q., Zhou, W., Liu, Y., and Wang, J. (2012). Epileptic seizure detection with linear and nonlinear features. *Epilepsy & Behavior*, 24(4):415–421.

Zawiślak-Fornagiel, K., Ledwoń, D., Bugdol, M., Romaniszyn-Kania, P., Małecki, A., Gorzkowska, A., and Mitas, A. W. (2023). The Increase of Theta Power and Decrease

of Alpha/Theta Ratio as a Manifestation of Cognitive Impairment in Parkinson' s Disease. *Journal of Clinical Medicine*, 12(4):1569.

Zeineh, M. M., Engel, S. A., Thompson, P. M., and Bookheimer, S. Y. (2003). Dynamics of the Hippocampus During Encoding and Retrieval of Face-Name Pairs. *Science*, 299(5606):577–580. Publisher: American Association for the Advancement of Science.

Zhang, G. Q., Cui, L., Mueller, R., Tao, S., Kim, M., Rueschman, M., Mariani, S., Mobley, D., and Redline, S. (2018). The National Sleep Research Resource: Towards a sleep data commons. *Journal of the American Medical Informatics Association*, 25(10):1351–1358.

Zisapel, N. (2007). Sleep and sleep disturbances: biological basis and clinical implications. *Cellular and Molecular Life Sciences*, 64(10):1174.

Åkerstedt, T. and Gillberg, M. (1990). Subjective and objective sleepiness in the active individual. *International Journal of Neuroscience*, 52(1-2):29–37.

Appendices

A. List of EEG time-series feature base used in the analysis

A.1. List of conventional EEG analysis features

Table A.1: List of conventional EEG analysis features with simple descriptions.

Names	***Descriptions***
Power-band features	
Delta power	The delta band (1 –4 Hz) power normalized to the entire spectrum power and log transformed
Theta power	The theta band (4 –8 Hz) power normalized to the entire spectrum power and log transformed
Alpha power	The alpha band (8 –12 Hz) power normalized to the entire spectrum power and log transformed
Beta power	The beta band (12 –30 Hz) power normalized to the entire spectrum power and log transformed
Delta to theta ratio	The log-transformed band power ratio between delta and theta band
Delta to alpha ratio	The log-transformed band power ratio between delta and alpha band
Delta to beta ratio	The log-transformed band power ratio between delta and beta band
Theta to alpha ratio	The log-transformed band power ratio between theta and alpha band
Theta to beta ratio	The log-transformed band power ratio between theta and beta band

Continued on next page

Table A.1: List of conventional EEG analysis features with simple descriptions. (Continued)

Alpha to beta ratio	The log-transformed band power ratio between delta and alpha band
Transition power ratio	The log-transformed transition power ratio defined as the low frequency band power (delta plus theta) divided by high frequency band power (alpha plus beta)
Peak delta band frequency	The peak frequency in the delta band (frequency bin with highest power)
Peak theta band frequency	The peak frequency in the theta band (frequency bin with highest power)
Peak alpha band frequency	The peak frequency in the alpha band (frequency bin with highest power)
Peak beta band frequency	The peak frequency in the beta band (frequency bin with highest power)
Temporal coherence features	
Temporal coherence delta band	The magnitude-squared temporal coherence in delta band
Temporal coherence theta band	The magnitude-squared temporal coherence in theta band
Temporal coherence alpha band	The magnitude-squared temporal coherence in alpha band
Temporal coherence beta band	The magnitude-squared temporal coherence in beta band
Phase-amplitude coupling features	
Phase amplitude coupling DT	The phase-amplitude coupling strength between delta band phase and theta band amplitude
Phase amplitude coupling DA	The phase-amplitude coupling strength between delta band phase and alpha band amplitude
Phase amplitude coupling DB	The phase-amplitude coupling strength between delta band phase and beta band amplitude
Phase amplitude coupling TA	The phase-amplitude coupling strength between theta band phase and alpha band amplitude

Continued on next page

Table A.1: List of conventional EEG analysis features with simple descriptions. (Continued)

Phase amplitude coupling TB	The phase-amplitude coupling strength between theta band phase and beta band amplitude
Phase amplitude coupling AB	The phase-amplitude coupling strength between alpha band phase and beta band amplitude

A.2. List of CATCH-22 highly comparative time-series features (Reused from Lubba et al. (2019))

Table A.2: List of Catch-22 time-series features. Table reused from Lubba et al. (2019) under Creative Commons Attribution 4.0 International License.

Names	***Descriptions***
Distribution	
DN HistogramMode 5	Mode of z-scored distribution (5-bin histogram)
DN HistogramMode 10	Mode of z-scored distribution (10-bin histogram)
Simple temporal statistics	
SB BinaryStats mean longstretch1	Longest period of consecutive values above the mean
DN OutlierInclude p 001 mdrmd	Time intervals between successive extreme events above the mean
DN OutlierInclude n 001 mdrmd	Time intervals between successive extreme events below the mean
Linear autocorrelation	
CO f1ecac	First crossing of autocorrelation function
CO FirstMin ac	First minimum of autocorrelation function
SP Summaries welch rect area 5 1	Total power in lowest 1/5 of frequencies in the Fourier power spectrum
SP Summaries welch rect centroid	Centroid of the Fourier power spectrum
FC LocalSimple mean3 stderr	Mean error from a rolling 3-sample mean forecasting
Nonlinear autocorrelation	
CO trev 1 num	Time-reversibility statistic, $\langle (x_{t+1} - x_t)^3 \rangle$
CO HistogramAMI even 2 5	Automutual information, $m = 2$, $\tau = 5$

Continued on next page

Table A.2: List of Catch-22 time-series features. Table reused from Lubba et al. (2019) under Creative Commons Attribution 4.0 International License. (Continued)

IN AutoMutualInfoStats 40 gaussian fmmi	First minimum of the automutual information function
Successive differences	
MD hrv classic pnn40	Proportion of successive differences exceeding 0.04σ
SB BinaryStats diff longstretch0	Longest period of successive incremental decreases
SB MotifThree quantile hh	Shannon entropy of two successive letters in equiprobable 3-letter symbolization
FC LocalSimple mean1 tauresrat	Change in correlation length after iterative differencing
CO Embed2 Dist tau d expfit meandiff	Exponential fit to successive distances in 2-d embedding space
Fluctuation Analysis	
SC FluctAnal 2 dfa 50 1 2 logi prop r1	Proportion of slower timescale fluctuations that scale with DFA (50% sampling)
SC FluctAnal 2 rsrangefit 50 1 logi prop r1	Proportion of slower timescale fluctuations that scale with linearly rescaled range fits
Others	
SB TransitionMatrix 3ac sumdiagcov	Trace of covariance of transition matrix between symbols in 3-letter alphabet
PD PeriodicityWang th0 01	Periodicity measure of Wang et al. (2007b)

A.3. List of inter-regional interaction features between EEG channels

Table A.3: List of inter-regional interaction features between EEG channels.

Names	***Simple descriptions***	***Directionality***
Correlation features		
Pearson	Pearson correlation coefficient between the 1s segments of signals X and Y and averaged across all segments	*Non-directional*
Cross correlation	The half-width of the cross-correlation between X and Y evaluated on 1s segments and averaged across all segments (Width as the lags to half of the maximum)	*Non-directional*

Continued on next page

Table A.3: List of inter-regional interaction features between EEG channels. (Continued)

Cross covariance	The half-width of the cross-covariance between X and Y evaluated on 1s segments and averaged across all segments (Width as the lags to half of the maximum)	*Non-directional*
Cross-channel power band features		
Band power ratio	The cross-channel band power ratios between X and Y, with power spectrum evaluated on 1s segments using Thompson's multi-taper approach. The power spectrums are firstly normalised to the total power for X and Y separately before computing the ratios. Frequency bands: 1-4 Hz (alpha), 4-8Hz (theta), 8-12Hz (alpha), 12-30Hz (beta).	*Non-directional*
Coherence	Cross channel magnitude squared coherence in different frequency bands between X and Y, evaluated on 1s segments and then averaged. Frequency bands: 1-4Hz (alpha), 4-8Hz (theta), 8-12Hz (alpha), 12-18Hz (beta1), 18-24Hz (beta2), 24-30Hz (beta3).	*Non-directional*
Granger-causality based metrics		
Directed transfer function (DTF)	Normalised directed transfer (DTF) function measured as the causal influence from X to Y at different frequency bands, with vector autoregression model estimated on all EEG channels. Normalised DTF shows the ratio between causal inputs from X to Y and the total inputs from all other channels to Y. Frequency bands: 1-4Hz (alpha), 4-8Hz (theta), 8-12Hz (alpha), 12-18Hz (beta1), 18-24Hz (beta2), 24-30Hz (beta3). Directionality is considered by computing both X to Y and Y to X.	*Bidirectional*
Partial directed coherence (PDC)	Normalised partial directed coherence (PDC) measured as the causal influence from X to Y at different frequency bands, with vector autoregression model estimated on all EEG channels. Normalised PDC shows the ratio between causal inputs from X to Y and the outflows of information from X. Frequency bands are the same as DTF. Directionality is considered by computing both X to Y and Y to X.	*Bidirectional*
Cross-channel frequency couplings		
Phase-amplitude coupling	Cross channel phase-amplitude coupling between the phases of lower frequency band from X channel to the amplitudes of a higher frequency band from Y channel. Frequency bands: 1-4 Hz (alpha), 4-8Hz (theta), 8-12Hz (alpha), 12-30Hz (beta). All combinations are considered by computing both X or Y providing phase or amplitude.	*Bidirectional*

Continued on next page

Table A.3: List of inter-regional interaction features between EEG channels. (Continued)

Cross bispectrum	The cross bispectrum is estimated from X to Y (or Y to X), and the average spectrum power is taken within each frequency band in the 2D spectrum. Frequency bands are the same as phase-amplitude coupling. Both directions are considered when computing the cross bispectrum. Cross bispectrum captures quadratic coupling between signals.	*Bidirectional*
General synchronisation		
Synchronization likelihood	Synchronization likelihood evaluated on 1s segments between X and Y and then averaged. Synchronisation likelihood estimates the linear and nonlinear inter-dependences between channels.	*Non-directional*
Information based metrics		
Transfer Entropy (TE)	Transfer entropy between X and Y signals evaluated on 1s segments and then averaged. Directionality is considered by computing both X to Y and Y to X. Transfer entropy measures the directed information transfer.	*Bidirectional*
Mutual Information (MI)	Mutual information between X and Y signals evaluated on 1s segments and then averaged. Mutual information measures the mutual dependence between two variables.	*Bidirectional*

B. Supplementary Materials

B.1. Sleep onset latency distributions before exclusion for MESA and Montreal datasets

Distribution stats for the MESA dataset: mean ± st.d. 81.72 ± 54.96 minutes; The Montreal dataset: mean ± st.d. 25.34 ± 23.90 minutes. The distribution and statistics after exclusion can be seen in the main text.

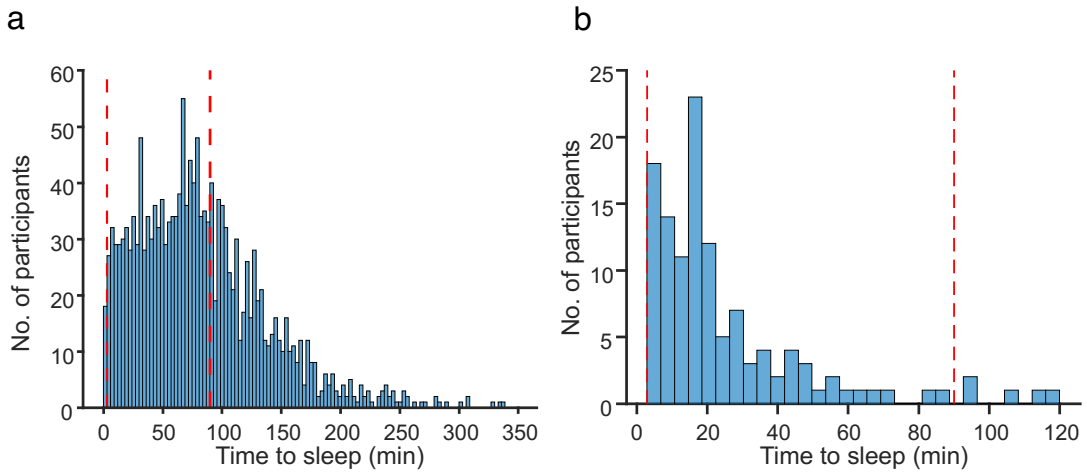


Figure B.1: The sleep onset latency distribution for MESA (a) and Montreal (b) datasets. Red lines are the cut-off threshold for limiting the latency to be within [3,90] minutes range.

B.2. Examplary EEG cardiac signature trace

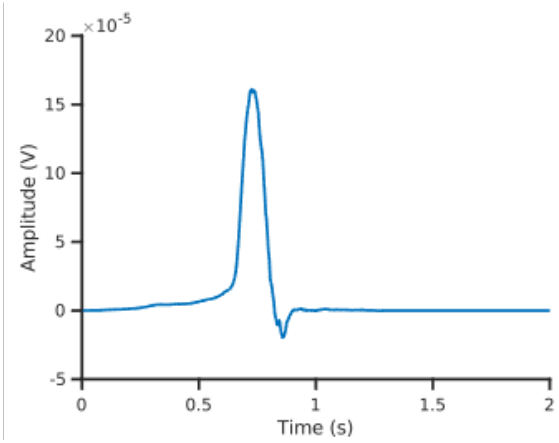


Figure B.2: An example EEG cardiac signature trace. The EEG cardiac signature trace is taken from one of the EEG channels and averaged across the entire overnight EEG period.

B.3. Examplary EEG power spectrum before and after differentiation filtering

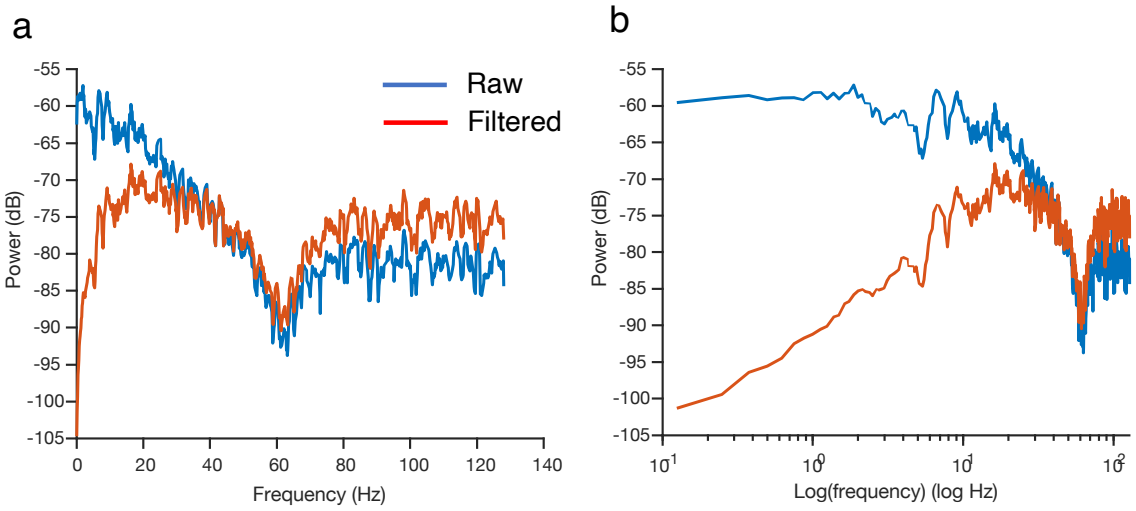


Figure B.3: Examplary EEG power spectrum before and after differentiation filtering. (**a**) shows the linear frequency scale, and (**b**) shows the log-scale. The power spectrum was evaluated on a 6-second EEG epoch, with blue showing the raw power spectrum, and red showing the power spectrum after differentiation filtering.

B.4. List of features in the hierarchical order in Figure 3.2 and the explanations of the key representative features selected

The features are arranged in the hierarchical order as in the figure, with features in the same cluster listed without borders. Features that are marked red and bold are the cluster centres with mean cluster accuracy over 65% and selected as key governing features. The remaining cluster centres are marked in bold only.

Table B.1: List of features in the hierarchical order in Figure 3.2. Key features selected are marked in red and bold, while other cluster centres not selected are marked in bold only.

Feature Name	**Accuracy**	**Mean Accuracy**
Peak alpha band frequency	52.80%	52.80%
Peak theta band frequency	53.81%	53.81%
SC FluctAnal 2 dfa 50 1 2 logi prop r1	54.90%	54.90%
Delta to theta ratio	52.61%	54.90%
Delta to alpha ratio	**57.19%**	
Theta to alpha ratio	61.91%	
Transition power ratio	**69.85%**	
Delta to beta ratio	70.95%	**71.24%**
delta power	76.08%	
theta power	74.10%	
Theta to beta ratio	74.59%	
SC FluctAnal 2 rsrangefit 50 1 logi prop r1	63.76%	63.76%
Peak beta band frequency	**66.25%**	**66.25%**
Alpha to beta ratio	**69.39%**	**68.77%**
alpha power	68.14%	
Peak delta band frequency	64.79%	64.79%
MD hr classic pnn40	63.25%	
SB BinaryStats diff longstretch0	65.06%	
PD PeriodicityWang th0 01	63.08%	**68.68%**
CO FirstMin ac	65.11%	
CO Embed2 Dist tau d expfit meandiff	72.45%	

Continued on next page

Table B.1: List of features in the hierarchical order in Figure 3.2. Key features selected are marked in red and bold, while other cluster centres not selected are marked in bold only. (Continued)

CO f1ecac	71.64%	
SB MotifThree quantile hh	72.15%	
SP Summaries welch rect centroid	73.62%	
SP Summaries welch rect area 5 1	**72.54%**	
FC LocalSimple mean3 stderr	70.63%	
SB BinaryStats mean longstretch1	74.18%	
FC LocalSimple mean1 tauresrat	60.50%	
IN AutoMutuallnfoStats 40 gaussian fmmi	55.64%	61.01%
CO HistogramAMI even 2 5	**66.49%**	
SB TransitionMatrix 3ac sumdiagco	57.11%	57.11%
Phase amplitude coupling TA	52.56%	52.56%
Phase amplitude coupling DT	56.17%	56.17%
Phase amplitude coupling DA	53.85%	53.85%
Phase amplitude coupling DB	55.29%	55.29%
Temporal coherence delta band	52.98%	52.98%
Temporal coherence theta band	**53.51%**	53.81%
Temporal coherence alpha band	54.10%	
Temporal coherence beta band	53.17%	53.17%
CO trev 1 num	52.07%	52.07%

There are four key representative features selected, which are listed and explained below.

Table B.2: List of key features extracted and their explanations.

Name in the figure	**Original name**	**Explanations**
Total EEG spectrum power	*SP Summaries welch rect area 5 1*	This feature captures the first fifth total spectrum power; Given that our EEG sampling rate is 256Hz, the first fifth of the spectrum ranges from 0 to around 25Hz; In the typical scenario of sleep EEG the spectrum power is mostly concentrated below 30Hz, so we interpret it as approximately the 'total spectrum power' .

Continued on next page

Table B.2: List of key features extracted and their explanations. (Continued)

Peak beta band frequency	*Peak beta band frequency*	This feature captures the frequency bin in beta band (13 to 30 Hz) with the highest power,
Alpha to beta band ratio	*Alpha to beta ratio*	This feature captures the band power ratio between alpha band (8 to 12 Hz) and beta band (13 to 30 Hz)
Low to high frequency band power ratio	*Transition power ratio*	The log-transformed transition power ratio defined as the low frequency band power (delta plus theta, 1 to 8 Hz) divided by high frequency band power (alpha plus beta, 8 to 30 Hz)

B.5. GLM statistics of all features at sleep onset to predict age, and explanations of the significant features in the model

The full GLM statistics are given in this table. Features marked in red are the significant features (p-value < 0.05) chosen and shown in the main text.

Table B.3: GLM statistics of all features at sleep onset to predict age. Significant features ($p < 0.05$) are marked in red.

Name	**Estimate Coefficient**	**Standard Error Estimate**	**t-Stat**	**p-value**
Intercept	147.32	22.06	6.68	4.21E-11
delta power	-10.50	32.93	-0.32	0.750
theta power	31.92	41.00	0.78	0.436
alpha power	50.62	35.85	1.41	0.158
beta power	-45.33	13.59	-3.34	0.001
Delta to theta ratio	-2.08	19.73	-0.11	0.916
Delta to alpha ratio	-22.41	29.66	-0.76	0.450
Delta to beta ratio	-8.33	41.39	-0.20	0.841
Theta to alpha ratio	-20.52	27.08	-0.76	0.449
Theta to beta ratio	-28.14	36.95	-0.76	0.446
Alpha to beta ratio	-60.33	31.02	-1.94	0.052
Transition power ratio	51.37	22.15	2.32	0.021
Peak delta band frequency	-6.86	2.37	-2.90	0.004

Continued on next page

Table B.3: GLM statistics of all features at sleep onset to predict age. Significant features ($p < 0.05$) are marked in red. (Continued)

Peak theta band frequency	-0.04	2.17	-0.02	0.984
Peak alpha band frequency	2.27	2.12	1.07	0.284
Peak beta band frequency	-2.51	2.70	-0.93	0.353
Temporal coherence delta band	-7.32	3.73	-1.96	0.050
Temporal coherence theta band	2.62	4.28	0.61	0.540
Temporal coherence alpha band	0.08	4.05	0.02	0.984
Temporal coherence beta band	-4.79	3.47	-1.38	0.167
Phase amplitude coupling DT	1.69	3.38	0.50	0.617
Phase amplitude coupling DA	4.30	3.39	1.27	0.206
Phase amplitude coupling DB	-1.32	3.44	-0.38	0.701
Phase amplitude coupling TA	2.76	3.26	0.85	0.397
Phase amplitude coupling TB	-3.78	3.13	-1.21	0.227
Phase amplitude coupling AB	1.56	2.97	0.52	0.600
DN HistogramMode 5	-3.40	3.34	-1.02	0.308
DN HistogramMode 10	0.69	3.70	0.19	0.853
CO f1ecac	-8.97	3.75	-2.39	0.017
CO FirstMin ac	-4.42	5.34	-0.83	0.408
CO HistogramAMI even 2 5	-12.89	2.82	-4.57	5.58E-06
CO trev 1 num	-0.20	2.52	-0.08	0.936
MD hrv classic pnn40	-6.15	4.10	-1.50	0.134
SB BinaryStats mean longstretch1	-3.51	5.36	-0.66	0.512
SB TransitionMatrix 3ac sumdiagcov	-0.69	2.96	-0.23	0.816
PD PeriodicityWang th0 01	0.55	3.36	0.16	0.870
CO Embed2 Dist tau d expfit meandiff	-2.06	4.56	-0.45	0.652
IN AutoMutualInfoStats 40 gaussian fmmi	5.65	3.30	1.71	0.087
FC LocalSimple mean1 tauresrat	-3.70	3.70	-1.00	0.317

Continued on next page

Table B.3: GLM statistics of all features at sleep onset to predict age. Significant features ($p < 0.05$) are marked in red. (Continued)

DN OutlierInclude p 001 mdrmd	-9.48	7.83	-1.21	0.227
DN OutlierInclude n 001 mdrmd	10.00	7.89	1.27	0.205
SP Summaries welch rect area 5 1	-15.56	11.79	-1.32	0.187
SB BinaryStats diff longstretch0	-3.69	3.35	-1.10	0.272
SB MotifThree quantile hh	-19.95	6.07	-3.28	0.001
SC FluctAnal 2 rsrangefit 50 1 logi prop r1	-2.05	3.19	-0.64	0.521
SC FluctAnal 2 dfa 50 1 2 logi prop r1	2.26	2.93	0.77	0.442
SP Summaries welch rect centroid	-4.50	9.75	-0.46	0.645
FC LocalSimple mean3 stderr	1.80	8.52	0.21	0.833

There are six features with significant p-values. Their names in the text, along with their explanations, are shown in the table following. One of the features, 'low to high frequency band power ratio', is the same as the four key features explained before in Table B.2.

Table B.4: Explanantions of the significant features in the GLM model to predict age at sleep onset.

Name in the text	**Original name**	**Explanations / Interpretations**
Beta band power	*beta power*	The log-transformed normalised band power in beta band (13-30 Hz).
Low to high frequency band power ratio	*Transition power ratio*	Same as previous.
Delta band peak frequency	*Peak delta band frequency*	The peak frequency at delta band (1-4 Hz).
Autocorrelation width	*CO f1ecac*	This CATCH-22 feature captures the 'first 1/e crossing of the autocorrelation function. This feature can be considered as a quantification of the autocorrelation width, where larger values imply a higher autocorrelation of the time series and thus, a wider range of self-similarity.

Continued on next page

Table B.4: Explanantions of the significant features in the GLM model to predict age at sleep onset. (Continued)

Auto mutual information	*CO HistogramAMI even 2 5*	A CATCH-22 features that captures the auto-mutual information of the time series. Higher feature values imply a higher auto mutual information, which means that the time series dynamics are repetitive and less noise-like (given the time lag $\tau = 5$) (Escudero et al., 2009).
Entropy	*SB MotifThree quantile hh*	This CATCH-22 feature captures the entropy of the time series by discretizing the time series values into 3 equal bins. Higher feature values imply a larger entropy, which means that the time series patterns are not repeating.

B.6. GLM statistics of all features at bedtime to predict age

Table B.5: GLM statistics of all features at bedtime to predict age. Significant features ($p < 0.05$) are marked in bold and red.

Names	**Estimate Coefficient**	**Standard Error Estimate**	**t-Stat**	**p-value**
Intercept	33.80	20.20	1.67	0.095
delta power	-19.80	22.55	-0.88	0.380
theta power	35.80	26.22	1.37	0.172
alpha power	-20.36	17.87	-1.14	0.255
beta power	-1.14	10.99	-0.10	0.918
Delta to theta ratio	19.11	13.50	1.42	0.157
Delta to alpha ratio	-11.60	18.81	-0.62	0.537
Delta to beta ratio	-3.82	29.92	-0.13	0.899
Theta to alpha ratio	-10.58	14.38	-0.74	0.462
Theta to beta ratio	-8.88	27.40	-0.32	0.746
Alpha to beta ratio	13.35	19.76	0.68	0.499
Transition power ratio	12.83	15.43	0.83	0.406
Peak delta band frequency	-1.98	2.38	-0.83	0.406
Peak theta band frequency	3.45	2.38	1.45	0.148
Peak alpha band frequency	2.09	2.06	1.01	0.312

Continued on next page

Table B.5: GLM statistics of all features at bedtime to predict age. Significant features ($p < 0.05$) are marked in bold and red. (Continued)

Peak beta band frequency	-1.54	2.51	-0.62	0.538
Temporal coherence delta band	3.30	3.60	0.92	0.360
Temporal coherence theta band	4.96	4.09	1.21	0.226
Temporal coherence alpha band	-5.69	3.89	-1.46	0.143
Temporal coherence beta band	3.70	3.52	1.05	0.292
Phase amplitude coupling DT	-0.95	3.45	-0.28	0.783
Phase amplitude coupling DA	-2.04	3.23	-0.63	0.527
Phase amplitude coupling DB	0.85	3.16	0.27	0.789
Phase amplitude coupling TA	-2.23	3.18	-0.70	0.482
Phase amplitude coupling TB	-0.75	3.17	-0.24	0.812
Phase amplitude coupling AB	-2.81	3.10	-0.91	0.364
DN HistogramMode 5	-0.78	3.12	-0.25	0.802
DN HistogramMode 10	2.46	3.91	0.63	0.530
CO f1ecac	1.00	6.81	0.15	0.883
CO FirstMin ac	-8.12	7.38	-1.10	0.272
CO HistogramAMI even 2 5	0.13	3.00	0.04	0.965
CO trev 1 num	1.36	2.07	0.66	0.512
MD hrv classic pnn40	3.93	4.32	0.91	0.363
SB BinaryStats mean longstretch1	0.65	6.50	0.10	0.921
SB TransitionMatrix 3ac sumdiagcov	-1.33	2.91	-0.46	0.646
PD PeriodicityWang th0 01	**8.36**	**3.38**	**2.47**	**0.014**
CO Embed2 Dist tau d expfit meandiff	1.34	5.52	0.24	0.808
IN AutoMutualInfoStats 40 gaussian fmmi	-2.30	3.69	-0.62	0.533
FC LocalSimple mean1 tauresrat	-3.44	5.52	-0.62	0.533
DN OutlierInclude p 001 mdrmd	4.90	9.92	0.49	0.621
DN OutlierInclude n 001 mdrmd	0.83	9.83	0.08	0.933

Continued on next page

Table B.5: GLM statistics of all features at bedtime to predict age. Significant features ($p < 0.05$) are marked in bold and red. (Continued)

SP Summaries welch rect area 5 1	**25.54**	**12.50**	**2.04**	**0.041**
SB BinaryStats diff longstretch0	-2.49	3.82	-0.65	0.515
SB MotifThree quantile hh	1.40	8.92	0.16	0.876
SC FluctAnal 2 rsrangefit 50 1 logi prop r1	-3.48	3.08	-1.13	0.259
SC FluctAnal 2 dfa 50 1 2 logi prop r1	-0.30	2.10	-0.15	0.884
SP Summaries welch rect centroid	**25.49**	**12.03**	**2.12**	**0.034**
FC LocalSimple mean3 stderr	13.26	9.97	1.33	0.184

B.7. Key feature statistics at bedtime and sleep onset for male and female participants in Figure 3.6

Table B.6: Key feature statistics at bedtime and sleep onset for male and female participants in Figure 3.6. Shown are mean ± st.d. for statistics, and p-values using the double-sided two-sample t-test.

Names	**Bedtime**			**Sleep onset**		
	Male stats	**Female stats**	**p-value**	**Male stats**	**Female stats**	**p-value**
Alpha to beta band ratio	0.41 ± 0.17	0.39 ± 0.17	0.21	0.59 ± 0.14	0.59 ± 0.14	0.60
Low to high frequency band power ratio	0.41 ± 0.14	0.41 ± 0.16	0.89	0.59 ± 0.13	0.59 ± 0.13	0.46
Peak beta band frequency	0.51 ± 0.18	0.50 ± 0.18	0.61	0.27 ± 0.17	0.30 ± 0.17	0.02
Total EEG spectrum power	0.25 ± 0.19	0.27 ± 0.19	0.04	0.54 ± 0.19	0.51 ± 0.20	0.06

B.8. List of features in the hierarchical order in Figure 3.7

This is the classification for the Montreal dataset as a whole.

Table B.7: List of features in the hierarchical order in Figure 3.7 along with their accuracy. Features marked in bold are cluster centres, with features marked in red as key features chosen (which satisfy the criterion). Single-cluster features will not be marked in bold.

Feature name	**Accuracy**	**Mean accuracy**
DN HistogramMode 5	51.38%	51.38%
Temporal coherence beta band	54.39%	54.39%
Temporal coherence alpha band	62.89%	63.56%
Temporal coherence theta band	**64.22%**	
Temporal coherence delta band	57.16%	57.16%
beta power	55.10%	55.10%
Peak delta band frequency	**72.19%**	**72.19%**
Peak theta band frequency	**68.17%**	**68.17%**
SC FluctAnal 2 rsrangefit 50 1 logi prop r1	60.89%	60.89%
Theta to beta ratio	77.99%	
theta power	76.90%	
Theta to alpha ratio	74.64%	
Transition power ratio	**80.59%**	**78.35%**
Delta to beta ratio	81.39%	
delta power	83.61%	
Delta to alpha ratio	73.34%	
Delta to theta ratio	61.27%	61.27%
alpha power	59.19%	60.55%
Alpha to beta ratio	**61.90%**	
Peak beta band frequency	**72.61%**	**72.61%**
FC LocalSimple mean1 tauresrat	58.41%	58.41%
PD PeriodicityWang th0 01	56.95%	
CO f1ecac	58.25%	
SB MotifThree quantile hh	63.69%	
CO Embed2 Dist tau d expfit meandiff	64.82%	
SP Summaries welch rect centroid	65.10%	63.18%
FC LocalSimple mean3 stderr	65.16%	

Continued on next page

Table B.7: List of features in the hierarchical order in Figure 3.7 along with their accuracy. Features marked in bold are cluster centres, with features marked in red as key features chosen (which satisfy the criterion). Single-cluster features will not be marked in bold. (Continued)

SP Summaries welch rect area 5 1	**69.67%**	
SB BinaryStats mean longstretch1	69.45%	
SB BinaryStats diff longstretch0	60.37%	
MD hrv classic pnn40	58.37%	
IN AutoMutualInfoStats 40 gaussian fmmi	**56.19%**	55.21%
CO HistogramAMI even 2 5	54.22%	
CO FirstMin ac	54.76%	54.76%
CO trev 1 num	51.38%	51.38%
SB TransitionMatrix 3ac sumdiagcov	54.96%	54.96%
Peak alpha band frequency	55.69%	55.69%
Phase amplitude coupling DB	57.00%	57.00%
Phase amplitude coupling DA	56.27%	56.27%
Phase amplitude coupling DT	56.84%	56.84%

B.9. List of features in the hierarchical order in Figure 3.9a and b

For classification in the Normal participants cohort:

Table B.8: List of features in the hierarchical order in Figure 3.9a along with their accuracy. Features marked in bold are cluster centres, with features marked in red as key features chosen (which satisfy the criterion). Single-cluster features will not be marked in bold.

Feature name	**Accuracy**	**Mean Accuracy**
Temporal coherence beta band	54.37%	54.37%
Temporal coherence alpha band	**61.85%**	61.72%
Temporal coherence theta band	61.58%	
Temporal coherence delta band	58.99%	58.99%
Peak alpha band frequency	54.90%	54.90%
beta power	56.11%	56.11%
Delta to theta ratio	60.25%	60.25%

Continued on next page

Table B.8: List of features in the hierarchical order in Figure 3.9a along with their accuracy. Features marked in bold are cluster centres, with features marked in red as key features chosen (which satisfy the criterion). Single-cluster features will not be marked in bold. (Continued)

Delta to alpha ratio	72.72%	
delta power	84.18%	
Delta to beta ratio	81.54%	
Transition power ratio	**82.45%**	**79.16%**
Theta to alpha ratio	76.22%	
theta power	77.92%	
Theta to beta ratio	79.09%	
SC FluctAnal 2 rsrangefit 50 1 logi prop r1	63.96%	63.96%
Peak theta band frequency	**68.58%**	**68.58%**
Peak delta band frequency	**73.88%**	**73.88%**
Peak beta band frequency	**75.46%**	**75.46%**
Alpha to beta ratio	**62.90%**	60.69%
alpha power	58.48%	
FC LocalSimple mean1 tauresrat	56.90%	56.90%
PD PeriodicityWang th0 01	60.66%	
FC LocalSimple mean3 stderr	65.66%	
SP Summaries welch rect centroid	66.87%	
SP Summaries welch rect area 5 1	**70.67%**	
SB BinaryStats mean longstretch1	70.33%	**64.57%**
CO Embed2 Dist tau d expfit meandiff	66.94%	
CO f1ecac	60.36%	
SB MotifThree quantile hh	64.67%	
SB BinaryStats diff longstretcho	61.27%	
MD hrv classic pnn40	58.29%	
IN AutoMutuallnfoStats 40 gaussian fmmi	**57.15%**	56.79%
CO HistogramAMI even 2 5	56.42%	
CO FirstMin ac	55.17%	55.17%
CO trev 1 num	58.30%	58.30%

Continued on next page

Table B.8: List of features in the hierarchical order in Figure 3.9a along with their accuracy. Features marked in bold are cluster centres, with features marked in red as key features chosen (which satisfy the criterion). Single-cluster features will not be marked in bold. (Continued)

SB TransitionMatrix 3ac sumdiagcov	55.74%	55.74%
Phase amplitude coupling DB	56.60%	56.60%
Phase amplitude coupling DA	56.21%	56.21%
Phase amplitude coupling DT	56.07%	56.07%
DN HistogramMode 5	53.74%	53.74%

For the classification in the MCI cohort:

Table B.9: List of features in the hierarchical order in Figure 3.9b for the MCI participants cohort, along with their accuracy. Features marked in bold are cluster centres, with features marked in red as key features chosen (which satisfy the criterion). Single-cluster features will not be marked in bold.

Feature name	**Accuracy**	**Mean Accuracy**
Phase amplitude coupling AB	52.20%	52.20%
Phase amplitude coupling TB	53.88%	53.88%
Phase amplitude coupling DT	56.62%	56.62%
Phase amplitude coupling DB	57.09%	57.09%
Phase amplitude coupling DA	56.14%	56.14%
Peak alpha band frequency	56.73%	56.73%
Temporal coherence alpha band	62.50%	64.39%
Temporal coherence theta band	**66.27%**	
Temporal coherence delta band	53.38%	53.38%
SC FluctAnal 2 rsrangefit 50 1 logi prop r1	57.36%	57.36%
Theta to beta ratio	75.85%	**75.01%**
theta power	75.03%	
Theta to alpha ratio	71.29%	
Transition power ratio	**77.85%**	
Delta to beta ratio	80.73%	**74.62%**
delta power	82.92%	
Delta to alpha ratio	**73.27%**	

Continued on next page

Table B.9: List of features in the hierarchical order in Figure 3.9b for the MCI participants cohort, along with their accuracy. Features marked in bold are cluster centres, with features marked in red as key features chosen (which satisfy the criterion). Single-cluster features will not be marked in bold. (Continued)

Delta to theta ratio	61.55%	
Peak theta band frequency	**68.55%**	**68.55%**
Peak delta band frequency	**69.14%**	**69.14%**
beta power	52.45%	52.45%
IN AutoMutuallnfoStats 40 gaussian fmmi	53.99%	53.99%
CO FirstMin ac	55.29%	55.29%
MD hrv classic pnn40	58.35%	62.53%
SB BinaryStats diff longstretch0	59.98%	
SB BinaryStats mean longstretch1	67.63%	
SP Summaries welch rect centroid	65.19%	
FC LocalSimple mean3 stderr	64.99%	
SP Summaries welch rect area 5 1	**68.71%**	
CO f1ecac	57.06%	
SB MotifThree quantile hh	63.70%	
CO Embed2 Dist tau d expfit meandiff	63.49%	
PD PeriodicityWang th0 01	56.18%	
FC LocalSimple mean1 tauresrat	61.11%	61.11%
Peak beta band frequency	**69.45%**	**69.45%**
Alpha to beta ratio	**60.15%**	59.23%
alpha power	58.30%	
SB TransitionMatrix 3ac sumdiagcov	52.81%	52.81%
Temporal coherence beta band	53.57%	53.57%
DN HistogramMode 5	52.54%	52.54%

B.10. Statistics of the feature distributions in Normal and MCI participants at bedtime and sleep onset (for Figure 3.10 and 3.11)

Table B.10: Statistics of the feature distributions in Normal and MCI participants at bedtime and sleep onset (for Figure 3.10 and 3.11). Shown are mean ± st.d., with p-values using the double-sided two-sample t-test.

Feature name	**Bedtime**			**Sleep onset**		
	Normal	**MCI**	**p-value**	**Normal**	**MCI**	**p-value**
Alpha to beta band ratio	0.45 ± 0.26	0.47 ± 0.26	0.18	0.56 ± 0.19	0.54 ± 0.18	0.09
Low to high frequency band power ratio	0.31 ± 0.17	0.38 ± 0.20	**1.27E-11**	0.63 ± 0.17	0.63 ± 0.13	0.99
Peak delta band frequency	0.27 ± 0.37	0.32 ± 0.39	**0.04**	0.65 ± 0.32	0.64 ± 0.34	0.40
Peak theta band frequency	0.83 ± 0.26	0.82 ± 0.28	0.21	0.54 ± 0.37	0.52 ± 0.37	0.34
Peak beta band frequency	0.53 ± 0.26	0.50 ± 0.26	**0.02**	0.23 ± 0.24	0.26 ± 0.26	**0.02**
Total EEG spectrum power	0.27 ± 0.21	0.26 ± 0.19	0.33	0.54 ± 0.26	0.47 ± 0.24	**3.06E-07**

B.11. List of features in the hierarchical order in Figure 3.15

Table B.11: List of features in the hierarchical order in Figure 3.15. The classification accuracy and mean cluster accuracy are also shown. Key features selected are marked in red and bold, while other cluster centres not selected are marked in bold only.

Feature name	**Accuracy**	**Mean accuracy**
Phase amplitude coupling TB	52.33%	52.33%
Phase amplitude coupling TA	52.14%	52.14%
Peak delta band frequency	52.93%	52.93%
Peak beta band frequency	54.23%	54.23%
MD hrv classic pnn40	52.55%	52.55%
PD PeriodicityWang th0 01	57.73%	
CO FirstMin ac	56.01%	
CO f1ecac	**58.52%**	58.90%
SB MotifThree quantile hh	60.29%	

Continued on next page

Table B.11: List of features in the hierarchical order in Figure 3.15. The classification accuracy and mean cluster accuracy are also shown. Key features selected are marked in red and bold, while other cluster centres not selected are marked in bold only. (Continued)

CO Embed2 Dist tau d expfit meandiff	61.97%	
SB BinaryStats diff longstretch0	57.66%	58.49%
CO HistogramAMI even 2 5	56.39%	
FC LocalSimple mean3 stderr	58.48%	
SB BinaryStats mean longstretch1	58.27%	
SP Summaries welch rect area 5 1	**60.70%**	
SP Summaries welch rect centroid	61.43%	
FC LocalSimple mean1 tauresrat	56.49%	
IN AutoMutualInfoStats 40 gaussian fmmi	55.38%	55.38%
Temporal coherence beta band	49.67%	49.67%
Temporal coherence alpha band	56.61%	57.31%
Temporal coherence theta band	**57.42%**	
Temporal coherence delta band	57.89%	
SC FluctAnal 2 rsrangefit 50 1 logi prop r1	53.37%	53.37%
Alpha to beta ratio	**55.07%**	54.88%
Theta to beta ratio	54.69%	
Peak alpha band frequency	50.65%	50.65%
Peak theta band frequency	57.80%	57.80%
Delta to alpha ratio	**61.00%**	57.37%
alpha power	54.98%	
Delta to theta ratio	56.90%	
beta power	53.21%	
Delta to beta ratio	58.25%	
Transition power ratio	59.88%	
Theta to alpha ratio	58.58%	59.27%
delta power	**59.96%**	
SC FluctAnal 2 dfa 50 1 2 logi prop r1	57.23%	57.23%
Phase amplitude coupling DB	50.51%	50.51%

Continued on next page

Table B.11: List of features in the hierarchical order in Figure 3.15. The classification accuracy and mean cluster accuracy are also shown. Key features selected are marked in red and bold, while other cluster centres not selected are marked in bold only. (Continued)

Phase amplitude coupling DA	50.54%	50.54%
CO trev 1 num	52.84%	52.84%
DN HistogramMode 5	50.76%	50.76%

B.12. The GLM statistics of all features to predict sleep onset latency at bedtime

Table B.12: GLM statistics of all features at bedtime to predict sleep onset latency.

Names	**Estimate Coefficient**	**Standard Error Estimate**	**t-Stat**	**p-value**
Intercept	33.38	53.10	0.63	0.530
delta power	117.25	59.27	1.98	0.048
theta power	-166.32	68.92	-2.41	0.016
alpha power	51.36	46.96	1.09	0.274
beta power	-5.84	28.88	-0.20	0.840
Delta to theta ratio	-7.68	35.49	-0.22	0.829
Delta to alpha ratio	35.82	49.44	0.72	0.469
Delta to beta ratio	-141.90	78.64	-1.80	0.071
Theta to alpha ratio	-59.75	37.79	-1.58	0.114
Theta to beta ratio	179.12	72.02	2.49	0.013
Alpha to beta ratio	-79.78	51.94	-1.54	0.125
Transition power ratio	62.20	40.56	1.53	0.125
Peak delta band frequency	9.18	6.25	1.47	0.142
Peak theta band frequency	10.61	6.26	1.69	0.091
Peak alpha band frequency	-1.14	5.43	-0.21	0.834
Peak beta band frequency	-3.31	6.59	-0.50	0.616
Temporal coherence delta band	0.99	9.45	0.10	0.917

Continued on next page

Table B.12: GLM statistics of all features at bedtime to predict sleep onset latency. (Continued)

Temporal coherence theta band	-8.38	10.76	-0.78	0.436
Temporal coherence alpha band	-6.39	10.21	-0.63	0.532
Temporal coherence beta band	16.21	9.24	1.75	0.080
Phase amplitude coupling DT	9.92	9.06	1.10	0.274
Phase amplitude coupling DA	-4.55	8.48	-0.54	0.592
Phase amplitude coupling DB	-3.55	8.30	-0.43	0.669
Phase amplitude coupling TA	-4.99	8.35	-0.60	0.550
Phase amplitude coupling TB	-5.49	8.34	-0.66	0.511
Phase amplitude coupling AB	-11.73	8.13	-1.44	0.150
DN HistogramMode 5	-2.63	8.19	-0.32	0.748
DN HistogramMode 10	15.21	10.27	1.48	0.139
CO f1ecac	-18.43	17.90	-1.03	0.304
CO FirstMin ac	-20.91	19.41	-1.08	0.282
CO HistogramAMI even 2 5	3.94	7.89	0.50	0.618
CO trev 1 num	-7.94	5.43	-1.46	0.144
MD hrv classic pnn40	-1.60	11.36	-0.14	0.888
SB BinaryStats mean longstretch1	-4.49	17.07	-0.26	0.793
SB TransitionMatrix 3ac sumdiagcov	0.54	7.65	0.07	0.944
PD PeriodicityWang th0 01	3.05	8.89	0.34	0.731
CO Embed2 Dist tau d expfit meandiff	-28.21	14.51	-1.94	0.052
IN AutoMutualInfoStats 40 gaussian fmmi	14.64	9.70	1.51	0.132
FC LocalSimple mean1 tauresrat	3.62	14.50	0.25	0.803
DN OutlierInclude p 001 mdrmd	-14.02	26.06	-0.54	0.591
DN OutlierInclude n 001 mdrmd	17.74	25.85	0.69	0.493
SP Summaries welch rect area 5 1	64.99	32.85	1.98	0.048
SB BinaryStats diff longstretch0	-5.02	10.03	-0.50	0.617
SB MotifThree quantile hh	-22.13	23.44	-0.94	0.345

Continued on next page

Table B.12: GLM statistics of all features at bedtime to predict sleep onset latency. (Continued)

SC FluctAnal 2 rsrangefit 50 1 logi prop r1	-1.77	8.09	-0.22	0.827
SC FluctAnal 2 dfa 50 1 2 logi prop r1	-1.72	5.51	-0.31	0.755
SP Summaries welch rect centroid	23.75	31.63	0.75	0.453
FC LocalSimple mean3 stderr	36.56	26.20	1.40	0.163

B.13. Confounding test for sleep state space variable and feature state space velocity (aligned to half sleep onset for each participant)

The confounding test is done by aligning the time series of each participant to half their sleep onset period and then averaged using the same bootstrapping approach with sample number $N = 200$ as well. Given that the alignment is to half the sleep onset latency, the time period for the time series is much shorter.

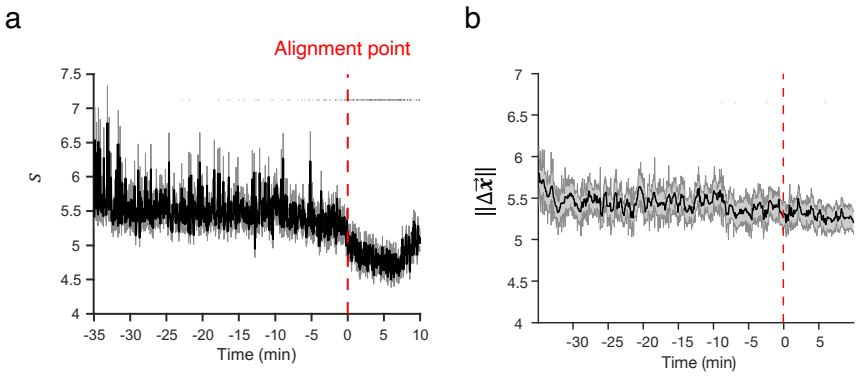


Figure B.4: Confounding test for sleep state space variable and feature state space velocity aligned to half sleep onset. (**a**) for the sleep state space variable $s(t)$ and (**b**) for the feature state space velocity $\|\Delta\vec{x}\|$. The time series for each participant are aligned to half the sleep onset period and then averaged. The horizontal line shows significant decreasing changes compared to the mean of the first 10 minutes, using the right-tailed two-sample t-test with a p-value threshold of 0.025 dividing the total number of timestamps. The red vertical dashed line shows the alignment point. The shaded area shows the standard error of the mean (s.e.m.) around the mean value (black line).

B.14. Sleep state space variable and feature state space velocity comparison between fast (< 40 minutes) and slow (> 60 minutes) falling asleep groups

All participants are divided into two cohorts, with one fast sleep onset (latency < 40 minutes) and one slow cohort (latency > 60 minutes). For each group, the group-averaged sleep state space variable $s(t)$ and feature state space velocity $\|\Delta\vec{x}\|$ are computed using the same bootstrapping approach with minimum sample number 150 (to reconcile with the reduced samples in each group). The fast cohort has a shorter time period available (starting from 19.5 minutes before sleep onset), and a two-sided two-sample t-test is done to compare the differences between them for the timestamps where both cohorts exist.

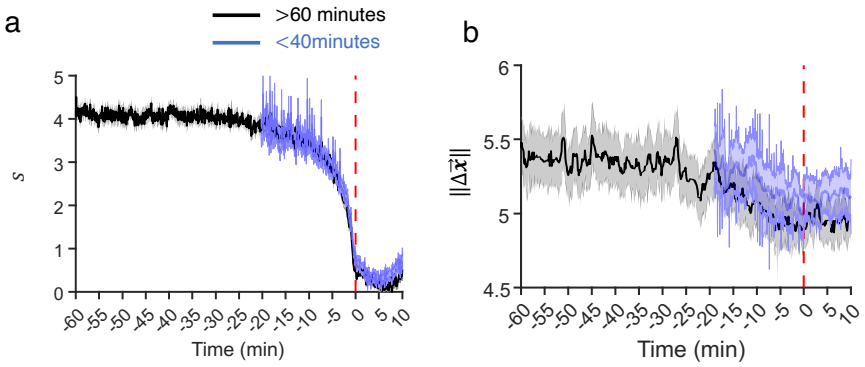


Figure B.5: Sleep state space variable and feature state space velocity comparison between fast (< 40 minutes) and slow (> 60 minutes) falling asleep groups. (**a**) for the sleep state space variable $s(t)$ and (**b**) for the feature state space velocity $\|\Delta\vec{x}\|$. The dynamics are compared between fast (< 40 minutes, blue traces) and slow (> 60 minutes, black traces) falling asleep groups. The differences between them are tested using the double-sided two-sample t-test with a p-value threshold of 0.05 dividing the total number of timestamps available for testing, no significance is detected. The red vertical dashed line shows the sleep onset. The shaded area shows the standard error of the mean (s.e.m.) around the mean value (solid line).

B.15. List of features in the hierarchical order in Figure 4.7

Table B.13: Feature names list in the hierarchical order in Figure 4.7. The names are original feature names and their interpretation can be found in previous explanation tables.

Original feature names (in the hierarchical order top down)
DN OutlierInclude p 001 mdrmd
DN OutlierInclude n 001 mdrmd

Continued on next page

Table B.13: Feature names list in the hierarchical order in Figure 4.7. The names are original feature names and their interpretation can be found in previous explanation tables. (Continued)

CO trev 1 num
DN HistogramMode 5
Phase amplitude coupling AB
DN HistogramMode 10
Peak alpha band frequency
Phase amplitude coupling TB
Temporal coherence beta band
SB TransitionMatrix 3ac sumdiagcov
Delta to theta ratio
SC FluctAnal 2 dfa 50 1 2 logi prop r1
beta power
Temporal coherence delta band
Peak theta band frequency
Temporal coherence theta band
Temporal coherence alpha band
Phase amplitude coupling DT
MD hrv classic pnn40
Phase amplitude coupling DB
Phase amplitude coupling DA
CO HistogramAMI even 2 5
Phase amplitude coupling TA
IN AutoMutualInfoStats 40 gaussian fmmi
Delta to alpha ratio
Theta to alpha ratio
Peak beta band frequency
FC LocalSimple mean3 stderr
FC LocalSimple mean1 tauresrat
SB MotifThree quantile hh

Continued on next page

Table B.13: Feature names list in the hierarchical order in Figure 4.7. The names are original feature names and their interpretation can be found in previous explanation tables. (Continued)

SP Summaries welch rect centroid
delta power
theta power
Theta to beta ratio
CO Embed2 Dist tau d expfit meandiff
SP Summaries welch rect area 5 1
alpha power
SB BinaryStats mean longstretch1
CO FirstMin ac
PD PeriodicityWang th0 01
Alpha to beta ratio
SB BinaryStats diff longstretch0
Delta to beta ratio
Transition power ratio
Peak delta band frequency
CO f1ecac
SC FluctAnal 2 rsrangefit 50 1 logi prop r1

B.16. Confounding test for feature dynamic time-series in Figure 4.7

The same confounding test as previously, but now for the 47 feature dynamics time series. For each participant, the feature time series are now aligned to half sleep onset period point and then averaged with boostrapping of 200 samples.

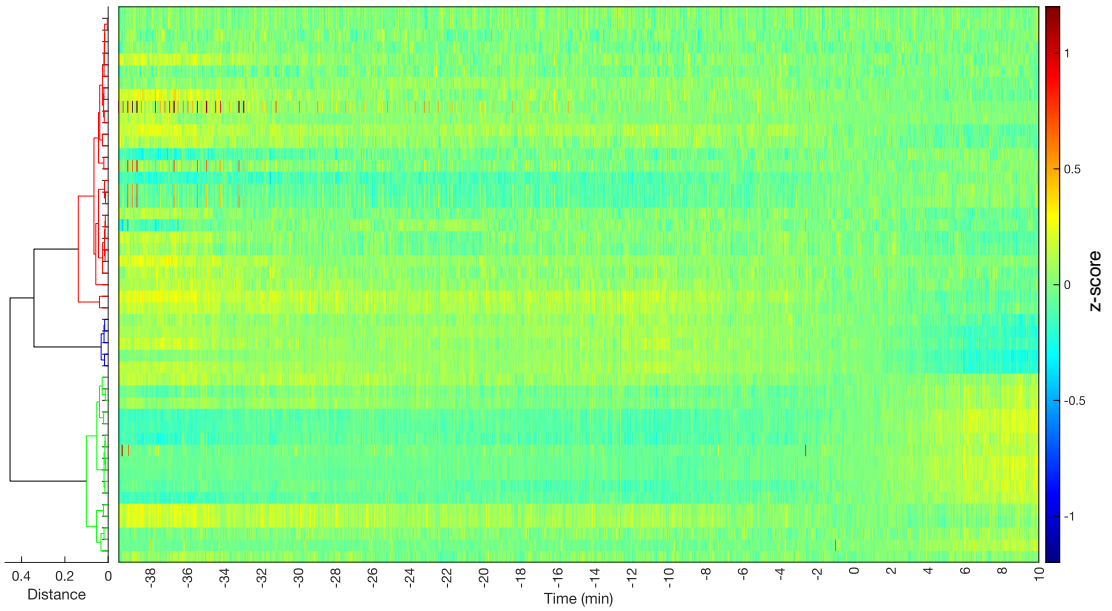


Figure B.6: Feature time series aligned to half of sleep onset point as a confounding test. The features are arranged in the same hierarchical order as in Figure 4.7 for easy comparison.

B.17. Similarity score between the first two FPCs and all 47 features, and explanation of the selected representative features

The similarity scores are computed between the FPCs and the features as in Equation 4.4. For each FPC, there are two features selected to represent the dynamics in both positive (maximum score) and negative (minimum score) directions.

Table B.14: Similarity score between the first two FPCs and all 47 features. The features selected with maximum and minimum values for each FPC are marked in red (for FPC1) and blue (for FPC2).

Feature names	**FPC1 Scores**	**FPC2 Scores**
delta power	-0.97	0.27
theta power	-1.00	0.06

Continued on next page

Table B.14: Similarity score between the first two FPCs and all 47 features. The features selected with maximum and minimum values for each FPC are marked in red (for FPC1) and blue (for FPC2). (Continued)

alpha power	-0.99	-0.15
beta power	-0.80	-0.51
Delta to theta ratio	0.48	0.94
Delta to alpha ratio	-0.46	0.97
Delta to beta ratio	-0.95	0.33
Theta to alpha ratio	-0.85	0.57
Theta to beta ratio	-0.99	0.12
Alpha to beta ratio	-0.99	-0.11
Transition power ratio	-0.94	0.37
Peak delta band frequency	-0.99	0.08
Peak theta band frequency	0.69	**-0.77**
Peak alpha band frequency	0.34	0.69
Peak beta band frequency	**1.00**	-0.04
Temporal coherence delta band	-0.92	-0.07
Temporal coherence theta band	0.68	-0.76
Temporal coherence alpha band	0.71	-0.73
Temporal coherence beta band	0.66	-0.21
Phase amplitude coupling DT	0.98	0.07
Phase amplitude coupling DA	0.96	0.17
Phase amplitude coupling DB	0.97	0.19
Phase amplitude coupling TA	0.90	0.30
Phase amplitude coupling TB	0.74	0.46
Phase amplitude coupling AB	-0.38	0.45
DN HistogramMode 5	-0.35	-0.11
DN HistogramMode 10	0.40	0.06
CO f1ecac	-1.00	-0.07
CO FirstMin ac	-1.00	-0.09
CO HistogramAMI even 2 5	0.96	0.18

Continued on next page

Table B.14: Similarity score between the first two FPCs and all 47 features. The features selected with maximum and minimum values for each FPC are marked in red (for FPC1) and blue (for FPC2). (Continued)

CO trev 1 num	-0.17	-0.14
MD hrv classic pnn40	0.99	0.03
SB BinaryStats mean longstretch1	-1.00	0.01
SB TransitionMatrix 3ac sumdiagcov	0.77	-0.02
PD PeriodicityWang th0 01	-0.99	-0.12
CO Embed2 Dist tau d expfit meandiff	-0.99	-0.12
IN AutoMutualInfoStats 40 gaussian fmmi	0.96	0.25
FC LocalSimple mean1 tauresrat	1.00	0.08
DN OutlierInclude p 001 mdrmd	-0.10	0.04
DN OutlierInclude n 001 mdrmd	-0.14	0.05
SP Summaries welch rect area 5 1	**-1.00**	-0.09
SB BinaryStats diff longstretch0	-0.99	-0.11
SB MotifThree quantile hh	1.00	0.09
SC FluctAnal 2 rsrangefit 50 1 logi prop r1	-0.99	0.13
SC FluctAnal 2 dfa 50 1 2 logi prop r1	-0.04	**1.00**
SP Summaries welch rect centroid	1.00	0.10
FC LocalSimple mean3 stderr	1.00	0.05

There are four features selected based on the scores: the peak beta band frequency and the total EEG spectrum power for FPC1, the peak theta band frequency and the proportion of slower timescale dynamics for FPC2. The first three features are straightforward to understand and have been interpreted before. The last feature, 'SC FluctAnal 2 dfa 50 1 2 logi prop r1', is a CATCH-22 feature, and the exact algorithmic details can be referred to in the paper Lubba et al. (2019). Briefly, detrended fluctuation analysis (Peng et al., 1995) is applied, which is a methodology to capture time series self-similarity and long-range correlations that exist in power-law scaling dynamics (log-log linear patterns). The feature captures the percentage of log-log linear scaling region. As an easy interpretation, the higher the output feature value, the input time series would have higher self-similarity (slower fluctuation dynamics) over a longer time range, or say, the signal is more predictable over a longer time range. From the neuroscientific perspective, however, there have not been any studies to explain what this feature exactly implies in EEG and how it is related to underlying neuronal mechanisms.

B.18. Similarity between the feature time series to the control variable time series $c(t)$

Table B.15: Similarity between the feature time series to the control variable time series $c(t)$. The four key dynamical features from FPCA are marked in bold, with the theta band peak frequency marked in red for best similarity. The theta-to-alpha band power ratio had the highest absolute similarity score (marked in blue). The theta band peak frequency had the smallest DTW distance.

	DTW distance	**Absolute similarity score**
delta power	66.35	0.93
theta power	66.30	0.86
alpha power	64.95	0.75
beta power	36.98	0.39
Delta to theta ratio	36.35	0.15
Delta to alpha ratio	54.17	0.86
Delta to beta ratio	66.09	0.94
Theta to alpha ratio	63.87	**0.96**
Theta to beta ratio	66.42	0.88
Alpha to beta ratio	65.40	0.78
Transition power ratio	65.82	0.95
Peak delta band frequency	64.57	0.87
Peak theta band frequency	**7.03**	**0.94**
Peak alpha band frequency	36.11	0.09
Peak beta band frequency	**14.21**	**0.86**
Temporal coherence delta band	52.37	0.75
Temporal coherence theta band	7.60	0.92
Temporal coherence alpha band	7.67	0.93
Temporal coherence beta band	15.73	0.70
Phase amplitude coupling DT	14.53	0.78
Phase amplitude coupling DA	17.44	0.71
Phase amplitude coupling DB	19.51	0.71

Continued on next page

Table B.15: Similarity between the feature time series to the control variable time series $c(t)$. The four key dynamical features from FPCA are marked in bold, with the theta band peak frequency marked in red for best similarity. The theta-to-alpha band power ratio had the highest absolute similarity score (marked in blue). The theta band peak frequency had the smallest DTW distance. (Continued)

Phase amplitude coupling TA	20.93	0.60
Phase amplitude coupling TB	28.23	0.35
Phase amplitude coupling AB	37.01	0.57
DN HistogramMode 5	33.58	0.23
DN HistogramMode 10	29.36	0.27
CO f1ecac	65.79	0.80
CO FirstMin ac	65.52	0.79
CO HistogramAMI even 2 5	18.58	0.70
CO trev 1 num	32.08	0.05
MD hrv classic pnn40	13.84	0.82
SB BinaryStats mean longstretch1	66.26	0.84
SB TransitionMatrix 3ac sumdiagcov	21.38	0.65
PD PeriodicityWang th0 01	65.28	0.77
CO Embed2 Dist tau d expfit meandiff	65.31	0.77
IN AutoMutualInfoStats 40 gaussian fmmi	22.82	0.66
FC LocalSimple mean1 tauresrat	18.07	0.80
DN OutlierInclude p 001 mdrmd	33.36	0.11
DN OutlierInclude n 001 mdrmd	33.82	0.15
SP Summaries welch rect area 5 1	**65.62**	**0.79**
SB BinaryStats diff longstretch0	65.38	0.78
SB MotifThree quantile hh	18.30	0.79
SC FluctAnal 2 rsrangefit 50 1 logi prop r1	66.33	0.89
SC FluctAnal 2 dfa 50 1 2 logi prop r1	**44.00**	**0.62**
SP Summaries welch rect centroid	18.93	0.78

Continued on next page

Table B.15: Similarity between the feature time series to the control variable time series $c(t)$. The four key dynamical features from FPCA are marked in bold, with the theta band peak frequency marked in red for best similarity. The theta-to-alpha band power ratio had the highest absolute similarity score (marked in blue). The theta band peak frequency had the smallest DTW distance. (Continued)

FC LocalSimple mean3 stderr	16.77	0.81

B.19. Table of p-values of the pair-wised Granger-causality test between the four key dynamical feature time series

Table B.16: Table of p-values of the pair-wised Granger-causality test between the four key dynamical feature time series, with sources in columns and destination in rows. Significant p-values are marked in red. There is no test for the feature time series itself, thus marked as NaN.

From / To	**Peak theta band frequency**	**Peak beta band frequency**	**Proportion of slower timescale fluctuations**	**Total EEG spectrum power**
Peak theta band frequency	NaN	0.002	0.054	0.041
Peak beta band frequency	0.004	NaN	0.104	0.003
Proportion of slower timescale fluctuations	8.57×10^{-7}	0.638	NaN	0.599
Total EEG spectrum power	0.114	0.176	0.149	NaN

B.20. Group averaged early warning signal for a non falling asleep transition period as controlled testing

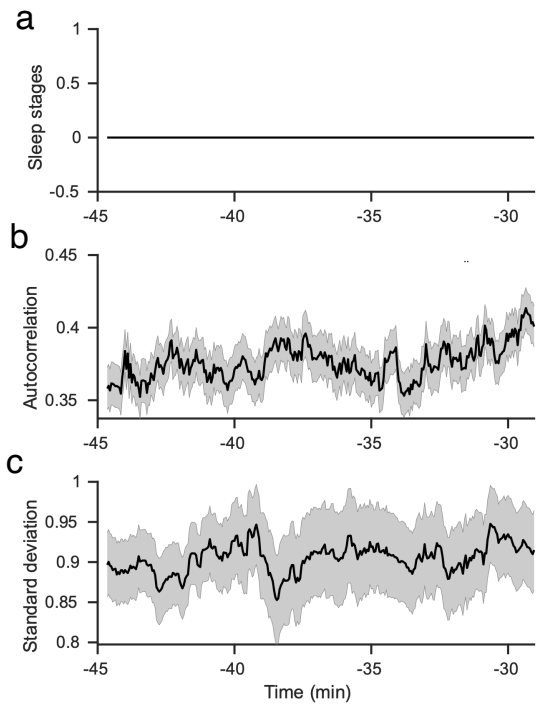


Figure B.7: Early warning signal in the sleep state variable for a control period without falling asleep transition. a. The group-averaged (across 118 participants selected) autocorrelation (ii) and standard deviation (iii) evaluated on a 20-minute non sleep onset period of the s time series for each participant. The average sleep stages (i) for these participants are also plotted for comparison. The vertical red dashed line shows the sleep onset. The shaded area shows the standard error of the mean (s.e.m.) around the mean value (black line).

The EWS results shown here are for the 20-minute non falling asleep transition periods (the 50 to 30 minutes period before sleep onset). A left-tailed two-sample t-test is also applied to detect significant increases (for all three) compared to the first 3 minutes, with a p-value threshold of 0.025, but no significance was detected thus, the plots do not have any horizontal lines.

B.21. The EWS results on exemplary participants and statistical methods for individual-level analysis

Three exemplary participants are chosen here to show their sleep state variable time series s, and the EWS evaluated for them. To test statistically the increase of autocorrelation and standard

deviation on the individual level, a surrogate-based method is used (Dakos et al., 2012). For each participant's time series, 1000 surrogate time series are generated from it using Ebisuzaki's method (Ebisuzaki, 1997). Basically, the algorithm will compute the time series power spectrum but randomly shuffle the phases and do the inversed Fourier transform to get the surrogated time series data. This is done 1000 times, where each surrogate time series will go through the same computation of the EWS. Kendall's tau coefficient is used to capture the correlations between the EWS and time. Since the aim is to test whether the increase of autocorrelation or standard deviation is due to chance, the tau values from the original time series are compared to the 1000 surrogates. Taking a p-value threshold of 0.05, it means that if the tau values from the original time series are greater than 95% of the values in the surrogates, the increase is very unlikely by chance. Of course, for stronger statistical power, more surrogate time series can be generated, but at the cost of a much heavier computation cost, as for one digit further, the computation time will be multiplied by 10. Aside from the computational cost, this approach is post hoc analysis where one needs decently long data to test on, therefore it is not feasible for online prediction, especially in the scenario of sleep.

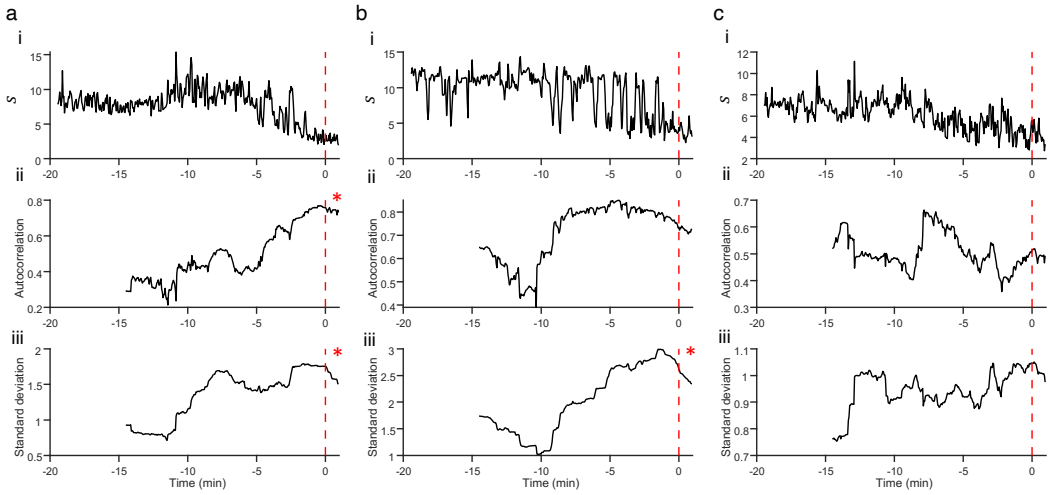


Figure B.8: The EWS results for three exemplary participants (a-c). For each participant, the sleep state variable s time series (i), the autocorrelation (ii) and standard deviation (iii) evaluated from s are plotted. * implies a significant Kendall's tau coefficient between autocorrelation or standard deviation with time, tested against the values of the EWSs of 1000 surrogated time series of s. The p-value threshold is 0.05.

B.22. An exploration of the sleep state variable between healthy and MCI cohorts in the Montreal dataset

This exploration in the Montreal dataset is preliminary. The results are constrained by the sample numbers (30 samples only for the computation of s variable shown below), therefore the dynamics might be a bit noisier; Further analyses have not continued yet on this database. The

s dynamics of both cohorts show clearly the bifurcation pattern, although there are no obvious differences between the dynamics of the two cohorts. The other methods used for evaluating the s parameter are the same.

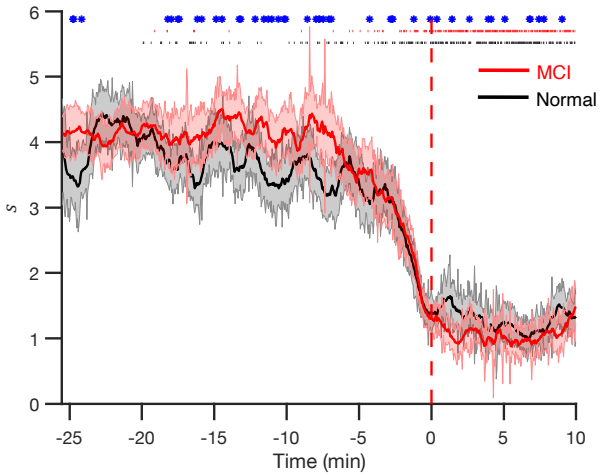


Figure B.9: The dynamics of sleep state variable compared between healthy and MCI cohorts in the Montreal dataset. The group-averaged temporal dynamics of sleep state variable s are computed in the healthy (black) and MCI cohorts separately. Shown are the mean ± standard error of the mean (s.e.m.) across 30 samples. The horizontal lines show timestamps with a significant decrease compared to the mean values of the first 5 minutes of each cohort respectively, using the right-tailed two-sample t-test, with a p-value threshold of 0.025. The blue vertical dashed line shows the sleep onset. The blue * marks significant differences between the healthy and MCI cohorts at the specific timestamps, using a double-sided two-sample t-test with a p-value threshold of 0.05.

B.23. Classification accuracy between bedtime and sleep onset for each region separately and compared to global accuracy

Table B.17: Classification accuracy between bedtime and sleep onset for all features in each individual region. The 4 key governing features are marked in red. The best accuracy of each feature among the 3 regions (channels) is marked in bold.

Feature names	***Frontal (%)***	***Posterior (%)***	***R-Central (%)***	***Global accuracy (%)***
delta power	67.22	**80.19**	79.75	76.08
theta power	67.78	77.13	**77.81**	74.1
alpha power	64.75	65.66	**69.84**	68.14

Continued on next page

Table B.17: Classification accuracy between bedtime and sleep onset for all features in each individual region. The 4 key governing features are marked in red. The best accuracy of each feature among the 3 regions (channels) is marked in bold. (Continued)

beta power	50.38	**50.78**	50.34	49.55
Delta to theta ratio	49.59	50.91	**55.09**	52.61
Delta to alpha ratio	54.88	**62.56**	55.88	57.19
Delta to beta ratio	63.56	**74.31**	73.84	70.95
Theta to alpha ratio	56.56	**65.66**	65.22	61.89
Theta to beta ratio	65.63	77.38	**80.31**	74.59
Alpha to beta ratio	64.47	67.84	**72.03**	69.39
Transition power ratio	62.63	**73.97**	73.34	69.85
Peak delta band frequency	57.28	67.09	**70.84**	64.79
Peak theta band frequency	51.03	**57.19**	54.81	53.81
Peak alpha band frequency	51.06	53.94	**54.88**	52.8
Peak beta band frequency	66.00	61.88	**68.94**	66.24
Temporal coherence delta band	52.25	53.66	**53.84**	52.98
Temporal coherence theta band	**54.34**	53.31	52.47	53.51
Temporal coherence alpha band	54.78	**55.31**	51.97	54.1
Temporal coherence beta band	52.69	**54.22**	53.03	53.17
Phase amplitude coupling DT	54.66	56.16	**57.03**	56.17
Phase amplitude coupling DA	54.03	51.78	**54.13**	53.85
Phase amplitude coupling DB	**57.09**	51.91	55.53	55.29
Phase amplitude coupling TA	50.53	52.25	**53.00**	52.56
Phase amplitude coupling TB	50.81	50.50	**51.22**	51.13
Phase amplitude coupling AB	50.56	50.69	**50.97**	50.57
DN HistogramMode 5	51.00	**52.78**	50.88	51.16
DN HistogramMode 10	**50.44**	50.16	49.09	49.15
CO f1ecac	71.84	**77.22**	65.47	71.64
CO FirstMin ac	65.69	**72.13**	58.41	65.12
CO HistogramAMI even 2 5	62.69	64.97	**70.63**	66.49
CO trev 1 num	57.47	**59.91**	50.38	52.07

Continued on next page

Table B.17: Classification accuracy between bedtime and sleep onset for all features in each individual region. The 4 key governing features are marked in red. The best accuracy of each feature among the 3 regions (channels) is marked in bold. (Continued)

MD hrv classic pnn40	59.56	**68.41**	59.19	63.25
SB BinaryStats mean longstretch1	71.03	**76.03**	74.00	74.18
SB TransitionMatrix 3ac sumdiagcov	**57.53**	56.31	56.91	57.11
PD PeriodicityWang th0 01	64.50	**69.91**	55.94	63.08
CO Embed2 Dist tau d expfit meandiff	74.09	**75.81**	69.69	72.45
IN AutoMutualInfoStats 40 gaussian fmmi	**55.97**	54.59	55.28	55.64
FC LocalSimple mean1 tauresrat	56.06	56.69	**66.00**	60.49
DN OutlierInclude p 001 mdrmd	**52.00**	51.41	51.09	51.32
DN OutlierInclude n 001 mdrmd	**52.34**	**52.34**	51.38	51.6
SP Summaries welch rect area 5 1	69.81	**73.16**	72.44	72.54
SB BinaryStats diff longstretch0	60.06	63.09	**64.22**	65.06
SB MotifThree quantile hh	71.84	**76.56**	68.75	72.15
SC FluctAnal 2 rsrangefit 50 1 logi prop r1	58.72	65.91	**68.97**	63.76
SC FluctAnal 2 dfa 50 1 2 logi prop r1	52.88	**64.88**	53.72	54.9
SP Summaries welch rect centroid	74.19	**76.34**	71.50	73.62
FC LocalSimple mean3 stderr	67.28	70.22	**71.81**	70.63
Mean Accuracy	59.18	**62.66**	61.53	61.12

B.24. Hierchical clustering results for each EEG region separately in the MESA dataset

The list of features in the hierarchical order for the results in the Posterior region (Cz-Oz channel) shown in Figure 5.4:

Table B.18: List of features in the hierarchical order in Figure 5.4. The classification accuracy and mean cluster accuracy are also shown. Key features selected are marked in red and bold, while other cluster centres not selected are marked in bold only.

Feature name	***Accuracy***	***Mean accuracy***
DN HistogramMode 10	50.22%	50.22%
DN HistogramMode 5	52.72%	52.72%
CO trev 1 num	59.91%	59.91%
Peak alpha band frequency	53.94%	53.94%
SB TransitionMatrix 3ac sumdiagcov	**56.44%**	60.72%
CO HistogramAMI even 2 5	65.00%	
IN AutoMutualinfoStats 40 gaussian fmmi	54.59%	54.59%
FC LocalSimple mean1 tauresrat	56.69%	56.69%
MD hr classic pnn40	68.53%	
SB BinaryStats diff longstretch0	63.09%	
PD PeriodicityWang th0 01	69.91%	
CO FirstMin ac	72.12%	
CO Embed2 Dist tau d expfit meandiff	75.81%	
CO f1ecac	77.22%	**72.64%**
SB MotifThree quantile hh	**76.56%**	
SP Summaries welch rect centroid	76.34%	
SP Summaries welch rect area 5 1	73.16%	
FC LocalSimple mean3 stderr	70.22%	
SB BinaryStats mean longstretch1	76.03%	
Peak beta band frequency	61.88%	61.88%
Alpha to beta ratio	**67.84%**	**66.75%**
alpha power	65.66%	
Peak delta band frequency	**67.09%**	**67.09%**
beta power	50.81%	50.81%
SC FluctAnal 2 rsrangefit 50 1 logi prop r1	65.91%	
Theta to beta ratio	77.44%	
theta power	77.16%	

Continued on next page

Table B.18: List of features in the hierarchical order in Figure 5.4. The classification accuracy and mean cluster accuracy are also shown. Key features selected are marked in red and bold, while other cluster centres not selected are marked in bold only. (Continued)

delta power Delta to beta ratio **Transition power ratio** Theta to alpha ratio	80.19% 74.38% **73.97%** 65.69%	
Delta to alpha ratio Delta to theta ratio	**62.59%** 50.78%	56.69%
SC FluctAnal 2 dfa 50 1 2 logi prop r1	64.88%	64.88%
Peak theta band frequency	57.19%	57.19%
Temporal coherence delta band	53.66%	53.66%
Temporal coherence theta band **Temporal coherence alpha band**	53.31% **55.31%**	54.31%
Temporal coherence beta band	54.25%	54.25%
Phase amplitude coupling DB	51.84%	51.84%
Phase amplitude coupling DA	51.84%	51.84%
Phase amplitude coupling DT	56.16%	56.16%
Phase amplitude coupling TA	52.25%	52.25%
Phase amplitude coupling TB	50.47%	50.47%
Phase amplitude coupling AB	50.72%	50.72%
DN OutlierInclude p 001 mdrmd **DN Outlierlnclude n 001 mdrmd**	51.47% **52.34%**	51.91%

The key features from this region have all been interpreted previously. The entropy feature was interpreted in Appendix Table B.4.

The hierarchical clustering result for the Frontal region (Fz-Cz channel):

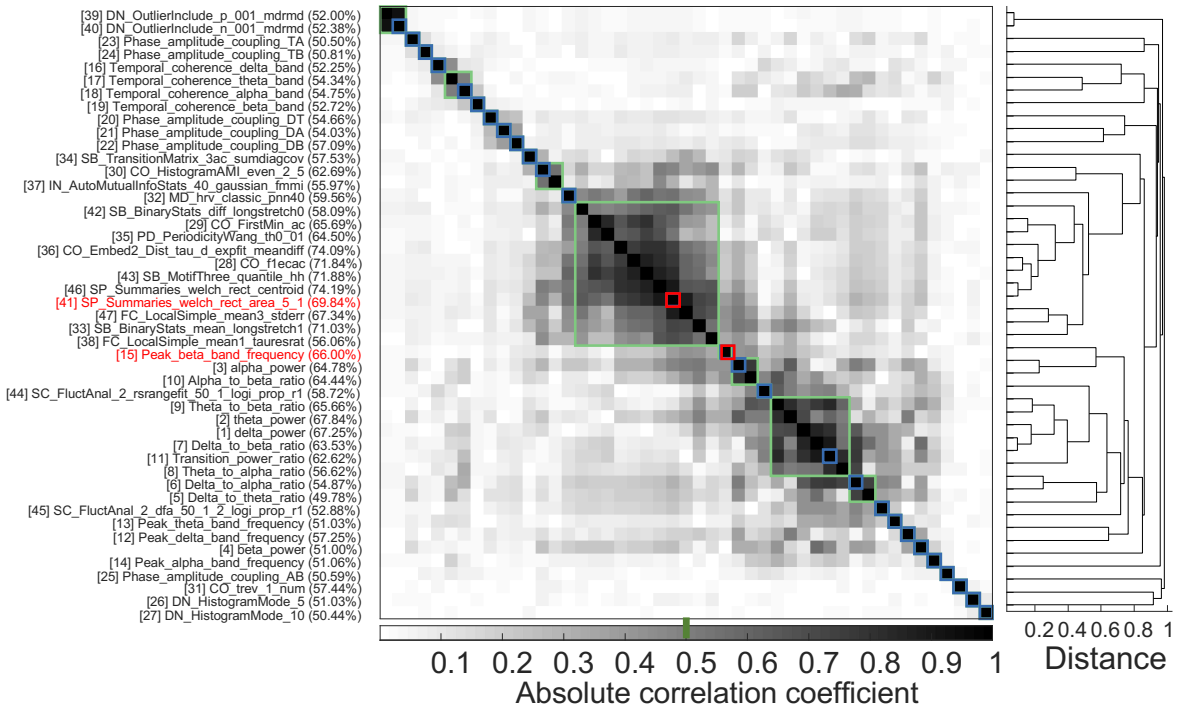


Figure B.10: The hierarchical clustering result of the features for classifying bedtime and sleep onset state in the Frontal region. The feature names, along with its SVM accuracy, are shown on the left side. The dendrogram of the hierarchical clustering is shown on the right with a clustering threshold of 0.5. The green boxes show the features in the same cluster, with blue squares showing the cluster centre. Red squares show the key features selected as the cluster centre with above 65% mean accuracy of all features in the cluster, and the selected key feature names are marked in red.

There are two key features from this region, which are all interpreted previously. Their distribution plots:

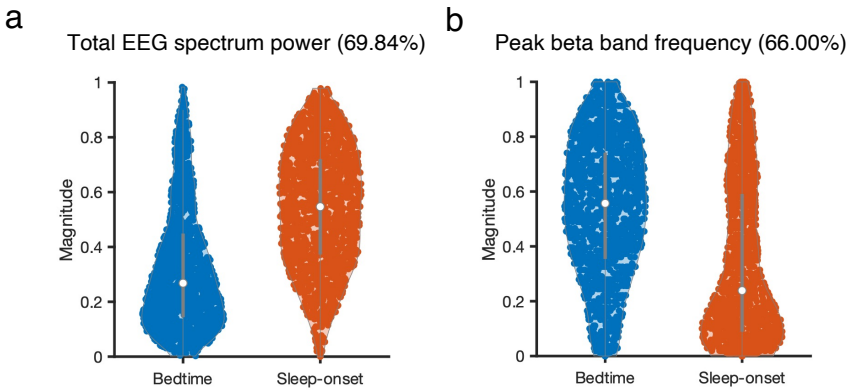


Figure B.11: The distribution comparison between bedtime and sleep onset of the key features selected from the Frontal region. (**a**) for the total EEG spectrum power with 69.84% accuracy; (**b**) for peak beta band frequency with 66.00% accuracy; Blue indicates bedtime, and red indicates sleep onset. Y-axes show the normalised feature values.

The hierarchical clustering result for the Right-Central region (C4-M1 channel):

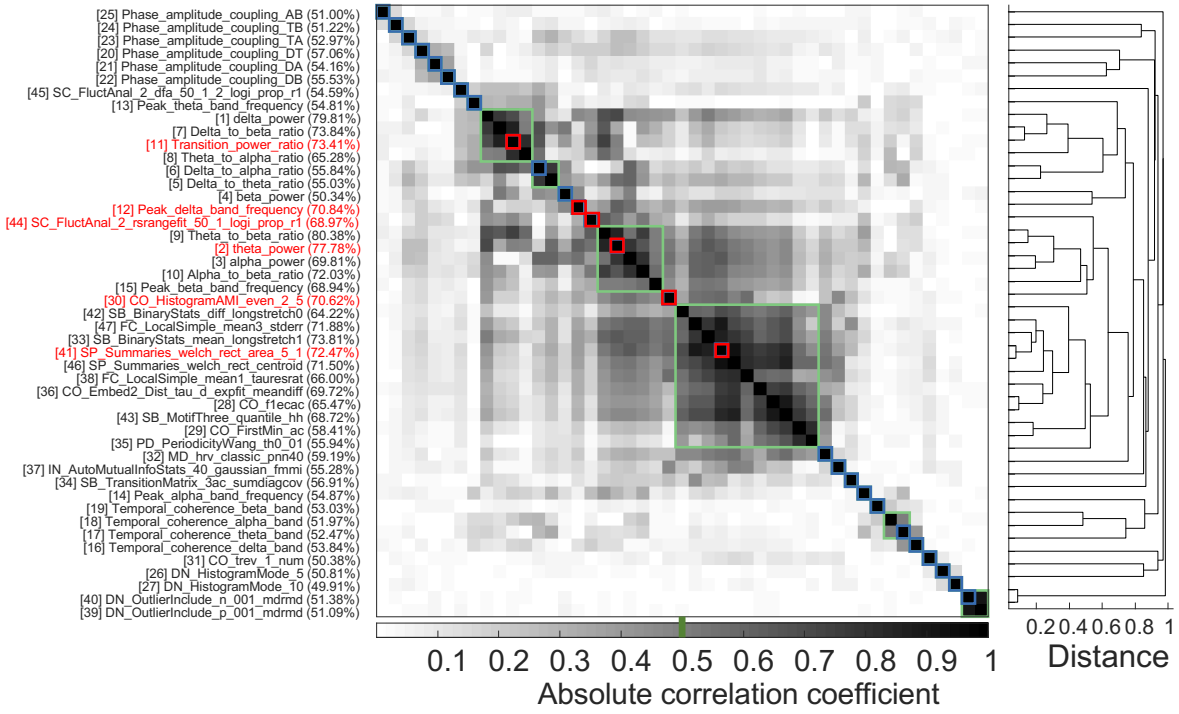


Figure B.12: The hierarchical clustering result of the features for classifying bedtime and sleep onset state in the Right-Central region. The feature names, along with its SVM accuracy, are shown on the left side. The dendrogram of the hierarchical clustering is shown on the right with a clustering threshold of 0.5. The green boxes show the features in the same cluster, with blue squares showing the cluster centre. Red squares show the key features selected as the cluster centre with above 65% mean accuracy of all features in the cluster, and the selected key feature names are marked in red.

There are six key features from this region, and most of them are pretty intuitive or interpreted previously. For the 'Auto mutual information' feature, see previous Appendix Table B.4. The 'Proportion of slower timescale dynamics' measures essentially the same property as previously interpreted in Appendix B.17, but simply using a different algorithm (linearly detrending rather than DFA, Lubba et al. (2019)). Their distribution plots:

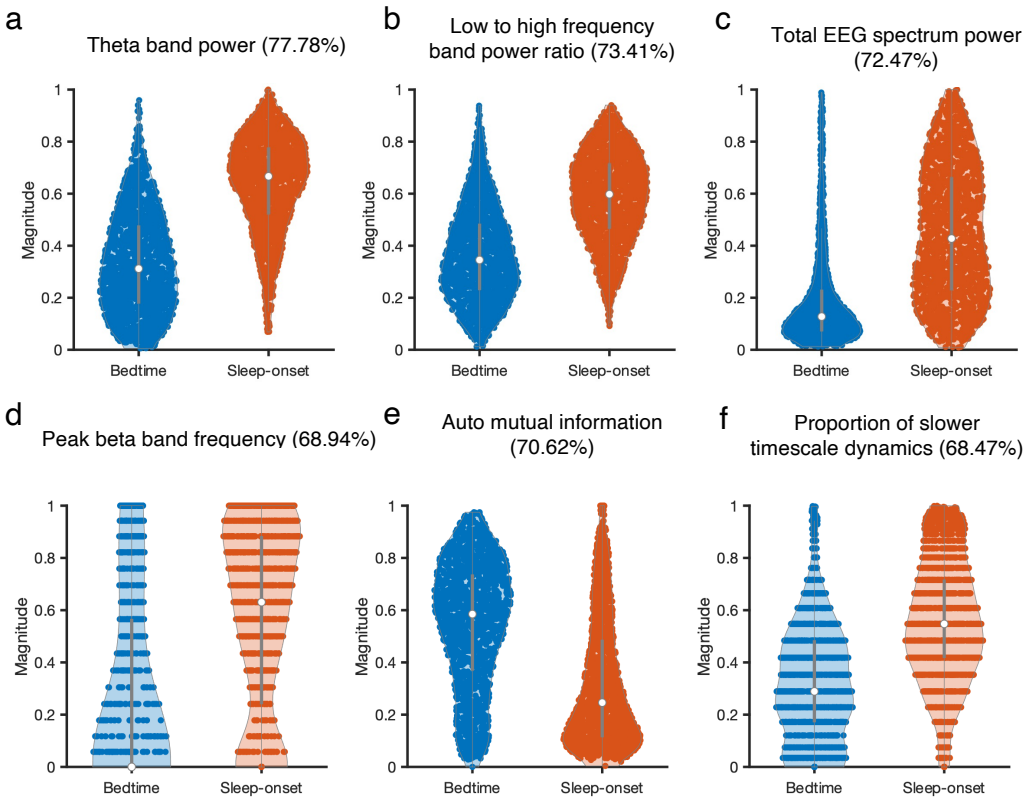


Figure B.13: The distribution comparison between bedtime and sleep onset of the key features selected from the Right-Central region. (**a**) theta band power with 77.78% accuracy; (**b**) low to high frequency band power ratio with 73.41% accuracy; (**c**) total EEG spectrum power with 72.47% accuracy; (**d**) peak beta band frequency with 68.94% accuracy; (**e**) auto mutual information 70.62% accuracy; (**f**) proportion of slower timescale dynamics with 68.47% accuracy. Blue indicates bedtime, and red indicates sleep onset. Y-axes show the normalised feature values.

B.25. The features added and statistics for the classification in Figure 5.5

Table B.19: The features added into the classification model in Figure 5.5 at each step, along with the accuracy and F-score at that step. Shown are the mean ± st.d. for the 10-fold leave-one-out cross-validation.

Feature	Accuracy	F-score
R-Central Theta band power	0.78 ± 0.02	0.78 ± 0.03
Posterior Entropy	0.83 ± 0.02	0.83 ± 0.02
Posterior Low-to-high frequency band power ratio	0.88 ± 0.02	0.87 ± 0.02
R-Central Low-to-high frequency band power ratio	0.89 ± 0.01	0.89 ± 0.01

Continued on next page

Table B.19: The features added into the classification model in Figure 5.5 at each step, along with the accuracy and F-score at that step. Shown are the mean ± st.d. for the 10-fold leave-one-out cross-validation. (Continued)

R-Central Total EEG spectrum power	0.89 ± 0.01	0.89 ± 0.01
R-Central Peak delta band frequency	0.89 ± 0.01	0.89 ± 0.01
R-Central Auto mutual information	0.89 ± 0.01	0.89 ± 0.01
Frontal Total EEG spectrum power	0.89 ± 0.02	0.89 ± 0.02
R-Central Proportion of slower timescale dynamics	0.89 ± 0.02	0.89 ± 0.02
Posterior Alpha to beta band ratio	0.89 ± 0.01	0.89 ± 0.02
Posterior Peak delta band frequency	0.90 ± 0.02	0.90 ± 0.02
Frontal Peak beta band frequency	0.90 ± 0.01	0.90 ± 0.01

B.26. The theoretical model fitting on the s time series dynamics of each region separately

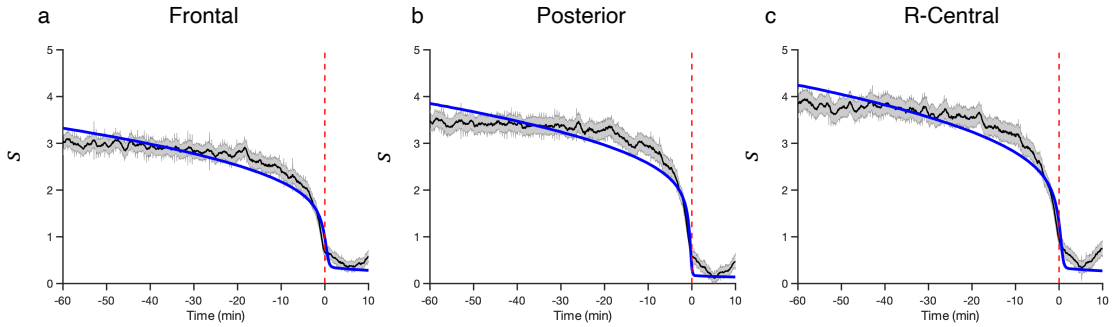


Figure B.14: The theoretical model fitting on the sleep state variable time series dynamics of each region separately. (**a**) for Frontal, (**b**) for Posterior and (**c**) for R-central. shown are mean ± standard error of the mean (s.e.m.) across 200 samples. The black lines are the s time series, with the shaded error bar showing the s.e.m. The blue lines show the results from the best-fitted theoretical model from Equation (4.1).

For Frontal region: $r = 6.73$, $K = 3.48$, $h = 0.52$ and $m = 0.095$, the best-fit model has an R-squared of 0.928 and root-mean-squared-error 0.05; For the Posterior: $r = 6.16$, $K = 4.06$, $h = 0.48$ and $m = 0.099$, the best-fit model has an R-squared of 0.933 and root-mean-squared-error 0.08; For the Right-Central: $r = 5.18$, $K = 4.46$, $h = 0.59$ and $m = 0.092$, the best-fit model has an R-squared of 0.930 and root-mean-squared-error 0.09.

B.27. The continuous feature dynamics for the three different regions in the MESA dataset

Using exactly the same methodology and algorithms (bootstrapping control), the feature dynamics are plotted for the feature data in each region separately. For the dynamics in each region, and for each feature, the feature values are firstly z-score normalised to all timestamps of all participants (all-region z-score normalisation) and then normalised to the sleep-onset values of each region respectively. Therefore one should take care when comparing directly the z-score values between different regions shown in the figure. The features are arranged in the same hierarchical order as previous global results (Appendix Table B.13), for easy visual comparison to global.

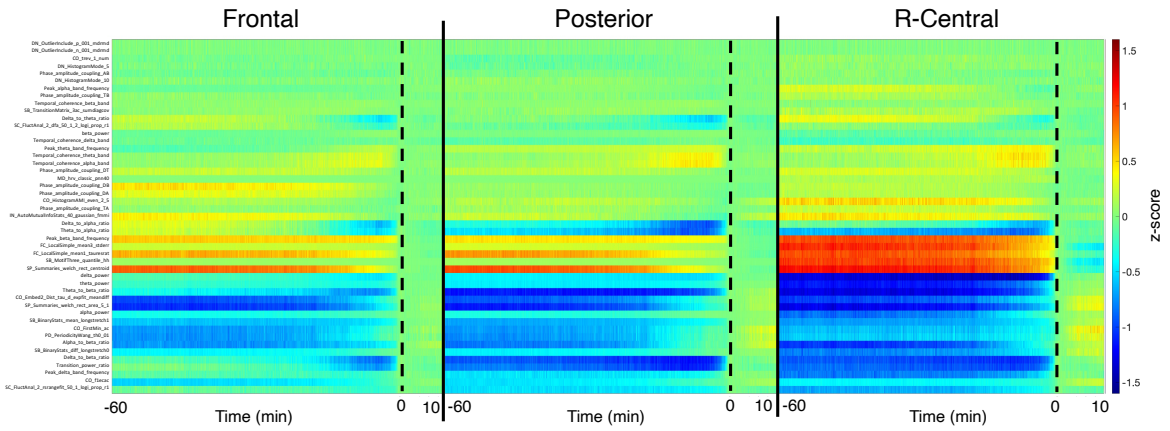


Figure B.15: The continuous feature dynamics for the 3 different regions (Frontal, Posterior and Right-Central) in the MESA dataset. The features are arranged in the same hierarchical order as in Appendix Table B.13 (the same as the group-level plots in Chapter 4), with the original feature names shown on the left side for easy checks. The feature values for each region are z-scores normalised to the sleep-onset values of each region itself, respectively.

In essence, the FPCA shows that the governing dynamics for the three regions are almost similar to each other, with slight differences in the amount of variance explained. The FPC scores for all the features in the different regions are shown in the following table:

Table B.20: The FPC similarity scores between all the features and the FPCs identified from the feature dynamics in each region, respectively. The selected most representative features for each FPC or each region are marked in red and bold.

Feature names	**Frontal**		**Posterior**		**R-Central**	
	FPC1	**FPC2**	**FPC1**	**FPC2**	**FPC1**	**FPC2**

Continued on next page

Table B.20: The FPC similarity scores between all the features and the FPCs identified from the feature dynamics in each region, respectively. The selected most representative features for each FPC or each region are marked in red and bold. (Continued)

delta power	0.84	0.57	-0.99	0.20	0.99	0.17
theta power	0.96	0.30	**-1.00**	0.02	1.00	-0.03
alpha power	1.00	-0.06	-0.96	-0.32	0.99	-0.19
beta power	0.81	-0.51	-0.19	-0.70	0.75	-0.43
Delta to theta ratio	-0.34	0.95	-0.44	**1.00**	-0.76	0.82
Delta to alpha ratio	-0.31	**1.00**	-0.85	0.61	0.65	**1.00**
Delta to beta ratio	0.72	0.72	-0.98	0.24	0.99	0.22
Theta to alpha ratio	-0.26	1.00	-0.94	0.40	0.97	0.29
Theta to beta ratio	0.90	0.45	-1.00	0.06	1.00	0.00
Alpha to beta ratio	1.00	0.05	-0.97	-0.26	0.99	-0.17
Transition power ratio	0.55	0.88	-0.97	0.27	0.99	0.21
Peak delta band frequency	0.87	0.38	-0.99	0.03	1.00	0.02
Peak theta band frequency	0.50	**-0.86**	0.95	-0.34	-0.89	-0.54
Peak alpha band frequency	0.93	-0.04	0.69	0.63	-0.83	0.61
Peak beta band frequency	-0.99	-0.13	**1.00**	-0.04	-1.00	0.01
Temporal coherence delta band	0.77	0.12	-0.87	-0.27	0.93	-0.16
Temporal coherence theta band	-0.70	-0.66	0.66	**-0.83**	-0.73	**-0.85**
Temporal coherence alpha band	-0.69	-0.69	0.72	-0.77	-0.72	-0.84
Temporal coherence beta band	-0.73	-0.24	0.77	-0.36	-0.44	-0.22
Phase amplitude coupling DT	-0.96	0.02	0.94	0.16	-0.97	0.10
Phase amplitude coupling DA	-0.97	0.11	0.67	0.42	-0.93	0.18
Phase amplitude coupling DB	-0.98	0.14	0.72	0.35	-0.94	0.25
Phase amplitude coupling TA	-0.50	0.34	0.75	0.44	-0.92	0.31
Phase amplitude coupling TB	-0.33	0.33	0.06	0.41	-0.80	0.47
Phase amplitude coupling AB	0.67	0.18	-0.32	0.13	-0.16	0.57
DN HistogramMode 5	0.31	0.08	-0.70	-0.20	-0.50	-0.02
DN HistogramMode 10	-0.08	-0.05	0.52	0.21	-0.20	0.07
CO f1ecac	1.00	-0.02	-0.99	-0.16	1.00	0.01

Continued on next page

Table B.20: The FPC similarity scores between all the features and the FPCs identified from the feature dynamics in each region, respectively. The selected most representative features for each FPC or each region are marked in red and bold. (Continued)

CO FirstMin ac	1.00	-0.05	-0.98	-0.20	0.99	0.01
CO HistogramAMI even 2 5	-0.88	0.28	0.90	0.23	-0.98	0.10
CO trev 1 num	0.48	-0.17	-0.75	-0.30	-0.66	0.12
MD hrv classic pnn40	-0.99	-0.01	0.99	0.15	-0.94	-0.17
SB BinaryStats mean longstretch1	0.99	0.11	-1.00	-0.06	1.00	0.00
SB TransitionMatrix 3ac sumdiagcov	-0.73	-0.16	0.63	-0.09	-0.86	-0.01
PD PeriodicityWang th0 01	0.99	-0.09	-0.98	-0.23	0.99	-0.01
CO Embed2 Dist tau d expfit meandiff	1.00	-0.07	-0.98	-0.21	0.99	-0.12
IN AutoMutualInfoStats 40 gaussian fmmi	-0.96	0.23	0.72	0.66	-0.98	0.10
FC LocalSimple mean1 tauresrat	-1.00	-0.01	0.99	0.19	-1.00	0.07
DN OutlierInclude p 001 mdrmd	0.10	0.03	-0.08	0.00	0.05	0.03
DN OutlierInclude n 001 mdrmd	0.11	0.03	-0.11	0.03	0.10	0.01
SP Summaries welch rect area 5 1	**1.00**	-0.02	-0.99	-0.20	**1.00**	-0.07
SB BinaryStats diff longstretch0	1.00	-0.06	-0.98	-0.22	1.00	-0.07
SB MotifThree quantile hh	**-1.00**	0.05	0.99	0.18	**-1.00**	-0.01
SC FluctAnal 2 rsrangefit 50 1 logi prop r1	0.76	0.58	-1.00	0.07	1.00	-0.03
SC FluctAnal 2 dfa 50 1 2 logi prop r1	-0.51	0.81	-0.74	0.68	0.11	0.94
SP Summaries welch rect centroid	-1.00	0.06	0.99	0.20	-1.00	0.05
FC LocalSimple mean3 stderr	-1.00	-0.07	0.99	0.16	-1.00	0.06

B.28. The distribution statistics of the two key feature types at bedtime and sleep onset between all region pairs (for Figure 5.9)

Table B.21: The distribution statistics of the two key feature types at bedtime and sleep onset between all region pairs (for Figure 5.9). Shown are the mean ± st.d. for the feature values distribution statistics, and the p-values are from the two-sided two-sample t-test.

Names	**Bedtime**	**Sleep Onset**	**p-value**

Continued on next page

Table B.21: The distribution statistics of the two key feature types at bedtime and sleep onset between all region pairs (for Figure 5.9). Shown are the mean ± st.d. for the feature values distribution statistics, and the p-values are from the two-sided two-sample t-test. (Continued)

Frontal vs Posterior Pearson	0.17 ± 0.26	-0.11 ± 0.22	6.35E-54
Frontal vs R-Central Pearson	0.07 ± 0.22	-0.19 ± 0.19	2.40E-63
Posterior vs R-Central Pearson	0.33 ± 0.22	0.50 ± 0.17	5.72E-33
Frontal to Posterior Transfer Entropy	0.12 ± 0.09	0.36 ± 0.10	8.02E-171
Posterior to Frontal Transfer Entropy	0.12 ± 0.08	0.37 ± 0.10	9.48E-177
Frontal to R-Central Transfer Entropy	0.09 ± 0.07	0.32 ± 0.11	3.88E-163
R-Central to Frontal Transfer Entropy	0.09 ± 0.07	0.31 ± 0.11	2.52E-158
Posterior to R-Central Transfer Entropy	0.08 ± 0.07	0.31 ± 0.10	1.23E-174
R-Central to Posterior Transfer Entropy	0.08 ± 0.07	0.28 ± 0.09	1.54E-169

B.29. Hierarchical clustering of all the inter-regional interaction features without combining into categories in the MESA dataset

The top 60 features (based on their individual classification accuracy between bedtime and sleep onset) are chosen out of the 240 interactional features, and the same hierarchical clustering pipeline is applied to extract the key features without consideration of the feature type. The result is shown in the below figure:

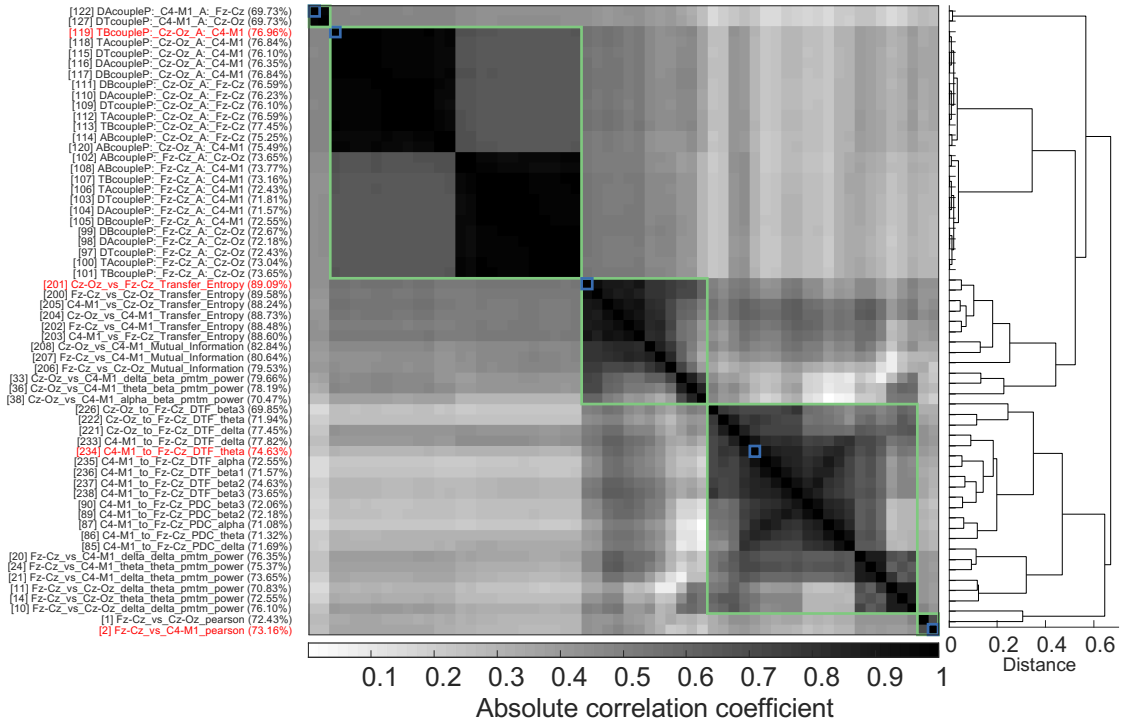


Figure B.16: The hierarchical clustering of the 60 top-performing interactional features in the MESA dataset. The feature names, along with their classification accuracy, are shown on the left side. The correlation map between all features is shown with the absolute correlation coefficient shown as the grey scales (colour bar at the bottom). The green boxes show the features clustered together, and the blue squares show the cluster centroid features. Red squares and feature names marked in red show the key features selected. The dendrogram for hierarchical clustering is shown on the right side.

There are four key features selected using a criterion of cluster centroid with a mean cluster accuracy above 70% accuracy: the coupling between theta band phase of Posterior and beta band amplitude of R-Central; the transfer entropy from Posterior to Frontal; the DTF at theta band from Right-Central to Frontal, and the Pearson correlation between Frontal and Right-Central. The distributions for the transfer entropy and Pearson correlation have already been shown in the main text, here the two other features are shown.

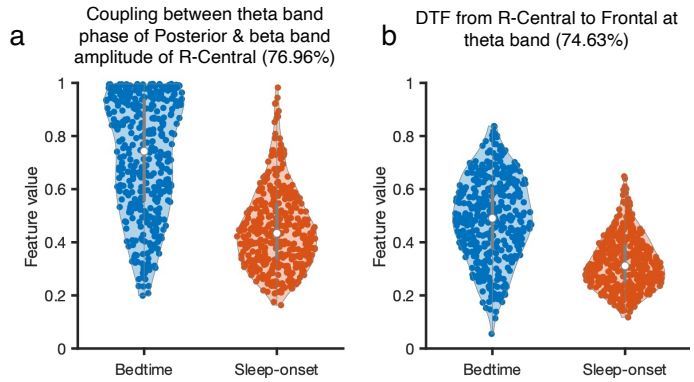


Figure B.17: The specific inter-regional features distribution differences between bedtime and sleep-onset (selected from hierarchical clustering). a. the coupling between theta band phase of Posterior and beta band amplitude of R-Central; **b**. the DTF at theta band from R-Central (Right-Central) to Frontal. The p-values using a two-sided two-sample t-test between bedtime and sleep onset show significance in both features ($p < 0.0005$). Blue shows bedtime and red shows sleep-onset. The accuracy for that specific feature is shown along with the feature name.

B.30. The distribution statistics of all the TE features at bedtime and sleep onset between all region pairs in the Montreal dataset (for Figure 5.11)

Table B.22: The distribution statistics of all the TE features at bedtime and sleep onset between all region pairs in the Montreal dataset (for Figure 5.11). Shown are the mean ± st.d. for the feature values distribution statistics, and the p-values are from the two-sided two-sample t-test.

Name	Bedtime	Sleep onset	p-value
Frontal to Central Transfer Entropy	0.11 ± 0.06	0.19 ± 0.07	1.86E-11
Central to Frontal Transfer Entropy	0.10 ± 0.06	0.21 ± 0.07	6.78E-17
Frontal to Parietal Transfer Entropy	0.15 ± 0.08	0.24 ± 0.09	2.93E-10
Parietal to Frontal Transfer Entropy	0.13 ± 0.07	0.24 ± 0.08	2.90E-12
Central to Parietal Transfer Entropy	0.11 ± 0.06	0.21 ± 0.07	8.91E-14
Parietal to Central Transfer Entropy	0.10 ± 0.06	0.18 ± 0.06	2.21E-10

B.31. Results for the re-referenced Montreal dataset and comparison to the previous results

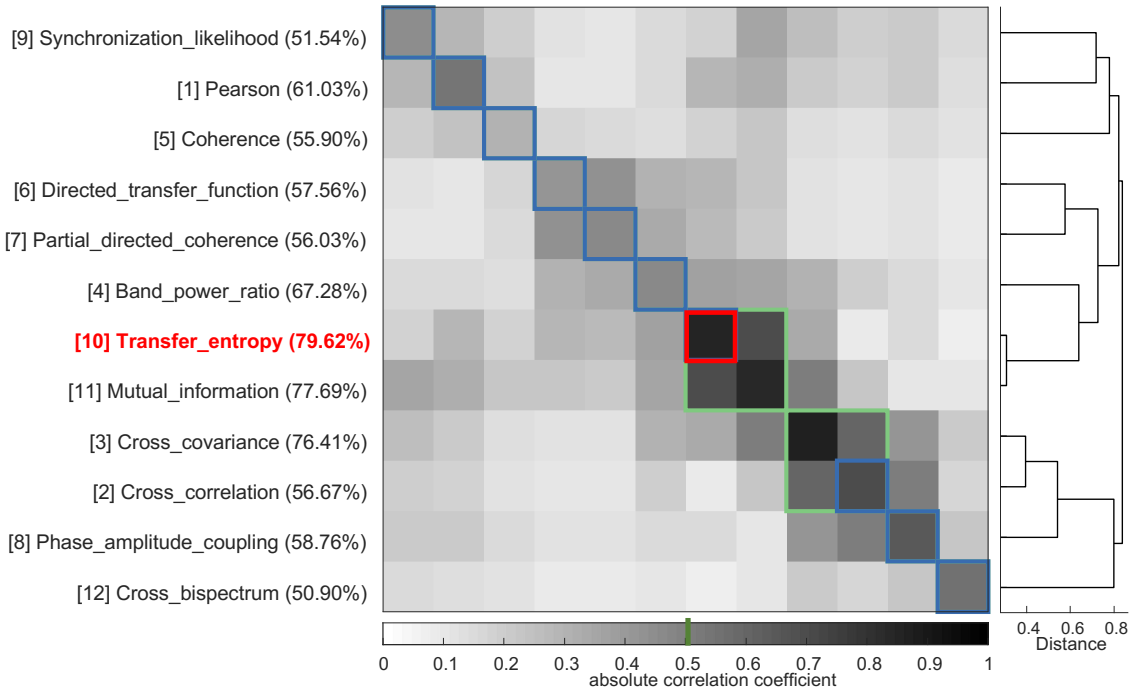


Figure B.18: Hierarchical clustering of the inter-regional interaction features (arranged in each category) that distinguish bedtime and sleep-onset in the re-referenced Montreal dataset (Fz-Cz, Cz-Pz, C4). The feature type name, along with the average accuracy, is shown on the left side. The average within-type correlation is shown on the left diagonal, with the between-type correlation shown between each pair, and the correlation strength is shown as the grey scale (colour bar at the bottom). The dendrogram is shown on the right side, with a clustering threshold of 0.5. The green boxes show the feature types clustered together, and the blue boxes show cluster centres. The selected key feature types are marked in red names and red squares.

The classification accuracies between bedtime and sleep onset for all feature types are compared in the below table, between the MESA dataset, the raw Montreal dataset (Fz, Cz and Pz), and the re-referenced Montreal dataset (Fz-Cz, Cz-Pz, and C4).

Table B.23: Table of classification accuracies for each inter-regional feature type in MESA and Montreal datasets (Raw and Re-referenced). Shown are the mean accuracies of each feature type. The two key selected feature types (in the MESA dataset) are marked in red.

Feature type	MESA	Montreal Raw	Montreal Re-Ref	Re-Ref – Raw
Pearson	71.20%	53.08%	61.03%	7.95%
Cross correlation	68.83%	61.03%	56.67%	-4.36%

Continued on next page

Table B.23: Table of classification accuracies for each inter-regional feature type in MESA and Montreal datasets (Raw and Re-referenced). Shown are the mean accuracies of each feature type. The two key selected feature types (in the MESA dataset) are marked in red. (Continued)

Cross covariance	64.13%	74.87%	76.41%	1.54%
Power band ratio	64.62%	67.56%	67.28%	-0.28%
Coherence	58.26%	60.00%	55.90%	-4.10%
Partial directed coherence	62.15%	57.59%	56.03%	-1.56%
Directed transfer function	62.84%	56.50%	57.56%	1.06%
Phase amplitude coupling	72.83%	62.76%	58.76%	-4.00%
Synchronization likelihood	51.12%	59.74%	51.54%	-8.20%
Cross bispectrum	50.26%	51.37%	50.90%	-0.47%
Transfer entropy	88.79%	74.87%	79.62%	4.75%
Mutual information	81.00%	52.05%	77.69%	25.64%

B.32. The continuous inter-regional interactional dynamics for all features during falling asleep

The exact feature names for this plot will not be listed due to the space limit and very little scientific value.

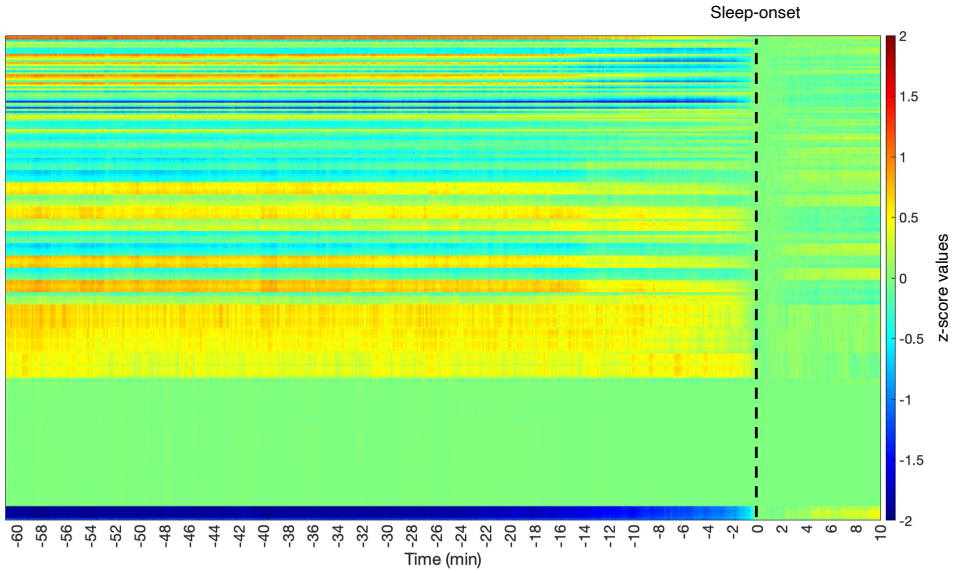


Figure B.19: The continuous inter-regional interactional dynamics for all features during falling asleep, from 60 minutes prior to sleep onset to 10 minutes after. The feature values are z-score normalised to the 1-minute sleep onset. Sleep onset is marked in the dashed vertical black line. The features are arranged (on the y-axis) in a sequential order based on the feature type: Pearson (3 features), Cross-correlation (3), Cross-covariance (3), Power band ratios (30), Coherence (21), DTF (36), PDC (36), PAC (36), Cross-bispectrum (60), Synchronisation likelihood (3), MI (3) and TE (6). In total, 240 features are shown.

B.33. The continuous inter-regional interactional feature dynamics plot for some specific feature categories

For all the six features in the TE type:

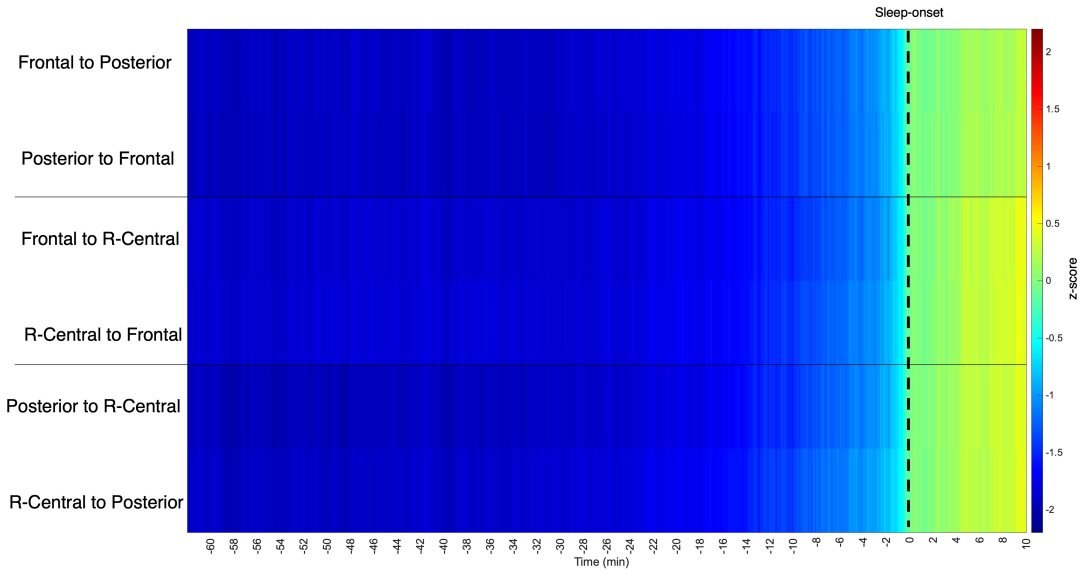


Figure B.20: The continuous feature dynamics of the six TE features in the MESA dataset from 60 minutes prior to sleep onset to 10 minutes after. The features are z-score normalised to the sleep onset (with colour scale showing z-score values), and the specific region pairs (including directions) are shown on the left. The vertical dashed line in the colourmap shows the sleep onset, and the horizontal lines divide the different region pairs.

For all the thirty features in the DTF type:

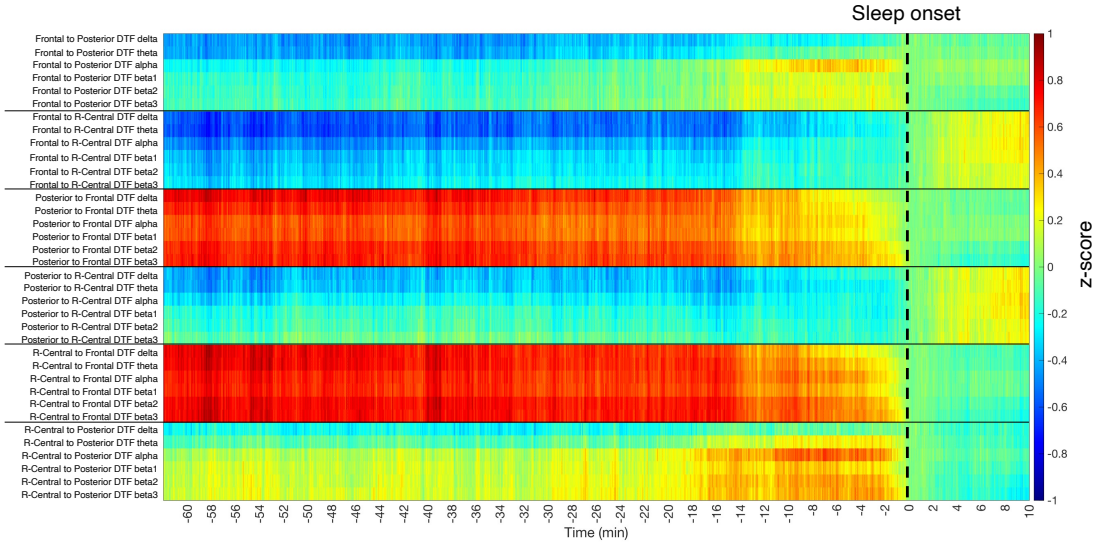


Figure B.21: The continuous feature dynamics of the thirty DTF features in the MESA dataset from 60 minutes prior to sleep onset to 10 minutes after. The features are z-score normalised to the sleep onset (with colour scale showing z-score values), and the specific region pairs (including directions) are shown on the left. The vertical dashed line in the colourmap shows the sleep onset, and the horizontal lines divide the different region pairs.

For all the thirty features in the cross-channel power band ratio type:

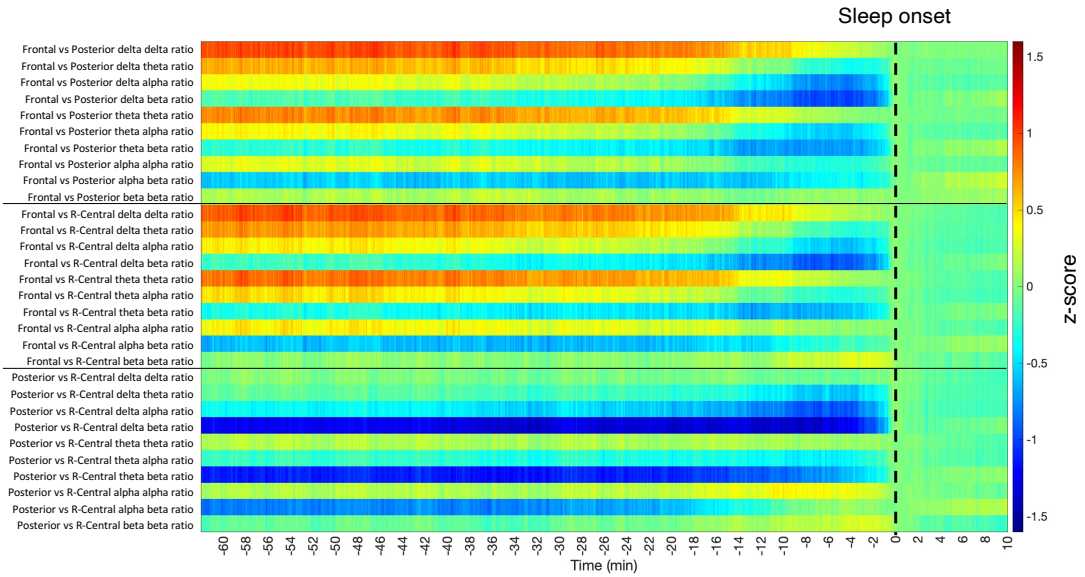


Figure B.22: The continuous feature dynamics of the thirty cross-channel power band ratio features in the MESA dataset from 60 minutes prior to sleep onset to 10 minutes after. The features are z-score normalised to the sleep onset (with colour scale showing z-score values), and the specific region pairs (including directions) are shown on the left. The vertical dashed line in the colourmap shows the sleep onset, and the horizontal lines divide the different region pairs.

The table of inter-regional FPC similarity scores for all the features mentioned previously (TE, DTF and power band ratios):

Table B.24: Table of the inter-regional FPC similarity scores for all the TE (6), DTF (30), and power band ratio (30) features (between all region pairs and directions).

Names	**FPC1**	**FPC2**
Transfer entropy		
Frontal to Posterior Transfer Entropy	1.00	-0.01
Posterior to Frontal Transfer Entropy	1.00	-0.01
Frontal to R-Central Transfer Entropy	1.00	0.01
R-Central to Frontal Transfer Entropy	1.00	0.00
Posterior to R-Central Transfer Entropy	1.00	-0.01
R-Central to Posterior Transfer Entropy	1.00	-0.03
DTF		
Frontal to Posterior DTF delta	0.98	-0.11

Continued on next page

Table B.24: Table of the inter-regional FPC similarity scores for all the TE (6), DTF (30), and power band ratio (30) features (between all region pairs and directions). (Continued)

Frontal to Posterior DTF theta	0.94	-0.32
Frontal to Posterior DTF alpha	0.39	-0.92
Frontal to Posterior DTF beta1	0.49	-0.82
Frontal to Posterior DTF beta2	0.35	-0.89
Frontal to Posterior DTF beta3	0.29	-0.88
Frontal to R-Central DTF delta	0.99	-0.08
Frontal to R-Central DTF theta	0.99	-0.08
Frontal to R-Central DTF alpha	0.98	-0.13
Frontal to R-Central DTF beta1	0.98	-0.09
Frontal to R-Central DTF beta2	0.97	-0.12
Frontal to R-Central DTF beta3	0.97	-0.07
Posterior to Frontal DTF delta	-0.99	0.17
Posterior to Frontal DTF theta	-0.99	0.12
Posterior to Frontal DTF alpha	-0.99	0.08
Posterior to Frontal DTF beta1	-0.99	0.10
Posterior to Frontal DTF beta2	-1.00	0.05
Posterior to Frontal DTF beta3	-1.00	0.04
Posterior to R-Central DTF delta	0.98	-0.04
Posterior to R-Central DTF theta	0.98	0.05
Posterior to R-Central DTF alpha	0.97	0.10
Posterior to R-Central DTF beta1	0.90	0.27
Posterior to R-Central DTF beta2	0.88	0.25
Posterior to R-Central DTF beta3	0.66	0.48
R-Central to Frontal DTF delta	-0.99	0.13
R-Central to Frontal DTF theta	-1.00	0.07
R-Central to Frontal DTF alpha	-1.00	0.01
R-Central to Frontal DTF beta1	-1.00	0.04
R-Central to Frontal DTF beta2	-1.00	0.00

Continued on next page

Table B.24: Table of the inter-regional FPC similarity scores for all the TE (6), DTF (30), and power band ratio (30) features (between all region pairs and directions). (Continued)

R-Central to Frontal DTF beta3	-1.00	-0.01
R-Central to Posterior DTF delta	0.92	-0.24
R-Central to Posterior DTF theta	0.07	-0.91
R-Central to Posterior DTF alpha	-0.71	-0.68
R-Central to Posterior DTF beta1	-0.85	-0.46
R-Central to Posterior DTF beta2	-0.83	-0.52
R-Central to Posterior DTF beta3	-0.85	-0.49
Power band ratios		
Frontal vs Posterior delta delta ratio	-0.99	0.16
Frontal vs Posterior delta theta ratio	-0.84	0.55
Frontal vs Posterior delta alpha ratio	-0.19	1.00
Frontal vs Posterior delta beta ratio	0.74	0.68
Frontal vs Posterior theta theta ratio	-0.96	0.26
Frontal vs Posterior theta alpha ratio	-0.42	0.92
Frontal vs Posterior theta beta ratio	0.87	0.49
Frontal vs Posterior alpha alpha ratio	-0.56	0.82
Frontal vs Posterior alpha beta ratio	0.99	0.07
Frontal vs Posterior beta beta ratio	-0.82	0.26
Frontal vs R-Central delta delta ratio	-0.98	0.19
Frontal vs R-Central delta theta ratio	-0.88	0.47
Frontal vs R-Central delta alpha ratio	-0.51	0.87
Frontal vs R-Central delta beta ratio	0.80	0.61
Frontal vs R-Central theta theta ratio	-0.97	0.22
Frontal vs R-Central theta alpha ratio	-0.78	0.63
Frontal vs R-Central theta beta ratio	0.93	0.36
Frontal vs R-Central alpha alpha ratio	-0.95	0.26
Frontal vs R-Central alpha beta ratio	0.99	-0.06
Frontal vs R-Central beta beta ratio	-0.48	-0.66

Continued on next page

Table B.24: Table of the inter-regional FPC similarity scores for all the TE (6), DTF (30), and power band ratio (30) features (between all region pairs and directions). (Continued)

Posterior vs R-Central delta delta ratio	0.35	0.35
Posterior vs R-Central delta theta ratio	0.75	0.64
Posterior vs R-Central delta alpha ratio	0.90	0.42
Posterior vs R-Central delta beta ratio	0.99	0.14
Posterior vs R-Central theta theta ratio	-0.91	0.00
Posterior vs R-Central theta alpha ratio	0.93	0.31
Posterior vs R-Central theta beta ratio	1.00	-0.01
Posterior vs R-Central alpha alpha ratio	-0.84	-0.48
Posterior vs R-Central alpha beta ratio	0.94	-0.33
Posterior vs R-Central beta beta ratio	0.23	-0.86

B.34. A zoom-in of specific regional band power dynamics for illustration

This appendix provides clear plots of the Frontal delta-band dynamics, and the alpha dynamics from Right-Central and Posterior; The exact dynamics have previously been shown in Appendix B.27 already but in as a colourmap.

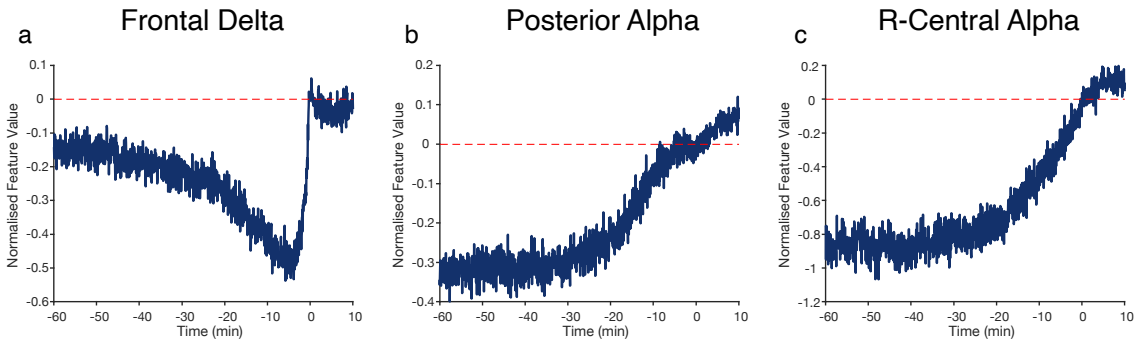


Figure B.23: The Frontal delta band dynamics (a) and the alpha dynamics in the Posterior (b) and Right-Central (c) region, from 60 minutes prior to sleep onset to 10 minutes after. The feature values are z-score normalised to the sleep onset, with the red dashed line showing the z-score value as 0. This plot is a zoom-in of the three specific features in Appendix Figure B.15.

B.35. The p-value tables for the Granger causality test between all the regional and inter-regional FPC dynamics

Table B.25: The p-value tables for the Granger causality test between all pairs of the regional and inter-regional FPC dynamics. The significant ones are marked in bold and red. The left diagonal is NaN, as causal interactions between oneself have no meaning.

From \ To	FPC1 Frontal	FPC1 Posterior	FPC1 R-Central	FPC1 Inter-Regional	FPC2 Frontal	FPC2 Posterior	FPC2 R-Central	FPC2 Inter-Regional
FPC1 Frontal	NaN	0.71	0.10	0.30	0.74	0.10	0.36	0.06
FPC1 Posterior	0.48	NaN	0.43	0.04	0.81	0.47	0.36	**4.85E-03**
FPC1 R-Central	0.04	0.09	NaN	0.15	0.96	0.67	0.83	0.24
FPC1 Inter-Regional	0.10	**0.01**	**8.74E-04**	NaN	0.56	0.97	0.95	0.18
FPC2 Frontal	0.18	0.56	0.51	0.06	NaN	0.46	0.38	**0.01**
FPC2 Posterior	0.49	0.22	1.00	0.45	**9.63E-04**	NaN	**0.01**	0.35
FPC2 R-Central	0.11	0.03	0.40	0.98	0.53	0.84	NaN	0.06
FPC2 Inter-Regional	0.51	0.36	0.80	**5.38E-04**	0.41	0.59	0.43	NaN

B.36. Phase histogram at stimulation onset for all 32 channels

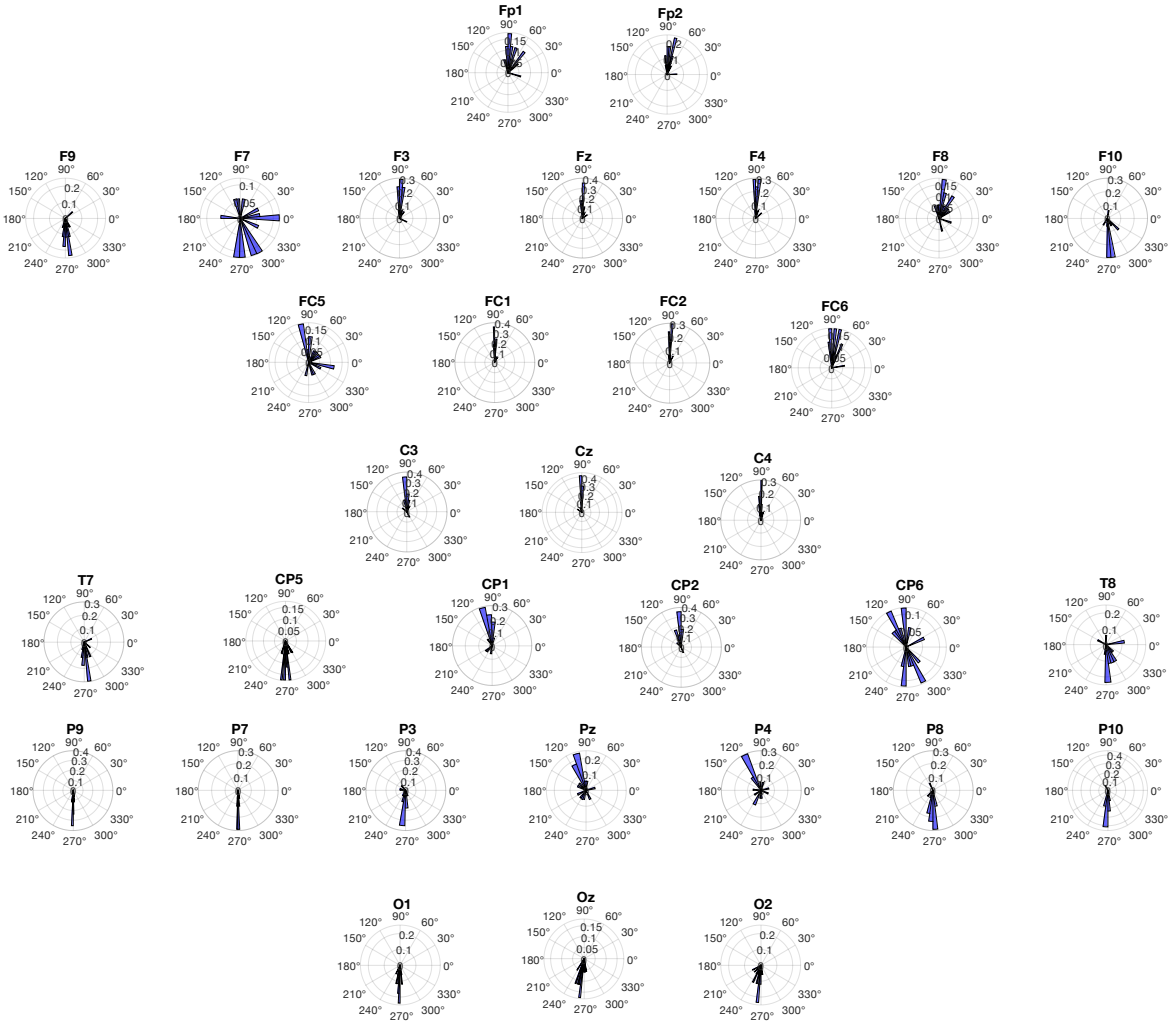


Figure B.24: Phase histogram at stimulation onset for all 32 channels, arranged in a frontal to posterior sequence. The channel names are marked above the histogram, with probability density also marked on the polar plots. The phase histograms are plotted with bin size 10°.

The phase distribution statistics of all 32 EEG channels (Notice again that Fz, P9 and P10 were interpolated):

Table B.26: The phase distribution statistics at stimulation onset for all 32 EEG channels, with the Fz, P9 and P10 signals interpolated.

Channel	Resultant	Rayleigh p-value	Mean phase (°)	St.d. phase
Fp1	0.90	1.84E-08	74.86	26.04

Continued on next page

Table B.26: The phase distribution statistics at stimulation onset for all 32 EEG channels, with the Fz, P9 and P10 signals interpolated. (Continued)

Fz	0.97	6.36E-11	84.93	13.55
F3	0.91	8.89E-09	82.26	24.52
F7	0.40	0.06	-39.93	62.71
F9	0.89	3.01E-08	-82.39	27.07
FC5	0.50	0.01	42.33	57.26
FC1	0.99	9.81E-12	87.41	8.30
C3	0.85	1.88E-07	89.28	30.90
T7	0.90	1.46E-08	-78.89	25.56
CP5	0.98	3.58E-11	-87.06	12.08
CP1	0.70	0.00	106.89	44.69
Pz	0.45	0.03	127.01	60.28
P3	0.79	4.15E-06	-108.00	37.50
P7	1.00	3.58E-12	-89.56	4.04
P9	1.00	3.81E-12	-89.31	4.41
O1	0.99	9.31E-12	-95.10	8.13
Oz	0.99	1.58E-11	-99.22	9.80
O2	0.96	2.80E-10	-104.26	17.05
P10	0.89	3.30E-08	-94.02	27.26
P8	0.87	7.07E-08	-97.06	28.85
P4	0.30	0.21	135.65	67.62
CP2	0.86	1.21E-07	100.36	29.97
CP6	0.03	0.99	48.01	79.80
T8	0.66	0.00	-64.56	47.06
C4	0.99	6.59E-12	89.40	6.89
Cz	0.97	5.31E-11	95.99	13.10
FC2	0.99	6.99E-12	87.90	7.11
FC6	0.95	5.89E-10	79.16	18.71
F10	0.84	3.46E-07	-78.92	32.18

Continued on next page

Table B.26: The phase distribution statistics at stimulation onset for all 32 EEG channels, with the Fz, P9 and P10 signals interpolated. (Continued)

F8	0.79	3.26E-06	64.35	36.97
F4	0.98	2.43E-11	85.15	11.03
Fp2	0.94	9.69E-10	79.85	19.80

B.37. Phase histogram at stimulation onset for Fz channel in the single-channel phase-locking EEG system

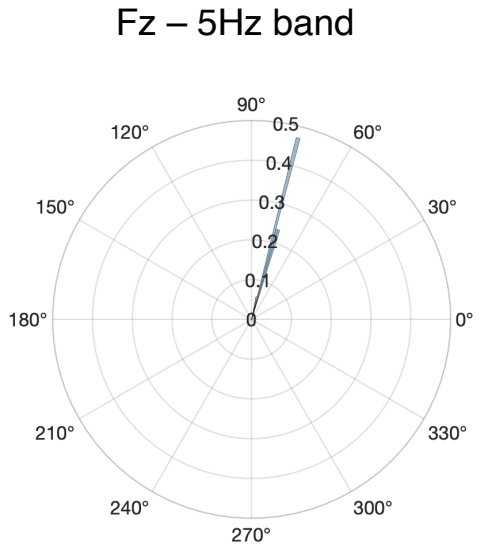


Figure B.25: Phase histogram at stimulation onset for Fz channel in the single-channel phase-locking EEG system. The histogram was plotted with a bin size of 10°, with probability density shown on the polar plot.

B.38. Feature-based classification between audio and control conditions in the last 3 minutes of nap

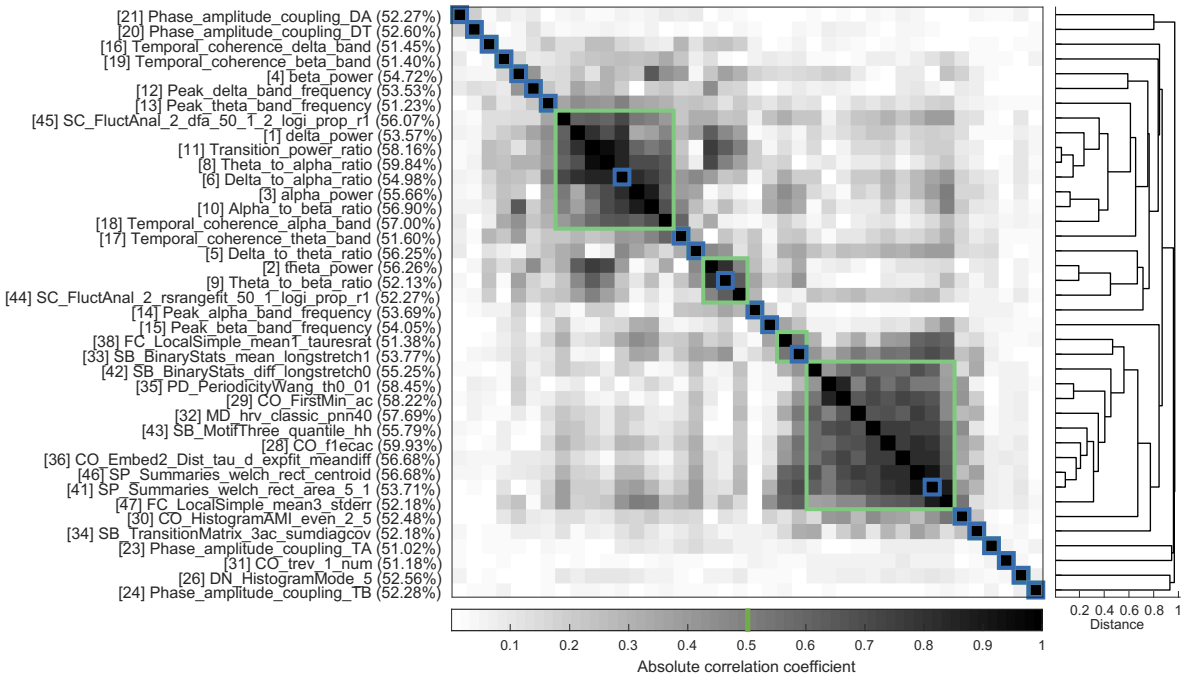


Figure B.26: Feature-based classification between audio and control conditions in the last 3 minutes of nap. The feature names, along with their accuracies (linear SVM model), are shown on the left side. The correlation map shows the inter-feature Pearson correlations (from 0 to 1) with colour scales shown on the bottom. The dendrogram for the hierarchical clustering is shown on the right side, with a clustering threshold of 0.5. Green boxes show the features that are clustered together; Blue squares show the cluster centroids.

B.39. The model statistics for FNAT performance comparison between audio and control groups in Figure 6.11

The percentage of each response type was compared and contrasted using the multi-nomial logistic regression model. The model was fitted in R using `multinom` function with the formula 'ResponseType ~ Condition', with 'Target' response type as the intercept. The model statistics were evaluated using the Type II Wald Chi-squared tests with Kenward-Roger correction for degrees of freedom, with results: Degree of freedom 2, $\chi^2 = 6.33$ and $p = 0.04$. The model coefficients denoting the relative probability (contrasted to the level Target): $\frac{P(Response=`Foil')}{P(Response=`Target')} = e^{b_1} = 0.86$ and $p = 0.01$, and $\frac{P(Response=`Distractor')}{P(Response=`Target')} = e^{b_2} = 0.92$ and $p = 0.32$. The post hoc contrasts from the estimated marginal means are shown in the table below.

Table B.27: The post hoc contrasts of the percentage of response in each response type between audio and control groups. SE: standard error; Df: Degree of Freedom; t-stat: t-statistic.

Contrasts in response percentage	**Estimate**	**SE**	**Df**	**t-stat**	**p-value**
Target, Control-Audio	-0.029	0.002	4	-2.391	0.075
Foil, Control-Audio	0.025	0.011	4	2.306	0.082
Distractor, Control-Audio	0.004	0.008	4	0.488	0.651

The GLME model was fitted for the prediction of response time (RT) using the formula: 'RT ∼ StimType:RespType + (1|S_ID) + (1|block)', where 'StimType' denotes the stimulation type, 'RespType' denotes the response type, 'S_ID' denotes the participant ID and 'block' denotes the block ID. The model statistics were evaluated using Type II Wald Chi-squared tests with Kenward-Roger correction for degrees of freedom.

Table B.28: The GLME model statistics for prediction of the response time (RT). Shown is the analysis of variance (ANOVA) table using Type II Wald Chi-squared tests with Kenward-Roger correction for degrees of freedom. Df: degree of freedom.

Fixed effect names	χ^2	Df	p-value
Stimulation Type (Audio or control)	0.13	1	0.714
Response Type (Target, Foil or Distractor)	326.82	2	<2E-16
Stimulation Type × Response Type	13.40	2	0.001

By estimating marginal means from the GLME model, the RT in each response type can be contrasted between conditions. The results are shown in the table below:

Table B.29: The post hoc contrasts of the response time (RT) in each response type between audio and control groups. SE: standard error; z-stat: z-statistic. The degree of freedom was set to infinite.

Contrast of RT	**Estimate**	**SE**	**z-stat**	**p-value**
Target, Control-Audio	0.005	0.220	0.022	0.98
Foil, Control-Audio	-0.293	0.235	-1.249	0.21
Distractor, Control-Audio	-0.374	0.255	-1.467	0.14

Similarly, the CLMM model was fitted to predict the confidence using the formula: 'Confidence ∼ StimType:RespType + (1|S_ID) + (1|block)', with the same denotation as in the GLME model. The model statistics were evaluated using Type II Wald Chi-squared tests with Kenward-Roger correction for degrees of freedom.

Table B.30: The CLMM model statistics for prediction of the confidence. Shown is the analysis of variance (ANOVA) table using Type II Wald Chi-squared tests with Kenward-Roger correction for degrees of freedom. Df: degree of freedom.

Fixed effect name	χ^2	**Df**	**p-value**
Stimulation Type (Audio or control)	0.70	1	0.40
Response Type (Target, Foil or Distractor)	1640.92	2	<2E-16
Stimulation Type × Response Type	7.03	2	0.03

The post hoc contrasts within each response type between the stimulation conditions by estimating marginal means:

Table B.31: The post hoc contrasts of the confidence in each response type between audio and control groups. SE: standard error; z-stat: z-statistic. The degree of freedom was set to infinite.

Contrast of confidence	**Estimate**	**SE**	**z-stat**	**p-value**
Target, Control-Audio	-0.144	0.296	-0.488	0.626
Foil, Control-Audio	-0.388	0.304	-1.274	0.203
Distractor, Control-Audio	-0.473	0.325	-1.456	0.145

The group-level statistics for the response type percentage, response time, and confidence are shown in the table below:

Table B.32: The group-level statistics for the response type percentage, response time, and confidence (mean ± s.e.m.) in each response type and stimulation condition.

Stimulation	**Response Type**	**Response percentage (%)**	**Response time (s)**	**Confidence**
Control	Target	58.30 ± 2.39	3.28 ± 0.13	2.96 ± 0.12
Control	Foil	24.33 ± 1.92	4.37 ± 0.22	1.86 ± 0.15
Control	Distractor	11.58 ± 1.13	4.65 ± 0.28	1.56 ± 0.11
Audio	Target	60.53 ± 2.93	3.27 ± 0.10	3.05 ± 0.12
Audio	Foil	21.82 ± 1.68	4.55 ± 0.17	2.09 ± 0.12
Audio	Distractor	11.07 ± 1.71	4.78 ± 0.25	1.71 ± 0.13

C. Permissions to reuse figures

Table C.1: List of reused figures and permissions.

Work type	**Source of work**	**Copyright holder and contact**	**Request date**	**Permission note**
Figure 1.1	Wright, K.P. (2009). EEG in Sleep States. In: Binder, M.D., Hirokawa, N., Windhorst, U. (eds) Encyclopedia of Neuroscience. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-29678-2_2892	Springer Nature springernaturesupport@copyright.com	04 Dec 2023	License number 5682140484440.
Figure 2.6	Schreglmann, S.R., Wang, D., Peach, R.L. et al. Non-invasive suppression of essential tremor via phase-locked disruption of its temporal coherence. Nat Commun 12, 363 (2021). https://doi.org/10.1038/s41467-020-20581-7	Schreglmann, S.R. et al.	08 Dec 2023	Creative Commons Attribution 4.0 International License
Figure 6.3	Violante, I.R., Alania, K., Cassarà, A.M. et al. Non-invasive temporal interference electrical stimulation of the human hippocampus. Nat Neurosci 26, 1994–2004 (2023). https://doi.org/10.1038/s41593-023-01456-8	Ines R. Violante et al.	08 Dec 2023	Creative Commons Attribution 4.0 International License
Figure 6.4	Sazgar, M., Young, M.G. (2019). Overview of EEG, Electrode Placement, and Montages. In: Absolute Epilepsy and EEG Rotation Review. Springer, Cham. https://doi.org/10.1007/978-3-030-03511-2_5	Springer Nature springernaturesupport@copyright.com	08 Dec 2023	License number: 5684210895018

Continued on next page

Table C.1: List of reused figures and permissions. (Continued)

Table A.2	Lubba, C.H., Sethi, S.S., Knaute, P. et al. catch22: CAnonical Time-series CHaracteristics. Data Min Knowl Disc 33, 1821–1852 (2019). https://doi.org/10.1007/s10618-019-00647-x	Lubba, C.H. et al.	08 Dec 2023	Creative Commons Attribution 4.0 International License