UNIVERSITY OF CALIFORNIA, SAN DIEGO

Tracking the Dynamic Brain: Modeling Nonstationarity in Human Electroencephalography

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Bioengineering

by

Sheng-Hsiou Hsu

Committee in charge:

Professor Tzyy-Ping Jung, Chair
Professor Gert Cauwenberghs, Co-Chair
Professor Todd P. Coleman
Professor Virginia de Sa
Professor Kenneth Kreutz-Delgado

2018

The dissertation of Sheng-Hsiou Hsu is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____
Co-Chair

_____
Chair

University of California, San Diego

2018

DEDICATION

To my Mom, Lan-Yu, for your unconditional love and unwavering support
to let me chase my dreams. Even though we are far apart,
you always live in my heart.

EPIGRAPH

Nothing truly valuable arises from a mere sense of duty;

it stems rather from love and devotion towards people and towards objective things.

—Albert Einstein

TABLE OF CONTENTS

Signature Page . . . iii

Dedication . . . iv

Epigraph . . . v

Table of Contents . . . vi

List of Figures . . . x

List of Tables . . . xiv

Acknowledgements . . . xv

Vita . . . xix

Abstract of the Dissertation . . . xxi

Chapter 1 Introduction . . . 1

Chapter 2 Monitoring Alert and Drowsy States by Modeling EEG Source Nonstationarity 9
2.1 Background . . . 10
2.2 Methods . . . 13
2.2.1 Modeling brain states with independent component analysis (ICA) . . . 13
2.2.2 The model deviation index (MDI) . . . 14
2.3 Materials . . . 16
2.3.1 Subjects and experiments . . . 16
2.3.2 EEG data acquisition and preprocessing . . . 16
2.3.3 EEG data analysis: learning multiple ICA models . . . 17
2.3.4 EEG data analysis: computing MDI and correlation analysis 19
2.3.5 Online framework for alertness estimation: the initial ICA model approach . . . 19
2.3.6 A comparison: the power-based approach . . . 20
2.4 Results . . . 20
2.4.1 Behavioral performance . . . 20
2.4.2 Source compositions . . . 22
2.4.3 Source nonstationarity . . . 23
2.4.4 Correlations between the MDI and the reaction speed . . . 25
2.4.5 Feature space of ICA models and the MDI . . . 27
2.4.6 The initial model . . . 29

2.4.7 Performance of online alertness estimation: ICA versus power-based approach . . . 29
2.4.8 Optimal parameters for training the ICA models . . . 30
2.5 Discussion . . . 31
2.5.1 Nonstationary source compositions and activities . . . 31
2.5.2 Distinct ICA models characterize different states . . . 32
2.5.3 Estimating levels of alertness . . . 33
2.5.4 Toward online alertness monitoring . . . 33
2.5.5 Neurophysiology or artifacts? . . . 35
2.5.6 Considerations in building ICA models . . . 35
2.5.7 Comparisons with existing works . . . 37
2.5.8 Limitations and future works . . . 38
2.5.9 Potential applications . . . 38
2.6 Conclusion . . . 39
2.7 Acknowledgments . . . 40

Chapter 3 Modeling Brain Dynamic State Changes with Adaptive Mixture Independent Component Analysis (AMICA) . . . 41
3.1 Background . . . 42
3.2 Materials and Methods . . . 45
3.2.1 Datasets and preprocessing . . . 45
3.2.2 Method description . . . 47
3.2.3 Validation and quantitative analyses . . . 51
3.3 Results . . . 55
3.3.1 Dataset I: Validation using simulated data . . . 55
3.3.2 Dataset II: Classify sleep stages . . . 58
3.3.3 Dataset III: Estimating behavioral alertness . . . 64
3.4 Discussion . . . 73
3.4.1 Unsupervised learning of brain dynamics by modeling source nonstationarity . . . 73
3.4.2 Classification of multiple brain states . . . 74
3.4.3 Estimation of rapid state changes . . . 75
3.4.4 Consistency across subjects . . . 76
3.4.5 Biological interpretation of AMICA models . . . 77
3.4.6 Choosing the model order . . . 78
3.4.7 Limitations and open questions . . . 80
3.5 Conclusions . . . 81
3.6 Acknowledgement . . . 82

Chapter 4 Real-time Adaptive EEG Source Separation Using Online Recursive ICA 83
4.1 Background . . . 84
4.2 Methods . . . 86
4.2.1 Online recursive-least-squares (RLS) pre-whitening . . . 87

4.2.2 Online recursive ICA (ORICA) . . . 88
4.2.3 Non-stationarity detection . . . 91
4.3 Materials . . . 92
4.3.1 Data collection . . . 92
4.3.2 The ORICA pipeline . . . 93
4.3.3 Data processing and analysis . . . 95
4.3.4 Performance evaluation . . . 96
4.4 Results . . . 98
4.4.1 Simulated 64-ch stationary EEG data . . . 98
4.4.2 Simulated 64-ch nonstationary EEG data . . . 102
4.4.3 Real 61-ch EEG data from the Flanker task . . . 102
4.5 Discussion . . . 105
4.5.1 Fast convergence speed . . . 105
4.5.2 Real-time processing . . . 106
4.5.3 Application to real EEG data . . . 107
4.5.4 Nonstationarity detection and adaptation . . . 107
4.5.5 Applications and future directions . . . 108
4.6 Conclusion . . . 109
4.7 Acknowledgment . . . 110

Chapter 5 Validating Online Recursive ICA on EEG data . . . 113
5.1 Background . . . 114
5.2 Methods . . . 115
5.2.1 Online recursive ICA and other ICA/BSS algorithms . . . 115
5.2.2 Evaluation methods . . . 116
5.3 Materials . . . 117
5.3.1 EEG data . . . 117
5.3.2 Data analysis . . . 117
5.4 Results . . . 117
5.5 Conclusion . . . 123
5.6 Acknowledgments . . . 123

Chapter 6 Tracking Non-stationary EEG Sources using Adaptive Online Recursive ICA 124
6.1 Background . . . 125
6.2 Adaptive forgetting factor for online ICA . . . 126
6.2.1 Online recursive ICA: revisit . . . 126
6.2.2 Design of adaptive forgetting factor . . . 127
6.3 Experiments . . . 129
6.3.1 Simulated EEG data . . . 129
6.3.2 Actual EEG data . . . 130
6.3.3 Data processing . . . 130
6.4 Results . . . 131
6.4.1 Simulated 16-ch non-stationary EEG data . . . 131

6.4.2 Actual 14-ch EEG data from the task-switching experiment 134
6.5 Conclusions . . . 134
6.6 Acknowledgments . . . 136

Chapter 7 Real-time EEG Source-Mapping Toolbox (REST): Online ICA and Source Localization . . . 137
7.1 Background . . . 137
7.2 REST: Architecture . . . 139
7.2.1 Preprocessing . . . 139
7.2.2 Source separation . . . 141
7.2.3 Source localization . . . 141
7.2.4 Graphical user interface . . . 142
7.3 Experiments . . . 142
7.3.1 Simulated EEG Data . . . 142
7.3.2 Actual EEG Data . . . 143
7.4 Results . . . 143
7.4.1 Synthetic 64-Channel Stationary EEG . . . 143
7.4.2 Real 14-Channel EEG during SSVEP . . . 146
7.5 Conclusion and future works . . . 146
7.6 Acknowledgments . . . 148

Chapter 8 Conclusion and Future Works . . . 149

Bibliography . . . 151

LIST OF FIGURES

Figure 2.1: The block diagram of the data preprocessing pipeline (blue) and the data analysis pipeline (orange) for the behavioral and EEG data. 18
Figure 2.2: (a) The reaction time and (b) the reaction speed, $RS = 1/RT$, of each trial over experimental time from a representative subject (Subject 1) is shown in the top panel. Their histograms are shown in the bottom panel. 21
Figure 2.3: Averaged power spectra and scalp maps of IC clusters of the four ICA models from 10 subjects. The numbers of subjects and ICs contributed to each cluster were specified. 24
Figure 2.4: The reaction speed and the model deviation index of all trials using the alertness models, the drowsiness models, the offline models and the initial models over time from two subjects. The reaction speed and the MDIs were smoothed with a 90-second sliding window. 26
Figure 2.5: Scatter plots of the reaction speed and the z-scored MDI of the test trials pooled across the 10 subjects. Each plot shows the MDI computed using the indicated subject-specific models. The linear least-square fitted lines and the correlation coefficients were shown. 28
Figure 2.6: Scatter plots of all trials, from (a) subject 1 and (b) subject 2, in the feature space spanned by the two model deviation indices computed using the drowsiness and alertness model as specified in Figure 2.4. Each trial was color-coded with its reaction speed. 29
Figure 2.7: The correlation magnitudes between the MDI of the initial models and reaction speed of test trials for various (a) training data lengths and (b) MDI window lengths, averaged across the 10 subjects. Error bars show the standard errors of the mean 31

Figure 3.1: Adaptive Mixture ICA (AMICA) in a nutshell. AMICA consists of three layers of mixing. The first layer is mixture of ICA models. The second layer is mixture of independent components. The third layer is mixture of generalized Gaussian distributions. 47
Figure 3.2: Mean changes in AMICA model probabilities clustered across AMICA decompositions of 100 repeated runs applied to the simulated quasi-stationary data. (a) AMICA decompositions using 3 models, (b) 4 models, (c) 5 models, and (d) 6 models. 56
Figure 3.3: (a) Model errors in the learned model unmixing matrices, (b) SIR of the decomposed model source activities, and (c) symmetric KL divergence of the learned source PDFs for AMICA decompositions using 1-6 models each averaged across 100 simulations. 57
Figure 3.4: The top panel shows the hypnogram, *i.e.* sleep stages annotated by a sleep expert, of a sleep session from a single subject. Bottom panels show probabilities of ICA models learned by 8-model AMICA decomposition. 59

Figure 3.5: Cross-subject mean (plus one standard deviation) model probabilities of 8 AMICA model clusters in six sleep stages. Model clusters were composed of best-matched models across subjects, as found by iterative template matching. 61
Figure 3.6: Scatter plot of window-mean model probabilities for AMICA model clusters A, B, and C (cf. Figure 3.5), each point representing mean model probability within a 30-sec data segment from sleep recordings of 7 healthy subjects and 10 patients. 63
Figure 3.7: (a) Means and standard deviations in accuracy of classification between 6 sleep stages across the 17 subjects. (b) Confusion matrix of 6-class classification across all the data using 8 cluster model probability features. 64
Figure 3.8: The top panel shows reaction speed changes in response to lane-departure challenges in one simulated driving session. The three bottom panels show shifting probabilities of the three ICA model learned by three-model AMICA decomposition of the whole EEG data session. 65
Figure 3.9: A scatter plot of lane-departure trials from all subjects in the feature space spanned by model probabilities of the three ICA models learned by 3-model AMICA. Each trial was color-coded with its reaction speed. 67
Figure 3.10: Across-subject mean correlation coefficients between reaction speed and model probabilities for fast-response versus slow-response models learned by unsupervised 2-to-4 model AMICA and by separate supervised single-model ICA. 68
Figure 3.11: Event-related changes in the dominant AMICA model in 3-model AMICA decompositions within data trial epochs sorted by driver reaction speed. The trials are time locked either to (a) driving challenge onsets, (b) driver response onsets, or (c) driver response offsets. 70
Figure 3.12: Average scalp maps and power spectra of independent component clusters from slow-response models versus clusters learned from fast-response models from separate three-model AMICA decompositions of data for each subject. 72

Figure 4.1: The ORICA pipeline for online EEG data processing. $X(t)$ is the input data vector at time t and L_B is the size of data in the online buffer. 94
Figure 4.2: (a) Evolution of component-wise correlation magnitudes between ORICA-decomposed ICs and ground truth on simulated stationary 64-ch EEG data. (b) Evolution of correlation magnitudes and component maps of a randomly selected IC. 97
Figure 4.3: Source dynamics and the corresponding component maps of four representative components reconstructed by ORICA at (a) 3 min, (b) 6 min, or (c) 9 min of simulated mixed 64-channel stationary EEG data. 99
Figure 4.4: Effect of ORICA pipeline parameters on convergence trajectory applied to 64-ch simulated stationary EEG data. (a) Decay rate γ of forgetting factor, (b) block size L_{ICA} of ORICA, (c) block size L_{white} of the online whitening, and (d) pre-assumed number n_{sub} of subgaussian sources. 100

Figure 4.5: Application of ORICA to 64-ch simulated non-stationary EEG data. (a) Electrode locations for each session. (b) Non-stationarity index changes. (c) The convergence curve. (d) Zoomed-in plots of log-scaled convergence curve. 103
Figure 4.6: Color-coded event-related potential (ERP) images of (a) fronto-central, (b) occipital, and (c) prefrontal components reconstructed by ORICA and RUNICA on real 61-ch EEG data from the flanker task, time-locked to the response. . . 111
Figure 4.7: (a) Evolution of component-wise correlation magnitudes between ORICA- and RUNICA-decomposed ICs on real 61-ch EEG data from the flanker Task. (b) Evolution of correlation magnitudes and spatial filters of three rapidly converged ICs: prefrontal, occipital, and fronto-central components. 112

Figure 5.1: Scalp maps and mean activity spectra of four ORICA component clusters accounting for non-artifact brain sources from four subjects. Components in each cluster are selected by visual inspection according to their scalp maps and power spectra. 118
Figure 5.2: Scalp maps of the five informative components - the occipital alpha, the frontal midline theta, the left mu, the right mu, and the horizontal eye movement (EOG) - decomposed by AMICA, Infomax, ORICA, FastICA, and SOBI from the dataset S2. 119
Figure 5.3: Percentage of the components (out of 71 ICs) with near-dipolar scalp maps (residual variance $<= 5\%$) and mean mutual information reduction (MIR) averaged across thirteen datasets for the five groups of the fourteen algorithms described in Section 5.2.1. 121
Figure 5.4: The log-scale computation time and the number of near-dipolar scalp maps (r.v. $<= 5\%$) of the five groups of the fourteen algorithms are shown. The dots and bars show the mean and standard deviation of the number of near-dipolar sources averaged across the thirteen datasets. 122

Figure 6.1: Source dipole locations and orientations (red dots and black arrows respectively in cortical head models), 10 second examples of source time-series, and component maps of select components in three sessions for the simulated 16-ch non-stationary EEG data. 129
Figure 6.2: Time evolution of (a) the non-stationarity index z_n, (b) the forgetting factor λ_n, and (c) cross-talk error of ORICA decompositions with adaptive, cooling, and constant forgetting factor profiles applied to the simulated EEG data. . 132
Figure 6.3: Component maps of select sources corresponding to those in Figure 6.1 (matched by correlation of component maps) decomposed and reconstructed by (a) Infomax ICA and (b) ORICA with an adaptive forgetting factor for the three sessions of simulated EEG data. 133
Figure 6.4: Adaptive ORICA applied to actual 14-ch EEG data from the task-switching experiment. Time evolution of (a) the adaptive forgetting factor and (b) the similarity (correlation) between two components decomposed by adaptive ORICA and Infomax ICA. 135

Figure 7.1: A. The pipeline used in the Real-Time EEG Source-Mapping Toolbox (REST). B. The toolbox GUI. The main window (left) shows the scrolling EEG channel or independent component (IC) activation data plus 8 (constantly updated) IC scalp maps. 140
Figure 7.2: The simulated data experiment. The simulated data source activations (Green). ORICA decomposes the simulated EEG data within REST (Blue). The source location is estimated as an equivalent current dipole or as a low-resolution cortical distribution. 144
Figure 7.3: Source localization accuracy in the simulated data experiment using an equivalent current dipole (ECD) model for each estimated independent component (IC). Each disk represents an IC. Disk size shows how well the recovered IC scalp map correlated with the simulated source scalp map. 145
Figure 7.4: Screen captures of the REST GUI during an actual EEG experiment. 147

LIST OF TABLES

Table 1.1: Proposed ICA-based frameworks for modeling nonstationary signals and tracking brain state changes 5

Table 2.1: The correlation coefficients between the 90-sec smoothed reaction speed and MDI of the test trials for the 10 subjects. Subject-specific alertness models, drowsiness models, offline models and initial models were used to compute the MDI. The means and standard errors of mean (SEM) were reported. . . 27

Table 2.2: The AUC of the ROC curve obtained by the two methods for all 10 subjects. "ICA" indicates the proposed method and Power indicates the method proposed by Pal *et al.* 30

Table 3.1: Parameters of Adaptive Mixture ICA for analyzing three datasets: I, simulated data; II, CAP sleep EEG data; III, drowsy driving EEG data. 51

Table 4.1: Comparison of online ICA methods, including the learning rules, the optimization approaches, and the pre-whitening methods. 85

Table 4.2: List of parameters for the ORICA pipeline: IIR high-pass filter (IIR), online RLS whitening filter (RLS) , and online recursive ICA filter (ICA). 95

Table 4.3: Averaged execution time (ms) for 1 sec (300 samples) 64-channel data using online RLS whitening and ORICA with different block sizes L_{white} and L_{ICA}. . 101

ACKNOWLEDGEMENTS

My deepest gratitude goes to my advisor Dr. Tzyy-Ping Jung, associate director of the Swartz Center for Computational Neuroscience (SCCN), UCSD. He provided me tremendous support and guidance through my PhD journey, and more importantly he gave me the freedom and trust to explore my own interests and to develop myself into an independent researcher. I enjoyed our unofficial meetings several times a week where he always has interesting ideas and exciting news. He is also a great mentor who teaches by actions than by words. Through my PhD years, I learned from him how to be truthful when encountering a problem, be faithful when facing a challenge, and be responsible when working with others.

I would also like to express my appreciation to my co-advisor Prof. Gert Cauwenberghs for his continuous support and numerous thought-provoking discussions. The ORICA project we collaborated on had been the most challenging yet rewarding project in my PhD. My thanks go to other members of my committee, Prof. Todd Coleman, Prof. Virginia de Sa, and Prof. Ken Kreutz-Delgado, for their contributions and fruitful discussions.

My PhD journey could not have been happier and more satisfying without the great people in SCCN. I would like to thank Dr. Scott Makeig, director of SCCN, for creating such a diverse and invigorating working environment, especially our daily tea time tradition. I always enjoy his deep insights into scientific and non-scientific (music, history, language, life, etc.) topics and his viola performances. My special thanks go to my collaborators, Dr. Tim Mullen, Luca Pion-Tonachini, Dr. Yuan-Pin Lin, Dr. Masaki Nakanishi, Dr. Jason Palmer, Dr. Makoto Miyakoshi, Dr. Ying Wu, and Chi-Yuan Chang. It is wonderful to work with enthusiastic scientists and talented engineers. I would also like to thank other wonderful members of SCCN and Institute of Neural Computation, Dr. John Iversen, Dr. Yu-Te Wang, Dr. Chun-Shu Wei, Dr. Mateusz Gola, Dr. Johanna Wagner, Dr. Ozgur Balkan, Dr. Zeynep Akalin Acar, Ramon Martinez-Cancino, Christian Kothe, Siddharth Siddarth, Robert Buffington, Rhonda McCoy, and many other visitors and scholars.

It has been a privilege and a great honor to work with fantastic researchers in excellent research groups in many places through my PhD years. I would like to thank Dr. Cory Stevenson, Dr. Chul Kim, Abraham Akinin, Bruno Pedroni, Jun Wang, and other members from Cauwenberghs Group at Department of Bioengineering at UCSD; Dr. Terry Sejnowski, Dr. Claudia Lainscsek, and the members for the Computational Neurobiology Laboratory at the Salk Institute for Biological Studies; Dr. Chin-Teng Lin, Dr. Chun-Hsiang (Michael) Chuang, Dr. Li-Wei Ko and their groups from University of Technology Sydney, Australia, and National Chiao Tung University, Taiwan; Dr. Vikash Gilja and his BMI research group from Department of Electrical and Computer Engineering at UCSD; Dr. Klaus Gramann, Dr. Thorsten Zander, Laurens Krol, Lukas Gehrke, Evelyn Jungnickel, Xixie Zhang, and their groups from Department of Psychology and Ergonomics at Berlin Institute of Technology, Germany.

I owe a debt of gratitude to those who financially supported my research and made this dissertation possible: Scholarship of Government Sponsorship for Overseas Study from Taiwan Ministry of Education, EFRI Program of NSF, UCSD Frontiers of Innovation Scholars Program, Army Research Laboratory, KreutzKamp TMS, Google, and Facebook.

Last by not least, I would not be able to overcome all the challenges in my PhD journey without the incredible support from my friends and family in Taiwan, in San Diego, and around the world. More importantly, I would like to thank my parents and my brother, Lan-Yu Cheng, Jung-Tsung Hsu, and Yi-Hsiou Hsu, for their unconditional love and unwavering support throughout my ups and downs in my life. This dissertation is dedicated to them.

Chapter 2, in part, is a reprint of the material as it appears in Hsu, S. H., and Jung, T. P. (2017). "Monitoring alert and drowsy states by modeling EEG source nonstationarity." *Journal of Neural Engineering*, 14(5), 056012. The dissertation author was the primary investigator and author of this paper. This study was sponsored in part by the US Army Research Laboratory and was accomplished under Cooperative Agreement Number W911NF-10-2-0022. The views and the conclusions contained in this document are those of the authors and should not be

interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S Government. The U.S Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein. We thank Scott Makeig, Luca Pion-Tonachini, Makoto Miyakoshi, and other members of the Swartz Center for Computational Neuroscience for the fruitful discussion and valuable suggestions.

Chapter 3 is currently being prepared for submission for publication by Hsu, S. H., Pion-Tonachini, L., Palmer, J., Miyakoshi, M., Jung, T. P., and Makeig. S. The dissertation author was the primary investigator and author of this paper.

Chapter 4, in part, is a reprint of the material as it appears in Hsu, S. H., Mullen, T. R., Jung, T. P., and Cauwenberghs, G. (2016). "Real-time adaptive EEG source separation using online recursive independent component analysis." *IEEE transactions on neural systems and rehabilitation engineering*, 24(3), 309-319. © 2016 IEEE. The dissertation author was the primary investigator and author of this paper. This study was sponsored in part by a gift by the Swartz Foundation (Old Field, NY), by the Army Research Laboratory under Cooperative Agreement Number W911NF-10-2-0022, by the NIH grant 1R01MH084819-03 and NSF EFRI-M3C 1137279. We thank Christian Kothe for help with BCILAB integration.

Chapter 5, in part, is a reprint of the material as it appears in Hsu, S. H., Mullen, T., Jung, T. P., and Cauwenberghs, G. (2015, April). "Validating online recursive independent component analysis on EEG data." In *7th International IEEE/EMBS Conference on Neural Engineering (NER)* (pp. 918-921). © 2015 IEEE. The dissertation author was the primary investigator and author of this paper. We thank Arnaud Delorme, Jason Palmer, Julie Onton, Robert Oostenveld and Scott Makeig for their datasets and MATLAB scripts for ICA comparisons (available at `http://sccn.ucsd.edu/wiki/BSSComparison`).

Chapter 6, in part, is a reprint of the material as it appears in Hsu, S. H., Pion-Tonachini, L., Jung, T. P., and Cauwenberghs, G. (2015, August). "Tracking non-stationary EEG sources using adaptive online recursive independent component analysis." In *37th Annual International*

Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), (pp. 4106-4109). © 2015 IEEE. The dissertation author was the primary investigator and co-author of this paper. Hsu, S. H. and Pion-Tonachini, L. contributed equally to this work. This study was sponsored by the Army Research Laboratory under Cooperative Agreement Number W911NF-10-2-0022, by NSF EFRI-M3C 1137279, and by NSF Graduate Research Fellowship under Grant No. DGE-1144086.

Chapter 7, in part, is a reprint of the material as it appears in Pion-Tonachini, L., Hsu, S. H., Makeig, S., Jung, T. P., and Cauwenberghs, G. (2015, August). "Real-time EEG source-mapping toolbox (REST): Online ICA and source localization." In *37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, (pp. 4114-4117). © 2015 IEEE. The dissertation author was the secondary investigator and co-author of this paper. Pion-Tonachini, L. and Hsu, S. H. made equal contribution to this work. This study was sponsored by the Army Research Laboratory under Cooperative Agreement Number W911NF-10-2-0022, by NSF EFRI-M3C 1137279, and by NSF Graduate Research Fellowship under Grant No. DGE-1144086.

VITA

2011	Bachelor of Science in Electrical Engineering National Taiwan University
2015	Master of Science in Bioengineering University of California, San Diego
2018	Doctor of Philosophy in Bioengineering University of California, San Diego

PUBLICATIONS

Journal Articles

- **Hsu, S. H.**, Pion-Tonachini, L., Palmer, J., Miyakoshi, M., Jung, T. P., and Makeig, S., "Modeling Brain Dynamic State Changes with Adaptive Mixture Independent Component Analysis." (in preparation)
- **Hsu, S. H.** and Jung, T.-P. (2017) "Monitoring alert and drowsy states by modeling EEG source nonstationarity." *Journal of Neural Engineering*, 14(5), 056012.
- **Hsu, S. H.**, Mullen, T. R., Jung, T. P., and Cauwenberghs, G. (2016). "Real-Time Adaptive EEG Source Separation Using Online Recursive Independent Component Analysis." *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 24(3), 309-319.

Conference Articles

- Pion-Tonachini, L., **Hsu, S. H.**, Chang, C. Y., Jung, T. P., and Makeig, S., (under review) "Automatic Artifact Rejection using the Real-time EEG Source-mapping Toolbox (REST)." *40th Annual International IEEE EMBS Conference*.
- Chang, C. Y., **Hsu S.-H.**, Pion-Tonachini, L., Jung T.-P., (under review) "Evaluating Artifact Subspace Reconstruction for Automatic Artifact Removal." *40th Annual International IEEE EMBS Conference*.
- Nakanishi M., Wang Y., **Hsu S.-H.**, Wang Y.-T., Jung T.-P., (2017) "Independent component analysis-based spatial filtering improves template-based SSVEP detection" *39th Annual International IEEE EMBS Conference*, pp.3620-3623.
- **Hsu, S. H.** and Jung, T. P. (2016). "Modeling and tracking brain nonstationarity in a sustained attention task." *International Conference on Augmented Cognition*, pp. 209-217, Springer.
- **Hsu, S. H.**, Pion-Tonachini, L., Jung, T. P., and Cauwenberghs, G. (2015). "Tracking non-stationary EEG sources using adaptive online recursive independent component analysis." *37th Annual International IEEE EMBS Conference*.

- Pion-Tonachini, L., **Hsu, S. H.**, Makeig, S., Jung, T. P., and Cauwenberghs, G. (2015). "Real-time EEG Source-mapping Toolbox (REST): Online ICA and source localization." *37th Annual International IEEE EMBS Conference*.
- Lin, Y. P., **Hsu, S. H.**, and Jung, T. P. (2015). "Exploring Day-to-Day Variability in the Relations Between Emotion and EEG Signals." *International Conference on Augmented Cognition*, pp. 461-469, Springer.
- **Hsu, S. H.**, Mullen, T., Jung, T. P., and Cauwenberghs, G. (2015). "Validating online recursive independent component analysis on EEG data." *7th International IEEE EMBS Conference on Neural Engineering*, pp. 918-921.
- **Hsu, S. H.**, Mullen, T., Jung, T. P., and Cauwenberghs, G. (2014). "Online recursive independent component analysis for real-time source separation of high-density EEG." *36th Annual International IEEE EMBS Conference*.

ABSTRACT OF THE DISSERTATION

Tracking the Dynamic Brain:
Modeling Nonstationarity in Human Electroencephalography

by

Sheng-Hsiou Hsu

Doctor of Philosophy in Bioengineering

University of California, San Diego, 2018

Professor Tzyy-Ping Jung, Chair
Professor Gert Cauwenberghs, Co-Chair

As the human brain performs cognitive functions or generates spontaneous mental processes within ever-changing, real-world environments, states of the brain are inevitably nonstationary. This calls for innovative approaches to both obtain objective and quantitative insights into hidden cognitive and mental states and study the dynamics of brain states that give rise to behaviors and mental disorders. Despite electroencephalography (EEG) offering a noninvasive, portable, real-time measurement of brain activity, an urgent need remains for computational tools to effectively decode brain states from continuous, unlabeled EEG data, to quantitatively assess state changes,

and to provide neuroscientific insights.

In this dissertation, I present three computational approaches for quantitative assessment of brain-state dynamics by modeling multichannel, nonstationary EEG data at the level of functional brain sources. I first describe a hypothesis-driven approach which uses independent component analysis (ICA) to model distinct source activities under different brain states, characterized as stationary processes that involve specific brain networks. Two ICA models corresponding to alert and drowsy states in a sustained-attention experiment are identified, and a quantitative measure is proposed to assess the fluctuations in drowsiness. Next, I present a data-driven approach - Adaptive Mixture ICA (AMICA) - as an unsupervised-learning method for exploring nonstationary dynamics of unlabeled data. AMICA effectively characterizes EEG dynamics during sleep for automatic staging, reveals transitions between alert and drowsy states at millisecond resolution, and provides neuroscientific insights into active brain sources in each brain state. In the remaining chapters, I introduce the Online Recursive ICA (ORICA) approach for adaptive tracking of nonstationary sources that underlie continuous state changes, which incrementally updates the ICA model upon presentation of new data. A comparison study demonstrates ORICAs fast convergence property and its feasibility for real-time and online processing. An adaptive forgetting factor is proposed to facilitate ORICA's adaptive capacity. For maximum utility, an open-source Real-time EEG Source-mapping Toolbox (REST) is built that integrates ORICA for the source-level analysis and interactive visualization of live-streaming data. These advances in computational tools provide a framework to study unlabeled, hidden cognitive or mental states and support developments toward objective, quantitative, continuous, and real-time assessments of the dynamic brain and its health.

Chapter 1

Introduction

How does the brain function? How can such understanding improve our lives? The field of neuroscience and neuroengineering in the 21st century has significantly advanced with technological innovations and scientific breakthroughs, evidenced and further supported by the launches of the BRAIN Initiative of the United States and the Human Brain Project of Europe in 2013. Yet most of such advancements and grand challenges focus on the study of brain mechanism at the molecular, cellular, and network levels. It is still challenging to translate the results from neuronal networks to brain systems, from structural imaging to functional decoding, from sensory and motor functions to cognitive and mental processes, and from animal to human studies. With the study of functional decoding of cognitive systems in human, the top-down approach may provide insights and hypotheses for closing the gap of understanding the brain at different levels and for developing real-world applications to improve our lives.

Within real-world, ever-changing environments, as a human brain performs various cognitive functions or undergoes different mental processes, states of the brain are inevitably nonstationary. This nonstationary brain dynamics [7, 59, 34] may arise from continual formation and dissolution of interacting brain regions whose activities can be described as dynamic systems featuring continuous transitions between intermittently stable states [18, 10]. This brain (func-

tional) state can be described as the state when the brain elicits stationary activities, *i.e.* statistical properties of such processes do not change over time, which usually correspond to predictable behaviors or cognitive performance. Examples of brain states include sleep stages [31] that was characterized 50 years ago and spontaneous covert cognitive states [69] during various input conditions or performance of mental tasks.

The dynamics of brain states that gives rise to fluidity and rapid adaptability in human cognitions and behaviors has become an expanding focus in neuroscience. Being able to decode and monitor human cognitive and mental states will significantly impact the fields of cognitive sciences, clinical diagnosis and monitoring, neuroergonomics, and brain-computer interfaces. For example, a neurofeedback system that estimates drowsiness of drivers can be applied to prevent catastrophic incidents [122, 72]. Automatic assessment of sleep dynamics and objective evaluation of the effectiveness of mindfulness training or rehabilitation can provide quantitative measures for clinical diagnosis and prognosis. In neuroergonomics, information of workers' cognitive workload can help reduce performance errors and increase productivity with adaptive scheduling and task assignment [35]. Estimation of mental fatigue [103] and operator's mindset [125] enables adaptive human-machine interaction in brain-computer interfaces (BCI) [124, 118]. Furthermore, successful decoding of emotional states and stress levels will help neuroscientists understand the high-level human cognitions and will revolutionize the fields in consumer electronics and wearable devices. However, much remains unknown regarding the large-scale neural dynamical systems that underlie the human cognitive and mental states, and a need remains for technological and computational innovations to objectively, quantitatively, robustly decode and monitor the dynamics of the brain states.

Advances in functional neuroimaging and electrophysiological techniques such as functional MRI (fMRI), magnetoencephalography (MEG), functional near-infrared spectroscopy (fNIR) and electroencephalography (EEG) have enabled researchers to measure large-scale brain activities in human noninvasively. Although none of these techniques have the spatial resolution

of single-neuron recording, they offer the ability to measure rapidly-changing temporal dynamics of the whole neural system at the spatial resolution of functional brain regions. Among these non-invasive modalities, EEG provides a more direct measurement of brain activity with millisecond temporal resolution. Recent technological advancements in EEG, including the design of wireless devices for portability, the invention of dry-sensors for fast preparation and convenience, and the low weight and bulk of its sensors for comfortability, have made EEG an appropriate modality for studying fast-changing human brain states in real-world environments.

While EEG recordings have long been available, most of conventional EEG methods were developed for the study of brain responses to the presentation of perceptual stimuli or cued actions, *i.e.* event-related potentials, and did not apply to spontaneous, endogenous, nonstationary brain activities reflecting dynamics of brain states that are unlabeled and continuous without precise timing. Data analytic methods and techniques for assessing the nonstationary activities recorded in EEG signals and studying the dynamics of brain states have only been an active research topic in the recent decades, thanks to the increase in computational power and the need for real-time applications like neurofeedback. Earlier works focused on monitoring power of EEG signals or spectra [7], information theory-based measures like entropy [87], or change points of nonparametric statistics [58] for detecting and segmenting piecewise stationary processes. These methods had not taken into account the spatial patterns of nonstationary signals. More recent studies such as microstate analysis [115, 61] and global functional connectivity analysis [18, 10] incorporated the spatial information for brain state identification by examining topologies of an electric potential across multichannel EEG electrode montage or measuring inter-electrode signal synchrony. However, none of the aforementioned methods offered hypothesis-driven computational models of nonstationary brain processes and provided interpretable connections to particular brain source activities that underlie the observed scalp phenomena.

Toward developing an EEG-based brain state monitoring system, a need remains for: computational models for effective and robust decoding of brain states from nonstationary EEG

signals; quantitative measures for assessing and tracking state changes; data-driven (unsupervised learning) approaches for addressing issues with unlabeled, continuous data; and hypothesis-driven approaches for neurophysiological interpretations and scientific insights. Desired properties of such approaches include: temporal specificity with sub-second resolution at the single-trial level; spatial specificity by measuring dynamics at a brain-sources level rather than a scalp-sensor level; adaptive capability for modeling nonstationarity in on-line live-streaming data; and real-time capability for real-world applications in which computational power and memory resources are limited. Finally, open-source software is in demand for deployment of such approaches for maximum impact and utility within and beyond the scientific community.

In this dissertation, I address these challenges and present three different computational models with quantitative measures for both hypothesis-driven analysis and unsupervised learning of brain state dynamics at the brain-sources level from continuous, nonstationary, unlabeled (and labeled) EEG data. Extensive validation of these methods on both simulation datasets and experimental EEG datasets are performed. Throughout the dissertation, I leverage these tools to study and address various forms of nonstationarity due to a shift of brain states or variability in recording devices (noises/artifacts), which are open challenges in the fields of cognitive neuroscience, clinical neuroscience, and brain-computer interfaces. These include monitoring drowsy and alert states in a driving experiment, characterizing sleep EEG dynamics for automatic sleep staging, tracking transitions of brain and non-brain sources in task-switching experiments, and adapting to abrupt signal changes due to displacements of sensors.

The overarching theme of the thesis is built upon the capability of independent component analysis (ICA) [57, 8], which has been widely applied to EEG as a means for blind source separation [74, 55, 78], in modeling activities of functional brain sources that underlie the observed scalp phenomena. The application of ICA to EEG data is justified by a reasonable assumption that multi-channel scalp EEG signals, x, arise as an instantaneous linear mixture, A, of latent independent sources dynamics s such that $x = A \cdot s$. ICA has demonstrated great success

Table 1.1: Proposed ICA-based frameworks for modeling nonstationary signals and tracking brain state changes

	Standard ICA	**Multi-model ICA**	**ICA mixture model**	**Adaptive ICA**
Model assumption	$\mathbf{x}(t) = \mathbf{A} \cdot \mathbf{s}(t)$	$\mathbf{x}_k(t) = \mathbf{A}_k \cdot \mathbf{s}_k(t)$	$\mathbf{x}(t) = \mathbf{A}_k \cdot \mathbf{s}_k(t)$	$\mathbf{x}(t) = \mathbf{A}(t) \cdot \mathbf{s}(t)$
Data assumption	stationary	quasi-stationary	quasi-stationary	non-stationary
State hypothesis	single	discrete (labeled)	discrete	continuous
Label requirement	no	yes	no	no
Online capability	no	no	no	yes
Computation cost	medium	medium	high	low
Thesis chapters	-	ch2	ch3	ch4, ch5, ch6

in modeling sources in EEG by separating artifacts such as eye blinks and muscle activities [54] and identifying neurophysiological sources [91, 90, 30, 78]. However, it is limited in modeling nonstationary changes in EEG source locations and activities because of its spatial-stationarity assumption produced by its limited model complexity (*e.g.* its use of a fixed spatial mixing matrix, *A*). Furthermore, standard ICA algorithms such as Infomax ICA [8] and FastICA [51] are only suitable for offline applications, where a large amount of training data are available.

This thesis aims to establish the use of three computational models, namely the multi-model ICA, the ICA mixture model, and the adaptive ICA, that relax the ICA assumption on the spatial stationarity of sources, extend ICA to model nonstationary sources, and equip with quantitative measures to enable tracking of brain state changes. Their assumptions and capabilities are summarized in Table 1.1. The main hypothesis, the source nonstationarity hypothesis [44], states that different brain or cognitive states may involve cortical macro- or meso-dynamics in networks of cortical brain areas that can be identified by distinct ICA models.

The multi-model ICA approach extends the standard ICA model by training distinct ICA models, *i.e.* mixing matrices $\mathbf{A}_k$ and source activities $\mathbf{s}_k$, on data recorded before and after the state transition, $\mathbf{x}_k$. The assumption is that the nonstationary signals arise from transitions between quasi-stationary data segments, each corresponds to the activities of a distinct brain state.

In Chapter 2, we introduce the multi-model ICA approach and test the hypothesis against EEG data from subjects performing a sustained-attention task in a driving simulator. We show that

EEG signals under alert versus drowsy states, indexed by reaction speeds to driving challenges, can be characterized by distinct ICA models. With ICA models of the brain states, *i.e.* fluctuation of alertness and drowsiness, can be tracked by measuring the goodness-of-fit of each ICA model to a sliding window of EEG data using the proposed model deviation index (MDI). Even though the ICA models are trained offline, we demonstrate an effective online framework by using a pre-trained ICA model on an initial reference state for tracking dynamics of the state changes. However, this hypothesis-driven approach requires labels of brain states to select training data for ICA models, which are usually hard to obtain in a real-world environment.

The ICA mixture model (ICAMM) approach is a general unsupervised learning approach that models data as a mixture of distinct source (ICA) models, *i.e.* $\mathbf{x}(t) = \mathbf{A}_k\mathbf{s}_k(t)$, with the similar assumption that distinct ICA models may better characterize different segments of quasi-stationary data. Different from the multi-model ICA approach, the ICAMM approach can learn multiple ICA models simultaneously on unlabeled data by jointly optimizing the data likelihood.

Chapter 3 explores the use of an ICAMM approach, namely the Adaptive Mixture ICA (AMICA), as a data-driven approach to studying dynamic changes in cognition and brain states. After validating AMICA on quasi-stationary simulated data, we demonstrate that AMICA can usefully characterize sleep EEG dynamics and effectively classify six sleep stages. Next, when applied to the driving EEG dataset described in Chapter 2, AMICA automatically identifies two source models accounting for alert and drowsy states. The degree of activation for each model, measured by the model likelihood, can serve as a quantitative estimator of levels of drowsiness. Furthermore, we show that the dynamic switching of the most active models can reveal the brain state changes at single-trail level with milliseconds resolution. In all of these applications, AMICA learns interpretable results including spatial distributions and temporal activities of active sources in each ICA model underlying distinct brain states.

Both multi-model ICA and ICA mixture model approaches model the nonstationary sources using a discrete number of ICA models, assuming quasi-stationary states, and reveal the

state changes as the deviation from or the activation of the corresponding state-associated models. The two approaches require a pre-defined number of ICA models, are computationally expensive, and are mainly designed for offline analysis.

In Chapter 4, Chapter 5, and Chapter 6, we study the adaptive ICA approach, which recursively and incrementally learns a time-varying ICA model $\mathbf{A}(t)$ upon presentation of new data $\mathbf{x}(t)$. Different from the first two approaches, the adaptive ICA approach assumes that brain states are continuously changing and attempts to track the state changes by adapting the time-varying ICA model to the nonstationary data. The adaptive nature and light-weight algorithm of this data-driven approach enable its online application for tracking state changes in near real-time.

Chapter 4 introduces the adaptive Online Recursive ICA (ORICA) combined with online whitening for blind source separation of high-density EEG data. We show that ORICA can rapidly extract principal brain and artifact sources in real 61-ch EEG data recorded by a wearable EEG system in a cognitive experiment. Furthermore, we demonstrate ORICA's capability to detect and adapt to nonstationarity in 64-ch simulated EEG data corresponding to abrupt displacements of electrodes, a common source of spatial nonstationarity in real-world mobile applications.

For further validation, Chapter 5 systematically compares ORICA with standard offline ICA algorithms on experimental EEG dataset as a benchmark. We demonstrate that ORICA reduces the computational runtime while still obtaining comparable source decomposition quality compared with most of the ICA methods. The properties of fast convergence and low computational complexity of ORICA enable the realization of real-time online ICA process.

In Chapter 6, we improve the adaptation performance of ORICA by proposing an adaptive forgetting factor, which dynamically adjusts the weight applied to new data in the ICA recursive update based on the measured nonstationarity of signals. We empirically show that the adaptive forgetting factors outperform other commonly-used non-adaptive rules when underlying source dynamics are changing both in the simulated nonstationary EEG dataset and in the actual EEG data recorded from a task-switching experiment.

To enable real-time processing for real-world applications, in Chapter 7, we introduce the Real-time EEG Source-mapping Toolbox (REST). The REST integrates ORICA into an online data processing pipeline and thereby allows separation and localization of nonstationary sources in near real-time. Furthermore, the REST provides an easy-to-use and interactive platform to perform source-level analysis on live-streaming EEG data with various visualizations options. It is an open-source toolbox, and the code is available at `https://github.com/goodshawn12/REST`.

Finally, Chapter 8 concludes the contributions and the findings of this dissertation. We also identify some key future steps toward building a real-time brain-state monitoring system and describe our current efforts in addressing those challenges. We end with a discussion of the potential impact of such system on applications such as continuous clinical monitoring and real-world brain-computer interfaces.

Chapter 2

Monitoring Alert and Drowsy States by Modeling EEG Source Nonstationarity

As a human brain performs various cognitive functions within ever-changing environments, states of the brain characterized by recorded brain activities such as electroencephalogram (EEG) are inevitably nonstationary. The challenges of analyzing the nonstationary EEG signals include finding neurocognitive sources that underlie different brain states and using EEG data to quantitatively assess the state changes.

In this chapter, we hypothesize that brain activities under different states, *e.g.* levels of alertness, can be modeled as distinct compositions of statistically independent sources using independent component analysis (ICA). This study presents a framework to quantitatively assess the EEG source nonstationarity and estimate levels of alertness. The framework was tested against EEG data collected from 10 subjects performing a sustained-attention task in a driving simulator.

Empirical results illustrate that EEG signals under alert versus drowsy states, indexed by reaction speeds to driving challenges, can be characterized by distinct ICA models. By quantifying the goodness-of-fit of each ICA model to the EEG data using the model deviation index (MDI), we found that MDIs were significantly correlated with the reaction speeds ($r = -0.390$ with

alertness models and $r = 0.449$ with drowsiness models) and the opposite correlations indicated that the two models accounted for sources in the alert and drowsy states, respectively. Based on the observed source non-stationarity, this study also proposes an online framework using a subject-specific ICA model trained with an initial (alert) state to track the level of alertness. For classification of alert against drowsy states, the proposed online framework achieved an averaged area-under-curve of 0.745 and compared favorably with a classic power-based approach.

This ICA-based framework provides a new way to study changes of brain states and can be applied to monitoring cognitive or mental states of human operators in attention-critical settings or in passive brain-computer interfaces.

2.1 Background

The concept of nonstationarity, as a change of ongoing patterns [7], in recorded brain activities such as electroencephalogram (EEG) has received increasing attention in the fields of neuroimaging and electrophysiology. The EEG nonstationarity might arise from different involvements of neurocognitive sources or brain processes underlying complex human cognition and behaviors [59]. Conventional EEG analyses that assume stationary signals such as averaging event-related potential may overlook rich information in inter-trial variability reflecting nonstationary brain processes, and brain-computer interfaces (BCI) having a stationary model may lead to deteriorating performance due to increased mental fatigue during the experiment [117, 103]. On the other hand, previous studies have suggested that stages of sleep can be characterized by EEG quasi-stationary patterns [14], the level of alertness fluctuates in a cycle of 4 minutes or longer [77], and routine spontaneous EEG is a stationary process over a short 3.5-second period [113]. Yet a big challenge is the lack of quantitative, objective, and online approach to assess the nonstationarity of the brain signals when humans show different cognitive and behavioral performance and to estimate and track the brain states.

EEG-based state monitoring has been an active research topic that focuses on the study of nonstationary EEG to continuously track human cognitive or mental states. Target states include mental workload [35, 104], fatigue [35], vigilance [111] and drowsiness [93, 122]. This has many applications in neuroergonomics [35] and passive brain-computer interfaces (BCI) [124] that can make use of the state information to prevent catastrophic incidents, *e.g.* drowsy driving, or to implicitly enhance human-machine interaction. This study centers on one of state changes known to impact human performance – namely changes between alertness and drowsiness.

Researchers have studied EEG correlates of alertness and drowsiness for the past two decades. In 1993 Makeig *et al.*[77] found coherence of fluctuations in EEG spectrum and subjects' levels of alertness inferred by performance in an auditory detection task. Several subsequent studies also reported that EEG alpha or theta band power at central, parietal, or occipital sites correlate with performance in sustain-attention tasks, which is associated with alert or drowsy state [56, 65, 98]. Hence these states can be monitored by tracking changes of these EEG biomarkers, *e.g.* fluctuations of their statistical properties or deviations from classification models [93, 122]. However, these studies focused on spectrum analyses of individual EEG channels, which required selection of discriminative channels and features or used machine-learning approaches that had limited neurophysiological validity and interpretation of the results. It is still desirable to develop approaches that could further our understanding in the dynamics underlying alertness and drowsiness fluctuations and could be generalized across sessions and subjects.

To model EEG dynamics and identify neurophysiological sources, one popular approach is to transform EEG data from channel-space to activities in a source space using source separation or source localization methods [74, 90, 24, 30]. For example, a widely used blind source separation method – Independent Component Analysis (ICA) [57, 8, 66] – assumes EEG channel data, $\mathbf{x}$, are linear mixtures, $\mathbf{A}$, of underlying independent sources activities $\mathbf{s}$, *i.e.* $\mathbf{x} = \mathbf{A}\mathbf{s}$, and learns a linear unmixing matrix $\mathbf{W}$ such that $\mathbf{y} = \mathbf{W}\mathbf{x}$ recovers the statistically independent source activities and $\mathbf{WA}$ ideally is a permuted and scaled identity matrix. ICA has demonstrated great

success in modeling EEG dynamics and identifying neurophysiological sources [90, 30, 78]. Previous studies have applied ICA to obtain a stationary model, *i.e.* **A** or **W** (ideally **W** is the pseudo-inverse of **A**), for drowsiness estimation, assuming the same sources were involved in different levels of alertness [93, 73, 19].

However, studies in electrophysiology and functional imaging have accumulated evidences that different brain states may involve neural activities arising from different brain areas and networks. Sarter *et al*.. [109] summarized that sustained attention and vigilance are mediated by a neuronal network consisted of an anterior attention system and a posterior attention system, which thereby facilitate or bias sensory processing in sensory-associational regions. Other studies suggested that activations of frontal and prefrontal cortical areas driven by anterior cingulate cortex are typically associated with executive functions and attention [53]. From alertness to drowsiness, fatigue or microsleep, brain activities seem to shift from high-frequency (beta and gamma waves) to low-frequency activities (alpha, theta, and delta waves) [65, 100], and spread through a larger brain area [65, 108]. For example, Lal *et al*.[65] reported that during the onset of fatigue, delta and theta activity were present mostly in the frontal, central, and parietal areas of the brain. Santamaria *et al*.[108] reported that during drowsiness there was a change in alpha distribution: occipito-parietal alpha spreads to anterior areas and becomes more centro-frontal and temporal.

Inspired and supported by these findings, this study assumed that different brain states, *i.e.* drowsy versus alert states, involve not only different activities of brain sources but also different networks of brain sources. This study further hypothesized that these distinct source compositions and activities associated with alertness or drowsiness can be modeled by different ICA models, *i.e.* (un-)mixing matrices **A** (**W**) and source activities **s**. The hypothesis was tested against EEG data collected from 10 subjects participating in a sustained-attention task within a driving simulator [72]. To demonstrate the utility of characterizing the EEG source nonstationarity, this study proposed to measure the goodness-of-fit while applying each state-associated ICA

model to EEG data, which was quantified by the model deviation index (MDI), an index for estimating the levels of alertness or drowsiness. Finally, an online framework for alertness estimation using the ICA model learned from an operator's initial (alert) state was proposed, and its performance was compared against a classic approach using alpha and theta power proposed by Pal *et al*.[93].

Compared to existing studies, the proposed ICA model-based approach has the following characteristics. (a) It is a generic approach that is not limited to assessing the levels of alertness or drowsiness. It can potentially be applied to monitoring other cognitive or mental states such as mental workload and emotions or brain states in clinical settings such as sleep stages and the depth of anesthesia. (b) It estimates continuous levels of deviations from a known state as opposed to classifying discrete states. (c) The ICA models of different brain states provide more insights into neurophysiological sources and artifactual components under these states. In addition, the proposed online framework for alertness estimation has the following advantages. (a) It is an unsupervised approach, which does not require the availability of data from another namely the drowsy state. (b) It does not require training data from other pilot sessions or subjects. (c) It learns an individual model for each session using a small amount of training data from the session, alleviating inter-session or inter-subject variability. In sum, this study aims to provide a new way to study complex human cognition and behaviors by assessing the underlying EEG source nonstationarity.

2.2 Methods

2.2.1 Modeling brain states with independent component analysis (ICA)

As a human shows different levels of alertness, we hypothesize that the EEG signals x_k under state k, *e.g.* alert or drowsy state, involve distinct EEG sources and can be modeled as projections of distinct source compositions onto the scalp A_k and source activities s_k of an ICA

model.

$$\mathbf{x}_k(t) = \mathbf{A}_k(t) \cdot \mathbf{s}_k(t) \tag{2.1}$$

The time-dependent ICA model, the mixing matrix $\mathbf{A}$ or the unmixing matrix $\mathbf{W}$, resembles an ICA mixture model [68, 55] and an adaptive ICA model [47] which assume data are quasi-stationary or non-stationary.

Based on the hypothesis, rather than learning the mixture or adaptive model as described in [68, 55, 47], we propose to learn individual ICA models from small segments of the training EEG data under alert or drowsy state, *i.e.* alertness and drowsiness ICA models, and then measure the goodness-of-fit while applying each model to a sliding window of EEG data over the whole session. We expect that the goodness-of-it would be better, *i.e.* smaller model-fitting error, while applying the alertness ICA model to other EEG data when a subject is alert than to the data when a subject is drowsy, and vice versa. The goodness-of-fit of ICA models is quantified by a model deviation index (MDI) to assess levels of alertness or drowsiness. The MDI is further explained in Section 2.2.2, the detailed implementation of the approach is described in Section 2.3, and alternative approaches are discussed in Section 2.5.6.

2.2.2 The model deviation index (MDI)

To quantify the goodness-of-fit of an ICA model, we start with the convergence criterion for ICA. It is worth noting that this framework is not limited to any specific ICA algorithm. Here we use Infomax ICA [66], which aims to maximize mutual information between source activities and EEG data and has been demonstrated to achieve good performance in separating dipolar neurophysiological sources in EEG data [30].

The Infomax ICA with natural gradient [66] has a general learning rule:

$$\mathbf{W} \leftarrow \mathbf{W} + \eta[\mathbf{I} - \mathbf{f}(\mathbf{y}) \cdot \mathbf{y}^T]\mathbf{W} \tag{2.2}$$

where $\mathbf{W}$ is an unmixing matrix that linearly projects data $\mathbf{x}$ into statistically independent sources $\mathbf{y} = \mathbf{W}\mathbf{x}$, $\mathbf{I}$ is an identify matrix, η is a learning rate, and f is a nonlinear activation function. The learning in 2.2 stops when $\mathbf{W}$ converges, *i.e.* the gradient term approaches zero or $\langle \mathbf{f} \cdot \mathbf{y}^T \rangle \approx \mathbf{I}$, where $\langle \cdot \rangle$ represents an average over the data. On the other hand, this convergence criterion can also be used to indicate the fitting error of an ICA model $\mathbf{W}$ applied to other data $\mathbf{x}$. Previous studies have defined the nonstationary index $\|\mathbf{I} - \langle \mathbf{f} \cdot \mathbf{y}^T \rangle\|_F$ as a quantitative measurement of the model fitting error and have demonstrated successful detection of abrupt electrode displacements [47] and changes of underlying sources [48] in simulated EEG data using the index.

Following our preliminary work in [43], we modify the index by dividing it with the Frobenius norm of source activity to account for the scale ambiguity of an ICA model and define the new index as model deviation index (MDI), as a function of an ICA model $\mathbf{W}_0$:

$$\mathrm{MDI}(\mathbf{W}_0) = \frac{\left\| \langle \mathbf{f} \cdot \mathbf{y}^T \rangle \right\|_{\text{F, off-diag}}}{\left\| \langle \mathbf{y} \cdot \mathbf{y}^T \rangle \right\|_F} \tag{2.3}$$

where $\mathbf{y} = \mathbf{W}_0\mathbf{x}$, $\mathbf{x}$ is the test data, $\mathbf{f} = \frac{1-\exp(-\mathbf{y})}{1+\exp(-\mathbf{y})}$ for the non-extended Infomax ICA used in this study, and $\|\cdot\|_{\text{F, off-diag}}$ is the Frobenius norm of the off-diagonal elements of the matrix. The numerator in Equation 2.3 quantifies cross-talk errors of sources in the ICA model W_0, indicating changes in underlying source composition in $\mathbf{x}$; the denominator is the covariance of source activations. Taking the ratio of the two terms normalizes the index such that the MDI value is bounded.

As mentioned in Section 2.2.1, the MDI enables quantitative measurement of EEG source nonstationarity, including both the spatial source composition and the temporal source activities. A small MDI corresponds to a small model fitting error and indicates that both the training and the test data arise from the similar set of sources with similar activities. On the other hand, a large MDI represents the model fails to fit the test data and thus the associated state has shifted away

from that of the training data. Note that the definition of MDI could be modified when other ICA algorithms or source separation methods are used, depending on their convergence criteria.

2.3 Materials

2.3.1 Subjects and experiments

Ten healthy volunteers participated in the 90-min experiment in an immersive VR-based driving simulator, performing a sustained-attention driving task. They were 20-28 year-old with normal or corrected-to-normal vision and had no history of neurological or sleep disorders and no drug or alcohol abuse. No subject reported sleep deprivation on the day before the experiment. The experiments were conducted in the early afternoon after lunch such that various levels of alertness and drowsiness were observed. All subjects practiced the driving tasks for 10 minutes to become familiar with the experimental procedures.

In the sustained-attention task, specifically an event-related lane-departure task [49], the subjects were presented with lane-departure events every 8-12 seconds randomly and were instructed to steer the car back to the cruising position quickly. On average a total of 550 trials were performed in the 90-min experiment. The duration from the onset of a lane-departure event to the onset of a steering action is defined as a reaction time (RT), which has been reported to be associated with alert and drowsy states [19, 72]. The reaction time was transformed to the reaction speed ($RS = 1/RT$) for further analyses, and the reason was discussed in Section 2.4.1. For the details of the subjects and the experiment, readers are referred to [19, 72].

2.3.2 EEG data acquisition and preprocessing

For each subject, 30-channel EEG data were recorded at a 500 Hz sampling rate using a NeuroScan NuAmp Express System (Compumedics Ltd., VIC, Australia). The electrodes

(Ag/AgCl) were arranged according to the international 10-20 system with a unipolar reference at the right earlobe. The impedance of each electrode was kept under $5k\Omega$ during the experiment.

Figure 2.1 displays the data preprocessing and analysis pipelines. The EEG data were band-pass filtered (1-50 Hz) and were down-sampled to 250Hz. Bad channels in the recordings such as flat channels due to poor contacts of electrodes and poorly correlated channels were removed (2-6 channels were identified for the 10 subjects) and interpolated using the PREP pipeline [11]. It is worth noting that the interpolation step is redundant for ICA but is what the PREP pipeline automatically performed. In addition, artifact subspace reconstruction (ASR) [83], implemented as a plug-in in EEGLAB toolbox [29], was applied to reduce the contaminations of high-amplitude artifacts with a mild threshold (burst repair $\sigma = 20$). These artifact-correction methods facilitate the convergence of an ICA model.

2.3.3 EEG data analysis: learning multiple ICA models

For each subject, the data from a 90-min session was split into 90-sec blocks with 30-sec overlaps, and four types of ICA models were learned from specific blocks of data. (1) Five alertness ICA models were individually trained with five 90-sec windows with the highest reaction speeds (*i.e.* alert states). (2) Five drowsiness ICA models were trained with five 90-sec windows with the lowest reaction speeds (*i.e.* drowsy states). (3) For a comparison, each subject's offline ICA model was learned from the EEG data of the entire session, *i.e.* how a conventional ICA is usually applied. (4) To enable online alertness estimation, a session-specific initial ICA model was learned from the first 90-sec window of EEG data at the beginning of the session.

All the ICA decompositions used the non-extended Infomax ICA implemented in EEGLAB toolbox [29]. Principle Component Analysis (PCA) was applied before ICA to facilitate ICA convergence and to account for the loss of data rank due to the removal of bad channels in the PREP pipeline [11]. On average, 24-28 sources were retained for the 10 subjects.

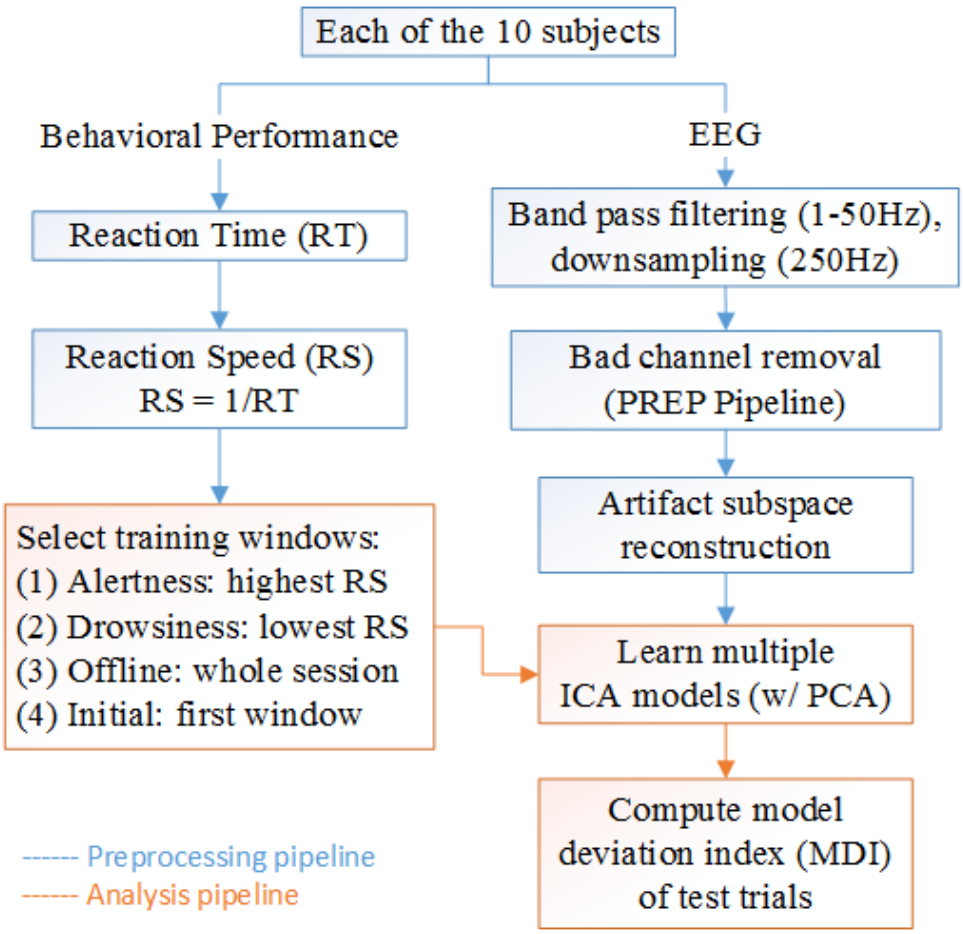

Figure 2.1: The block diagram of the data preprocessing pipeline (blue) and the data analysis pipeline (orange) for the behavioral and EEG data.

2.3.4 EEG data analysis: computing MDI and correlation analysis

Each ICA model was applied to the EEG data of the same session. For each trial in the session, its MDI was computed by Equation 2.3 with a 90-second window of continuous data right before the lane-deviation onset. The Pearson correlation coefficient was computed between the reaction speed and the MDI of test trials, *i.e.* all except the trials selected to train the ICA model. To illustrate the global fluctuations of the reaction speed and the MDI over an entire driving session, a 90-second smoothing window was applied where their median values in the window were taken. In addition, the effect of different MDI window lengths was also tested and reported in Section 2.4.8.

2.3.5 Online framework for alertness estimation: the initial ICA model approach

An online framework for alertness estimation was proposed and quantitatively evaluated to fully demonstrate the usefulness of the proposed approach. In an online session, we first trained the initial ICA model with the first 90-sec of EEG data at the beginning of the session. The initial model was then applied to 90-sec sliding windows of EEG data across the rest of the online session to obtain their MDIs.

For the drowsiness classification, the MDIs were compared with a threshold value: if the MDI of the window before a lane-departure event was larger than the threshold, the trial is classified as a non-alert trial; otherwise, it is classified as an alert trial. The true alert trials were defined as the test trials whose reaction time $RT_{alert} \leq 1.5 \cdot \overline{RT}_{train}$, where $\overline{RT}_{train}$ is the mean value of RT of the training trials, and the non-alert trials were defined as those with $RT_{non-alert} \geq 2.5 \cdot \overline{RT}_{train}$. This criterion follows the definition in [119] to deal with the highly skewed distribution of reaction times.

We varied the threshold, classified the test trials, computed the corresponding true positive

rate (TPR = number of correctly classified non-alert trials / number of total non-alert trials) and false positive rate (FPR = number of incorrectly classified alert trials / number of total alert trials), plotted the Receiver Operating Characteristics (ROC) curve, and reported the area under curve (AUC) as the effectiveness of the alertness estimation.

2.3.6 A comparison: the power-based approach

A classic approach proposed by Pal *et al*.[93] was implemented for a comparison. Similar to the approach described in Section 2.3.5, the Pal's approach is also an unsupervised approach that trains an alert model and measures the distance of the test data from the model as an estimate of alertness levels. Firstly, exactly the same first 90-sec of EEG data at Oz were used as the training data, which were divided into 2-sec non-overlapping segments with their log power of alpha band (8-12Hz) and theta band (4-8Hz) computed. The means μ_i and covariance matrices Σ_i of the two band-power (i indicates alpha or theta) of the training segments were estimated. Secondly, the Mahalanobis distance (MD) was used to measure the deviation of the alpha and theta power of 2-sec test data from that of the training data $\mathbf{x}$: $MD_i(\mathbf{x}) = \sqrt{(\mathbf{x}-\mu_i)^T\Sigma_i^{-1}(\mathbf{x}-\mu_i)}$. The final index was a linear combination: $MD(\mathbf{x}) = a \cdot MD_{alpha} + (1-a) \cdot MD_{theta}$, as an estimate of the alertness level. According to [93], the optimal $a = 0.3$ was used. The same evaluation procedure such as computation of the AUC of the ROC curve was applied to enable a fair comparison.

2.4 Results

2.4.1 Behavioral performance

Figure 2.2a shows, for a representative subject, 73% (430 out of 587) of the lane-departure trials had the reaction time smaller than 1.26 seconds, illustrating the subject mostly stayed

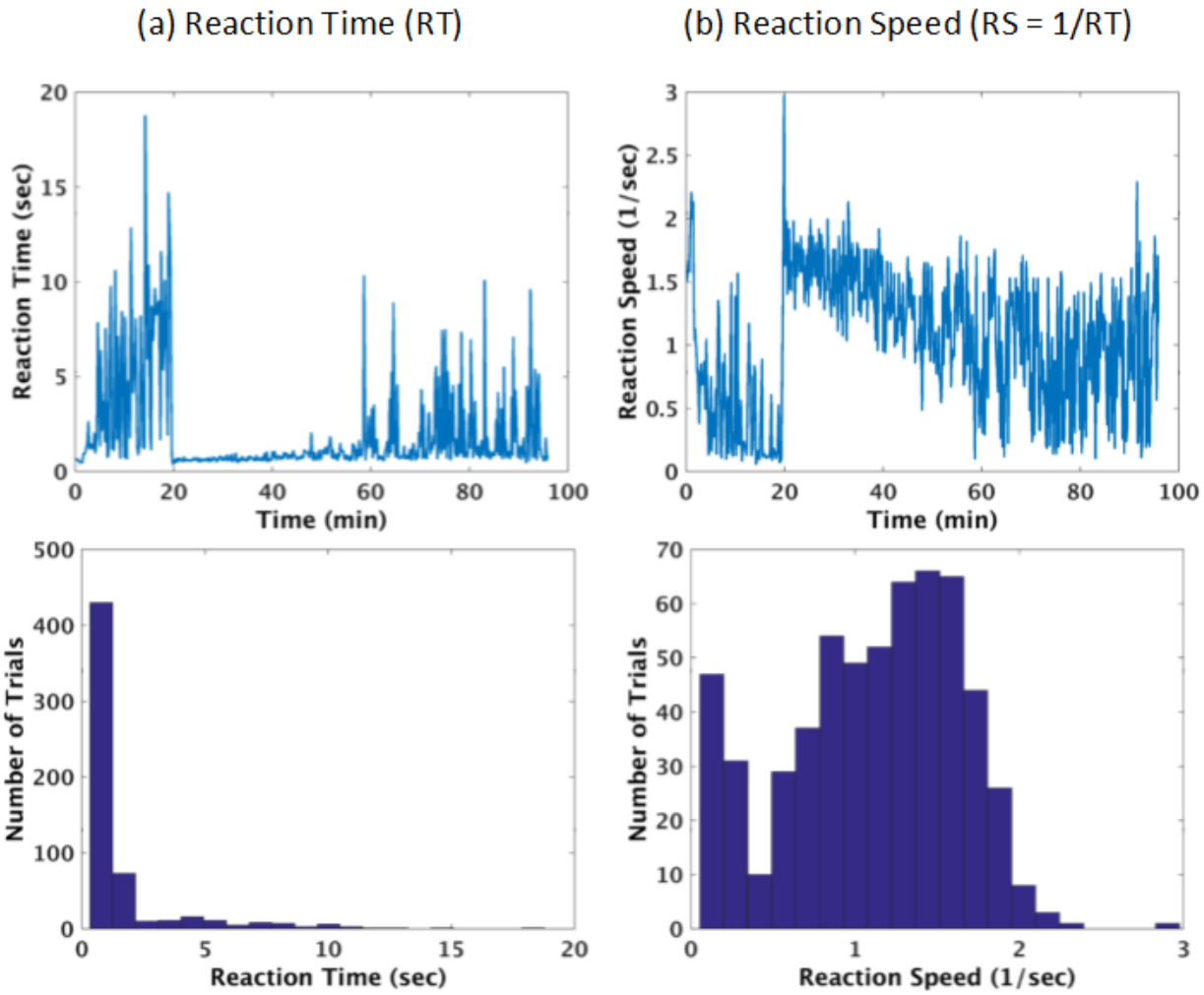

Figure 2.2: (a) The reaction time and (b) the reaction speed, $RS = 1/RT$, of each trial over experimental time from a representative subject (Subject 1) is shown in the top panel. Their histograms are shown in the bottom panel.

alert and reacted quickly to the lane-departure events in the experiment. To avoid bias in the correlation analysis, the reaction time was transformed to the reaction speed to produce a less skewed distribution as shown in figure 2.2b.

The averaged mean reaction speed of all the trials across subjects was 1.014 ± 0.164 (1/sec) and the averaged standard deviation was 0.428 ± 0.099 (1/sec), indicating the subjects' levels of alertness fluctuated during the course of the experiments. For the selected most-alert windows described in Section 2.3.3, 76% were within the first 30 minutes of the experiments,

16% were in the 30-60 minutes, and 8% were from the 60 minute to the end of the 90 minute session. For the selected most-drowsy windows, the distribution was 36% (0-30 min), 18% (30-60 min), and 46% (60 min - end). The subjects were mostly alert at the beginning of the sessions and became drowsy over time-on-task.

2.4.2 Source compositions

Figure 2.3 shows the averaged power spectra and the scalp maps, *i.e.* the contribution or projection on the scalp, of clusters of independent components (ICs) from all of the 10 subjects for each of the four ICA models. The IC clusters were obtained using CORRMAP function in EEGLAB [116] based on the component similarity measured by correlations of the inverse weights of ICA, *i.e.* columns of $\mathbf{W}^{-1}$ in (1). Manually selected component templates and a correlation threshold of 0.85 (0.9 for eye-blink and eye-movement clusters) were used, and the number of ICs contributed by each subject was limited to two. The alertness (drowsiness) models trained from the data windows with the highest (lowest) reaction speeds were used for reporting the clustering results.

First of all, the results illustrate that the initial, alertness and drowsiness models, which were trained with only 90-sec of EEG data, were able to reach converged solutions and find dipolar components. Comparable ocular and neurophysiological IC clusters were identified in all of the four models. For example, the eye-blink and lateral eye-movement clusters were characterized by the distribution of scalp maps and the high power at low frequencies. The frontal clusters with high theta or alpha power and the parietal and occipital clusters with high alpha power were reported to be associated with alertness or drowsiness and were consistent with the previous study involving the same driving task [19].

Although comparable IC clusters were found, the source compositions and power spectra across the four ICA models as well as the numbers of contributed ICs and subjects in the clusters still varied. From alert or initial state to drowsy state, fewer ocular components and more frontal

and occipital components were observed, and the alpha power spread from the occipital and parietal regions to a larger brain area including the frontal and even prefrontal regions. These results were consistent with previous studies [65, 108]. In addition, lateral parietal clusters were absence in the drowsiness models, and fewer frontal and central components were present in the initial models.

Furthermore, the offline models contained the combination of ICs found in other three models with almost all of the subjects contributed ICs to each category of the clusters. However, the alpha peaks were not clearly observed in the frontal and central clusters because alpha activities were not always present over the course of the experiment.

2.4.3 Source nonstationarity

Figure 2.4 plots the model deviation indices (MDI) of the four ICA models and the measured reaction speed over the experimental time from the two representative subjects. Firstly, the MDI curves of the alertness models significantly inversely correlated with the reaction speed curves (correlation coefficients of -0.828 and -0.534 for the two subjects). The alertness models could model the data segments with high reaction speed trials (alert state) and yielded small MDI values, but they failed to fit the data with low reaction speed trials (drowsy state) and thus yielded large MDI values. On the other hand, the MDIs of the drowsiness models could track the dynamics of the reaction speed with significantly positive correlation coefficients of 0.768 and 0.724 for the subjects. The results indicate that the drowsiness models well characterized the source compositions and dynamics of the data with low reaction speed trials (drowsy state), leading to smaller MDI values.

As a comparison, the offline models showed smaller averaged MDI values across the session but had the smallest correlation coefficient between the MDI and the reaction speed among the four models. Since the offline models attempted to learn the stationary sources over the entire session, they did not favor modeling the data segments with high or low reaction speed

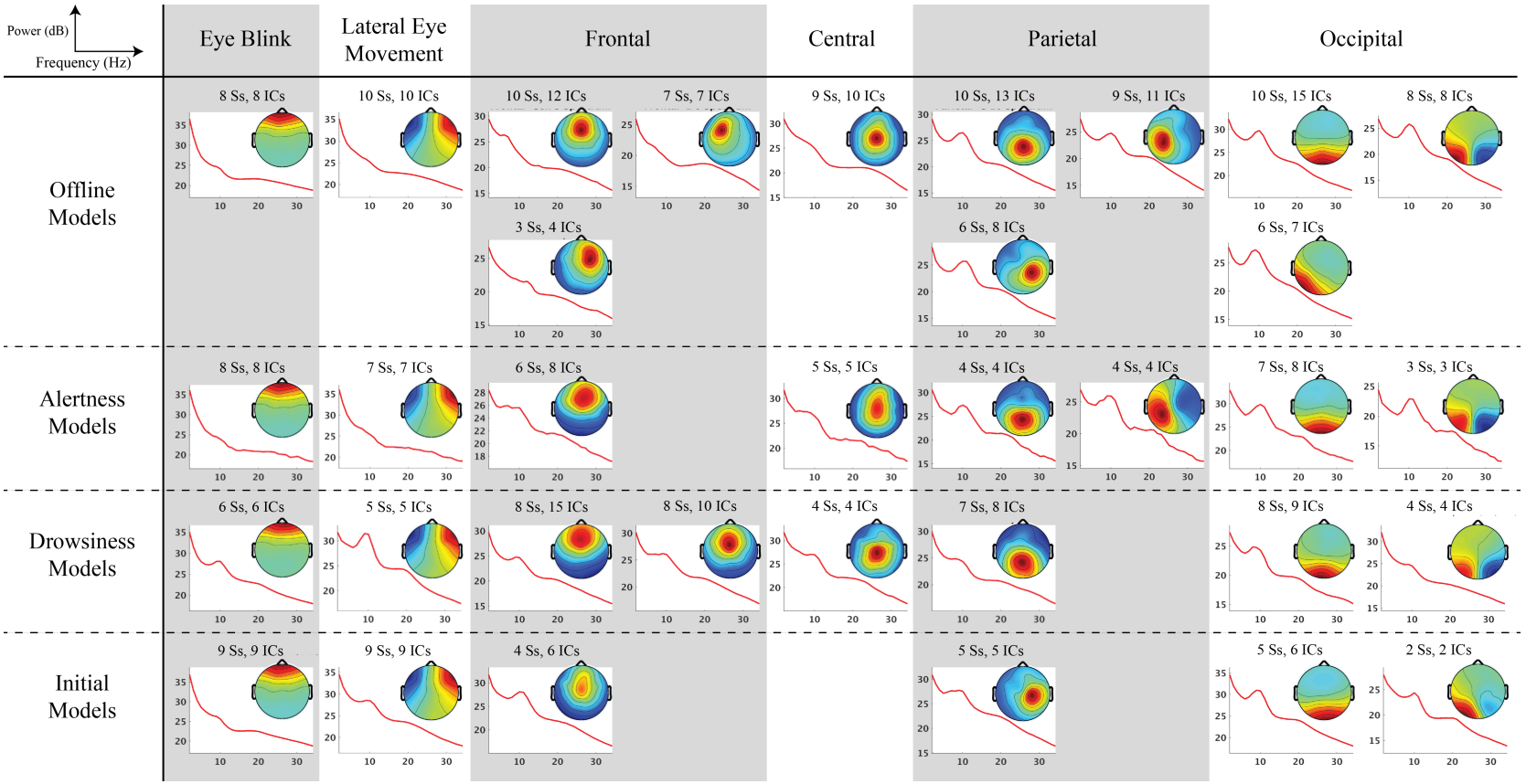

Figure 2.3: Averaged power spectra and scalp maps of IC clusters of the four ICA models from 10 subjects. The numbers of subjects and ICs contributed to each cluster were specified.

trials. The results of the initial models were discussed in Section 2.4.6.

2.4.4 Correlations between the MDI and the reaction speed

Table 2.1 shows the correlation coefficients between the smoothed reaction speeds and the MDI of the test trials for all of the 10 subjects, where MDIs were computed with the subject-specific offline, alertness, drowsiness, and initial models separately. Paired t-tests were performed to test the significance levels of the correlation magnitudes between the different models.

First, the selected alertness models yielded significantly negative correlation compared to the offline models ($p < 1 \times 10^{-4}$) and the corresponding drowsiness models ($p < 1 \times 10^{-6}$). Similarly, the selected drowsiness models yielded significantly positive correlation compared to the offline models ($p < 0.05$). In addition, no significant differences in the correlation coefficients were found between the results of the top one alertness model and the averaged results of the top five alertness models ($p = 0.52$), and between the results of the top one drowsiness model and that of the top five drowsiness models ($p = 0.76$). This provides the evidence for the robustness of the method to training window selection. However, inter-subject variability existed. For example, the top one alertness model for Subject 2 and the top one drowsiness models for Subjects 4 and 8 did not show significant correlations. The sources of the variability may come from the selection of training windows and from the percentage of time where the subjects were drowsy.

To summarize and visualize the results of the goodness-of-fit of each of the four models across subjects, figure 2.5 shows the smoothed z-scored MDI and the reaction speed of all test trials from the 10 subjects. The consistent results were observed: significant negative correlations between MDI and reaction speed for the alertness models and the initial models, significant positive correlations for the drowsiness models, and significantly weaker correlations for the offline models. These results demonstrate that the z-scored MDI of the alertness, drowsiness, and initial models could characterize the source nonstationarity underlying the alertness and drowsiness fluctuations and could be used to estimate levels of alertness across subjects. It is

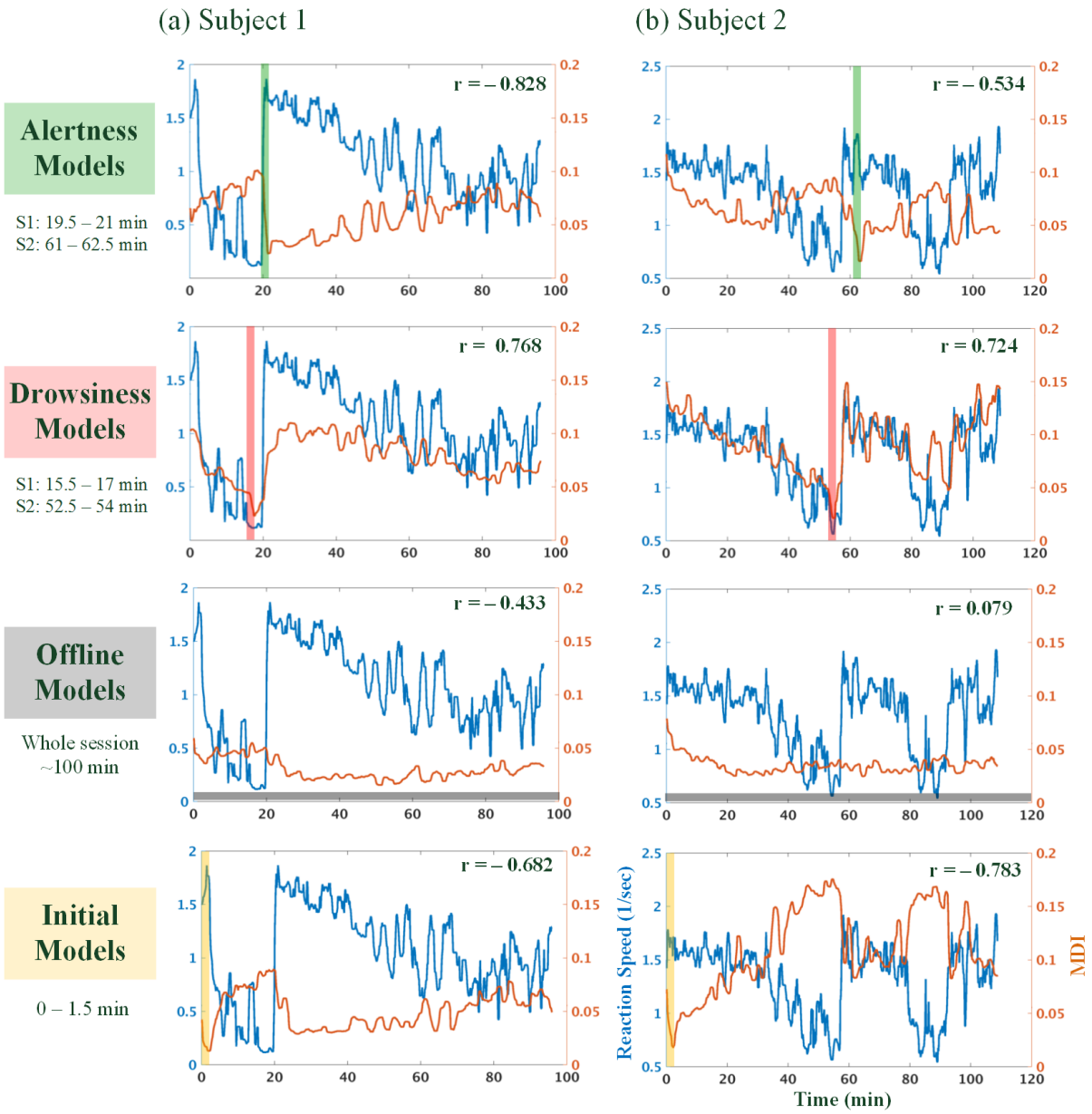

Figure 2.4: The reaction speed (blue) and the model deviation index (orange) of all trials using the alertness models (1st row), the drowsiness models (2nd row), the offline models (3rd row) and the initial models (4th row) over time from two subjects. The colored windows were selected for training the ICA models, with their time intervals specified on the left. The selected alertness and drowsiness models were the ones with the highest correlation among the five models presented in Table 2.1. The reaction speed and the MDIs were smoothed with a 90-second sliding window, and their correlation coefficients were shown.

Table 2.1: The correlation coefficients between the 90-sec smoothed reaction speed and MDI of the test trials for the 10 subjects. Subject-specific alertness models, drowsiness models, offline models and initial models were used to compute the MDI. For alertness (drowsiness) models, the results of both the top one and the averaged results of the top five windows with the highest (lowest) reaction speed were reported. The means and standard errors of mean (SEM) were reported.

	Alertness models		Drowsiness models		Offline	Initial
Subjects	Top 1	Top 5	Top 1	Top 5	model	model
1	-0.681	-0.673	0.694	0.740	-0.433	-0.682
2	-0.039	-0.273	0.384	0.549	0.079	-0.783
3	-0.388	-0.358	0.503	0.485	0.097	-0.569
4	-0.485	-0.343	-0.074	0.066	0.031	-0.480
5	-0.454	-0.378	0.377	0.425	0.036	-0.390
6	-0.242	-0.190	0.469	0.474	0.475	0.003
7	-0.679	-0.617	0.809	0.819	0.630	-0.564
8	-0.688	-0.676	-0.041	-0.041	0.091	-0.701
9	-0.510	-0.068	0.340	0.347	0.572	-0.425
10	-0.348	-0.326	0.642	0.629	0.122	-0.602
Mean	-0.451	-0.390	0.410	0.449	0.170	-0.519
SEM	0.066	0.065	0.091	0.086	0.099	0.070

worth noting that the z-scoring was performed to enable a comparison across subjects but would change the corresponding correlation coefficients.

2.4.5 Feature space of ICA models and the MDI

Using the alertness and the drowsiness models defined in figure 2.4, we constructed a feature space spanned by MDIs of the two models, as shown in figure 2.6. We found that low reaction-speed trials (red) distributed toward the top left corner in the feature space, illustrating that those trials could be modeled by the drowsiness model and deviated from the alertness model. On the contrary, the high reaction speed trials (blue) distributed toward the bottom right corner. We also noted that some trials in figure 2.6b located in the top right corner, indicating that neither the alertness model nor the drowsiness model fitted those data.

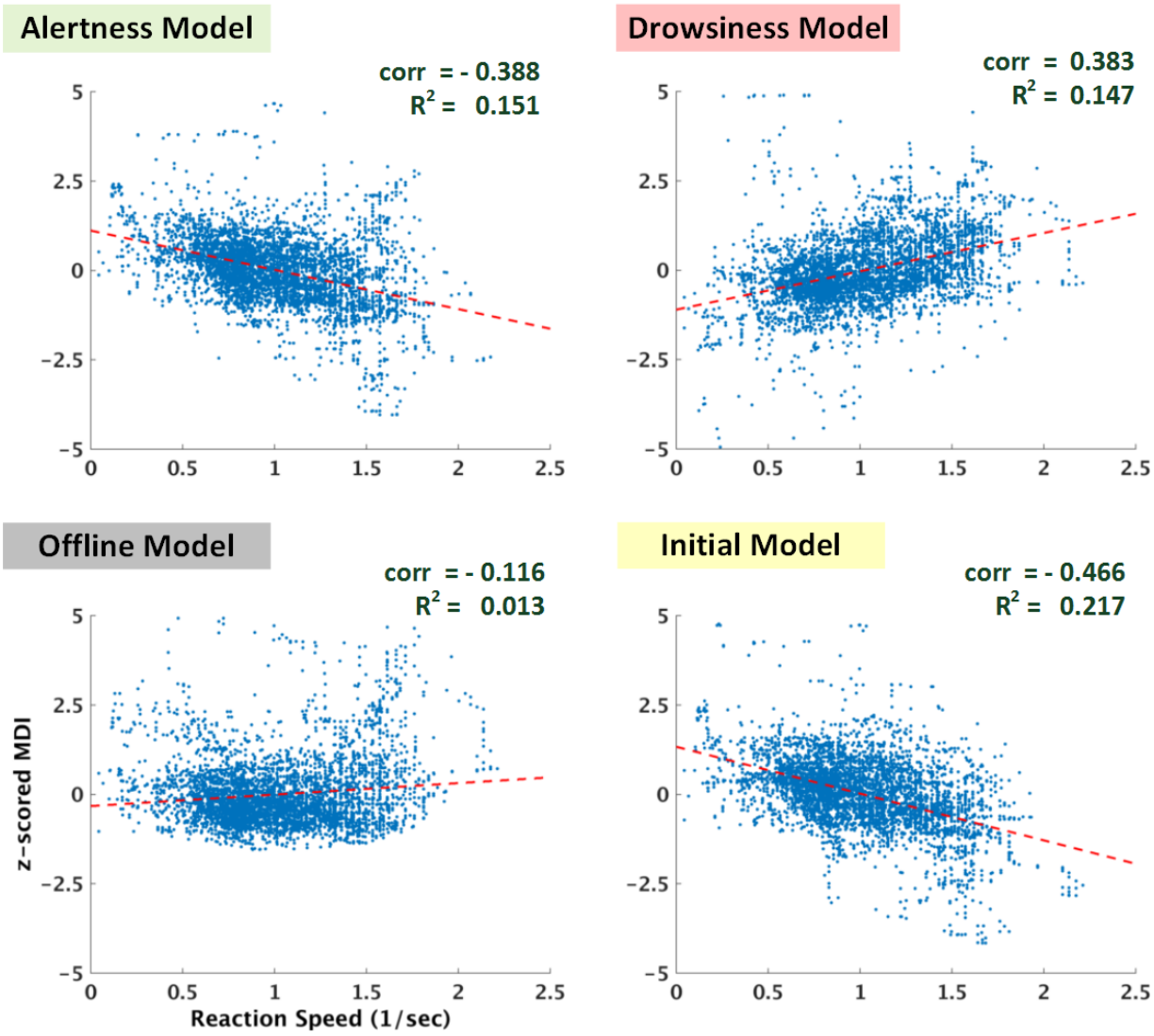

Figure 2.5: Scatter plots of the reaction speed and the z-scored MDI of the test trials pooled across the 10 subjects. Each plot shows the MDI computed using the indicated subject-specific models. The reaction speed and the MDI of the trials were smoothed with a 90-second sliding window, and the MDIs were z-scored within each subject. The linear least-square fitted lines and the corresponding correlation coefficients and R-squared values were shown.

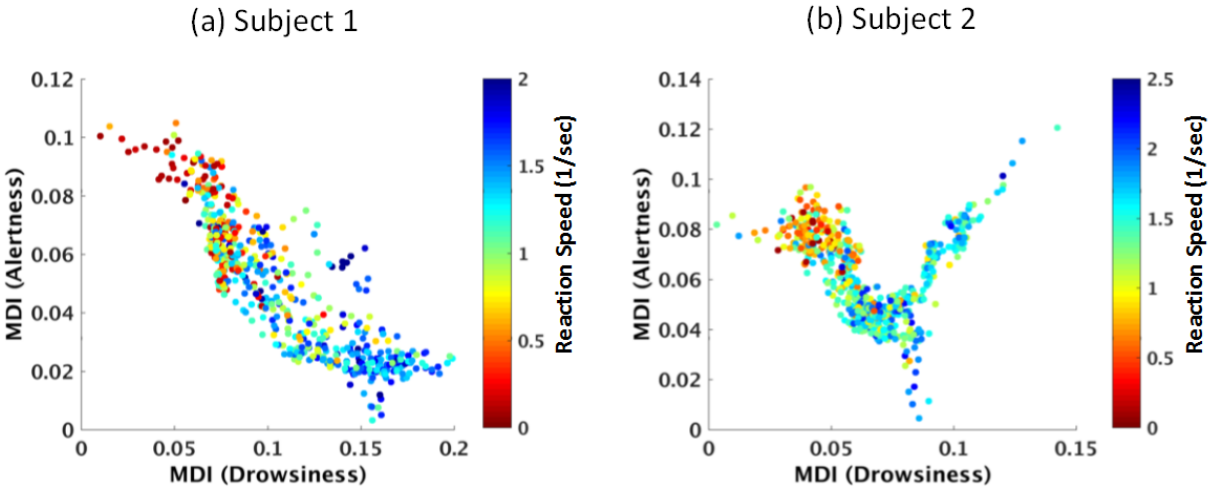

Figure 2.6: Scatter plots of all trials, from (a) subject 1 and (b) subject 2, in the feature space spanned by the two model deviation indices computed using the drowsiness and alertness model as specified in Figure 2.4. Each trial was color-coded with its reaction speed.

2.4.6 The initial model

Figure 2.4 illustrates that the MDI of the initial models negatively correlated with the reaction speed. The findings could be generalized across the subjects as shown in Figure 2.5 and Table 2.1. Firstly, the initial models show significantly higher correlations than the offline models do ($p < 3 \times 10^{-5}$ by a paired t-test) and are significantly different from the results of the drowsiness models ($p < 1 \times 10^{-6}$). Secondly, no significant difference of the correlations was observed between the initial models and the top 1 alertness models ($p = 0.49$), suggesting the initial models resembled the alertness models in that they both fitted the high reaction speed trials and deviated from the low reaction speed trials.

2.4.7 Performance of online alertness estimation: ICA versus power-based approach

As shown in Table 2.2, the proposed ICA approach achieved higher than chance AUCs (0.5) on all except Subject 6, with an average of 0.745 and half of the subjects had AUC higher

Table 2.2: The AUC of the ROC curve obtained by the two methods for all 10 subjects. "ICA" indicates the proposed method and Power indicates the method proposed by Pal *et al*.

Subject	ICA	Power
1	0.8312	0.7173
2	0.9167	0.6615
3	0.8229	0.6063
4	0.6241	0.5860
5	0.7817	0.4649
6	0.4646	0.6976
7	0.6166	0.8359
8	0.9222	0.6557
9	0.6186	0.7809
10	0.8513	0.7091
Mean	0.7450	0.6715
Std.	0.1537	0.1042

than 0.8. This result demonstrates the effectiveness of using the initial ICA models for alertness estimation. In comparison, the Pal's approach using alpha and theta power achieved a lower averaged AUC of 0.6715. However, the power approach obtained much higher AUCs for a few subjects (Subject 7 and 9) where the ICA approach performed poorly. Overall, the AUCs of the two approaches showed no significant difference. It is worth noting that the two approaches have variable success across different subjects and might be combined to improve the online alertness estimation.

2.4.8 Optimal parameters for training the ICA models

We tested the effects of the data length for training the model (figure 2.7a) and the window length for computing the MDI (figure 2.7b) on correlation magnitudes between the single-trial MDI and the reaction speed across the 10 subjects. Figure 2.7a illustrates the correlation magnitudes decreased as the training data exceeded 3 minutes and the optimal training data length was around 2 minutes. It is worth noting that 0.5 minutes of the data was insufficient for a converged ICA solution with 30 sources as described in Section 2.4.2. However, even this model

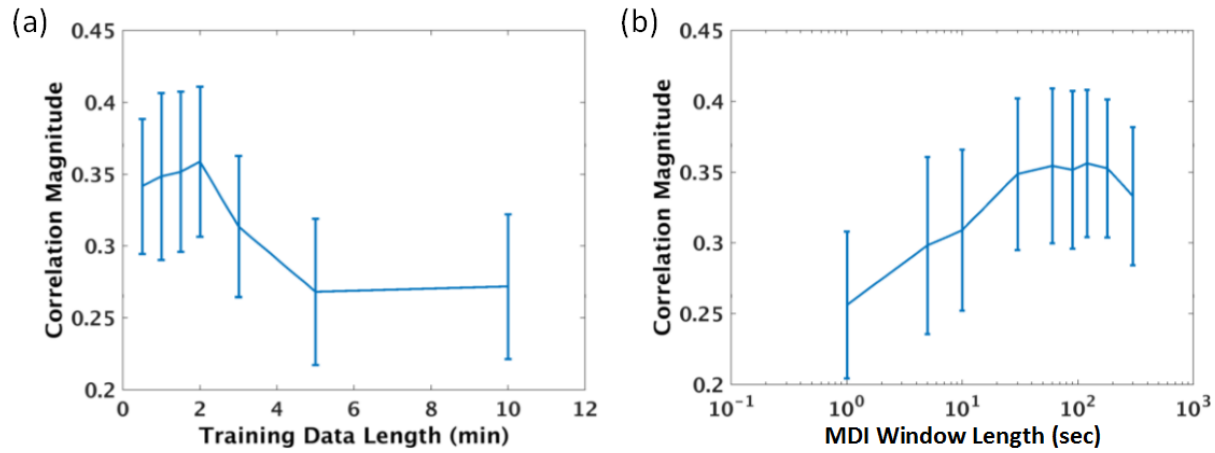

Figure 2.7: The correlation magnitudes between the MDI of the initial models and reaction speed of test trials for various (a) training data lengths and (b) MDI window lengths, averaged across the 10 subjects. The MDI window length in (a) and the training data length in (b) are both 90 seconds. Error bars show the standard errors of the mean.

yielded the significant correlation. Figure 2.7b illustrates that the correlation magnitudes were significantly smaller when the MDI window was shorter than 10 seconds ($p < 0.05$ by paired t-test) and also slightly decreased when the length is 300 seconds, and the optimal length was around 60 to 120 seconds. Hence this study selected the length of 90 seconds for both the training data and the MDI window. More details were discussed in Section 2.5.6.

2.5 Discussion

2.5.1 Nonstationary source compositions and activities

This study showed that ICA models trained with data segments when subjects showed different levels of alertness, indexed by reaction speed to driving challenges, had distinct source compositions and activities. From alert to drowsy state, there were fewer ocular components and more frontal and occipital components, and the alpha power spread from the occipito-parietal region into a larger brain area including the frontal and prefrontal regions, which were consistent with previous studies [65, 108]. Furthermore, the offline model contained the combination of ICs

found in the alertness and drowsiness models, indicating that it made a compromise between the nonstationary source compositions and activities over the entire session.

Previous studies have also reported nonstationary source compositions. Jung *et al*.[55] have shown that different ICA models, obtained by an ICA mixture model, accounted for different time periods of data when subjects' level of alertness fluctuated. Similarly, an adaptive ICA model revealed changes in both source compositions and activities in an alternating eye-open and eye-close experiment [48].

2.5.2 Distinct ICA models characterize different states

This study hypothesizes that neurocognitive processes or EEG sources that underlie different states, *e.g.* levels of alertness, can be characterized by distinct compositions and activities of independent sources obtained by ICA. Study results showed that the goodness-of-fit of the alertness ICA models, quantified by the model-deviation index, had significantly negative correlations with the reaction speed, demonstrating that the models were a good fit only when the subjects were alert, *i.e.* with high reaction speed. However, when the EEG data deviated from the alertness models where MDIs were high and subjects were drowsy, *i.e.* with low reaction speed, it only meant the data were dissimilar from that of the alert state. It did not necessarily mean the data were under the same (drowsy) state.

The story is only complete if we can model those data where the alertness models could not explain. This was achieved by the drowsiness ICA models. We found that when subjects were drowsy their EEG data shared the similar source composition and activation characterized by the same drowsiness models. This led to significantly positive correlations between the MDI and reaction speed, the opposite to the results of the alertness models. In summary, we have demonstrated that the alertness and drowsiness ICA models, learned from the small segments of the data based on the reaction speed, could characterize EEG data under alert and drowsy states respectively. This supports our hypothesis.

2.5.3 Estimating levels of alertness

If the hypothesis holds, we can estimate subjects' levels of alertness by quantifying the deviation of their EEG signals from each state-associated ICA models. This is based on the observations that the degree of model deviation, *i.e.* the MDI value, correlated with the levels of alertness inferred by the reaction speed. In other words, if different ICA models were applied to a data segment containing higher reaction-speed trials, the resulting MDI value would be lower for the alertness models and higher for the drowsiness model. Hence we can construct a feature space using the MDIs of the alertness and the drowsiness ICA models, and Figure 2.6 clearly demonstrates that trials with different reaction speeds can be separated in the feature space. The combination of MDIs of the alertness and the drowsiness models, *e.g.* their ratio, serves as a direct estimator of the alertness level.

So far we have focused on alertness and drowsiness models, but were there only two ICA models and states in the driving experiment? The MDI feature space of Subject 2 (Figure 2.6) shows that there were some trials that could not be modeled by either the alertness or the drowsiness model. These trials turned out to be at the beginning of the session. This indicates that even though this subject performed well in both time periods, *e.g.* $t = 0 - 15$ min (beginning of the session) and $t = 60 - 75$ min (awaken from the previous drowsy state), their underlying ICA compositions and activities differed. Hence by exploring clusters in the MDI feature space, it is possible to identify the existence of distinct states characterized by ICA models. It is worth noting that this framework can be applied to model more cognitive or mental states beyond just alertness and drowsiness.

2.5.4 Toward online alertness monitoring

While the MDI of ICA models allow the estimation of alertness levels, to apply this approach as an online framework, either the models need to be transferable across sessions or the

models need to be learned from limited amount of online training data. It is possible to use the ICA models for each class, *i.e.* the alertness and drowsiness models, trained on other sessions and apply both models to estimation of alertness levels in a test session. However, the generalizability of the ICA models across sessions with different electrode placements or subjects with different brain anatomies requires extra cautions and a significant amount of work. It is an open question to the field of ICA, but beyond the scope of this study.

Nevertheless, we have demonstrated the feasibility of learning a session-specific "initial" ICA model from limited training data, *i.e.* the first 90-sec of the online session, to monitor level of alertness for the rest of the session. This online framework is based on the assumption that at the beginning of the session the subjects are likely to be alert and thus their initial models resemble their alertness models, which is supported by the empirical results across subjects. By learning the initial models and measuring the corresponding MDIs of the test data, we can continuously and quantitatively measure the deviation of the subjects' current states from the initial (alert) state.

This ICA model approach achieved the averaged AUC of 0.745 across 10 subjects, which is higher than that of another state-of-the-art approach proposed by Pal *et al.*[93] using alpha and theta power of the Oz channel (AUC = 0.6715). But the difference was not statistically significant because of the inter-subject variability. It is worth nothing that we used the ROC curves and AUC measures to characterize the performance of alertness estimation rather than simply reporting the classification accuracy because levels of alertness should be continuous values, which is intrinsically lack of a clearly defined threshold for binary classification. For a truly online framework where a threshold is required for alertness/drowsiness detection, the threshold can be determined by the mean and variance of MDIs of the training trials.

2.5.5 Neurophysiology or artifacts?

ICA models not only assess neurocognitive or physiological brain sources but could also separate non-brain sources such as eye-blink, eye-movement, and possible muscle artifact components. Previous studies have demonstrated that eye and muscle activities were associated with sustained attention [65]. To test the effects of EOG/EMG sources on the prediction of the alertness levels, we firstly identified the artifactual components such as eye-blink and eye-movement components in each ICA model for each subject by template matching, where the templates were constructed by averaging the manually identified components in a few subjects. Secondly, we computed the MDIs without the artifactual ICs and performed the correlation analysis between the MDI and the reaction speed.

We found that the alertness models had reduced correlations after artifactual ICs were excluded, but were still significantly higher than the correlations of the offline ICA models. On the other hand, the drowsiness models with and without artifactual ICs showed no significant difference in correlations. A possible explanation is that the alertness models had more identified artifactual components and activities that were also associated with the alertness levels. These results confirmed that the EOG/EMG components combined with the neurophysiological sources in the ICA models could affect the prediction of the alertness levels. Nevertheless, even with the EOG/EMG sources removed, both models still demonstrated high correlations between MDI and reaction speed, suggesting the contribution of other brain sources in both models to the prediction of the alertness levels.

2.5.6 Considerations in building ICA models

To build an effective ICA model, we found that both the optimal training data length and MDI window length were around 90 seconds. An ideal training data length should be long enough to warrant the convergence of the ICA solution and short enough to avoid including

data from different brain states; the ideal MDI window length would be long enough to obtain a robust estimate of the MDI and short enough to obtain a locally stationary EEG signals. Previous studies have reported a 90-second window for measuring a global level of alertness [19, 72] and a 90-second alertness fluctuation cycle [77]. Our results were consistent with the previous findings and implied that EEG source dynamics associated with alert or drowsy states also have a locally stationary window around 90 seconds.

This ICA model-based framework is not limited to a specific ICA algorithm. We used Infomax ICA since it has been demonstrated to achieve good performance in separating dipolar neurophysiological sources in EEG data [30]. However, a limitation of the Infomax ICA algorithm is that learning an ICA solution with N stable sources requires $k \cdot N^2$ samples, where empirically $k \geq 25$ [90]. For $N = 30$, $k = 25$ and 250 Hz sampling rate, this corresponds to 90 seconds of data. The non-extended Infomax ICA has less computational complexity while the extended Infomax ICA [66] can separate both supergaussian and subgaussian sources and can be applied when the amount of training data are limited. The empirical results showed that both algorithms returned comparable results. It is worth noting that an alternative class of source separation algorithms based on second-order statistics (SOS) can also be applied. Congedo *et al*.[24] have suggested that SOS algorithms such as SOBI [9] and AJDC [24] may achieve better performance for shorter data segments and are more robust with respect to artifacts.

The selection of an ICA algorithm affects the derivation of the model deviation index (MDI). The MDI associated with the non-extended Infomax ICA defined in equation 2.3 is computationally efficient yet the normalization is not exact since f(y) is a nonlinear function. A more accurate but computationally expensive approach is to measure the likelihood of the ICA model applied to the data, as described in [68].

2.5.7 Comparisons with existing works

This study aimed to test the hypothesis of source nonstationarity and evaluated the feasibility of using multiple state-associated ICA models to monitor changes of the brain states. Compared to existing works, the proposed approach has the following advantages. (a) It is a generic framework that is not limited to assessing the levels of alertness or drowsiness. It can be applied to monitoring other brain states such as sleep stages, depth of anesthesia, etc. (b) It estimates continuous levels of deviations from a known state as opposed to classifying discrete states. (c) The quantitative model deviation indices are informative and discriminative features for state monitoring which are designable and interpretable. (d) The MDI feature space could reveal the possible existence of other states where their source compositions and activities could not be characterized by the existing models. (e) It does not require pre-selection of features such as channels, components, and frequency bands that may vary widely across sessions and subjects.

Previous studies chose specific frequency bands, *e.g.* alpha- and theta-band power of pre-selected channels, *e.g.* Oz [93], or independent components, *e.g.* bilateral occipital source [72], and showed the selected features were highly correlated with changes in the subject's drowsiness levels. However, the optimal frequency bands and EEG channels could vary widely across different subjects [93, 56].

We also proposed an online framework for alertness estimation using the initial ICA models. This online framework has the following additional advantages: (a) it is an unsupervised approach, which does not rely on the availability of data from both states (namely alertness and drowsiness), (b) it does not require training data from other sessions or subjects, (c) it builds an individual model for each session such that it avoids the problem of inter-session or inter-subject variability, and (d) it only uses a small amount of training data namely 1.5 minutes for 30-channel EEG, at the beginning of the session.

To make a fair comparison, we chose the Pal's approach [93] because it is among a few approaches that satisfied the training and testing requirements described above. Charbonnier *et*

al.[16] have also proposed a similar online framework for mental fatigue estimation by measuring the distance between the EEG spatial covariance matrices of 20-sec epochs to a mean covariance matrix learned during an initial reference state. Most of the classic approaches for online alertness/drowsiness estimation, which use discriminative machine-learning models or regression, would not work in this case because they require a large amount of data from both states. Other approaches that train classifiers and models on data from other sessions also do not apply because they require additional training data.

2.5.8 Limitations and future works

One challenge of this approach is to select data window for training an ICA model of a known brain state. In this study we selected training data based on the subjects' reaction speed, assuming that the reaction speeds reflect their alert and drowsy states. For unlabeled data, one solution is to learn an ICA mixture model and perform unsupervised data segmentation as described in [68, 55, 97]. However, these approaches are offline analyses.

Another challenge is to make this approach online capable. This study proposed the use of the initial model to track the alertness level online. However, this still requires a small amount of training data from the same session and assumes the initial reference state is known, which might not be generalizable. An alternative approach could be applying the ICA models learned offline from other sessions or subjects and using transfer learning to select or adapt the models to the online session. Important future works include studying the generalizability of the ICA models across subjects and building unifying source-level models associated with brain states.

2.5.9 Potential applications

This ICA model-based framework is a generic approach for assessing EEG source non-stationarity associated with cognitive or psychophysiological processes or behaviors. With the

association between the source composition/activity and different brain states, we can continuously monitor the state changes. This EEG-based state monitoring approach has several important applications: (1) in monitoring the cognitive or mental states of humans in real-world situations, *e.g.* mental workload [35, 104], mental fatigue [35], emotional responses, attention and drowsiness [93, 122], (2) in monitoring patients' brain states in clinical settings, *e.g.* seizure activities, sleep stages, anesthesia level, etc., and (3) in neuroergonomics [35] and passive brain-computer interfaces (BCI) [124] to prevent catastrophic incidents, *e.g.* drowsy driving and fatigue working, and to implicitly enhance the effectiveness of human-machine interaction.

2.6 Conclusion

This study presents a generic framework to quantitatively assess EEG source nonstationarity based on ICA and the model deviation index and to track the performance fluctuation in a sustained-attention task. Empirical results illustrate that both EEG source composition and activity were nonstationary as subjects' levels of alertness fluctuated. In general, when subjects are alert, *i.e.* performing the task at high reaction speed, the source composition of the recorded EEG signals shared the comparable source composition and activity characterized by the alertness ICA model. On the contrary, the EEG signals collected when the subjects were drowsy could be well modeled by the drowsiness ICA model. These results provide strong evidence for the hypothesis that brain activities under different levels of alertness can be characterized by distinct compositions and activations of independent sources, *i.e.* ICA models. Based on the idea of the source nonstationarity, this study also proposes an online framework using a subject-specific ICA model trained with an initial reference (alert) state to track the level of alertness. The proposed online framework compared favorably with a state-of-the-art approach using alpha and theta power for the classification of alert against drowsy states. This ICA-based framework can be applied to monitoring cognitive or mental states of human operators in attention-critical settings

or in passive brain-computer interfaces.

2.7 Acknowledgments

Chapter 2, in part, is a reprint of the material as it appears in Hsu, S. H., and Jung, T. P. (2017). “Monitoring alert and drowsy states by modeling EEG source nonstationarity.” *Journal of Neural Engineering*, 14(5), 056012. The dissertation author was the primary investigator and author of this paper. This study was sponsored in part by the US Army Research Laboratory and was accomplished under Cooperative Agreement Number W911NF-10-2-0022. The views and the conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S Government. The U.S Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein. The authors would like to thank Scott Makeig, Luca Pion-Tonachini, Makoto Miyakoshi, and other members of the Swartz Center for Computational Neuroscience for the fruitful discussion and valuable suggestions.

Chapter 3

Modeling Brain Dynamic State Changes with Adaptive Mixture Independent Component Analysis (AMICA)

In this chapter, we explore the use of adaptive mixture independent component analysis (AMICA) to model multichannel electroencephalographic (EEG) data with a set of ICA models, each of which decomposes an adaptively learned portion of the data into statistically independent sources. We first show that AMICA can segment simulated quasi-stationary EEG data and accurately identify ground-truth sources and source model transitions. Next, we demonstrate that AMICA decomposition, applied to 6-13 channel scalp recordings from the CAP Sleep Database, can characterize sleep stage dynamics, allowing 75% accuracy in identifying transitions between six sleep stages without use of EEG power spectra. Finally, applied to 30-channel data from subjects in a driving simulator, AMICA identifies models that account for EEG during faster and slower response to driving challenges, respectively. We show changes in relative probabilities of these models allow effective prediction of subject response speed and moment-by-moment characterization of state changes within single trials. AMICA thus provides a generic

unsupervised approach to identifying and modeling changes in EEG dynamics. Applied to continuous, unlabeled multichannel data, AMICA may likely be used to detect and study any changes in cognitive states.

3.1 Background

An expanding focus in neuroscience has been on endogenous temporal dynamics of neural network activity that gives rise to fluidity and rapid adaptability in cognition and behavior. A growing body of evidence suggests that these temporal dynamics may arise from continual formation and dissolution of interacting cortical and allied subcortical source activities in large-scale brain regions whose joint electrical activities can be described as dynamic systems featuring continuous transitions between intermittently stable states [18, 10].The temporal dynamics and network topology of these "brain states" can be identified using functional neuroimaging techniques including invasive electrophysiological recordings, functional MRI (fMRI), magnetoencephalography (MEG), and electroencephalography (EEG) [34, 18]. Among noninvasive modalities, EEG provides a more direct measurement of brain activity with millisecond resolution that, because of the low weight and bulk of its sensors, is appropriate for studying fast-changing brain states in real-world environments.

Earlier methods applied nonparametric statistical approaches that used EEG power spectral density, autocorrelation function, and entropy measures [87] to detect change points allowing segmentation of EEG into piecewise stationary processes [58]. Microstate analysis (see Khanna *et al.*[61] for a review) takes the spatial distribution of electrodes into account and attempts to define quasi-stable "microstates" in terms of unique electric potential patterns across the multichannel EEG scalp electrode montage during behavioral states or resting states [70, 115]. The global functional connectivity approach [18, 10] measures inter-electrode channel signal synchrony to attempt to characterize brain states as stable functional networks. However, both the microstate

and global connectivity models analyze scalp electrode signals that in themselves are highly correlated through common volume conduction and summation at the electrodes of potentials arising from brain and also non-brain sources (eye movements, ECG, etc.). The results of both methods have few or no interpretable connections to particular brain source activities that underlie the observed scalp phenomena.

To study source-resolved EEG activities, independent component analysis (ICA) [57, 8] has been widely applied as a means for blind source separation [74, 55, 78]. ICA models data, x, as an instantaneous linear mixture of statistically independent source processes, s ($x = A \cdot s$, where A is a mixing matrix). A physiological interpretation of ICA applied to scalp EEG recordings can be found in [91] and [30]. In short, this research has clarified: 1) functional independence across brain regions, similar to the regional dependence and independence now measured by functional brain mapping using fMRI, should be accompanied by temporal independence of the source EEG activities; 2) linear and instantaneous mixing of source EEG activities is produced by volume conduction and scalp mixing. While ICA works well in separating out localized cortical sources of event-related potentials and eye-movement related activity, it is limited in modeling nonstationary changes in EEG source locations and activities because of its spatial-stationarity assumption produced by its limited model complexity (*e.g.* its use of a fixed spatial mixing matrix in which the number of learned temporally independent components is equal to the number of channels measured).

Our recent study [44] hypothesized that transitions to a different cognitive state may involve cortical macro- or meso-dynamics in new networks of cortical brain areas that can be identified by distinct ICA models trained on data recorded before and after the state transition, respectively. This hypothesis motivates the application of an extension of ICA – the ICA mixture model (ICAMM) [67] – an unsupervised learning approach to modeling EEG activities in different brain states and detecting brain dynamic state changes associated with cognitive state changes [55]. The ICAMM assumes distinct ICA models may better characterize different segments

of nonstationary data, i.e. $x(t) = A_h s_h(t)$ where h is the model index. By allowing multiple ICA models to focus simultaneously on different parts of the data, ICAMM relaxes the spatial-stationarity assumptions and allows more total sources to be learned than the number of channels. ICAMM is thereby capable of modeling nonstationary, multi-state data and thus is a promising approach to studying dynamic changes in cognition and brain states. While the few prior attempts to apply ICAMM to EEG data were able to monitor attention [55], to detect microarousals during sleep [106], and to detect mental state changes during a memory test [105], the full power of the ICAMM approach has not yet been demonstrated, including modeling of multiple brain states, tracking of state transitions in continuous recordings, consistency of the learned models across subjects, and more precise physiological interpretation of those models.

Here we report an EEG study using an unsupervised ICAMM to investigate dynamics of cognitive states. For this we chose an adaptive mixture ICA (AMICA), proposed by Palmer *et al*.[97] that adaptively learns individual source probability density functions (PDFs) as well as source scalp projection patterns. Palmer *et al*.[97] has also provided an efficiently optimized algorithm for learning an ICAMM from multichannel data using a parallel implementation (the code is available at `https://sccn.ucsd.edu/~jason/amica_web.html` and also as an open source plug-in for EEGLAB [28] at `https://sccn.ucsd.edu/wiki/EEGLAB_Extensions_and_plug-ins`). In the following sections, we will show: 1) AMICA learns the ground truth in simulated quasi-stationary data – we test the effect of numbers of ICA models on AMICA performance; 2) AMICA usefully characterizes sleep EEG dynamics, producing consistent results across subjects that can be applied to classify six sleep stages; 3) AMICA can quantitatively assess subjects' continuous changes in attention and drowsiness levels during simulated driving and thereby can track brain dynamic state changes at single-trial level with millisecond resolution; and 4) AMICA provides interpretable models allowing computation of the spatial distribution and frequency content of active sources in each brain state.

3.2 Materials and Methods

3.2.1 Datasets and preprocessing

Dataset I: simulated quasi-stationary data

To systematically validate AMICA, we use the EEG data simulator in the Source Information Flow Toolbox (SIFT) [29] to simulate a quasi-stationary dataset in which underlying sources are alternatingly active and inactive. We first generate 28 source time-series using EEG-like autoregressive models driven by super-Gaussian noise (simulated 250-Hz sampling rate, 135k samples). The nine total minutes of simulated source data are partitioned into three 3-min segments, each with a different set of 16 active sources. For each of the $2^3 = 8$ possible combinations of active/inactive sources over three segments, *e.g.* active-inactive-active, etc., we simulate 4 sources that follow such a pattern, excluding the inactive-inactive-inactive pattern in which no sources are active. This results in 16 active sources in each segment with 8 active sources changing between segments. Each source is assigned to the center of a randomly selected cortical region in the MNI brain atlas, with its dipole moment oriented normal to the cortical surface. Finally, the simulated source-level activities are projected through a zero-noise three-layer boundary element method (BEM) forward model (MNI Colin27), yielding simulated 16-channel EEG data. More details are included in [48].

Dataset II: CAP sleep database

We used 17 human EEG recordings, each consisting of 6-10 hours of sleep, from the CAP sleep database [112] on the PhysioNet [37]. Excluding subjects whose recordings had less than 5 channels gave 7 EEG datasets from healthy subjects. We also used EEG recordings from 10 patients with nocturnal frontal lobe epilepsy (NFLE), selected on the basis of data quality, i.e. longer data length, higher number of channels and more balanced numbers of sleep labels. We included NFLE patient recordings in an attempt to test ability of the proposed approach to

generalize across subjects and patients.

The EEG data comprise 6-13 bipolar channels (*e.g.* F3-C3, C3-P3, P3-O1, O1-A1, without common reference) affixed at scalp sites in the International 10-20 System and recorded with a sampling rate of 128 Hz using a Galileo System (Esaote Biomedica). After collection, the EEG signals were band-pass filtered between 0.5 Hz to 25 Hz. The hypnograms had been annotated by expert neurologists at 30-second intervals using standard R&K rules into 6 sleep stages: wake (W), rapid eye movement (REM), and 1 to 4 non-REM sleep stages (N1, N2, N3 and N4). More detailed description of the data and their hypnograms can be found at `https://physionet.org/pn6/capslpdb/`.

Dataset III: drowsiness fluctuation in simulated driving

Ten healthy volunteers participated in a 90-min experiment in an immersive VR-based driving simulator, performing an event-related lane-departure task [49]. The subjects experienced visually presented lane-departure events every 8-12 seconds (with randomized event onset asynchronies) and were instructed to steer the car back to the cruising position quickly using a steering wheel. The duration between the onset of a lane-departure event to the onset of a responsive steering action was defined as subject reaction time (RT), which can be used to index degree of subject alertness / drowsiness [72]. The RT data were transformed to reaction speed (RS = 1/RT) to partially normalize the highly skewed RT distribution. For more details on the subjects and the experiment, refer to [72].

For each subject, 30-channel EEG data were recorded with a 500-Hz sampling rate using a NeuroScan System (Compumedics Ltd., VIC, Australia) with electrode sites according to the International 10-20 System. The EEG data were band-pass filtered (1-50 Hz) and downsampled to 250 Hz. Using the PREP pipeline [11], poorly recorded channels in the recordings, such as channels with flat signals arising from poor electrode contacts, and channels whose signals were poorly correlated with those of neighboring channels were removed. Two to six channels were so

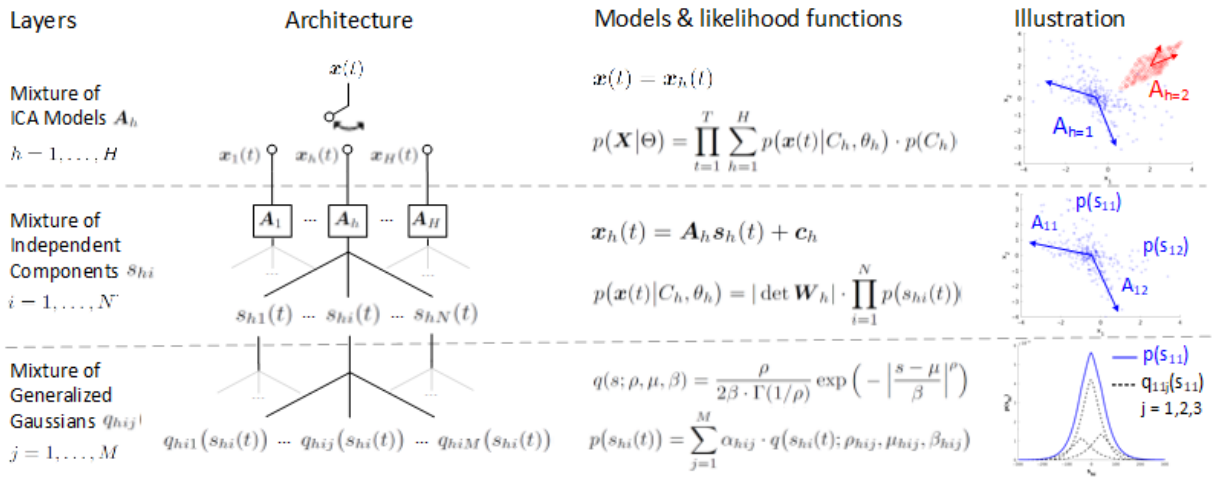

Figure 3.1: Adaptive Mixture ICA (AMICA) in a nutshell. AMICA consists of three layers of mixing. As shown in illustration, the first layer is mixture of ICA models A_1 and A_2 that learn the underlying data clusters, simulated based on Laplace and uniform distribution respectively. The second layer is mixture of independent components A_{11} and A_{12} that decompose the data cluster into statistically independent sources' activations s_{11} and s_{12}. The third layer is mixture of generalized Gaussian distributions q_{11j} that approximate the probability distribution of the source activation $p(s_{11})$.

identified and removed for each of the 10 subjects. In addition, artifact subspace reconstruction (ASR) [82], implemented as a plug-in to the EEGLAB environment [28], was applied using a mild threshold (burst repair $\sigma = 20$) to reduce data contamination by high-amplitude artifacts. These artifact-correction methods were chosen to facilitate convergence of the ICA models. Detailed description of the data pre-processing can be found in [44].

3.2.2 Method description

Comprehensive formulation of the ICAMM problem and detailed derivation of the AMICA algorithm have been presented in [67] and [97] respectively. The following sections give a brief summary of the multi-model AMICA approach in an attempt to provide intuition and facilitate readers' understanding.

Adaptive mixture ICA (AMICA)

Figure 3.1 gives a schematic overview of the architecture of AMICA and its models. AMICA is, conceptually, a three-layer mixing network: the top two layers constitute one or more ICA mixture models and the bottom layer, specific to AMICA, focuses each learned model on accounting for a subset of the data.

Starting from the top layer, the key assumption of multiple mixture models is that data $X = \{x(t)\}$ (N-channel by T-time samples) are nonstationary, so that different models may be dominant in characterizing the data at different times, i.e. $x(t) = x_h(t)$ where h is the model index. Previous studies have provided evidence that EEG activities during different brain states (*e.g.* alert versus drowsy) are nonstationary and can be modeled by a finite set of distinct ICA models [44].

In the top two layers of the AMICA network, a standard ICA model is employed to model the data x as an instantaneous linear mixture A ($N \times N$ matrix) of statistically independent components s, i.e. $x = As$. The first two layers consist of the ICA mixture model:

$$x(t) = x_h(t) = A_h s_h(t) + c_h, \quad h = 1, \ldots, H \tag{3.1}$$

where $h = h(t)$ and A_h is the dominant or active model at time t with source activities $s_h(t)$ and bias c_h. For simplicity, it is assumed that only one of the H models is active at each time and that the model index h and the data $x(t)$ are temporally independent. Hence the likelihood of data given the ICA mixture model can be written as:

$$p\big(X\big|\Theta\big) = \prod_{t=1}^{T} \sum_{h=1}^{H} p\big(x(t)\big|C_h, \theta_h\big) \cdot p(C_h) \tag{3.2}$$

where $\Theta = \{\theta_1, \ldots, \theta_H\}$ contains the parameters of ICA models and $p(C_h) = \gamma_h$ is the probability of the h-model being active that satisfies $\sum_{h=1}^{H} \gamma_h = 1$.

Given the assumption of statistical independence between components $s_{hi}(t)$ ($i = 1, \ldots, N$),

the likelihood of the data given the active ICA model $W_h = A_h^{-1}$ is:

$$p\big(x(t)\big|C_h,\theta_h\big) = |\det W_h| \cdot \prod_{i=1}^{N} p\big(s_{hi}(t)\big) \tag{3.3}$$

In the third layer of the AMICA network, the probability density function (PDF) of each component $p\big(s_{hi}(t)\big)$ is approximated by a mixture ($j = 1,\ldots,M$) of generalized Gaussian distribution $q(s)$ [95, 97]:

$$p\big(s_{hi}(t)\big) = \sum_{j=1}^{M} \alpha_{hij} \cdot q\big(s_{hi}(t);\rho_{hij},\mu_{hij},\beta_{hij}\big) \tag{3.4}$$

where α_{hij} is the weight for each PDF. The generalized Gaussian distribution parameterized by shape ρ, scale β and location μ is defined as:

$$q(s;\rho,\beta,\mu) = \frac{\rho}{2\beta \cdot \Gamma(1/\rho)} \exp\Big(-\Big|\frac{s-\mu}{\beta}\Big|^{\rho}\Big) \tag{3.5}$$

It is worth noting that most standard ICA mixture models, in contrast to AMICA, assume pre-defined PDFs for sub-Gaussian and super-Gaussian sources [67]. A previous study has shown that by adaptively learning the PDFs for each source, AMICA is able to achieve higher decomposition quality than other ICA approaches [30] when applied to 70-channel EEG data.

In the three-layer AMICA mixing network, the parameters to be estimated are $\Theta = \{W_h, c_h, \gamma_h, \alpha_{hij}, \beta_{hij}, \rho_{hij}, \mu_{hij}\}$ that correspond to the model index $h = 1,\ldots,H$, the component index $i = 1,\ldots,N$ and the PDF index $j = 1,\ldots,M$. The next section describes an efficient approach to estimating these parameters.

Parameter estimation and interpretation

The expectation-maximization (EM) algorithm is employed to estimate the parameters $\hat{\Theta}$ that maximize the data likelihood function in Equation 3.2. The algorithm consists of two-step

iterative learning involving alternating E-steps and M-steps. The E-step uses equations 3.3 and 3.4 to construct the expectation of the likelihood function in equation 3.2 using current estimates of the parameters $\hat{\Theta}^l$. The M-step maximizes the likelihood function returned by the preceding E-step. Instead of using standard or natural gradient approaches [67, 107], AMICA uses the Newton approach as derived by Palmer *et al.*[97] based on the Hessian (matrix of second-order derivatives) to achieve quadratic, and thus faster, convergence. For a detailed derivation and learning rules, see [97].

As an unsupervised approach with generative models, the Θ parameters learned by AMICA provide rich information about the underlying data clusters and their temporal dynamics. As illustrated in Figure 3.1, ICA models W_h and c_h can characterize distinct data clusters that represent different quasi-stationary states in the data. In addition, the corresponding source activations s_{hi} can be better estimated by α_{hij}, β_{hij}, ρ_{hij}, and μ_{hij} instead of assuming a fixed PDF as in many other ICA models including the original Infomax ICA [8]. Furthermore, the activation of each ICA model $h(t)$ can be represented as the likelihood of the data sample $x(t)$ given the estimated parameters of the model θ_h, using equations 3.2, 3.3 and 3.4:

$$L_{h(t)} = p(C_h) \cdot |\det W_h| \cdot \prod_{i=1}^{N} \sum_{j=1}^{M} \alpha_{hij} \cdot q\big(s_{hi}(t); \rho_{hij}, \mu_{hij}, \beta_{hij}\big) \tag{3.6}$$

Therefore the probability of activation of each ICA model at time t can be calculated by normalizing $L_{h(t)}$ across all models:

$$p\big(h(t)\big) = L_{h(t)} \Bigg/ \sum_{h=1}^{H} L_{h(t)} \tag{3.7}$$

This value, $p\big(h(t)\big)$, characterizes the temporal dynamics of activations of distinct states modeled by ICA and is referred to as "ICA model probability" in following sections.

Table 3.1: Parameters of Adaptive Mixture ICA for analyzing three datasets: I, simulated data; II, CAP sleep EEG data; III, drowsy driving EEG data.

Dataset	**I**	**II**	**III**
# of models (H)	2-6	8	2-4
# of sources (N)	16	6-13	24-28
# of PDFs (M)	3	3	3
# of rejection steps	0	15	15
Rejection thresholds	N/A	3	3
Max learning steps	2500	2000	2000

Application of multi-model AMICA

Multi-model AMICA decompositions were applied to all datasets described in Section 3.2.1 with the parameters specified in Table 3.1. For datasets II and III, rejection of data samples based on their posterior probabilities was applied to alleviate the effects of transient artifacts, such as data discontinuities, that might disrupt ICA learning. An efficient implementation of AMICA with parallel computing capability by Palmer *et al.*[97] was used in this study. The code for that implementation is available at `https://sccn.ucsd.edu/~jason/amica_web.html` and also as an open source plug-in for EEGLAB [28] at `https://sccn.ucsd.edu/wiki/EEGLAB_Extensions_and_plug-ins`.

3.2.3 Validation and quantitative analyses

Decomposition errors of ICA models

To determine whether AMICA could accurately decompose the simulated quasi-stationary data, three different measures were employed: model errors for unmixing matrices W_h, the signal-to-interference ratios (SIR) for source activities s_h, and the symmetric Kullback-Leibler (KL) divergence [64] for parameters of the source probability densities $\Theta = \{\alpha, \beta, \mu, \rho\}$. These measures are defined as follows:

Model errors For a known ground truth (N-by-N) mixing matrix A, the model error $E^{(h)}$ of the learned unmixing matrix W_h and sphering matrix S can be defined [32]:

$$E^{(h)} = \frac{1}{N-1}\left[N - \frac{1}{2}\sum_{i=1}^{N}\left(\frac{\max_{1\le j\le m}|C_{ij}^{(h)}|^2}{\sum_{j=1}^{N}|C_{ij}^{(h)}|^2} + \frac{\max_{1\le j\le m}|C_{ji}^{(h)}|^2}{\sum_{j=1}^{N}|C_{ji}^{(h)}|^2}\right)\right] \tag{3.8}$$

where $C^{(h)} = W_h \cdot S \cdot A$. This measure is a normalized total cross-talk error and accounts for scale and permutation ambiguities. In the case of perfect separation, $E^{(h)}$ equals zero. For three quasi-stationary segments and the three ground-truth mixing matrices A_k, we report the sum of logarithm of the model error for each segment k, i.e. total log error $= \sum_{k=1}^{3} \log E_k$, where $E_k = \min_{h=1\ldots H} E_k^{(h)}$.

SIR SIR estimates the log-scaled normalized mean-squared errors of the decomposed time series of the ith component, $\hat{s}_i(t)$, compared with ground-truth source activities, $s(t)$, defined as $\mathrm{SIR}_i = \min_{j=1\ldots N} 10\log_{10}\left(\sum_{t=1}^{T} s_j(t)^2 / \sum_{t=1}^{T}\left(\hat{s}_i^2(t) - s_j^2(t)\right)\right)$ [107] where T is the number of data samples. The final SIR is averaged across all components and then summarized across three quasi-stationary segments.

KL divergence The KL divergence quantifies the difference between the estimated PDF $P_i(x)$ of the ith source and the ground-truth source PDF $Q(x)$, which is defined as: $\text{KL-Div}_i = \min_{j=1\ldots N} \frac{1}{2}\left[D_{KL}(P_i||Q_j) + D_{KL}(Q_j||P_i)\right]$ where $D_{KL}(P||Q) = \sum_x P(x)\log\frac{P(x)}{Q(x)}$. $Q(x)$ was approximated by histograms because the ground truth source activities was generated by an autoregressive process instead of a parametric model.

Classification of sleep stages

To quantitatively assess results of unsupervised segmentation of the sleep EEG data by AMICA decomposition, we used ICA model probabilities (equation 3.7) as features and applied a Gaussian Bayes classifier to 30-second data windows to classify the data into six sleep stages.

The Gaussian Bayes classifier models the features of each class as a D-dimensional multivariate Gaussian distribution:

$$G(x;\mu,\Sigma) = \frac{1}{\sqrt{(2\pi)^D \det|\Sigma|}} \exp\Big(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\Big) \tag{3.9}$$

where μ is the mean and Σ is the covariance matrix estimated from the training data with the same class. To classify a test data window, the classifier compares the posterior probabilities of each class given the test data x:

$$C_k = \underset{k}{\operatorname{argmax}}\, p(C_k|x) = \underset{k}{\operatorname{argmax}}\, p(x|C_k)\cdot p(C_k) = \underset{k}{\operatorname{argmax}}\, G(x;\mu_k,\Sigma_k)\cdot p(C_k) \tag{3.10}$$

where $p(C_k)$ is the prior distribution of class k. In this study, the relative proportion of labels in each class is used as the prior distribution.

Five-fold cross-validation was performed for each subject data set. To ensure each fold had enough training data for each class, the data were first pooled according to their labels and then divided into five folds. The cross-validation accuracy and the confusion matrix were computed and the results summarized across subjects. The effect of the number of features used, i.e. model probabilities, on classification accuracy was also tested. It is worth noting that a generative classifier like the Gaussian Bayes classifier was here employed not to produce optimal classification accuracy but to illustrate the separability of EEG activities into six sleep stages using the feature space learned by AMICA decomposition.

Distinguishing alert versus drowsy behavior

A relational analysis was performed for dataset II (Section 3.2.1) to quantitatively evaluate the relationship between ICA model probabilities and drowsiness level as indexed by decreased reaction speed to driving challenges introduced into a simple driving simulation. Here AMICA model probabilities were first computed for 5-second data windows immediately preceding onsets

of lane-departure event (as might be produced during actual driving by unseen crosswinds). Pearson correlation coefficients were computed between preceding model probabilities and reaction speeds across all driving challenge trials. To assess longer-lasting fluctuations in behavioral drowsiness level over a driving session, a 90-sec smoothing window was applied [77]. Median reaction speed and model probabilities were computed across the 5-10 trials in each 90-sec window. The effects of model probability smoothing length is discussed in [44].

Clustering ICA models across subjects

To examine the consistency of the learned AMICA models across subjects, the models were clustered using iterative template-matching according to the relation between model probabilities and data labels (sleep stages). For each subject, mean model probabilities were obtained for each model in each sleep stage, giving a matrix of 6-stages by 8-models, P_i, for each subject (i). Each column of P_i was normalized to sum to one for further analysis. Next, a typical (template) subject was selected and his/her P_0 was used as an initial condition. The ICA models for each subject (i) and the template subject were matched. Model pair Pearson correlation coefficients (between models in columns of P_i and columns of P_0) were computed using $matcorr()$ in EEGLAB [28]. The sorted model probabilities $\hat{P}_i$ were averaged across the N subjects to obtain the new template $P_0^{(1)} = \frac{1}{N}\sum_i^N \hat{P}_i$ in place of P_0. The above template-matching process was iterated until the difference between new and old templates was smaller than a pre-defined threshold, i.e. $\sum_i \sum_j |P_{0,ij}^{(t)} - P_{0,ij}^{(t-1)}| \leq \varepsilon$ for t-th iteration. This study used $\varepsilon = 0.1$ to ensure that the results were consistent regardless of the choice of template subject.

Clustering independent components across subjects

Clustering of independent components (ICs) was performed to identify across-subject IC equivalences within model classes. The IC clusters were obtained using the $corrmap()$ plug-in in EEGLAB [116] using component similarity assessed by scalp map correlations. An IC scalp

map is the set of relative contributions or projections of the IC source to the scalp channels, i.e. a column of W^{-1} in equation 3.3. Component templates were selected visually. The number of ICs contributed by each subject was limited to two for the centro-occipital cluster and one for the other clusters. The following correlation thresholds were used: 0.9 for eye-blink and eye-movement clusters; 0.85 for the other clusters. These parameters were carefully chosen to avoid assignments of near-duplicate ICs to different clusters and to reduce variability produced by template selection.

3.3 Results

3.3.1 Dataset I: Validation using simulated data

Automatic data segmentation by ICA model probability

Figure 3.2 shows mean and upper/lower-bound model probabilities of the model clusters at each time point, across 100 repeated runs each decomposed using 3-, 4-, 5- and 6-model AMICA. All the 3-, 4-, 5- and 6-model AMICA decompositions successfully segregated data within the three simulated quasi-stationary segments, assigning them distinct ICA models (those with the highest probabilities, here labeled M1-M3).

Variability in model cluster probabilities across simulations, indicated by the heights of the shaded regions representing the 90 and 10 percentiles of the probability distributions, increased as the numbers of models used were larger than the simulated ground truth (3 models). For these (over-complete) mixture model decompositions, model clusters M4-M6 were more probable than model clusters M1-M3 only in small portions (3% to 7%) of the data. Under-complete 1-model and 2-model AMICA decompositions (Figure 3.2) tended to model ground truth in one or two of the three simulated data segments. Overall, complete and over-complete AMICA decompositions accurately segmented the nonstationary simulated data in an unsupervised manner.

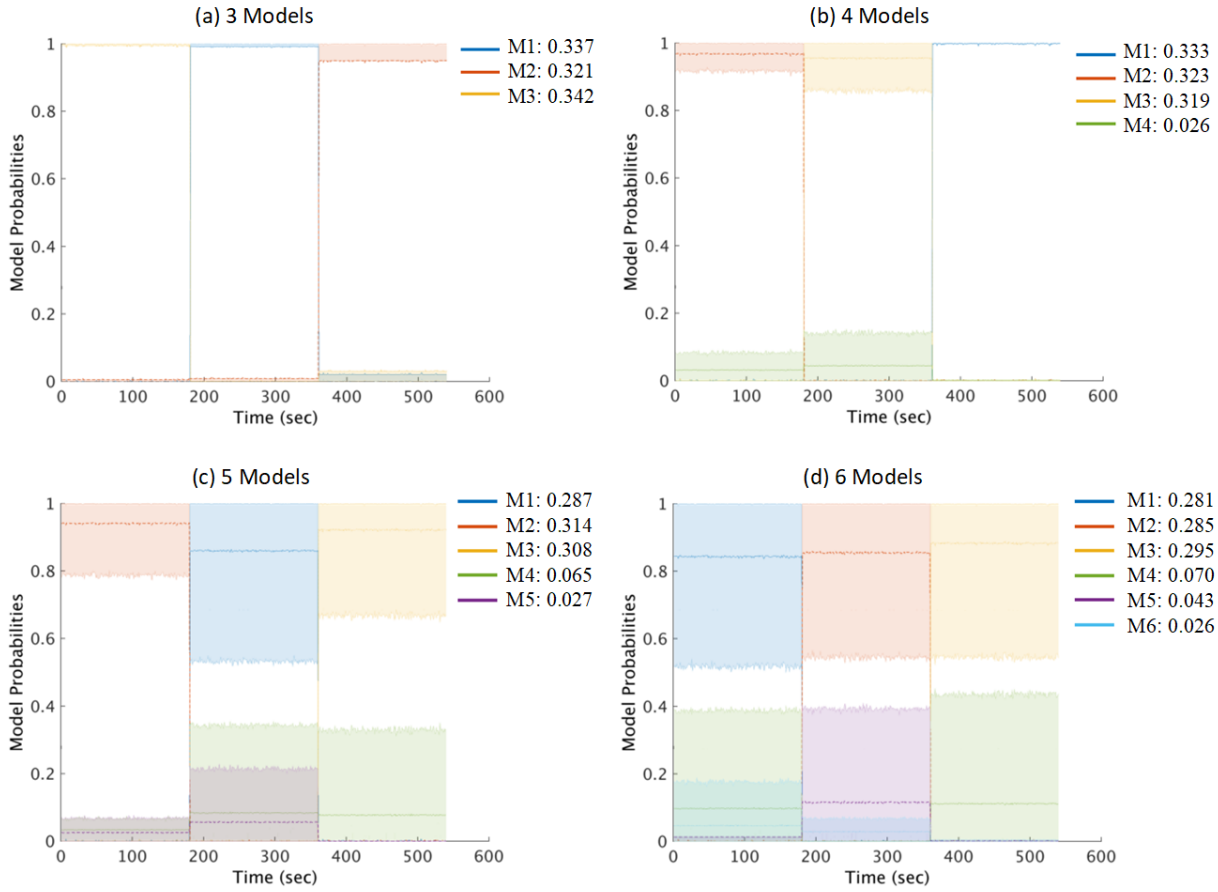

Figure 3.2: Mean changes in AMICA model probabilities clustered across AMICA decompositions of 100 repeated runs applied to the simulated quasi-stationary data. (a) AMICA decompositions using 3 models, (b) 4 models, (c) 5 models, and (d) 6 models. Upper and lower edges of the shaded regions represent the 10th and 90th percentiles of the cluster normed probability distribution. Figure legends give the mean probabilities $p(C_h)$ for each model cluster.

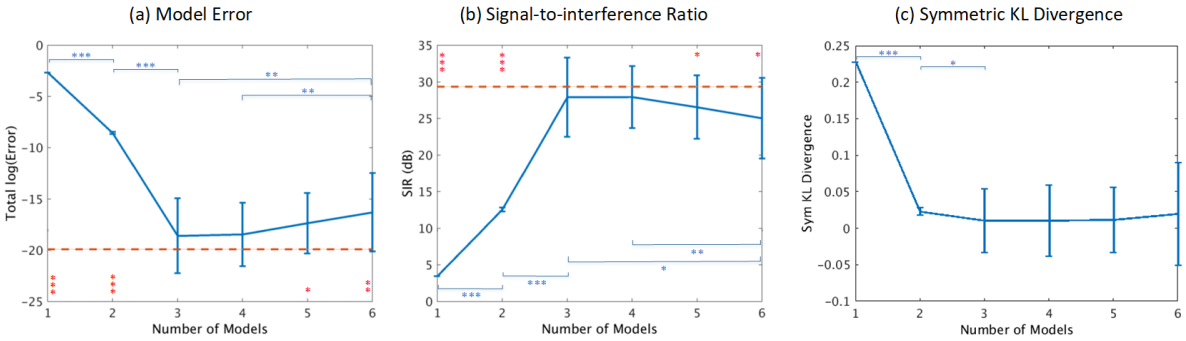

Figure 3.3: (a) Model errors in the learned model unmixing matrices versus simulated ground truth, (b) signal-to-interference ratios (SIR) of the decomposed model source activities, and (c) symmetric KL divergence of the learned source probability densities for AMICA decompositions using 1-6 models each averaged across 100 simulations. Red dashed line indicate the performance of single-model Infomax ICA applied to each single-model data segment. Significant differences in unpaired t-tests are shown (* $p < 0.01$, ** $p < 1 \times 10^{-4}$, *** $p < 1 \times 10^{-6}$). Red asterisks denote comparisons between AMICA and Infomax ICA models; blue asterisks denote comparisons between AMICA model orders. Overall, model errors were lowest for veridical (3-model) and slightly over-complete (4-model) AMICA decompositions.

AMICA decomposition errors

Figure 3.3a shows that 3- and 4-model AMICA decompositions achieved model errors comparable to the combined results of Infomax ICA decompositions of the single-model data segments (difference probability, $p = 0.11$ by unpaired t-test), demonstrating the three ground-truth mixing matrices could be learned accurately by AMICA without identified model boundaries. By comparison, the performance of 5-model AMICA was slightly worse and model errors for 6-model AMICA were significantly higher for 3- and 4-model AMICA decompositions. Nevertheless, 6-model AMICA still outperformed under-complete 1-model and 2-model AMICA decompositions applied to the three data model segments.

Figure 3.3b shows that 3- and 4-model AMICA decompositions gave the highest SIR, 5- and 6-model decompositions marginally lower and 1- and 2-model AMICA decompositions still lower SIR. Both 3- and 4-model AMICA decompositions achieved SIR results comparable to Infomax ICA run on the single-model data segments ($p = 0.24$ and 0.14 respectively). These results show that the ground-truth source activities for each model segment were well reconstructed

by complete (or, here, slightly over-complete) AMICA decompositions.

Figure 3.3c shows that 3- to 6-model AMICA decompositions produced the smallest (on average, near-zero) KL divergence values, suggesting that the source probabilities densities were also properly approximated. Here, 2-model AMICA performed slightly worse ($p < 0.05$) and 1-model AMICA much worse.

In summary, 3-model AMICA decomposition could simultaneously and accurately learn the true mixing matrices, source activities, and probability densities for three independent component models used to simulate 3-segment quasi-stationary data. AMICA performance using an unsupervised learning approach was comparable to Infomax ICA applied to each segment separately in a supervised fashion. Further, slightly over-complete (4-model) AMICA decompositions produced nearly comparable results, and performance only marginally decreased as the number of AMICA models was further increased.

3.3.2 Dataset II: Classify sleep stages

We applied 8-model AMICA to 17 sleep EEG datasets to evaluate AMICA performance applied to actual EEG data and to assess its capability to distinguish the 6 conventional sleep stages from the data themselves without regard to changes in spectra or other time series properties.

Model probabilities characterize sleep dynamics

To illustrate the temporal dynamics learned by 8-model AMICA from the sleep EEG data, Figure 3.4 shows the sleep stages annotated by experts and the probabilities of AMICA models ordered by overall data likelihood in one sleep session. Four distinct patterns of model probability changes were observed: (1) Models M1 and M2 were relatively active, i.e. had high model probabilities, during light sleep (N1 and N2) and had low probabilities during deep sleep (N4). Model M1, however, was more probable in rapid eye movement (REM) sleep than model M2. (2) In contrast, both models M3 and M5 were active only during deep sleep (N3 and N4).

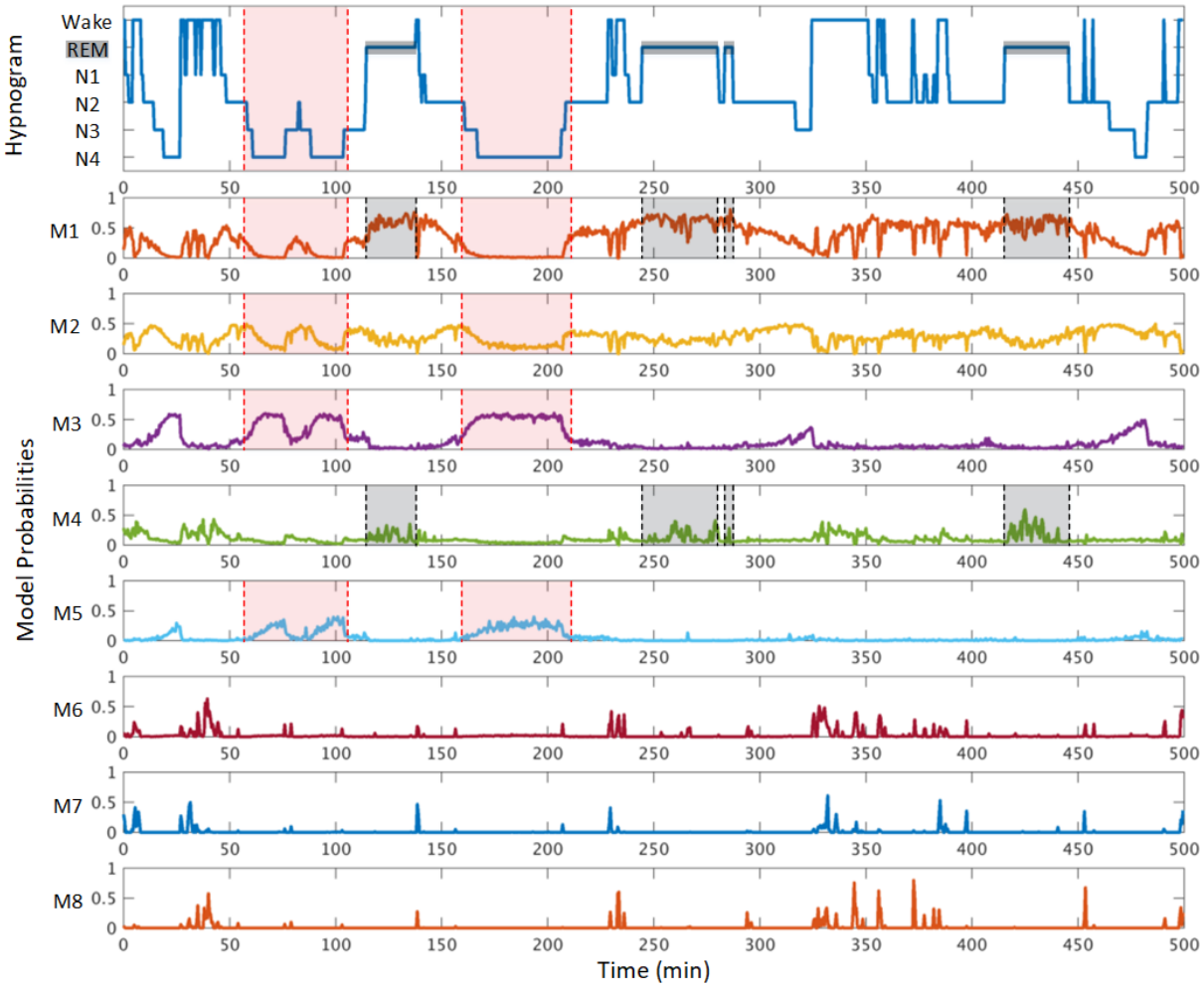

Figure 3.4: The top panel shows the hypnogram, *i.e.* sleep stages annotated from the EEG record by a sleep expert, of a sleep session from a single subject. Bottom panels show probabilities of ICA models learned by 8-model AMICA decomposition applied to the EEG record. Red-shaded regions highlight changes in model probabilities for relevant models during transitions to and periods of deep sleep (N4). Gray-shaded regions highlight probability value changes for relevant models during REM sleep.

These first two patterns sufficiently characterized changes from light sleep to deep sleep and back again (red-shaded regions) over the course of the sleep session. (3) Model M4 was most probable (gray-shaded regions) during REM sleep and in the wake state. (4) Probabilities of models M6, M7 and M8 rose only sporadically, mainly in the wake state.

Thus, the probabilities of the 8 learned ICA models for this session had notable relationships to the annotated sleep stages, but ICA model probabilities could not be mapped one-to-one with sleep stages. Some ICA models appeared to jointly characterize a sleep stage (*e.g.* M1 and M2 for N2, and M3 and M5 for N4), while probabilities for other models rose in different sleep stages (*e.g.* M1 probability rose briefly during N1, N2 and REM stages).

The dynamics of the model probabilities suggested that the changes in EEG activities during transitions between sleep stages were continuous as opposed to discrete – unlike as indicated by the hypnogram (scored by convention in successive 30-second intervals). Transition times varied across sleep stages. For example (red-shaded regions), major model probability shifts for models M1, M2, M3 and M5 had slower transitions (5-10 min) from stage N2 to N4 than from N4 to N2 (2-5 min). Some model probability transitions began before changes in the annotated sleep-stage labels. These results provide compelling evidence that AMICA model probabilities might be used to study the dynamics of EEG changes during sleep at much finer (*e.g.* approaching sample-by-sample) temporal resolution than offered by standard sleep scoring.

Relationships between ICA models and sleep stages across subjects

Next, we explored relationships between ICA models and sleep stages to assess if these relationships could be generalized across subjects using iterative template-matching of models from different subjects (Section 3.2.3). Figure 3.5 shows that ICA model clusters across subjects could be built based on relationships between data-driven model probabilities and annotated sleep stages. Resulting standard deviations of cluster model probability in each sleep stage were surprisingly small. Furthermore, each AMICA model cluster probability profile across

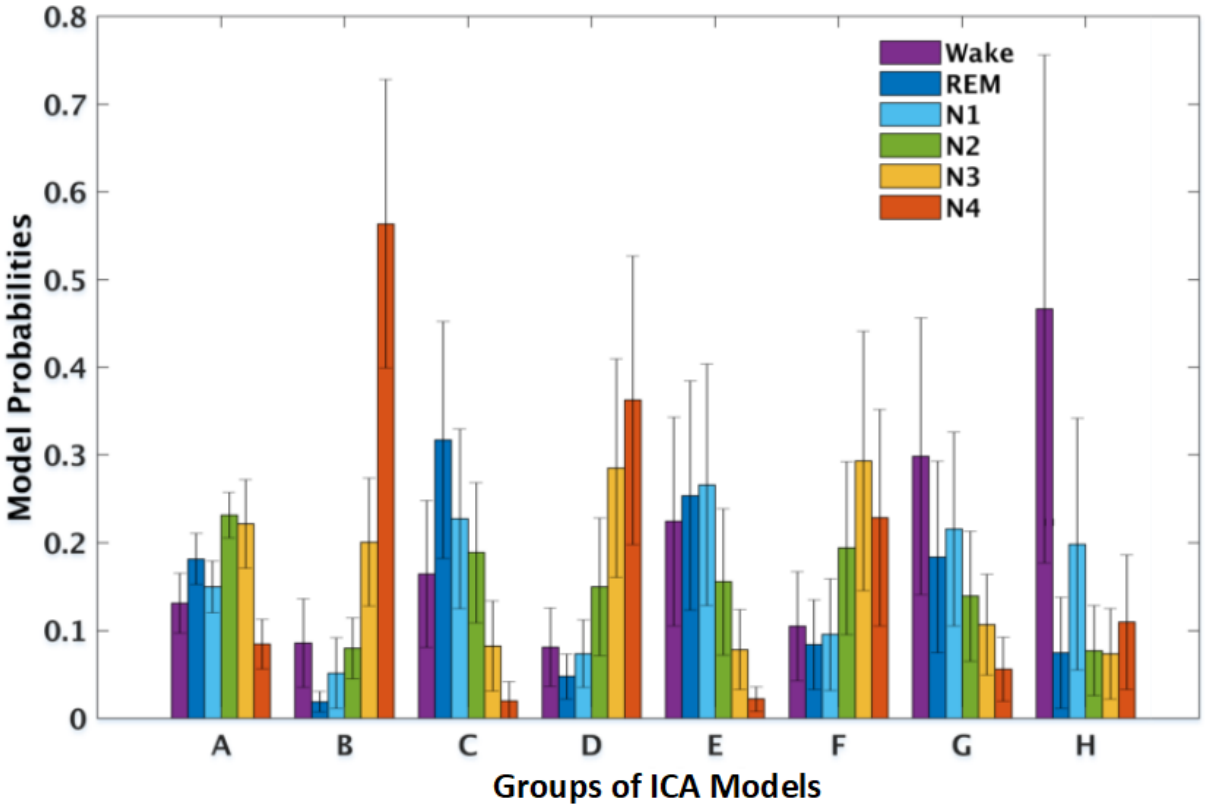

Figure 3.5: Cross-subject mean (plus one standard deviation) model probabilities of 8 AMICA model clusters in six sleep stages. Model clusters were composed of best-matched models across subjects, as found by iterative template matching.

sleep stages was distinct. For example, model A was relatively active in lighter sleep (N2 and N3), models B and D in deep sleep (N4), models C, E, and F in REM and stages N1 and N3, respectively. Models G and H were most probable during the wake state.

To visualize relationships between ICA model probabilities and sleep stages, Figure 3.6 presents 30-sec window-mean model probabilities for model clusters A, B, and C (cf. Figure 3.5) for all 17 subjects. Model probability values in the different (color-marked) sleep stages are clearly separated in this feature space. The progression from light sleep (N2, green) to deeper sleep (N3, yellow, to N4, red) is associated with smooth changes in cluster model probabilities. Model probabilities in the (purple) wake state were mostly low (near the *(0,0,0)* corner). These characteristics were consistent across AMICA models from 7 healthy subjects and 10 patients with nocturnal frontal lobe epilepsy.

Quantitative analysis: classification accuracy

To quantitatively assess the potential utility of model probabilities for separating sleep stages, we entered the window-mean model probabilities from the 8-model AMICA decomposition into a Gaussian Bayes classifier that fits a Gaussian distribution of 8-model probability vectors for each of the six annotated sleep stages (Section 3.2.3), and measured classification accuracy using 5-fold cross validation for each subject.

Figure 3.7a shows classification accuracy across all subjects. Accuracy improved when the number of model clusters was increased up to the use of the first three clusters. For all data (blue curve), mean accuracy was 74% - 76% when using three or more cluster model probabilities as features (no significant difference was observed by paired t-test). Classification accuracy was much lower (to 45% - 49%, yellow curve) for 30-sec data windows near a state change (*e.g.* when the sleep-stage label was different from that of the previous or succeeding windows). Accuracy was higher (78% - 80%, red curve) when the window was not near a state change.

Note that classification accuracy was biased by the unbalanced class sample sizes. Fig-

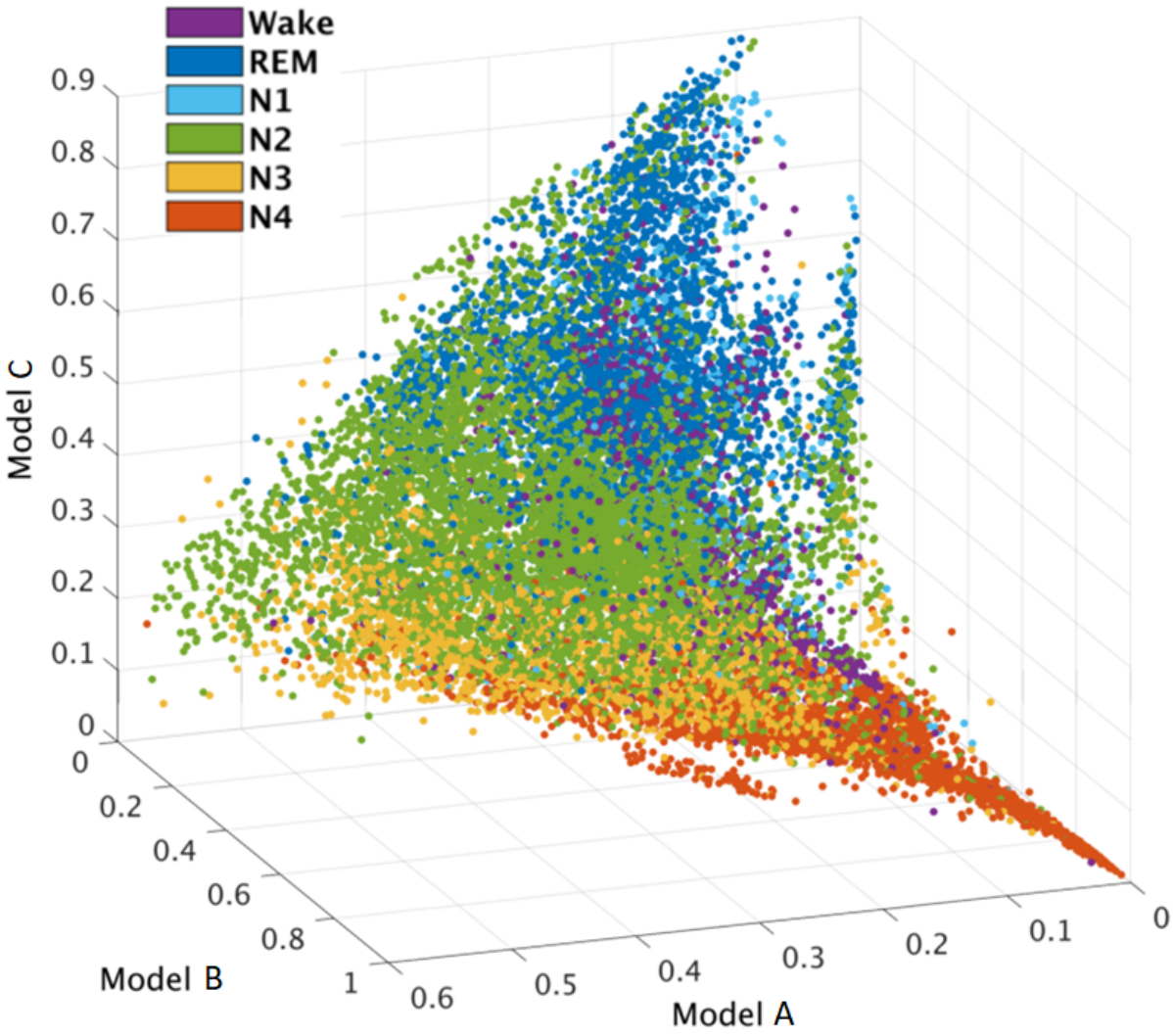

Figure 3.6: Scatter plot of window-mean model probabilities for AMICA model clusters A, B, and C (cf. Figure 3.5), each point representing mean model probability within a 30-sec data segment from sleep recordings of 7 healthy subjects and 10 patients. Colors represent expert designated sleep-stage labels for the same data segments. Note the distinct deep sleep (N4) pattern and the relative closeness of wake and REM sleep characteristics.

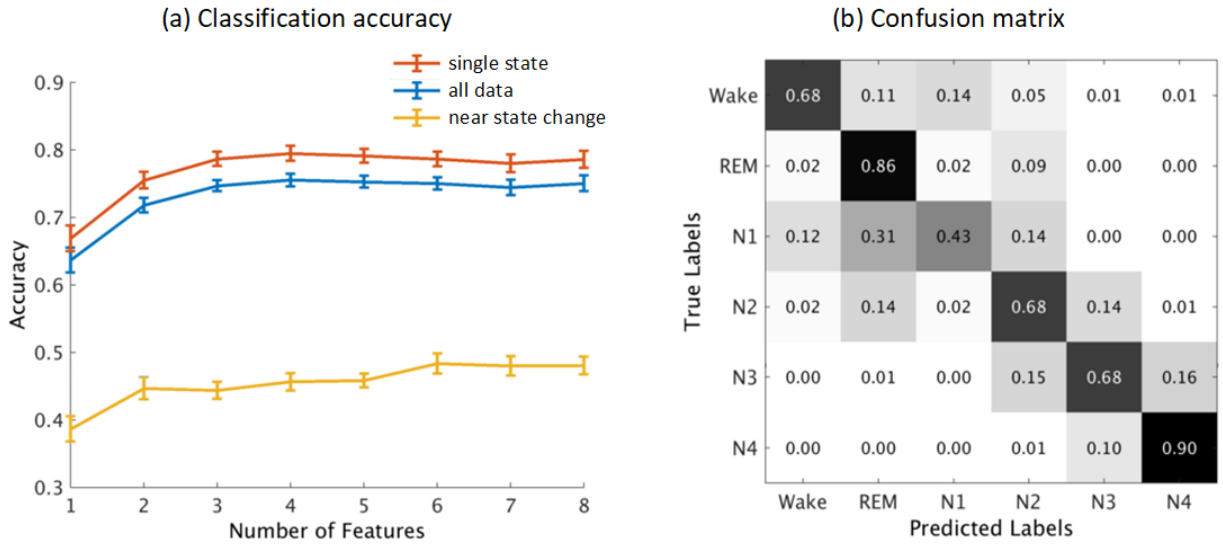

Figure 3.7: (a) Means and standard deviations in accuracy of classification between 6 sleep stages across the 17 subjects using cluster model probabilities for different numbers of models as features. Results were separated into two conditions, depending on whether the data window was or was not near a sleep state change. (b) Confusion matrix of 6-class classification across all the data using 8 cluster model probability features.

ure 3.7b shows the sleep-stage confusion matrix for the classification using all 8 model cluster probabilities. For the most distinctive sleep stages (REM and N4), the sensitivity (true positive rates) were 86% and 90%. For sleep entry stage N1 (with fewer class samples), sensitivity was significantly lower (43%), in line with clinical expectation. In addition, misclassification between sleep stages shown as nearest neighbors in Figure 3.7b accounted for 87% of the total errors.

3.3.3 Dataset III: Estimating behavioral alertness

Given the results using multi-model AMICA decompositions on sleep stage classification, described in the previous section, we assessed whether nonstationary AMICA decomposition can be used to estimate more continuous state transitions, *e.g.* changes in drowsiness level defined by changes in behavior in a continuous performance task.

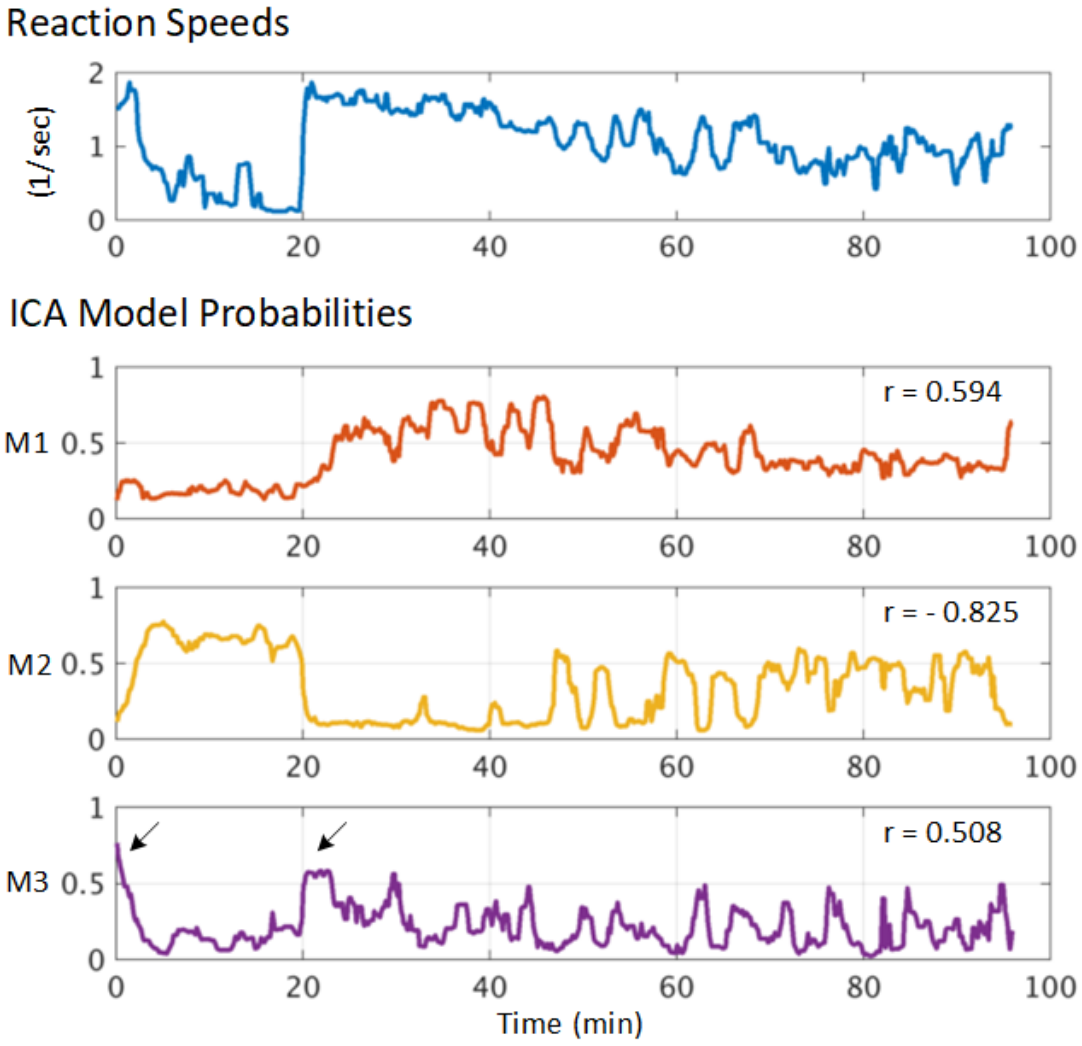

Figure 3.8: The top panel shows reaction speed changes (inverse of reaction times) in response to lane-departure challenges in one simulated driving session. The three bottom panels show shifting probabilities of the three ICA model learned by three-model AMICA decomposition of the whole EEG data session. Correlation coefficients (r) between each model probability time course and reaction speed are indicated. Black arrows in the lower panel mark brief (alert) periods when model M3 was dominate and reaction speed high.

Model probability shifts accompanying changes in behavioral alertness level

Figure 3.8 plots model probability time courses for a three-model AMICA decomposition of data from one subject, with the subject's reaction speed in response to driving challenges. The probability of model M1 correlated positively with reaction speed ($r = 0.594$), implying that this model was dominant during (more alert) periods when the subject responded quickly to driving challenges. In contrast, the probability of model M2 was strongly negatively correlated with reaction speed ($r = -0.825$), rising when subject reaction speed was low (less alert or drowsy periods). Surprisingly, model M3 was active at the beginning of the experiment and during quick transitions from slower to faster responding (arrows in Figure 3.8). These single-subject results provide evidence that model probabilities learned by three-model AMICA may co-vary with changes in reaction speed (often used, in long experiment sessions, as an index of behavioral alertness), and that the three models each accounted for EEG activity under a different set of performance conditions. Below, we will call models whose model probabilities have the most positive and negative correlations to reaction speed as "fast-response models" and "slow-response models" respectively. The remaining models may be dubbed "intermediate-response models".

To summarize the relation between model probabilities and reaction speeds for all subjects, we applied 3-model AMICA to individual subjects, aggregated ICA models according to their categories (fast response, slow response and intermediate models) across subjects, and plotted their model probabilities of 5-second windows before onset of the lane-departure trials, as shown in Figure 3.9. The trials with different levels of reaction speeds appeared to be organized in the feature space spanned by the model probabilities. The data were distributed on a probability simplex because model probabilities are compositional data, which sum up to one. The trials with slow reaction speeds (red) tend to distribute toward the bottom right corner of the triangle, where the probability was high for slow response model. In contrast, the trials with fast reaction speeds (blue) appeared to scatter at the other two corners where either fast response model or the third intermediate model was active. It is worth noting that the ICA model probabilities can be directly

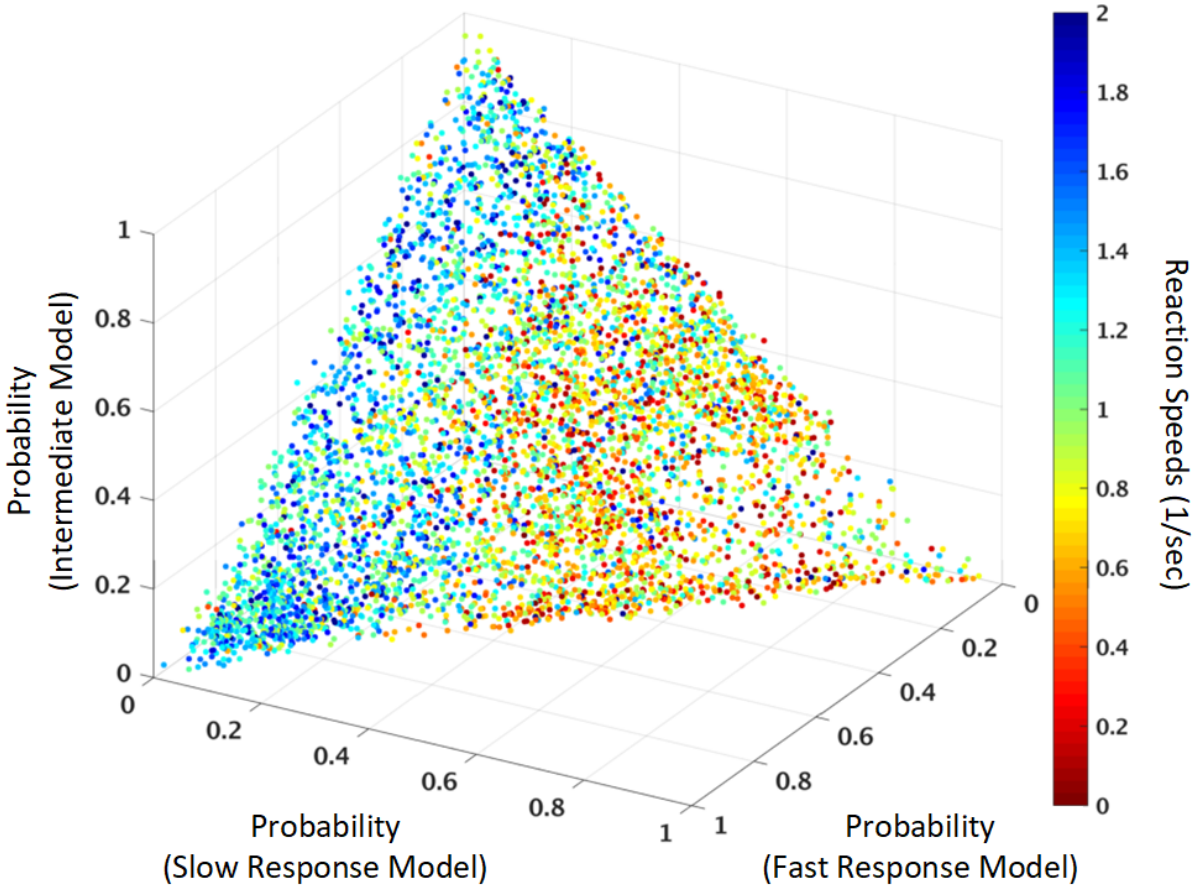

Figure 3.9: A scatter plot of lane-departure trials from all subjects in the feature space spanned by model probabilities of the three ICA models learned by 3-model AMICA. Each trial was color-coded with its reaction speed.

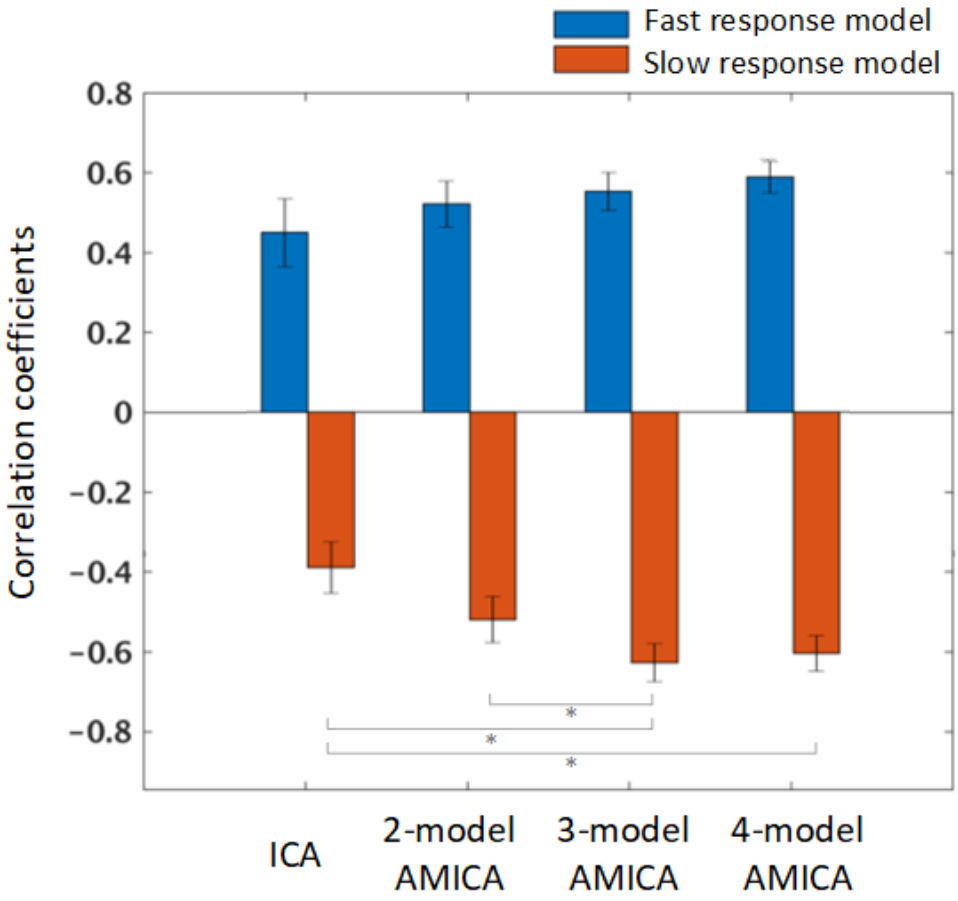

Figure 3.10: Across-subject mean correlation coefficients between reaction speed and model probabilities for fast-response versus slow-response models learned by unsupervised 2-to-4 model AMICA and by separate (supervised) decompositions of fast-response and slow-response periods using separate single-model ICA [44]. Standard errors of the mean (I-bars) and (* $p < 0.05$) results of paired t-tests are shown.

summarized across subjects without normalization, which is consistent with the results shown in Figure 3.6.

Relationships of model probabilities to performance changes

In Figure 3.10, we report subject mean correlations between model probabilities and reaction speed to study inter-subject variability, enable significance tests between AMICA decompositions with different numbers of models, and compare results against a previous study [44]. For

all subjects, AMICA decompositions with 2 to 4 models always included at least one fast-response model and one slow-response model, i.e. models whose model probability correlations to reaction speed were significantly positive and negative, respectively. This is a striking result: AMICA, an unsupervised learning approach, automatically and consistently identified two linearly unmixed source models of EEG data acquired when subjects were producing faster and slower responses, respectively.

A two-way repeated measures ANOVA with bootstrap significance testing showed ($p < 0.05$) significant interaction between model types (fast-response vs. slow-response) and different decomposition types. Post-hoc paired t-test results revealed no significant differences between behavior / model probability correlations for the fast-response models. However, model probabilities of slow-response models in 3-model and 4-model AMICA decompositions were significantly more negatively correlated with reaction speed than measures derived using the ICA-based approach proposed in Hsu *et al.*[44] in which fast- and slow-response models were learned from 90-sec data segments in which reaction speeds were fastest and slowest, respectively. Significantly more negative correlations were produced by slow-response models from 3-model than from 2-model AMICA decompositions; there was no significant correlations difference between models from 3-model and 4-model AMICA decompositions. This observation supports our choice to focus further analysis on the 3-model AMICA decompositions.

Rapid model switching dynamics during driving challenges

Changes in model probabilities can also characterize moment-by-moment state changes within single trials. The Figure 3.11 plots, for each latency across trials sorted by driver reaction speed, the index of the highest probability AMICA model time locked to the driving challenge onset, the driver's response onset or response offset. The results for the same subject as in Figure 3.8) are shown above the results for all the trials from the ten drivers to demonstrate that the results generalize across subjects and across (vertical) smoothing of smaller (top) or larger

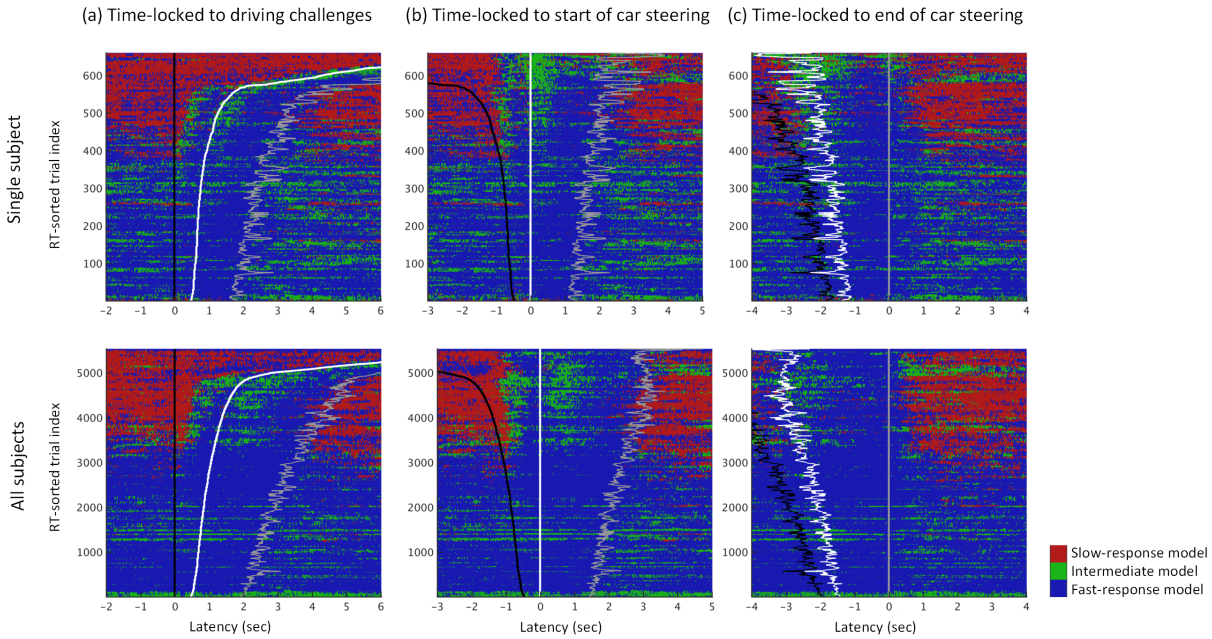

Figure 3.11: Event-related changes in the dominant AMICA model in 3-model AMICA decompositions within data trial epochs (horizontal colored lines) sorted by driver reaction speed. Model probabilities were computed in non-overlapping 20-msec windows. The same trials in the same top-to-bottom order shown are time locked either to (a) driving challenge onsets (black traces), (b) subsequent driver response onsets (white traces), or (c) driver response offsets (gray traces). Top panels show results for 600+ epochs for one subject (same as in Figure 3.8). AMICA models associated with fast, slow, and intermediate response speeds, respectively, were found among each subject's AMICA models. Bottom panels merge model cluster results for all 5000+ available epochs from all 10 subjects. Results shown are smoothed across trials (vertically) using a (single subject) 3-trial or (all subjects) 50-trial sliding window. Note the dominance of the (red) "slow-response" AMICA model (top panels) or model cluster (lower panels) results preceding and following driving challenge onsets in trials in which drivers responded relatively slowly. Notice also the transient dominance of the (green) "intermediate" models following driving-challenge and driver-response onsets in slower-response trials.

(bottom) numbers of trials.

Figure 3.11a shows that in trials with faster responses, before driving challenge onset, the (blue) fast-response model best fit the data, while before driver challenges in slow-response trials, the (red) slow-response model best fit the data. Switching between the two models occurs as driver response onsets increase from 0.9 sec to 1.1 sec (single subject, top) and from 1 sec to 1.2 sec for all drivers (bottom).

The dynamic switching between best-fitting AMICA models documented in Figure 3.11 thus measure brain dynamic changes preceding behavior on a near-millisecond time scale. Plotting the same trials time locked to driver response onsets (Figure 3.11b) shows that from 0.9 sec to 1.2 sec before response onset (white vertical trace), and again in the 1 sec following response onset, the third, (green) "intermediate" model became dominant briefly, possibly indicating brief hypnagogic ("dreamy") periods moving into and again out of relative alertness. Note that circa 0.5 sec spent by drivers in the relative (blue) alert state preceding response onsets in slow-response (upper) trials is close to the minimum time required by the drivers to respond to driving challenges in fastest-response (lower) trials. All these details are consistent with the driver challenge (lane deviation) and driver response (car-steering action) constituting a (briefly) arousing event sequence. Figure 3.11c shows that in (upper) slower-response trials the slow-response model learned by AMICA dominated for less than 0.5 sec after the offset of the car-steering action, suggesting that the drivers then relapsed into a more drowsy state, *e.g.* as soon as attention could safely be withdrawn from the task for some seconds.

Clustering ICs within AMICA models

So far we have demonstrated that shifting AMICA model probabilities can accompany changes in EEG dynamics supporting different cognitive and brain states. Another substantial advantage of the AMICA approach is that it learns generative models, *i.e.* sets of independent components and their activities and probability density functions (pdfs), that can be related to neu-

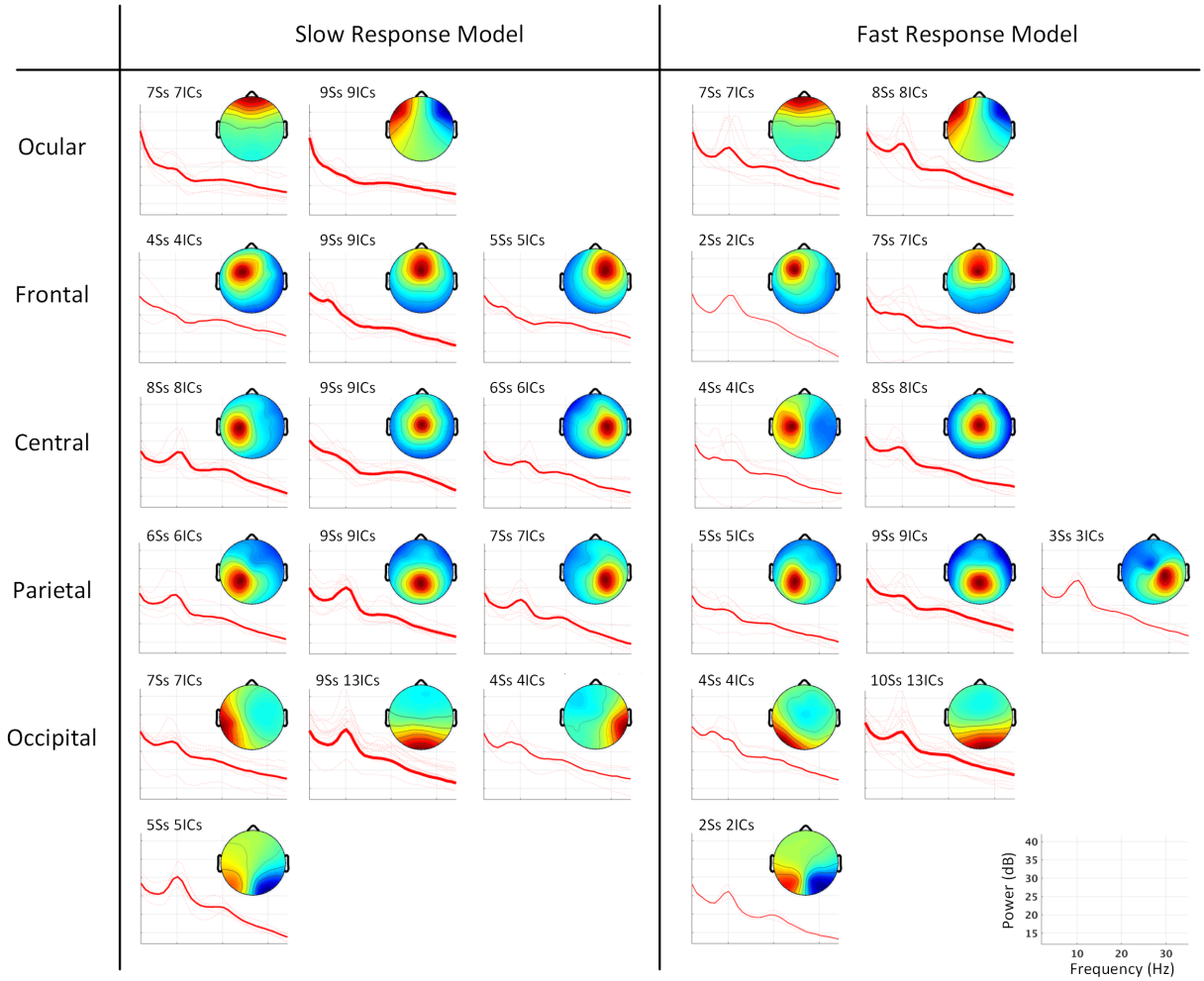

Figure 3.12: Average scalp maps and power spectra of independent component (IC) clusters from slow-response models versus clusters learned from fast-response models from separate three-model AMICA decompositions of data for each subject. The number of subjects and ICs contributing to each cluster are specified and are also indicated by the width of the power spectral traces.

rophysiological locations and functions, thereby enabling biologically plausible interpretations.

Figure 3.12 shows IC clustering results for fast-response and slow-response models across the 10 subjects (clustering details are described in Section 3.2.3). Both model class clustering solutions included ocular, frontal, central, parietal, and occipital clusters. Slow-response models included more dipolar sources and source clusters (108 ICs, 15 clusters) compared to fast-response models (72 ICs, 12 clusters). This difference appears most notable in right lateral clusters. found only among ICs in the slow-response models. By contrast, the slow-response model left central, parietal, and occipital clusters included 25 ICs, while the corresponding clusters for fast-response models included only 15 ICs.

3.4 Discussion

3.4.1 Unsupervised learning of brain dynamics by modeling source nonstationarity

This study aims to demonstrate the utility of AMICA as a general, unsupervised approach to assessing nonstationary dynamics of cortical dynamic states from nonstationary multichannel EEG signals. Our underlying hypothesis is that the ever-changing formation and dissolution of locally synchronous (or near-synchronous) cortical effective source activities and the network interactions they reflect and support give rise to the fluidity of cognition and behavior. Our results show that these nonstationary dynamics in cortical and cognitive state may be effectively modeled using an ICA mixture model and, specifically, by multi-model AMICA decomposition. Here we applied multi-model AMICA decomposition to one simulated and two actual EEG data sets to evaluate the efficacy of AMICA to estimate abrupt and continuous state changes, to classify multiple sleep stages, and to reveal moment-to-moment cortical (and likely cognitive state) dynamics supporting performance in a simulated driving task. In so doing, we tested the

capability of AMICA to estimate continuous state changes, to return consistent model sets across subjects, and to return models suitable for biological interpretation. We also tested the effects of the number of ICA models used. The following subsections discuss these topics in more detail.

3.4.2 Classification of multiple brain states

Our results demonstrate the capability of AMICA decomposition, applied to low channel-count (6- to 13-channel) sleep data, to separate six recognized sleep stages with high classification accuracy based only on changes in the likelihoods of the models AMICA learned from the data. Although the relationship between sleep stages and dominant ICA models was not a one-to-one mapping, the ICA models each captured different source dynamics that jointly characterized differences in EEG activities during the six sleep stages. Hence, in the feature space of model probabilities shown in Figure 3.6, EEG activities from different sleep stages could be clearly separated. Applying a simple Gaussian Bayes classifier to quantitatively assess state separability, we found that based on multi-model AMICA decomposition and using only 4 to 8 data features, we could achieve an average cross-validation accuracy of 75%, significantly higher than chance (17% for a general 6-class problem, 38% taking into account the unbalanced numbers of class labels). This sensitivity was higher for REM and N4 stages and lower for stage N1, in alignment with clinical expectations.

Furthermore, classification errors occurred more frequently near sleep stage transitions, and particularly between more strongly related stages (3.7ab). This may in part reflect the relatively coarse grain (30-sec) of the manual sleep staging, and possible lower inter-scorer consistency in distinguishing strongly related stages. Figure 3.4 shows that during stage changes in model probabilities and thus in EEG activities were not discreet or regular but were continuous and irregular. In particular, in REM or between progressive stages N1 to N4, changes in model probabilities were distributed continuously (Figure 3.6). Thus, the AMICA results suggest that transitions between sleep stages were more continuous, across both time and AMICA “feature

space", than as measured by standard sleep stage scoring.

3.4.3 Estimation of rapid state changes

While AMICA assumes and learns discrete ICA models, the relative probabilities of each model measure the "fitness" or likelihood of each model at each data point or group of neighboring data points, that can be effective estimators of moment-to-moment cognitive and behavioral state changes. Applied to the drowsy driving dataset, AMICA automatically and consistently learned fast-response and slow-response trial models for each subject whose model probability changes across time were positively and negatively correlated, respectively, with drowsiness level as indexed by driver speed in reacting behaviorally to occasional lane-deviation driving challenges. These strong and opposite correlations signified that higher likelihood for the fast-response model predicted higher reaction speed, while higher likelihood for the slow-response model predicted lower reaction speed in response to an immediately upcoming driving challenge. Further, rapid (sub-second scale) patterns of shifts between most probable models were consistent with interpretations that appearance of driving challenges induced brief changes from less alert to more alert EEG dynamics, and that during less alert (slow-response model) periods, EEG dynamics typically shifted back to less alert model a second or less after the offset of the drivers behavioral response. Further, these brief transitions between less alert and more alert state often involved momentary transitions through a third ("intermediate") AMICA model. The models returned by AMICA decompositions exhibited these close relationships to the behavioral data record despite not using any direct information about the nature or timing of experimental events and behavioral responses.

Our previous studies have employed other measures to quantify EEG state changes during simulated driving and sleep, including a nonstationary index [44] and relative likelihoods [80] of separately-trained ICA models. Compared with these studies, AMICA here learned multiple ICA models that proved able to better characterize the EEG dynamics and could be generalized to

follow both irregular and transient shifts between more than two brain states. Instead of training separate ICA models on separate sets of data segregated by behavior [44], AMICA, an unsupervised learning approach, here automatically learned distinctions between EEG activities occurring in different brain/behavioral states. More importantly, as shown in Figure 3.10, unsupervised AMICA outperformed the previous supervised multiple ICA decomposition approach [44] in estimating drowsiness levels.

By examining switching between dominant models within single trials with sub-second temporal resolution, we found a consistent sequence and timing of brain state changes immediately before and after driver responses to experimental driving challenges. When drivers were drowsy, *i.e.* exhibiting EEG best fit by their "slow-response" model, they were slow in detecting lane-departure events. In many trials, drivers began their behavioral response to these challenges within about a second (0.9 to 1.2 secs) after their EEG exhibited a very brief transition to "intermediate" model dynamics, their motor response appearing about half a second after their faster-response model then became dominant. Following the end of these motor responses, drivers relapsed into the slow-response model dynamics after only about a second. These results demonstrate capability of multi-model AMICA decomposition to track cortical dynamic state changes on the sub-second time scale.

3.4.4 Consistency across subjects

Although AMICA, as an unsupervised learning approach, need not give learned ICA models that are similar across subjects, applied to actual experiment data AMICA here produced results that were surprisingly consistent across subjects in three senses:

(1) Consistent relations between ICA models and brain states were clearly observed in both applications (sleep and driving challenges). In the sleep dataset, Figure 3.5 shows that ICA models with similar probability distributions over sleep stages were found across all subjects. In other words, for each subject some ICA models were dominant during specific sleep stages

(*e.g.* group B model during stage N4). Similarly, Figure 3.10 shows that slow-response and fast-response models were consistently learned for all subjects in the drowsy driving experiment.

(2) Results included consistent differences in AMICA model probabilities across subjects. As shown in Figure 3.6 and Figure 3.9, although the model probabilities were here based on subject-specific ICA models, their values could be directly summarized across all subjects without normalization. These results provide strong evidence that model probabilities, intrinsically bounded from 0 to 1, can be global indices that generalize across subjects.

(3) Differences between independent component (IC) clusters (here based on IC scalp maps) for different model classes appeared and are discussed in the next subsection.

3.4.5 Biological interpretation of AMICA models

Besides its unsupervised segmentation of nonstationary data into putative brain dynamic and function states, another benefit of the AMICA approach is that it learns a generative model that characterizes a complete set of active, statistically maximally independent components (ICs) in each state, plus a set of time series giving the probability of each model at each time point based on a probability density function (PDF) learned for each model IC from the data. During iterative training, each model becomes adapted to time points at which it is most probable. We validated this characterization by applying multi-model AMICA decomposition to simulated quasi-stationary data, showing that multi-model AMICA can accurately learn the ground-truth source IC scalp projection patterns, activities, and PDFs. A growing amount of evidence suggests an association between many ICs and localized biological and functional processes in cortex [78, 91]. By constructing an individualized subject electrical forward model from an MR head image, a subset of (brain source) ICs can be further localized using either single or dual equivalent dipole or distributed cortical patch models [2, 40, 1].

Applied to the drowsy driving dataset, IC processes learned by AMICA were generally consistent across subjects. ICs compatible with a compact cortical source area or eye movement

artifact could be clustered into similar fast-response and slow-response model source clusters based on scalp map correlations. The identified IC clusters, including frontal-scalp projecting clusters with high theta or alpha power and occipital and parietal clusters with high alpha power, were consistent with previous studies applying single-model ICA decomposition to these data [19, 44]. There, differences in the dynamics of similar ICs were shown to be associated with alert and drowsy states respectively. Interestingly, more dipolar sources were returned by AMICA in the slow-response models. This result may be related to the fact that brain activities, especially alpha waves, spread through larger cortical areas during drowsiness as reported in [108] and [65]. In our results, stronger alpha activities appear in frontal and prefrontal (indicated as ocular) fast-response cluster ICs. These clusters may be driven by anterior cingulate activity [53]. Compared to the results of the supervised multi-model ICA decomposition approach in our previous report (Figure 3 in [44]), an larger number of dipolar ICs compatible with a localized cortical source were identified in the data, suggesting that unsupervised multi-model AMICA may be a more effective approach to learning state-related ICA models and model changes than a supervised multi-model ICA approach.

We could not study the cortical origins of the ICs learned from these sleep data as the CAP sleep database offers only low-density sleep EEG data recorded using bipolar channels. The application of AMICA decomposition to high-density sleep EEG data could be of interest to sleep research exploring changes in effective EEG sources and source network activities in each sleep stage.

3.4.6 Choosing the model order

One of the most important parameters required to apply AMICA is the number of ICA models, i.e. H in equation 3.1. Since the ground-truth model order of the data is typically unknown, the present work focuses on examining the effects of assumed model order on AMICA performance. Applied to simulated 3-segment quasi-stationary data, complete 3-model AMI-

CA decomposition and over-complete (4- to 6-model) AMICA decomposition all successfully segmented the data and accurately learned the ground-truth sources, suggesting that in many applications choice of model order might not crucially affect the validity of AMICA results in particular when a complete (ground-truth) number of models, or at most only a few excess models are learned. Typically, excess models only account for a small portion of data not well modeled by the other ICA models, *e.g.* data points at which many sources are unusually co-activated (in the presence of adventitious artifact, for example). When applied to the simulated driving data, AMICA decomposition using 2, 3, or 4 models consistently returned "fast-response" and "slow-response" models accounting for the EEG data in alert and drowsy behavioral conditions, respectively. A third ("intermediate") model (M3 in Figure 3.8) accounted for EEG activities not well fit by the two dominant models, *e.g.* during brief transitions between the two dominant EEG states.

The above results provide evidence that choosing a precise number of models is not critical to the information value of AMICA decomposition (including model probabilities and brain source characteristics). For example, applying 2-model through 10-model AMICA decomposition to the sleep data from a single subject, we found that that adding or eliminating one model typically returned models with almost identical model probability dynamics. As with other clustering analyses, adding another model may give a model accounting for lower-probability data points of one or two of the existing models and leave other existing models intact.

Several approaches have been proposed to help select the number of nonstationary data models. For example, one may compare the marginal likelihood for different candidate models by adding a penalty on model complexity, for example the Akaike information criterion (AIC) [3] or Bayesian information criterion (BIC) [110]. Some adaptive approaches including variational Bayesian learning [15] and online adaptive learning [71] have also been proposed. However, these methods are computationally expensive and also require heuristic setting of thresholds for splitting or merging source clusters.

3.4.7 Limitations and open questions

Like most unsupervised-learning and data-driven approaches, successful AMICA decomposition has data and computation requirements. Source-level analyses such as ICA require multichannel EEG data to achieve meaningful source separation. They also implicitly assume that the number of data channels is as least as large as the number of effective sources. Although AMICA relaxes the ICA assumption by learning multiple ICA models, AMICA still requires a minimum number of data channels to produce interpretable results. For example, applied to the sleep dataset, AMICA achieved an average accuracy of 75% in 6-class classification using EEG data with only 6-13 channels. However this accuracy dropped to 68% for subjects with only 5 EEG channels available.

Another requirement for successful AMICA decomposition is a reasonable number of data samples. Learning H ICA models, each with N stable sources, requires approximately $k \cdot H \cdot N^2$ samples, where empirically $k \geq 25$ [90]. For $H = 6$, $N = 16$, $k = 25$, and a 250-Hz sampling rate, this corresponds to $\sim$2.5-min of data; hence here we generated 3-minute stationary segments in the simulated data. Lastly, AMICA decomposition requires significant computation time to run on a personal computer. For 13-channel sleep EEG data from 9-hour recordings with a 512-Hz sampling rate, multi-model AMICA decomposition required 13-15 hours on a 2.40-GHz CPU. However, AMICA computation time can be significantly reduced through parallelization, as featured in the AMICA code made available (`https://sccn.ucsd.edu/~jason/amica_web.html`) by its author, Jason Palmer, and interested users might explore use of the Neuroscience Gateway (`www.nsgportal.org`) to run AMICA decompositions on larger data sets.

Although this study focuses on AMICA decomposition, the results might be able to generalize to other ICAMM approaches that may have other desirable properties. For example, different approximations of source probability density functions (PDFs) can be used to better match the underlying source activity in the data, such as a generalized exponential model [102], a mixture of Gaussians [15], and a nonparametric model [107]. In addition, a sequential ICAMM

approach which considers temporal dependence using a hidden Markov model [106] may be suitable for modeling sleep data in which state transitions are continuous.

Results of this study support the use of multi-model AMICA decomposition for assessing brain state changes by validating its performance on sleep stage classification and alert versus drowsy performance estimation. These results provide evidence to support the application of multi-model AMICA decomposition as a general unsupervised-learning approach to study the continuous, endogenous, and nonstationary brain dynamics in either EEG, MEG [52], or electroencephalographic (ECoG) data [120]. For example, AMICA decomposition might be applied to multichannel brain electrical signals to explore brain dynamics during rest, movie watching, or hypnotherapy, to identify the nonstationary, task-irrelevant brain source activity changes during performance of a complex cognitive task, or even to study mental strategy or emotional shifts using a brain-computer interface.

3.5 Conclusions

Here we have demonstrated that AMICA decomposition provides a general unsupervised approach to mining changes in effective source dynamics in nonstationary multichannel EEG signals. The underlying hypothesis here is that different brain states may involve different active effective sources (each typically compatible with an emergent area of locally-synchronous cortical field activity), and that the locations and source-level probability density functions (PDFs) of these state-specific effective source activities can be well modeled by transitions between ICA data models.

We showed that, applied to simulated quasi-stationary data, AMICA decomposition could accurately learn the ground truth sources and source activities, either when directed to return complete or (mildly) over-complete model sets. Applied to some sleep EEG data, multi-model AMICA decompositions could be used to meaningfully characterize sleep dynamics, giving

consistent results across subjects and allowing 75% cross-validation accuracy in classifying data from six sleep stages validated by expert sleep scoring.

Applied to EEG datasets recorded during simulated driving, AMICA automatically identified two models accounting for EEG activity in slow- and fast-response trials respectively. The corresponding model probability differences could be used as an effective estimator of reaction speed in single trials and appeared to track brain dynamic state changes on the sub-second scale. In addition, AMICA decomposition also learned physiologically interpretable results including the spatial distribution and temporal activity pattern of the effective brain sources in each ICA model.

Thus multi-model AMICA decomposition can be applied to continuous and unlabeled EEG (or other electrophysiological) data to study, for example, non-stationarities in brain dynamics during resting states, accompanying mental strategy changes, or through different states of emotion, fatigue, and arousal.

3.6 Acknowledgement

Chapter 3 is currently being prepared for submission for publication by Hsu, S. H., Pion-Tonachini, L., Palmer, J., Miyakoshi, M., Jung, T. P., and Makeig. S. The dissertation author was the primary investigator and author of this paper.

Chapter 4

Real-time Adaptive EEG Source Separation Using Online Recursive ICA

In this chapter, we study and validate an optimized online recursive ICA algorithm (ORICA) with online recursive least squares (RLS) whitening for blind source separation of high-density EEG data, which offers instantaneous incremental convergence upon presentation of new data. Empirical results of this study demonstrate the algorithm's: (a) suitability for accurate and efficient source identification in high-density (64-channel) realistically-simulated EEG data; (b) capability to detect and adapt to non-stationarity in 64-ch simulated EEG data; and (c) utility for rapidly extracting principal brain and artifact sources in real 61-channel EEG data recorded by a dry and wearable EEG system in a cognitive experiment. ORICA was implemented as functions in BCILAB and EEGLAB and was integrated in an open-source Real-time EEG Source-mapping Toolbox (REST), supporting applications in ICA-based online artifact rejection, feature extraction for real-time biosignal monitoring in clinical environments, and adaptable classifications in brain-computer interfaces.

4.1 Background

Independent Component Analysis (ICA), as a means for blind source separation (BSS), has enjoyed great success in telecommunications and biomedical signal processing [50]. In biomedical applications, such as scalp electroencephalography (EEG), ICA methods have been widely used to separate artifacts such as eye blinks and muscle activities [54] and to study brain activities [78]. For example, ICA can extract fetal electrocardiography (ECG) from maternal abdomen electrode recordings [27], and it can also isolate pathological activities associated with disease states of epilepsy [84]. In addition, applying ICA to remove task-irrelevant activities and reduce dimensionality of data can improve the performance of Brain-Computer Interfaces (BCI) [118].

The application of ICA to EEG data is justified by a reasonable assumption that multi-channel scalp EEG signals arise as a mixture of weakly dependent latent non-Gaussian sources [30]. Although several ICA algorithms have been developed [50] to learn these sources from channel mixtures, most of the algorithms require access to large amount of training data and are only suitable for offline applications. Furthermore, the offline ICA algorithms commonly assume spatiotemporal stationarity of the data, as in the widely used Infomax ICA [8] and FastICA [51] algorithms. For a few ICA methods that allow non-stationarity such as Adaptive Mixture ICA [96], they are computationally expensive. In many real-world applications, including real-time functional neuroimaging [83], artifact rejection and adaptive BCI [118], online (sequential) source separation methods are needed. Desirable properties of an online method include fast convergence, real-time computational performance, and adaptivity to non-stationary data.

Many existing online ICA methods are listed in Table 4.1. Two major learning rules are least-mean-squares (LMS) and recursive-least-squares (RLS) methods. LMS-type algorithms use stochastic gradient descent approaches and are computationally simple, but they require careful selection of an appropriate learning rate for stable convergence. Examples include

Table 4.1: Comparison of online ICA methods, including the learning rules, the optimization approaches, and the pre-whitening methods.

Name	Author	Year	Learning rule and optimization method	Pre-whitening
EASI	Cardoso *et al*.[12]	1996	Relative gradient-based (LMS) that max. kurtosis	LMS
NG	Amari *et al*.[6]	1996	Natural gradient-based (LMS) that min. mutual information	no
RLS-NPCA	Karhunen *et al*.[60]	1997	RLS that min. LSE of NPCA criterion	PCA
Iterative Inversion	Cruces-Alvarez *et al*.[26]	2000	Quasi-Newton method with iterative inversion that decorrelates high order statistics	no
NG-RLS	Zhu *et al*.[126]	2004	RLS with natural gradient that min. LSE of NPCA criterion	RLS
ORICA	Akhtar *et al*.[4]	2012	Recursive rule with iterative inversion from fixed-point solution of Infomax with natural gradient	no

Equivariant Adaptive Separation via Independence (EASI) [12] and Natural Gradient (NG) [6] methods. RLS-type algorithms accumulate past data in an exponentially decaying fashion and use Sherman-Morrison matrix inversion to achieve higher convergence rate and better tracking capability, yet require complex computation [36, 127]. This category includes the RLS approach of Nonlinear PCA (RLS-NPCA) [60], Iterative Inversion [26], and Natural Gradient-based RLS (NG-RLS) [126]. Alternatively, Online Recursive ICA (ORICA) [4] gives an RLS-type recursive rule by solving a fixed-point approximation and has been shown to exhibit fast convergence and low computational complexity [45]. Readers can refer to [127], [26] and [121] for theoretical relationships and comparisons between the above methods.

The aforementioned papers focused on theoretical derivations and proofs of convergence and only demonstrated applications of the methods to low-density data (fewer than 10 channels) and simulated "toy" examples such as sinusoidal and square waves. When the number of channels and sources increases, many existing algorithms exhibited slow convergence and poor real-time performance [51]. In a recent study, a real-time online ICA method for high-density EEG was proposed [45]. The method was compared with other offline ICA methods, and its stability and steady-state performance were analyzed in [46].

Additionally, an important advantage of online ICA methods is their ability to adapt to spatiotemporal non-stationary data, a common occurrence in real-world applications. For instance, spatial non-stationarity in the ICA (un)mixing matrix can arise as a consequence of location shifts in either sensors or sources, or changes in electrode impedances. However, few of

the online ICA methods have been carefully studied under non-stationary conditions. Further investigation is needed to characterize algorithmic performance and optimal parameter selection using non-stationary simulated and real EEG data.

In this study, we extend ORICA as formulated in [4] and [45], and the contributions are three-fold. Firstly, we demonstrate ORICA's suitability for accurate and efficient source identification in a realistic simulation of stationary 64-channel EEG data. Specifically, we include a serial orthogonalization step of the unmixing matrix in the ORICA algorithm, and we systematically examine the impact of parameters such as the forgetting factor and block sizes for pre-whitening and ORICA on algorithmic performance. Secondly, we examine ORICA's capability to adaptively decompose spatially non-stationary 64-channel EEG data corresponding to abrupt displacements of electrodes, a common source of spatial non-stationarity in real-world mobile applications. We introduce a non-stationarity index and an adaptation approach for non-stationarity detection and online adaptation. Thirdly, we evaluate ORICA's real-world applicability for rapidly extracting principal brain and artifact sources using 61-channel real EEG data recorded from a subject performing an Eriksen flanker task [81]. We demonstrate that ORICA and offline Infomax ICA [90] obtain comparable results in terms of extracting informative independent components (ICs) and their single trial and averaged event-related potentials (ERPs), yet ORICA can learn the ICs online with less than half of the data. Finally, the proposed ORICA pipeline is made freely available as functions supported in BCILAB [63] and EEGLAB [29], and it is also integrated in an open-source Real-time EEG Source-mapping Toolbox (REST) [101].

4.2 Methods

Standard ICA assumes a linear generative model $\mathbf{x} = \mathbf{A}\mathbf{s}$, where $\mathbf{x}$ represents scalp EEG observations, $\mathbf{s}$ contains unknown sources, and $\mathbf{A}$ is an unknown square mixing matrix. The objective is to learn an unmixing matrix $\mathbf{B} = \mathbf{A}^{-1}$ such that the sources are recovered exactly, up

to an unknown permutation and scaling matrix, by $\mathbf{y} = \mathbf{Bx}$, where $\mathbf{y}$ represent the recovered source activations. A column of $\mathbf{B}^{-1}$ represents the spatial distribution of a source over all channels, often referred to as a "component map."

It is desirable to optimize the ICA contrast function, a measurement of the degree of independence between sources such as kurtosis or mutual information, under the decorrelation constraint $\mathbf{R}_y = E[\mathbf{y}\mathbf{y}^T] = \mathbf{I}$. Hence the separating process can be factored into two stages as $\mathbf{B} = \mathbf{WM}$, where $\mathbf{M}$ is the whitening matrix that decorrelates the data and $\mathbf{W}$ is the weight (preferably orthogonal) matrix that optimizes the ICA contrast function [12, 126]. Serial update rules of $\mathbf{M}$ and $\mathbf{W}$ and detailed features are presented in the following subsections.

4.2.1 Online recursive-least-squares (RLS) pre-whitening

Pre-whitening (decorrelating) the data reduces the number of independent parameters an ICA update must learn, and can improve convergence [50]. Pre-whitening may be efficiently carried out in an online RLS-type learning rule [126]:

$$\mathbf{M}_{n+1} = \mathbf{M}_n + \frac{\lambda_n}{1-\lambda_n}\left[\mathbf{I} - \frac{\mathbf{v}_n \cdot \mathbf{v}_n^T}{1+\lambda_n(\mathbf{v}_n^T \cdot \mathbf{v}_n - 1)}\right]\mathbf{M}_n, \tag{4.1}$$

where n is the number of iterations, $\mathbf{M}_n$ is the whitening matrix, $\mathbf{v}_n = \mathbf{M}_n\mathbf{x}_n$ is the decorrelated data, λ_n is a forgetting factor, and $\mathbf{I}$ is the identity matrix. A non-overlapping block of data $\mathbf{x}_n$ with a block size L_{white} is used at each iteration to reduce the computational load and to increase the robustness of the estimated correlation matrix $\mathbf{v}_n\mathbf{v}_n^T$. This RLS-type whitening rule exhibits faster convergence than LMS whitening methods [126].

4.2.2 Online recursive ICA (ORICA)

The ORICA algorithm can be derived from an incremental update form of the natural gradient learning rule of Infomax ICA [66]:

$$\mathbf{W}_{n+1} = \mathbf{W}_n + \eta[\mathbf{I} - \mathbf{f}(\mathbf{y}_n) \cdot \mathbf{y}_n^T]\mathbf{W}_n \tag{4.2}$$

where $\mathbf{y}_n = \mathbf{W}_n \mathbf{v}_n$, η is a learning rate, and $f(\cdot)$ is a nonlinear activation function. In the limit of a small η and assuming a fixed $f(\cdot)$, the convergence criterion $\langle \mathbf{f}(\mathbf{y}) \cdot \mathbf{y}^T \rangle = \mathbf{I}$ leads to a fixed-point solution in an iterative inversion form [4]:

$$\mathbf{W}_{n+1}^{+} = (1 - \lambda_n)\mathbf{W}_n^{+} + \lambda_n \mathbf{v}_n \cdot \mathbf{f}^T(\mathbf{y}_n) \tag{4.3}$$

where $\mathbf{W}_n^{+}$ is the Moore-Penrose pseudoinverse of $\mathbf{W}_n$ and λ_n is a forgetting factor for an exponentially weighted series of updates. It should be noted that λ_n differs from η, which is the step size for stochastic gradient optimization.

Following [4], applying the Sherman-Morrison matrix inversion formula to Equation 4.3, the final online recursive learning rule becomes:

$$\mathbf{W}_{n+1} = \mathbf{W}_n + \frac{\lambda_n}{1 - \lambda_n}\left[\mathbf{I} - \frac{\mathbf{y}_n \cdot \mathbf{f}^T(\mathbf{y}_n)}{1 + \lambda_n(\mathbf{f}^T(\mathbf{y}_n) \cdot \mathbf{y}_n - 1)}\right]\mathbf{W}_n \tag{4.4}$$

The near-identical forms of Equation 4.4 and Equation 4.1 allow us to understand ORICA as a nonlinear (or kernel) form of the RLS whitening filter: ORICA's use of non-linearity $f(\cdot)$ allows for independence of sources for moments above second order.

Following the Equation 4.4, the orthogonality of the weight $\mathbf{W}$ is not guaranteed. To preserve the decorrelation property of recovered source activities $\mathbf{y} = \mathbf{WMx}$ and maintain learning

stability, we apply an orthogonal transformation to $\mathbf{W}$ after each ICA update:

$$\mathbf{W}_{n+1} \leftarrow (\mathbf{V}\mathbf{D}^{-1/2}\mathbf{V}^{-1})\mathbf{W}_{n+1}, \tag{4.5}$$

where $\mathbf{D}$ and $\mathbf{V}$ contain, respectively, the eigenvalues and eigenvectors of $\mathbf{W}_{n+1}\mathbf{W}_{n+1}^T$. Note that this orthogonalization step is costly compared to the ORICA update. A possible alternative would be to reformulate the ORICA update rule under the orthogonal constraint or combine the serial whitening and weight updates into a single update rule [12].

Block-update rule

The typical single measurement vector approach [17] requires application of the update rule (Equation 4.4) for each data sample, which can be computationally expensive, particularly for the commonly-used MATLAB (The Mathworks, Natick, MA) runtime environment. To reduce the computational load and ensure consistent real-time performance, we may adopt an multiple measurement vector approach [17] and perform updates on short blocks of samples. To achieve this without loss of accuracy, we solve Equation 4.4 for time index $l = n$ to $l = n+L-1$, assuming $\mathbf{y}_l$ is approximated as $\mathbf{W}_n\mathbf{v}_l$ and λ_l is small. This leads to a block-update rule [45]:

$$\mathbf{W}_{n+L} \approx \Big(\prod_{l=n}^{n+L-1} \frac{1}{1-\lambda_l}\Big) \cdot \Big[\mathbf{I} - \sum_{l=n}^{n+L-1} \frac{\mathbf{y}_l \cdot \mathbf{f}^T(\mathbf{y}_l)}{\frac{1-\lambda_l}{\lambda_l} + \mathbf{f}^T(\mathbf{y}_l)\cdot \mathbf{y}_l}\Big]\mathbf{W}_n \tag{4.6}$$

In this form, the sequence of updates can be vectorized for fast computation. Note that Equation 4.6 appropriately accounts for the decaying forgetting factor at each time point. This keeps the approximation error to a minimum.

Forgetting factor

The forgetting factor λ determines an effective length of a time window wherein data are aggregated. A large value of λ corresponds to a short window length. In this case, much heavier

weights are applied to new data than past data, yielding fast adaptation and convergence yet large errors and variability. As a general rule, a large λ is preferred during initial learning to promote fast convergence; a small λ is suggested at convergence to minimize variance. To this end, we adopt the forgetting factor with time-varying annealing defined in [4]:

$$\lambda_n = \frac{\lambda_0}{n^\gamma} \tag{4.7}$$

where λ_0 is a fixed initial forgetting factor and γ determines the rate of exponential decay of λ towards zero as a function of time. The same forgetting factor is applied to the RLS whitening filter, although theoretically it can be different.

Nonlinear function

The choice of non-linearity $f(\cdot)$ depends on the probability distribution of the sources. Lee *et al.*[66] proposed an extended Infomax ICA algorithm that adopts distinct activation functions to separate subgaussian and supergaussian sources based on an estimate of source kurtosis. Here we follow [50, 4] and choose the component-wise non-linearity $f(y) = -2\tanh(y)$ for supergaussian sources and $f(y) = \tanh(y) - y$ for subgaussian sources.

Number of sub- and super-gaussian sources

While approaches for adaptively selecting $f(y)$ within ORICA have been proposed [4], these are heuristic and presently lack convergence proofs. In practice, we found that both convergence speed and run-time performance were improved by preassuming a fixed number of sub- and super-Gaussian sources. The detail of selecting an appropriate number was discussed in Section 4.5 and the effect of inaccurate assumptions was explored in Section 4.4.

4.2.3 Non-stationarity detection

In this study, the non-stationarity refers to any changes in the ICA model $\mathbf{x} = \mathbf{As}$, including spatial non-stationarity of the mixing process between sources and the measurements $\mathbf{A}$ and temporal non-stationarity of probability distributions of source activities $\mathbf{s}$. Hence the non-stationarity might arise from switching of active brain sources, transient muscle activities, sensor displacement, or impedance changes. Our goal is to propose a generic approach to detect and adapt to the non-stationarity in EEG data.

Non-stationarity index

As previously described, ORICA is derived from a fixed point solution to the convergence criterion $\langle \mathbf{y} \cdot \mathbf{f}(\mathbf{y})^T \rangle = \mathbf{I}$, reflecting independence of sources for moments above second order. Violation of this criterion once an algorithm reaches steady-state can be interpreted as a change in the latent source or mixture distribution, which leads us to define the following heuristic non-stationarity index:

$$\delta_{ns} = ||\mathbf{y}_n \mathbf{f}^T(\mathbf{y}_n) - \mathbf{I}||_F \tag{4.8}$$

where $||\cdot||_F$ represents the Frobenius norm and n is the current sample point. After ICA decomposition converges, δ_{ns} would remain small when data are stationary, while δ_{ns} would increase and fluctuate when data are non-stationary.

Adaptation of the forgetting factor

If the non-stationarity index δ_{ns} increases above a threshold, we may interpret this as evidence of a change in the latent mixing matrix and increase the RLS forgetting factor allowing ORICA to more rapidly adapt to the new mixing matrix. In this study, the threshold value was heuristically chosen to be a percentage (*e.g.* 1-10%) of initial δ_{ns}, which was several standard deviations above the mean of δ_{ns} at convergence in the simulated stationary data. Once δ_{ns}

reached the threshold, we increased the forgetting factor to its initial value.

4.3 Materials

4.3.1 Data collection

The performance of ORICA was evaluated under simulated and real-world conditions. Previous works of online ICA mostly used "toy" simulations with artificially constructed sources (sinusoids, i.i.d. random data, etc), stationary mixing matrices, and relatively small numbers of channels and sources. Here we generated high-density EEG (64 channels and 64 sources) under more realistic conditions, including use of auto-correlated stochastic sources, realistic source locations and mixing matrices derived from Boundary Element Method (BEM) modeling, and spatial non-stationarity. Simulated data were generated using the EEG simulation module in Source Information Flow Toolbox (SIFT) [85], using an approach similar to [42].

Simulated spatially stationary EEG data

To test ORICA's performance in separating stationary EEG sources, we generated 64 supergaussian independent source time-series from stationary and random-coefficient order-3 autoregressive (AR-3) models (300Hz sampling rate, 10-min), assigned each source a random cortical dipole location, and projected these through a zero-noise 3-layer BEM forward model (MNI "Colin27") with standard 10-20 electrode locations matching the 64-channel Cognionics montage used subsequently for real-world ORICA evaluation. This yielded 64-channel EEG data.

Simulated spatially non-stationary EEG data

To evaluate ORICA's capability to adapt to spatial non-stationarity, we simulated abrupt shifts of the electrode montage during continuous recording. We first generated 30 minutes of temporally stationary AR-3 source data, as described above. The data was partitioned into thirds.

For each 10 minute segment, 64 channel EEG data was generated using a unique BEM forward (mixing) matrix, corresponding respectively to (a) the standard electrode montage, (b) a 5 degree anterior cap rotation, and (c) a subsequent 10 degree posterior cap rotation (5 degree posterior rotation from standard position). The procedure is illustrated in Figure 4.5a.

Real EEG data

One session of high-density EEG data was collected from a 24 year-old right-handed male subject using a 64-channel wearable and wireless EEG headset with dry electrodes (Cognionics, Inc) [41]. In the 20 minutes session, the subject performed a modified Eriksen flanker task [81] with a 133 ms delay between flanker and target presentation. The subject was asked to press buttons according to the target stimuli as quickly as possible. Flanker tasks are known to produce robust error-related negativity (ERN, Ne) at frontal-central electrode sites. The goal here is to extract these event-related potential (ERP) components from high-density EEG data in a real-world setting using the proposed ORICA pipeline.

4.3.2 The ORICA pipeline

As shown in Figure 4.1, the ORICA pipeline continuously fetched the streamed data with variable size L_B in the online buffer and processed the data with the three filters in sequence: a Butterworth IIR high-pass filter, an online RLS whitening filter, and an ORICA filter. The high-pass filter removed the trend and low-frequency drift, ensuring the zero-mean criterion for ICA was satisfied. For each update, the pipeline computed and output the whitening matrix $\mathbf{M}$ and the weight matrix $\mathbf{W}$ according to Equation 4.1 and Equation 4.6. The next non-overlapping data chunk was then used for the subsequent update.

The pipeline was implemented and analyzed in a simulated online environment using BCILAB, an open source MATLAB toolbox designed for BCI research [83, 63]. It was initialized with the first second of data segment from the dataset. For the simulated 64-ch stationary data,

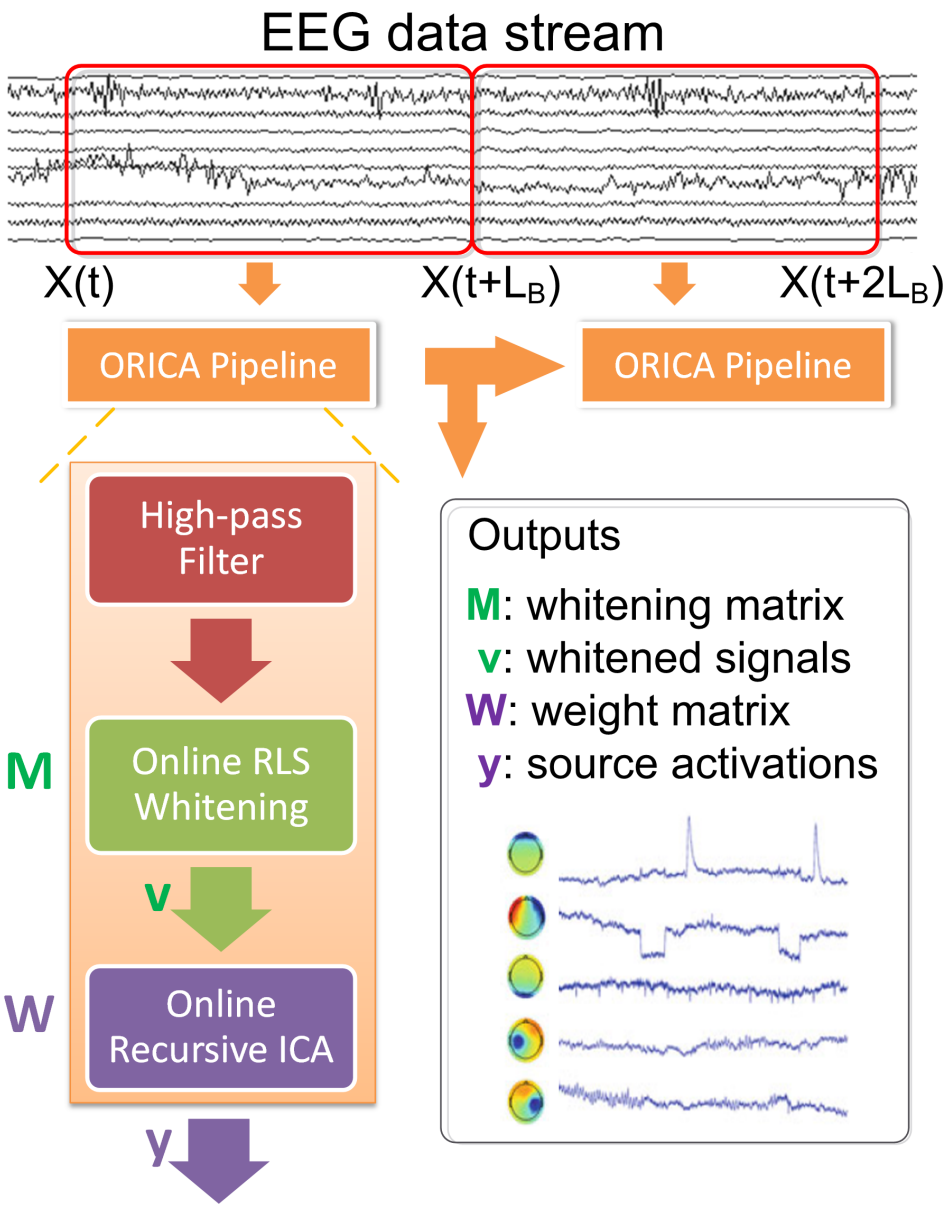

Figure 4.1: The ORICA pipeline for online EEG data processing. $X(t)$ is the input data vector at time t and L_B is the size of data in the online buffer.

Table 4.2: List of parameters for the ORICA pipeline: IIR high-pass filter (IIR), online RLS whitening filter (RLS) , and online recursive ICA filter (ICA).

Filters	Parameters	Values	Description
IIR	BW	0.2−2 Hz	Transition bandwidth
RLS	L_{white}	1∼16	Block-average size
RLS	λ_0	0.995	Initial forgetting factor
ICA	γ	0.60	Decay rate of forgetting factor
ICA	L_{ICA}	1∼16	Block-update size
	n_{sub}	0	Number of subgaussian sources

we investigated the effect of the parameters on empirical convergence, and we used block sizes $L_{white} = L_{ICA} = 16$ as an example to show the decomposed components since the block sizes returned satisfactory results in the shortest computational time. For the simulated 64-ch non-stationary data and real 61-ch data, we used $L_{white} = 8$ and $L_{ICA} = 1$, which were found to be optimal for the simulated stationary data. Table 4.2 summarizes the parameters of the three filters.

4.3.3 Data processing and analysis

We applied additional procedures to process and analyze the recorded EEG data from the subject. Firstly, an automatic removal of bad (*e.g.* flatlined or abnormally correlated) channels was applied prior to the ORICA pipeline using BCILAB routines, which removed 3 out of 64 channels. Secondly, following application of the ORICA pipeline, the source activities were epoched in a -400 to 600 msec window time locked to subject's responses (button press), yielding 693 epochs (104 error and 589 correct trials). The epochs were averaged to produce ERPs and were analyzed offline in EEGLAB [29].

4.3.4 Performance evaluation

Performance index

If the ground truth (N-by-N) mixing matrix $\mathbf{A}$ is known, a performance index PI for assessing quality of source separation can be defined as [32]:

$$PI = \frac{1}{N-1}\left[N - \frac{1}{2}\sum_{i=1}^{N}\left(\frac{\max\limits_{1\le j\le m}|\mathbf{C}_{ij}(n)|^2}{\sum_{j=1}^{N}|\mathbf{C}_{ij}(n)|^2} + \frac{\max\limits_{1\le j\le m}|\mathbf{C}_{ji}(n)|^2}{\sum_{j=1}^{N}|\mathbf{C}_{ji}(n)|^2}\right)\right] \tag{4.9}$$

where $\mathbf{C}(n) = \mathbf{W}_n\mathbf{M}_n\mathbf{A}$. This measures a normalized total cross-talk error of the estimated whitening matrix $\mathbf{M}$ and weight matrix $\mathbf{W}$, accounting for scale and permutation ambiguities. For perfect separation at convergence, PI approaches zero.

Best-matched correlation coefficients and Hungarian algorithm

PI reflects ORICA's global performance across all components. However, it is also useful to evaluate convergence of individual independent components (ICs), *i.e.* rows of $\mathbf{W}$. One metric is the Pearson correlation between an estimated IC and its counterpart in a "ground truth" weight matrix, $\mathbf{W}^*$. Due to permutation ambiguities, a matching algorithm is required for optimal pairing of rows of $\mathbf{W}$ and $\mathbf{W}^*$. This study used the Hungarian method [13] to maximize the sum of absolute pairwise correlations. We used Niclas Borlin's implementation in EEGLAB's $matcorr(\cdot)$ function.

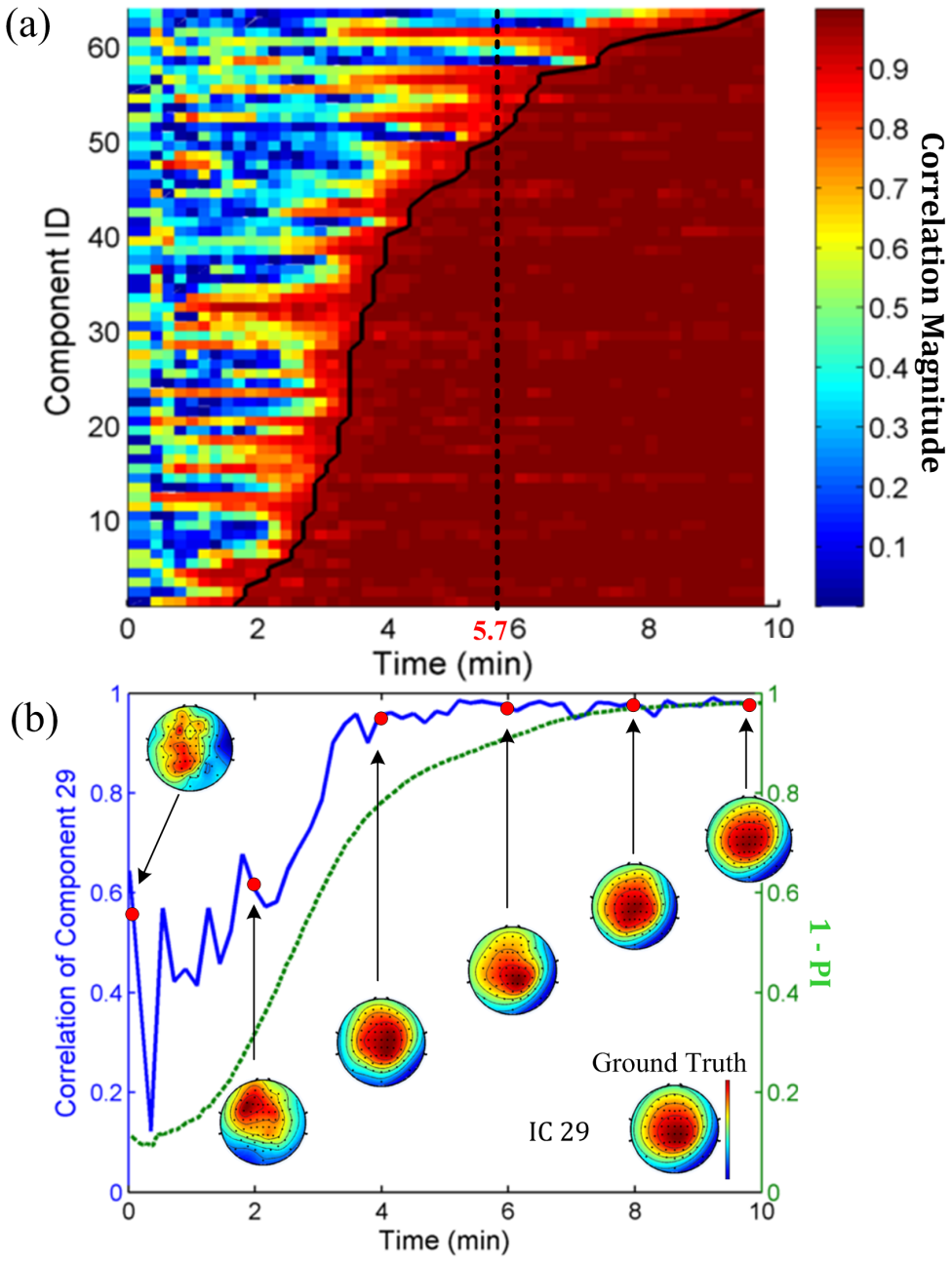

Figure 4.2: (a) Evolution of component-wise correlation magnitudes between ORICA-decomposed ICs and ground truth on simulated stationary 64-ch EEG data [45]. ICs sorted with respect to time required to reach a correlation magnitude of 0.95 (solid curve). The dotted line is the heuristic time for separating 64 stable ICs. (b) Evolution of correlation magnitudes (blue) and component maps of a randomly selected IC #29. One minus the performance index (green) is superimposed.

4.4 Results

4.4.1 Simulated 64-ch stationary EEG data

Evaluation of the decomposed components

Figure 4.2a plots the correlation magnitudes between ORICA components and their ground-truth counterparts. The components were sorted such that smaller component ID represented faster convergence. We observed that all components converged, *i.e.* correlation magnitudes approached 1, by the end of the 10-minute session. A common empirical heuristic for the number of training samples required for separating N stable ICA sources using Infomax ICA was kN^2, where $k > 25$ [90]. For 64 channels, the heuristic time required for convergence amounted to $64^2 \times 25 = 102400$ samples $= 5.7$ minutes with a 300 Hz sample rate. By 5.7 minutes, 77% (91%) of ICs reached a correlation magnitude of 0.95 (0.8); by 3.5 minutes, more than half of the ICs reached a correlation magnitude of 0.95. Figure 4.2b shows the evolution of the component maps of a randomly selected IC #29 and its correlation magnitudes with ground truth. This IC converged to a steady-state correlation magnitude of 0.95 under 4 minutes. The superimposed global performance index ($1 - PI$, in green) exhibited a similar convergence trajectory. Figure 4.3 shows 300-msec time-series of four representative ICs reconstructed by ORICA at 3, 6, and 9 minutes. At 9 minutes, ORICA correctly reconstructed all the source dynamics with the errors approached 0. At the heuristic time 5.7 minutes, only ID #58 had not converged. Interestingly, components such as IC #3 converged within 3 minutes. Both decomposed component maps and recovered source dynamics demonstrated ORICA's suitability for accurate and efficient decomposition of high-density (64-channel) data, albeit with systematic variation in convergence speed.

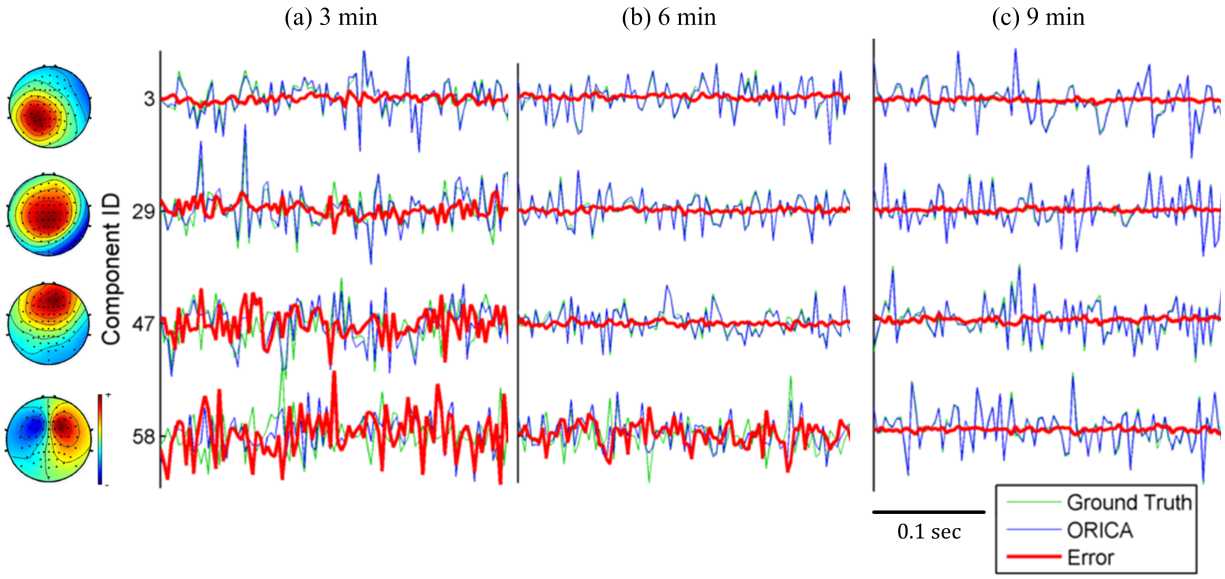

Figure 4.3: Source dynamics and the corresponding component maps of four representative components reconstructed by ORICA at (a) 3 min, (b) 6 min, or (c) 9 min of simulated mixed 64-channel stationary EEG data. Reconstructed source dynamic (blue) is superimposed on ground truth (green) with error, *i.e.* difference, (red). The oscillatory (autocorrelated) and burst-like source dynamics, as well as homogeneous component maps of the four depicted ICs, are characteristic of real EEG source dynamics.

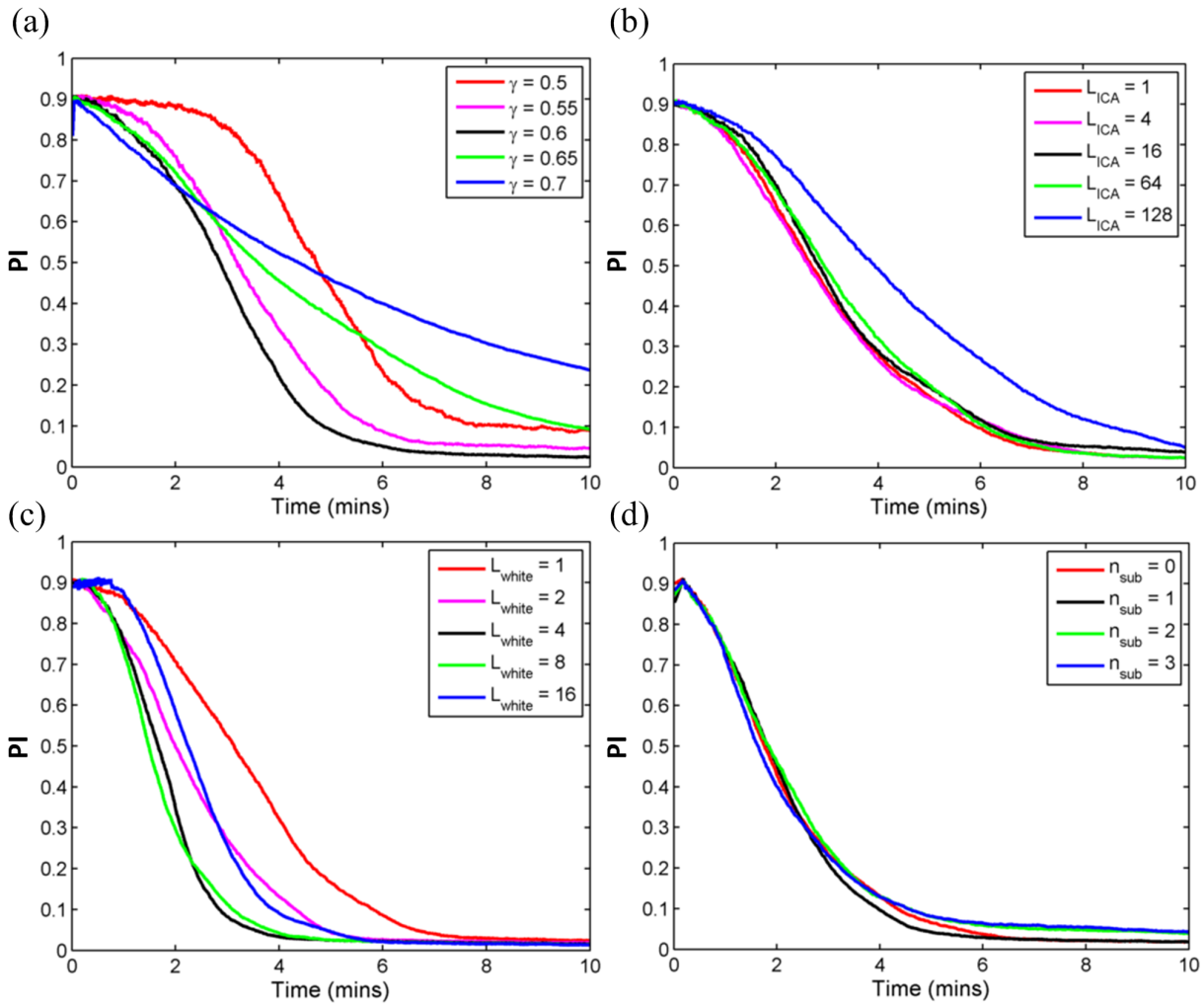

Figure 4.4: Effect of ORICA pipeline parameters on convergence trajectory, *i.e.* performance index over time, applied to 64-ch simulated stationary EEG data. (a) Decay rate γ of forgetting factor, (b) block size L_{ICA} of ORICA, (c) block size L_{white} of the online whitening, and (d) pre-assumed number n_{sub} of subgaussian sources.

Table 4.3: Averaged execution time (ms) for 1 sec (300 samples) 64-channel data using online RLS whitening and ORICA with different block sizes L_{white} and L_{ICA}.

Time (ms)	Block Size							
Algorithm	1	2	4	8	16	32	64	128
RLS	35.2±5.2	23.8±7.9	13.5±6.0	8.6±4.3	5.3±1.5	4.1±4.5	3.0±4.2	2.5±4.9
ORICA	332±29	174±37	79.5±14.5	36.3±11.8	21.0±5.2	15.1±7.2	10.6±9.4	6.6±2.6

Run in MATLAB 2012a on a dual-core 2.50GHz Intel Core i5-3210M CPU with 8GB RAM.

Effect of ORICA parameters

As shown in Figure 4.4, we systematically evaluated the effects of four ORICA parameters on convergence. The decay rate of forgetting factor γ had a significant impact on convergence speed, with the fastest convergence for $\gamma = 0.6$ (Figure 4.4a). A sigmoidal profile for the convergence trajectory was observed for $\gamma \leq 0.6$, while an exponential decay profile was observed for $\gamma > 0.6$. The ORICA block size L_{ICA} had negligible effect on the convergence for $L_{ICA} \leq 64$ (Figure 4.4b). This demonstrated the approximation error of the block update rule (Equation 4.6) was negligible for small to moderate block sizes. The block size of online whitening L_{white} significantly affected the ICA convergence (Figure 4.4c). L_{white} between 4 and 8 samples achieved the best performance. Interestingly, $L_{white} = 1$ was not the optimal value, mainly due to the effect of variable time scale of adjustments in the whitening matrix on the convergence of the subsequent ORICA. The pre-assumed number of subgaussian sources n_{sub} had little effect on the convergence, within the range of $n_{sub} = 0 \sim 3$, with the true number $n_{sub} = 0$. The performance of ORICA was rather insensitive to assumptions on the kurtosis of the sources. For the above results, we set $\gamma = 0.6$, $L_{ICA} = 1$, $L_{white} = 1$, and $n_{sub} = 0$ unless otherwise noted.

Quantification of computational load

Table 4.3 shows the average execution time required to apply the ORICA pipeline to 1-second of data, computed by averaging the processing rates (data size divided by time) of the incoming data chunks for 1 minute. Runtime was uniformly less than one second, illustrating the

64-ch data were processed faster than accumulated in the input buffer, and thus the pipeline was capable of real-time operation. The online whitening filter ran 3 to 9 times faster than ORICA did and the runtime monotonically decreased as the block size increased. The execution time of the ORICA filter was nearly halved as block size doubled when $L \leq 8$, with diminishing returns for $L \geq 16$. This allows us to balance the trade-off between runtime of the pipeline and accuracy of the block-update rule.

4.4.2 Simulated 64-ch nonstationary EEG data

Figure 4.5 plots ORICA's performance in tracking spatial non-stationarity in simulated 64-channel EEG data, with the simulated abrupt shifts of the EEG cap (Figure 4.5a). Figure 4.5b plots the non-stationarity index δ_{ns} as a function of time. The index robustly identified changes in the mixing matrix due to cap displacements. Figure 4.5c plots the performance index of ORICA's decomposition as a function of time. Following the detection of non-stationarity, ORICA's forgetting factor was reset to its initial value, and ORICA smoothly adapted to the new mixing matrix. Figure 4.5d plots the ground truth and the estimated component maps for a representative IC at different time-points, superimposed on a plot of log-transformed PI. For this IC, suitable convergence was obtained within 15 minutes, and improved further over time. The effect of cap rotation was captured by the concomitant shift of the component maps, indicating ORICA's capability to detect and adapt to the spatial abrupt non-stationarity.

4.4.3 Real 61-ch EEG data from the Flanker task

Since the ground truth was unknown for real EEG data, we adopted the offline Extended Infomax ICA algorithm [66], as implemented in the EEGLAB [29] function RUNICA, as a "gold standard". The robustness and stability of the algorithm on high-density EEG data had been shown to outperform most blind source separation algorithms [30].

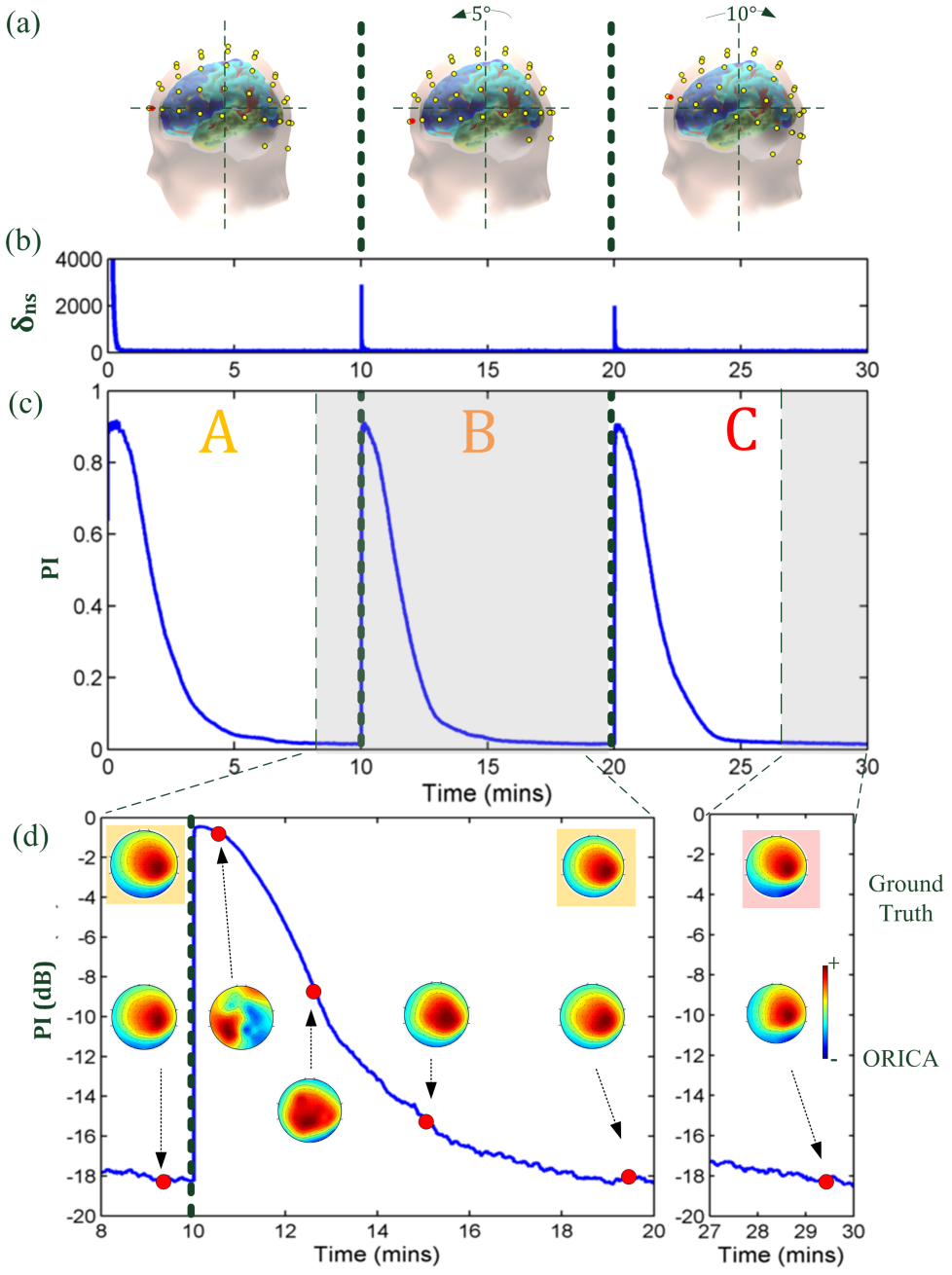

Figure 4.5: Application of ORICA to 64-ch simulated non-stationary EEG data consisting of three concatenated 10-minute sessions which simulate a 5-deg forward EEG cap rotation followed by a 10-deg backward cap rotation. (a) Electrode locations for each session. (b) non-stationarity index $\delta_{ns} = ||\mathbf{y}\mathbf{f}^T - \mathbf{I}||_F$ detects the abrupt change between sessions. (c) The *PI* convergence curve shows the adaptation behavior. (d) Zoomed-in plots of log-scaled convergence curve with ground truth component maps (top row) and reconstructed IC at different time points.

To confirm whether ORICA yielded a comparable solution at convergence (average over the last minute) as RUNICA, we investigated event-related activities of three sets of components with stereotypical fronto-central, occipital, and prefrontal spatial topographies. Figure 4.6a revealed that the ORICA- and RUNICA-decomposed fronto-central ICs and their characteristic ERN were comparable and consistent with the results from previous studies [83, 33]. Figure 4.6b shows that averaged occipital visual-evoked potentials (VEP) elicited by flanker (0-50 msec window after the onset of stimulus) and target (100-150 msec window after) presentation were clearly observed using both methods. Figure 4.6c shows eye blinks time-locked to response (click of button), as previously described in [33]. The ERN and VEP shown in Figure 4.6 were representative, with comparable results obtained from all subjects, indicating the reproducibility of the ORICA pipeline. In summary, the empirical results demonstrated comparable performance of ORICA to RUNICA in separating informative ICs and resolving single trial and averaged ERPs. Furthermore, ORICA required significantly less computation time than RUNICA.

Adopting a procedure similar to Figure 4.2, Figure 4.7a plots the correlation magnitudes between all ICs learned by ORICA and their best-matched RUNICA counterparts. Firstly, only 10% (26%) of the ORICA components reached a correlation magnitude of 0.9 (0.8) at the end of the session (average over the last minute). Secondly, 8% (21%) of the components reached a correlation magnitude of 0.9 (0.8) within 3-4 minutes, only half of the empirical heuristic time suggested by [90].

Among those ICs with the highest correlation magnitudes (larger than 0.95), we found a number of ICs with stereotypical and plausible component maps. Figure 4.7b plots the convergence profile (evolution of correlation magnitudes) for three such informative ICs: prefrontal (IC 1, accounting for eye-blink), occipital (IC 12, accounting for VEP), and fronto-central (IC 7, accounting for ERN) components. These ICs converged to their RUNICA counterparts within 3-4 minutes. The correlation magnitude curves of both the occipital and fronto-central components fluctuated across time and eventually reached a steady-state. In contrast, the correlation magnitude

time series of the prefrontal component was relatively stable across the whole session.

Figure 4.2a and Figure 4.7a exhibit significant performance differences in decomposing simulated versus real EEG data, which can be attributed to the differences in the quality of the gold standard. Those ICs producing poor correlation corresponded to non-dipolar sources in the gold standard, *i.e.* , sources with high residual variance in dipole fitting, such as mixtures of sources or noise [30]. This phenomenon is commonly observed and reported when applying ICA methods to real EEG data that are inevitably noisy and likely non-stationary [50], [118].

4.5 Discussion

4.5.1 Fast convergence speed

The ORICA pipeline was capable of accurately decomposing 64-channel simulated EEG data within the required heuristic convergence time (kN^2) [90]. The fast convergence could be attributed to three important factors: (a) combining online RLS pre-whitening and ORICA, (b) choosing an optimal forgetting factor profile and parameters, and (c) fixing the numbers of modeled sub- and super-gaussian sources.

Simulation results showed that faster convergence of the whitening matrix facilitated ICA convergence. This was consistent with the findings in previous studies [50, 126], which suggested pre-whitening could significantly improve the ICA convergence by reducing the dimensionality of the parameter space. For online whitening, a local block-average approach could provide a more robust estimate than a stochastic (single sample) update approach. Besides RLS whitening, online LMS whitening method [12] can also be considered, which has lower computational complexity, but slower convergence [126].

The forgetting factor, especially its decay rate γ, had a significant effect on ORICA's convergence. For highly non-stationary data, a large γ and thus a shorter effective window size were preferred. Factors including data dimension (number of channels) and underlying

stationarity of the data affected the choice of optimal parameters. Alternative approaches for adapting the forgetting factor were suggested in [26] and [94].

Fixing the numbers of sub- and super-gaussian sources increased both the stability and the speed of convergence, especially for high-density data. The experiments with synthetic data showed that the choice was not critical, and a model mismatch in these numbers could be tolerated. This study further examined the kurtosis of the real 61-ch EEG data and found that all sources were supergaussian distributed, which was consistent with previous studies [54, 29] that EEG signals from most brain activities and non-brain artifacts were primarily supergaussian. For greater accuracy under general conditions, online kurtosis estimation as described in [4] and [22] can be incorporated into the ORICA pipeline. This may lead to better steady-state performance but potentially decreased stability of convergence.

Several approaches not yet implemented in this study can further improve the convergence speed of ORICA. For example, dimensionality reduction methods such as PCA, or selecting a subset of channels prior to ORICA decomposition, can reduce the empirical convergence time. Another promising approach is to pre-process the data using artifact reduction methods such as Artifact Subspace Reconstruction [83] to mitigate sensitivity to transient artifacts in noisy high dimensional EEG data.

4.5.2 Real-time processing

ORICA was implemented as a BCILAB function with block-update vectorization, and could easily perform real-time processing with a user-defined block-size for 64-channel EEG on a standard laptop. The block update rule, while approximate, incurred negligible loss in accuracy up to $L_{ICA} = 64$ for 64-channel simulated data. Even without the block-update ($L_{white} = L_{ICA} = 1$), the ORICA pipeline still ran in real-time. The block update may be most valuable when computational resources are constrained; for instance, when applying multiple data processing operations in serial or operating on a low-power mobile device.

4.5.3 Application to real EEG data

Empirical results on the 61-channel EEG data collected in the flanker task experiment demonstrated that ORICA could decompose brain sources and artifact ICs that resembled results from standard RUNICA. We observed that the most informative ICs, such as VEP and ERN brain sources and eye-blink artifacts, had the highest correlation magnitudes among all ICs and converged much quicker than the heuristic convergence time—a fortunate circumstance for real-time applications in mobile EEG BCI. We speculate that those components exhibit robust and frequently occurring statistical patterns which facilitate ICA separation. These observed phenomena support applications of ORICA for rapid decomposition of high-density data as much less time is required to decompose brain and artifact components.

The ORICA pipeline also revealed non-stationarity in the experimental data, captured by the dynamics of component maps and the non-stationarity index. One challenge for ORICA and RUNICA is the order switching of ICs, especially for non-stationary real EEG data. This hinders the identification of informative ICs over time, *i.e.* tracking the same set of components regardless of the weight matrix permutations. One solution to the problem is to sort the current weight matrix $\mathbf{W}(t)$ based on the correlation matching with the previous weight matrix $\mathbf{W}(t-1)$ using the Hungarian method described in section 4.3.4 to keep track of the same components. This is useful for online identifying and separating artifact components from noisy EEG recordings.

4.5.4 Nonstationarity detection and adaptation

This study proposed using the Frobenius norm of ICA error matrix, $||\mathbf{yf}(\mathbf{y})^T - \mathbf{I}||_F$, for the ORICA pipeline as an index to detect non-stationarity events in the data, identified as abrupt changes in the mixing matrix. This non-stationarity index measurement captured the fitness of the ICA model to the current data, *i.e.* the degree to which nonlinear decorrelation was achieved. Alternative forms for the non-stationarity index can be used depending on applications, for

instance mutual information reduction (MIR) in windowed data provides a measure of statistical independence between sources [30] and thus can capture the changes in data statistics as MIR varies.

This study also presented a method of non-stationary adaptation by increasing the forgetting factor when the non-stationarity index exceeds a hard threshold, *e.g.* 1-10% of the initial value when ICA had not converged. However, this method required prior knowledge of the hard threshold and did not address continuous variation in degree of non-stationarity. For selection of the threshold, it is possible to design an adaptive threshold that depends on the online estimated mean and standard deviation of δ_{ns}. For adaption of the forgetting factor, one possible solution is to adopt the strategy similar to the adaptive learning rate proposed by Murata *et al.*[86] for a gradient-based algorithm in an online environment. An adaptive forgetting factor for ORICA, as an RLS-like recursive online algorithm, is crucial for its ability to track non-stationarity, calling for further investigation.

4.5.5 Applications and future directions

The proposed online ICA method for real-time processing of high-density EEG opens up new opportunities for the following potential applications: (1) ICA-based real-time artifact removal (especially for sporadic muscle activities), (2) ICA-based brain activity monitoring (*e.g.* epilepsy, etc) for clinical practice, and (3) adaptable ICA-based features for brain state (*e.g.* cognitive functions, fatigue level, etc) classification in real-time brain-computer interfaces.

A significant next step is to leverage ORICA for real-time source localization, for instance using anatomically constrained low resolution electrical tomographic analysis (LORETA) [114]. A recent study [101] attempted to combine online ICA and source localization, yet further validation of sources' reliability were needed. Knowledge of the source locations in the brain can be used to assess the reliability of the sources (*e.g.* validate consistency of the source locations over time and with anatomical expectations), to provide biological interpretation of the decomposed

sources, and to integrate with other real-time source-level methods such as connectivity analysis in SIFT [83, 85].

4.6 Conclusion

This study proposed and demonstrated an efficient computational pipeline for real-time, adaptive blind source separation of EEG data using Online Recursive ICA. The efficacy of the proposed pipeline was demonstrated on three datasets: a simulated 64-channel stationary dataset, a simulated 64-channel non-stationary dataset, and a real 61-channel EEG dataset collected under an Eriksen flanker task. Through application of ORICA to simulated stationary data we (a) systematically evaluated the effects of key parameters on convergence; (b) characterized the convergence speed, steady state performance, and computational load of the algorithm; and (c) quantitatively compared the proposed ORICA method with a standard offline Infomax ICA algorithm. Our analysis of a simulated non-stationary 64-channel EEG dataset demonstrated ORICA's ability to adaptively track changes in the mixing matrix due to electrode displacements.

Applied to 61-channel experimental EEG data, we demonstrated ORICA's ability to decompose brain and artifact subspaces online, with comparable performance to offline Infomax ICA. Furthermore, we found that subspaces of biologically plausible ICs (*e.g.* eye, occipital and frontal midline sources) could be reliably learned in much less time than required by the kN^2 empirical heuristic for ICA convergence. To serve the EEG and BCI communities, the proposed pipeline has been implemented as BCILAB [63] and EEGLAB compatible functions, it has also been integrated into an open-source Real-time EEG Source-mapping Toolbox (REST) [101]. Future work will focus on further validation of this promising method as well as application to artifact rejection, clinical monitoring, and brain-computer interfaces [83].

4.7 Acknowledgment

Chapter 4, in part, is a reprint of the material as it appears in Hsu, S. H., Mullen, T. R., Jung, T. P., and Cauwenberghs, G. (2016). "Real-time adaptive EEG source separation using online recursive independent component analysis." *IEEE transactions on neural systems and rehabilitation engineering*, 24(3), 309-319. © 2016 IEEE. The dissertation author was the primary investigator and author of this paper. This study was sponsored in part by a gift by the Swartz Foundation (Old Field, NY), by the Army Research Laboratory under Cooperative Agreement Number W911NF-10-2-0022, by the NIH grant 1R01MH084819-03 and NSF EFRI-M3C 1137279. The authors would like to thank Christian Kothe for help with BCILAB integration.

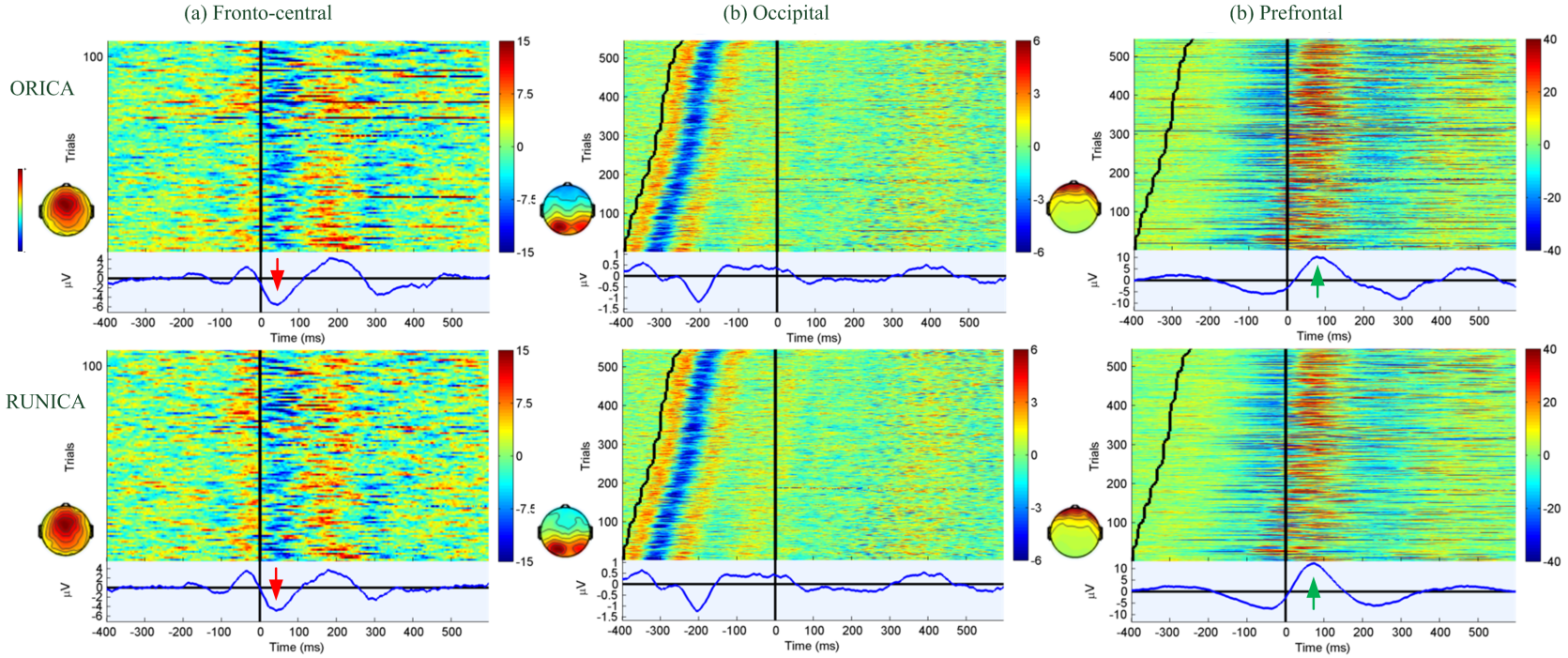

Figure 4.6: Color-coded event-related potential (ERP) images (trials by time) of (a) fronto-central, (b) occipital, and (c) prefrontal components reconstructed by ORICA (top row) and RUNICA (bottom row) on real 61-ch EEG data from the flanker task, time-locked to the response at time 0 (vertical straight line). Averaged ERP traces are shown in bottom panel. Only error trials are included in (a) such that error-related negativity (ERN) can be observed as red arrows indicate. In (b) and (c), all trials are sorted based on reaction time, *i.e.* onset of flanker stimulus (sawtooth line) to response. A visually evoked potential (VEP) is clearly observed in (b). Green arrows in (c) indicate to eye blinks.

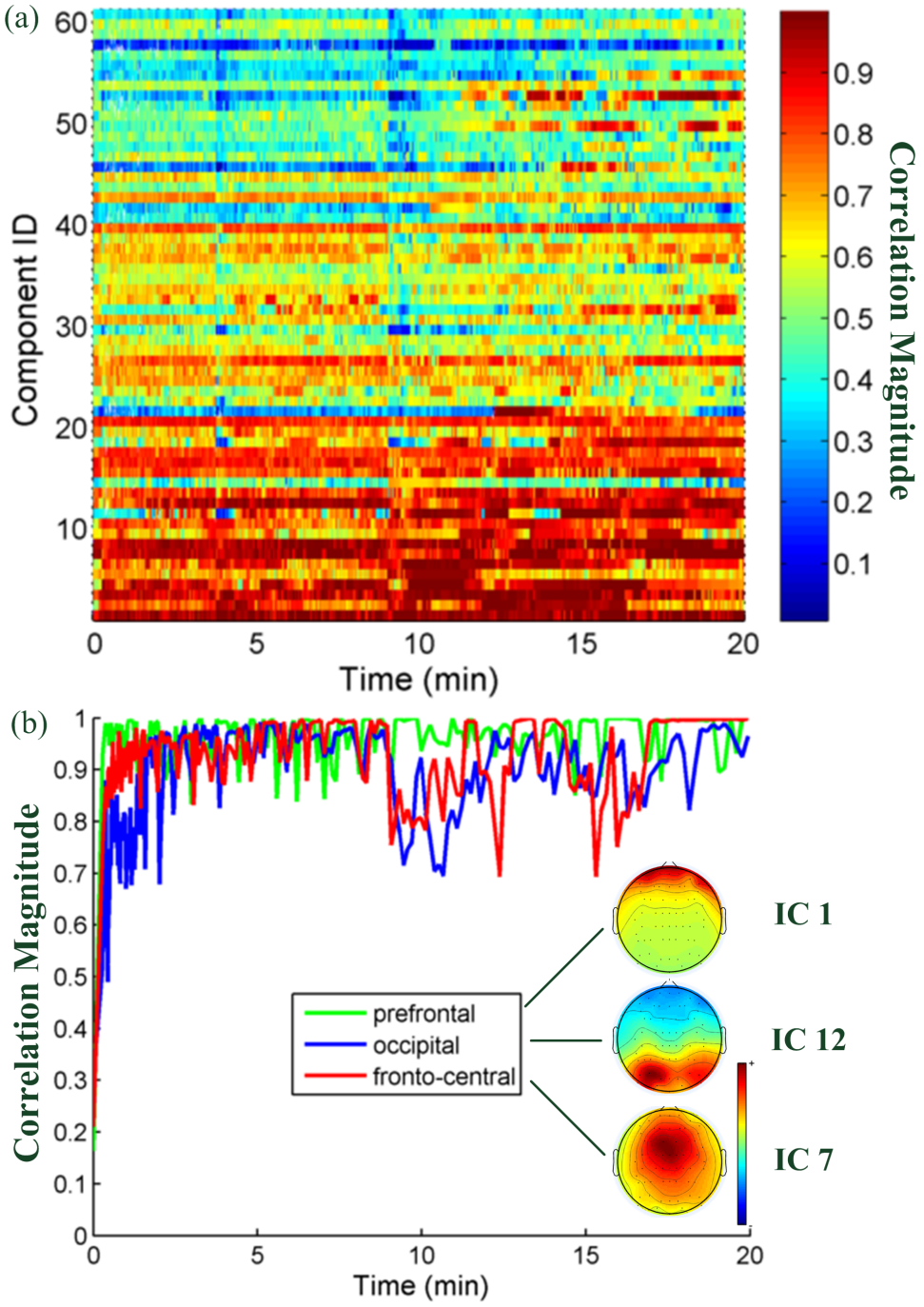

Figure 4.7: (a) Evolution of component-wise correlation magnitudes between ORICA- and RUNICA-decomposed ICs on real 61-ch EEG data from the flanker Task. (b) Evolution of correlation magnitudes and spatial filters of three rapidly converged ICs: prefrontal (eye-blinks), occipital (VEP), and fronto-central (ERN) components.

Chapter 5

Validating Online Recursive ICA on EEG data

The needs for online Independent Component Analysis (ICA) algorithms arise in a range of fields such as continuous clinical assessment and brain-computer interface (BCI). Among the online ICA methods, online recursive ICA algorithm (ORICA) has attractive properties of fast convergence and low computational complexity. However, there hasn't been a systematic comparison between an online ICA method such as ORICA and other offline (batch-mode) ICA algorithms on real EEG data.

In this chapter, we compared ORICA with ten ICA algorithms in terms of their decomposition quality, validity of source characteristics, and computational complexity on the thirteen experimental 71-ch EEG datasets. Empirical results showed that ORICA achieved higher mutual information reduction (MIR) and extracted more near-dipolar sources than algorithms such as FastICA, JADE, and SOBI did while the performance of ORICA approached that of the best-performed Infomax-based algorithms. Furthermore, ORICA outperforms most of ICA methods in terms of the computational complexity. The properties of fast convergence and low computational complexity of ORICA enable the realization of real-time online ICA process, which has further

applications such as real-time functional neuroimaging, artifact reduction, and adaptive BCI.

5.1 Background

Many ICA algorithms have been developed to separate the sources from the channel signals [50], with the assumption of spatiotemporal stationarity of the data, such as widely used Infomax ICA [8] and FastICA [51]. For non-stationary EEG signals, ICA methods such as Adaptive Mixture ICA (AMICA) [96] can generate good component decomposition but are computationally expensive. However, the aforementioned ICA algorithms are all offline analyses. In many real-world applications, including real-time functional neuroimaging [83], artifact rejection, and adaptive brain-computer interface (BCI), online (sequential) source separation methods are needed. Desirable properties in these applications are fast convergence and low computational complexity.

Among the online ICA methods, online recursive ICA (ORICA) [4] has such attractive properties of fast convergence and low computational complexity because it employs a recursive-least-squares (RLS) type method derived from a fixed point solution to Infomax ICA. Subsequently, a computationally real-time and online ICA pipeline has been proposed [45] which incorporates the whitening step prior to the application of ORICA to improve the convergence speed and employs a block-update rule to accelerate the computational speed. It has been shown that the ORICA method can decompose 64-ch simulated EEG data and extract informative components from the 61-ch experimental EEG data. However, some questions remained unanswered: how consistent are the ORICA components across experimental EEG datasets and how is its performance compared to other ICA methods.

This study tested the ORICA algorithm on thirteen experimental 71-ch EEG datasets and compared the results to those obtained by ten popular batch-mode ICA or BSS algorithms, following the ICA comparison approach published by [30]. Both quantitative and qualitative per-

formance of ORICA are given in terms of decomposition quality, validity of source characteristics, and computational complexity.

5.2 Methods

With the standard ICA model $\mathbf{x} = \mathbf{A}\mathbf{s}$, complete decomposition is performed, *i.e.* estimate unmixing matrix $\mathbf{W}$ such that all sources are recovered by $\mathbf{y} = \mathbf{W}\mathbf{x}$, where $\mathbf{x}$ are scalp EEG observations, $\mathbf{A}$ is an unknown mixing matrix, and $\mathbf{s}$ and $\mathbf{y}$ are original and reconstructed source activation time courses respectively.

5.2.1 Online recursive ICA and other ICA/BSS algorithms

ORICA is an online ICA algorithm with a recursive update rule derived from a fixed-point solution to the natural-gradient Infomax ICA, proposed by [4]. The final learning rule is extended to a block-update form [45] to reduce the computational load, as shown in Equation 4.6. Followed the optimized parameters used in [45], we chose the block size $L = 8$, the heuristic time-varying forgetting factor $\lambda_n = 0.995/n^{0.6}$, and the nonlinear projection function $f(y) = -2\tanh(y)$ with the assumption that all EEG sources are supergaussian. The assumption is verified on the datasets and is generally hold for filtered, *e.g.* line-noise removed, EEG data.

A whitening step is added prior to ORICA to decorrelate the data and improve the ICA convergence. Pre-whitening provides a fair comparison between ORICA and other batch-mode ICA algorithms where data are also pre-whitened. ORICA with pre-whitening (denoted as *ORICA*) is highlighted in pink in Section 5.4). ORICA with two to three passes over the data (denoted as *ORICA-2 or ORICA-3*) are also tested to see the improvement of performance as more data points are available.

Ten ICA or BSS algorithms from three groups plus principle component analysis (PCA) are used for comparisons. The first group is the natural gradient ICA algorithms, including

AMICA [96], Infomax [8], and Extended Infomax [66], highlighted in yellow in Section 5.4. This group is known for its ability to extract dipolar sources from EEG data while it is computational expensive. The second group includes FastICA [51] that maximizes negentropy, JADE [12] that maximizes the fourth-order cumulants, and ThinICA (TICA) [25], highlighted in green. The third group belongs to the second-order time-delay approach [20] to the BSS problem including SOBI, icaMS, FOBI, and EVD, highlighted in blue. PCA is also tested to characterize the effect of pre-whitening, highlighted in purple. The detailed selection of parameters and references of the above algorithms are described and listed in [30].

5.2.2 Evaluation methods

Mutual Information Reduction (MIR)

ICA aims at maximizing the independence and mutual information reduction (MIR) provides an estimation of independence in terms of mutual information between sources, defined as

$$MIR = I(\mathbf{x}) - I(\mathbf{y}) = \log|\det \mathbf{W}| + [h(\mathbf{x}_1) + \cdots + h(\mathbf{x}_N)] - [h(\mathbf{y}_1) + \cdots + h(\mathbf{y}_N)] \tag{5.1}$$

where I is mutual information and h is marginal entropy. Better ICA decomposition will have more independent sources and higher MIR.

Equivalent dipole modeling

It has been demonstrated that independent EEG sources are dipolar [30]. Hence to quantify the quality of the decomposed individual sources, the surface spatial distribution of each IC is fitted into a single equivalent dipole in a spherical four-layer head model using the DIPFIT plug-in in the EEGLAB toolbox [29]. The error of the dipole fitting is the residual variance (r.v.). Near-dipolar sources are defined as r.v. $< 5\%$. More near-dipolar components correlates with

better ICA decomposition for EEG data.

5.3 Materials

5.3.1 EEG data

The EEG datasets used in this study are available at http://sccn.ucsd.edu/wiki/BSSComparison, from thirteen subjects performing the visual working memory task described in [89]. Each dataset has EEG signals from 71 channels with sampling rate of 250Hz and data length of 296,000 to 315,000 samples (around 20 minutes of recording). The data have been high-pass filtered with cutoff frequency at 0.5Hz to remove trend, and the epochs with high-amplitude or high-frequency abnormalities have been removed [30].

5.3.2 Data analysis

Data were loaded into MATLAB and analyzed by the open source EEGLAB toolbox [29]. We ran each algorithm on each dataset, recorded the runtime, solved for unmixing matrix, and evaluated the decomposition performance with the methods described in Section 5.2.2. The MATLAB scripts for running the comparisons can be found at http://sccn.ucsd.edu/wiki/BSSComparison. The MATLAB functions of the algorithms in Section 5.2.1 can be downloaded from EEGLAB [29] (http://sccn.ucsd.edu/eeglab) and ICALAB [21] (http://www.bsp.brain.riken.jp/ICALAB).

5.4 Results

The informative independent components (ICs) decomposed by ORICA from each dataset are manually selected based on their scalp maps and mean activity spectra and grouped into the

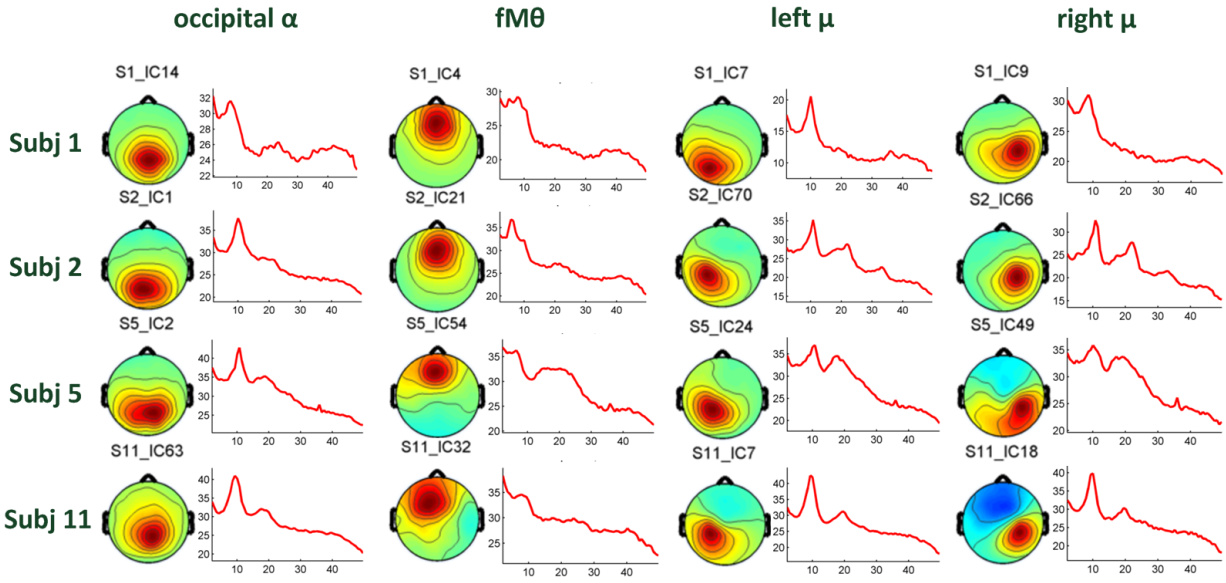

Figure 5.1: Scalp maps and mean activity spectra of four ORICA component clusters (central occipital alpha, frontal midline theta, left mu and right mu) accounting for non-artifact brain sources from four subjects. Components in each cluster are selected by visual inspection according to their scalp maps and power spectra.

following four non-artifact brain source clusters: central occipital alpha (8-12Hz), frontal midline theta (near 6Hz), left and right motor mu (near 10Hz with harmonics) clusters.

Each ORICA cluster contains similar ICs from 5-8 subjects, demonstrating the consistency of ORICA decomposition across the datasets. The results from the four selected subjects are shown in figure. 5.1. Due to variability across subjects, not every dataset returns the same informative ICs. However, the scalp maps and the source dynamics of the ORICA components are consistent with those from the previous study [30] using AMICA, which has shown to be the ICA algorithm with the best performance.

Figure 5.2 gives the qualitative comparisons across algorithms by looking at the five informative ICs from ORICA on the dataset S2 and their counterparts decomposed by AMICA, Infomax, FastICA, and SOBI algorithms from the same dataset. Similar components were found by best-matching the ICs, in terms of the correlation of the scalp maps, of ORICA and the other algorithms using the Hungarian method.

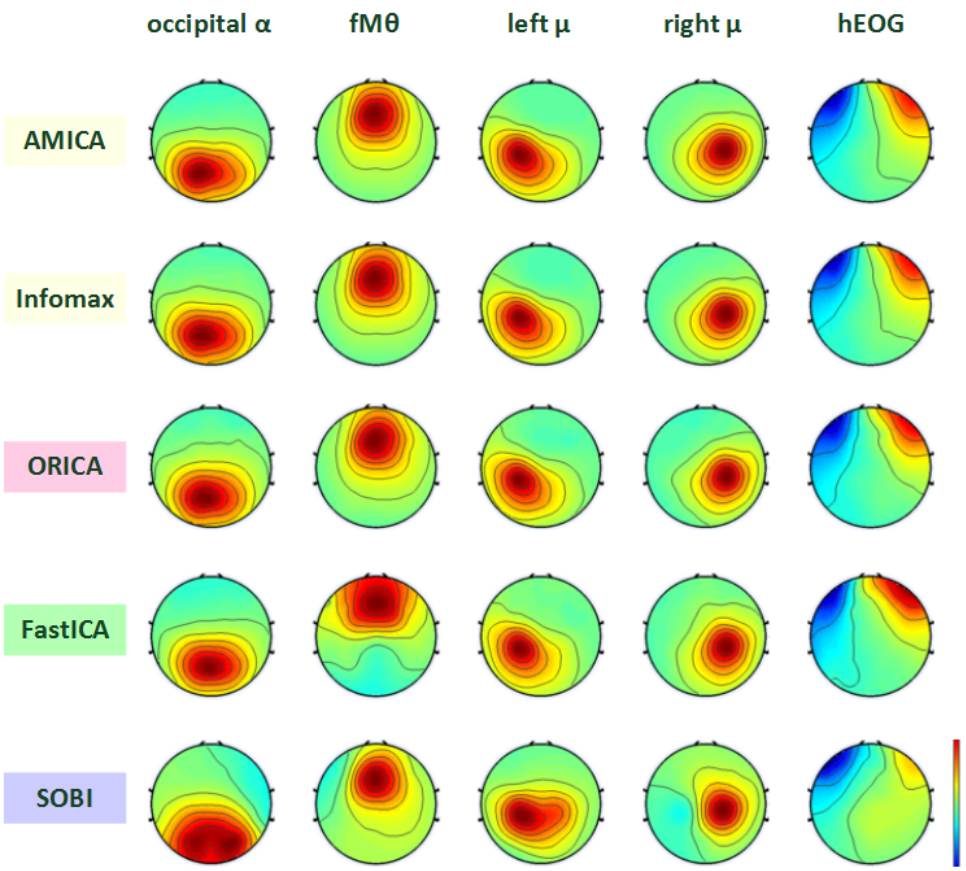

Figure 5.2: Scalp maps of the five informative components - the occipital alpha, the frontal midline theta, the left mu, the right mu, and the horizontal eye movement (EOG) - decomposed by AMICA, Infomax, ORICA, FastICA, and SOBI from the dataset S2. The similar components are found by best-matching ICs with the results of ORICA.

ORICA components strongly resemble the ICs from the best-performed ICA methods such as Infomax and AMICA. The FastICA returned the similar components but some ICs such as fMθ and hEOG are less homogeneous. Spatial distribution of the SOBI components are even worse. Results of the PCA components are not shown because PCA failed to decompose similar components. The components from AMICA, Infomax, and ORICA have lower r.v. values than those from FastICA, and SOBI.

Figure 5.3 provides the quantitative comparisons of the five groups of the fourteen algorithms, assessed by mutual information reduction (MIR) and percentage of near-dipolar components with r.v. $< 5\%$ averaging across the thirteen datasets. Along the linear regressed curve, algorithms at top right have better decomposition performance and more near-dipolar sources. ORICA (pink) has worse performance than the standard Infomax-based algorithms (yellow), whereas it outperforms the other groups of algorithms such as FastICA (green), SOBI (blue), and PCA (purple).

It should be emphasized that even the ORICA with single-pass performs better than several batch-mode algorithms. With more passes through the dataset, the performance of ORICA-n improves and approaches the performance of the leading Infomax group, which has a few hundreds passes. This suggests ORICA has fast convergence characteristic.

Figure 5.4 compares the computation time, *i.e.* complexity, and shows the number of near-dipolar sources averaged across the thirteen datasets for all algorithms. ORICA runs four times faster than Infomax does and is able to decompose 80% of Infomax's near-dipolar sources. Furthermore, ORICA outperforms FastICA (green) in both computation time (three times faster) and performance (13% more near-dipolar sources). SOBI and PCA run faster than ORICA but have poor decomposition quality. The result of AMICA is not shown because it requires parallel processing and runs for hours.

Algorithms that decompose more near-dipolar components require more computation time. In the performance-complexity relations, ORICA stands out with its low computational

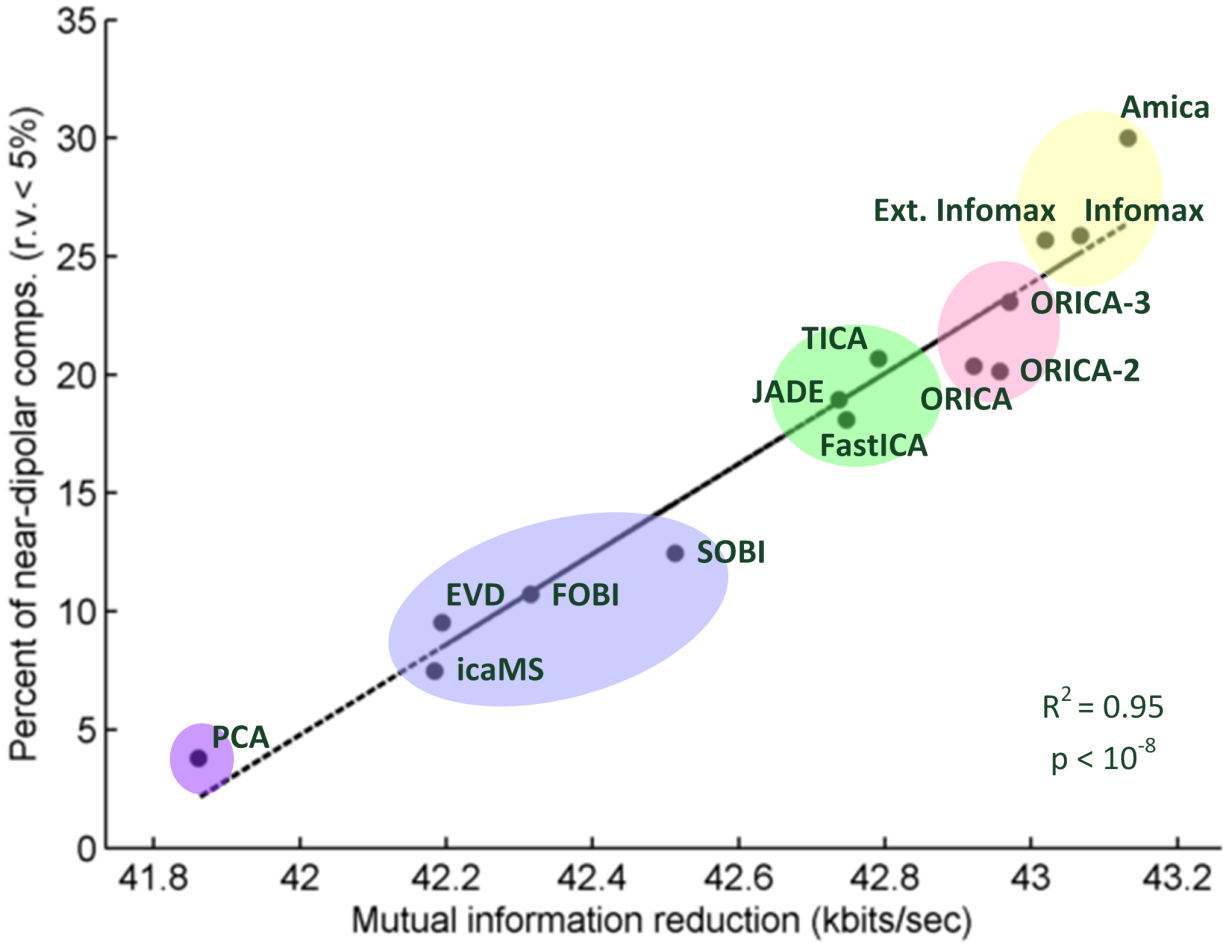

Figure 5.3: Percentage of the components (out of 71 ICs) with near-dipolar scalp maps (residual variance $<= 5\%$) and mean mutual information reduction (MIR) averaged across thirteen datasets for the five groups of the fourteen algorithms described in Section 5.2.1. The color of each group is manually added. Results of ORICA with n-passes over the data (ORICA-n) are shown.

complexity while the performance remains at high level. It is worth noting that ORICA can on average decompose a 71-ch dataset with 300,000 samples in about 40 seconds on a laptop without the feature of parallel processing. The even more powerful feature of ORICA is that the computation time can be further reduced by increasing the block-update size [45].

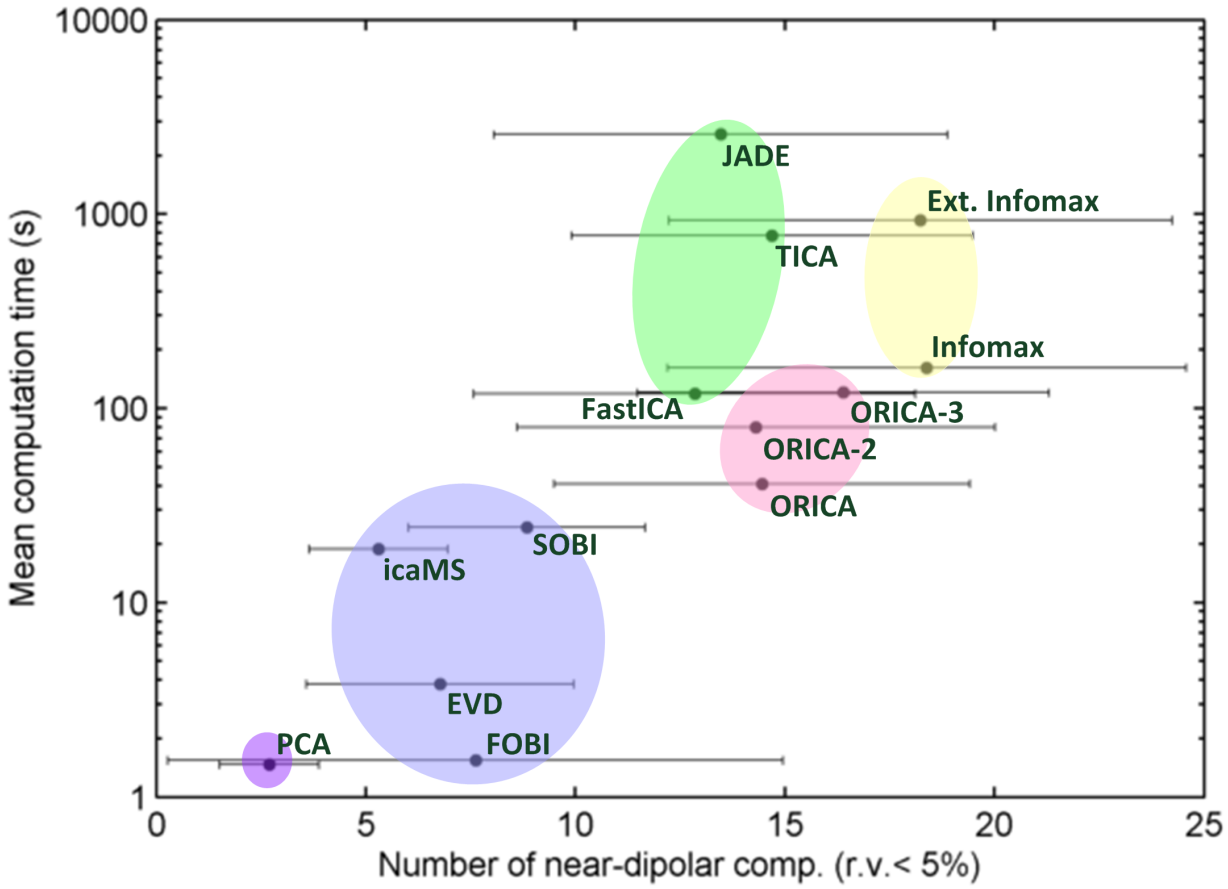

Figure 5.4: The log-scale computation time and the number of near-dipolar scalp maps (r.v. $<= 5\%$) of the five groups of the fourteen algorithms are shown. The dots and bars show the mean and standard deviation of the number of near-dipolar sources averaged across the thirteen datasets. The standard deviation for the computation time is too small to plot. Algorithms at lower right are better. The datasets are processed in MATLAB 2012a on a dual core 2.5GHz Intel i5-3210M CPU with 8GB RAM.

5.5 Conclusion

This study provides both qualitative and quantitative evaluation of the performance of the ORICA method. Qualitatively, the results confirm that ORICA can extract the informative ICs (brain and artifact sources) from the 71-ch experimental EEG data. The decomposed ICs are consistent across the datasets from different subjects and also agree with the previous study using AMICA. Quantitatively, in terms of decomposition performance (MIR) and validity of source characteristics (number of near-dipolar sources), ORICA's performance approaches the best-performed Infomax-based method as number of passes over data increases and is uniformly better than other algorithms. Furthermore, ORICA outperforms most of ICA methods in terms of computational complexity, except the algorithms with poor decomposition such as SOBI and PCA.

The results of this study indicate ORICA can achieve good decomposition with low computational complexity, which can be an alternative for ICA decomposition of high-density EEG data, compared to the standard ICA methods such as Infomax. These attractive properties of ORICA enable the realization of real-time online ICA process, which has further applications such as real-time functional neuroimaging, artifact rejection, and adaptive BCI.

5.6 Acknowledgments

Chapter 5, in part, is a reprint of the material as it appears in Hsu, S. H., Mullen, T., Jung, T. P., and Cauwenberghs, G. (2015, April). "Validating online recursive independent component analysis on EEG data." In *7th International IEEE/EMBS Conference on Neural Engineering (NER)* (pp. 918-921). © 2015 IEEE. The dissertation author was the primary investigator and author of this paper. We thank Arnaud Delorme, Jason Palmer, Julie Onton, Robert Oostenveld and Scott Makeig for their datasets and MATLAB scripts for ICA comparisons (available at `http://sccn.ucsd.edu/wiki/BSSComparison`).

Chapter 6

Tracking Non-stationary EEG Sources using Adaptive Online Recursive ICA

In Chapter 4 and 5, we show that Online Recursive ICA (ORICA) can achieve fast convergence in decomposing high-density EEG data for real-time applications. However, its adaptation performance has not been fully explored due to the difficulty in choosing an appropriate forgetting factor: the weight applied to new data in a recursive update which determines the trade-off between the adaptation capability and convergence quality.

In this chapter, we propose an adaptive forgetting factor for ORICA (adaptive ORICA) to learn and adapt to non-stationarity in the EEG data. Using a realistically simulated non-stationary EEG dataset, we empirically show adaptive forgetting factors outperform other commonly-used non-adaptive rules when underlying source dynamics are changing. Standard offline ICA can only extract a subset of the changing sources while adaptive ORICA can recover all. Applied to actual EEG data recorded from a task-switching experiments, adaptive ORICA can learn and re-learn the task-related components as they change. With an adaptive forgetting factor, adaptive ORICA can track non-stationary EEG sources, opening many new online applications in brain-computer interfaces and in monitoring of brain dynamics.

6.1 Background

A variety of analytical tools for biosignals such as EEG have successful uncovered functional features related to human cognitive functions and non-brain activities [29]. In order to obtain robust estimates from noisy measurements, most of the analyses are done offline and assuming stationary data. However, in practical applications such as brain-computer-interfaces and clinical monitoring of brain dynamics, batch learning is rather infeasible in the online environment due to the large number of samples required and high computational load [86]. Furthermore, assumptions of stationarity often do not apply to real-life situations, where humans organize their behavior with respect to a complex, multi-scale and ever-changing environment. Therefore, EEG source activities are inevitably non-stationary.

To address this, several online learning methods have been proposed which take into account non-stationarity in the data to study brain dynamics or mitigate non-stationarity problem [86][117][7]. Among these methods, Akhtar *et al*.[4] have proposed Online Recursive ICA (ORICA), derive from iterative inversion of the fixed point solution to natural gradient Infomax ICA rule, yielding the fast convergence of a recursive-least-squares (RLS) type filter and low computational load. Previous studies have shown that ORICA is capable of real-time processing of high-density EEG data and extracting informative sources from experimental EEG data comparable to those of standard ICA approaches [45, 46]. However, one important advantage of online ICA—adaptation performance—has not been explored. The difficulty lies in choosing a forgetting factor: the weight applied to new data in a recursive update rule determining the trade-off between adaptation capability and convergence quality.

This study proposes an adaptive forgetting factor for ORICA (adaptive ORICA), modified from the adaptive rule of [86] and [23], to learn and adapt to the non-stationary data. To compare the performance of adaptive and non-adaptive forgetting factors, we use a realistically simulated non-stationary EEG dataset with the underlying sources activating and deactivating alternatively.

We also compare the results of adaptive ORICA to those of a standard offline ICA (Infomax) [29]. Finally, we demonstrate the real-world applicability of adaptive ORICA using 14-ch EEG data recorded from a subject performing a task-switching experiment.

6.2 Adaptive forgetting factor for online ICA

6.2.1 Online recursive ICA: revisit

Comprehensive formulation of the online ICA problem and detailed derivation of the Online Recursive ICA algorithm have been presented in Section 4.2. This section gives a brief summary in an attempt to provide intuition and facilitate readers' understanding.

With the assumption that scalp EEG signals, $\mathbf{x}$, are linear mixtures of underlying independent sources, $\mathbf{s}$, *i.e.* $\mathbf{x} = \mathbf{A}\mathbf{s}$, ICA aims to learn an unmixing matrix $\mathbf{B}$ such that the estimated source activations $\mathbf{y}$ are recovered, with permutation and scaling ambiguities, by $\mathbf{y} = \mathbf{B}\mathbf{x}$. An online framework is proposed which separates the unmixing process into two stages $\mathbf{B} = \mathbf{W}\mathbf{M}$ [45]. The whitening matrix $\mathbf{M}$ is updated by an online RLS whitening update [126] to decorrelate the signals:

$$\mathbf{M}_{n+1} = \mathbf{M}_n + \frac{\lambda_n}{1-\lambda_n}\Big[\mathbf{I} - \frac{\mathbf{v}_n \cdot \mathbf{v}_n^T}{1+\lambda_n(\mathbf{v}_n^T \cdot \mathbf{v}_n - 1)}\Big]\mathbf{M}_n, \tag{6.1}$$

where $\mathbf{v}_n = \mathbf{M}_n\mathbf{x}_n$ are the decorrelated signals and λ_n is a time-varying forgetting factor. This is followed by ORICA update on the weight matrix $\mathbf{W}$ [4]:

$$\mathbf{W}_{n+1} = \mathbf{W}_n + \frac{\lambda_n}{1-\lambda_n}\Big[\mathbf{I} - \frac{\mathbf{y}_n \cdot \mathbf{f}_n^T}{1+\lambda_n(\mathbf{f}_n^T \cdot \mathbf{y}_n - 1)}\Big]\mathbf{W}_n \tag{6.2}$$

where $\mathbf{y}_n = \mathbf{W}_n\mathbf{v}_n$ are the recovered source activations and $f(y) = -2\tanh(y)$ are the component-wise nonlinear functions for super-gaussian sources and $f(y) = \tanh(y) - y$ for sub-gaussian sources. An orthogonalization step is applied after each update to ensure $\mathbf{W}$ matrix remains orthogonal. Following [45], we assume all EEG sources are super-gaussian to improve convergence

and we concurrently update the $\mathbf{M}$ and $\mathbf{W}$ matrix for a block of samples (here we used 8 samples to a block) to reduce the computational load. The forgetting factors in Equations 6.1 and 6.2 are chosen to be the same.

6.2.2 Design of adaptive forgetting factor

In adaptive RLS filtering, the forgetting factor λ determines the effective length for a time window wherein data are exponentially weighted. A small value of λ corresponds to a long window length which yields small errors and good stability at the expense of tracking capability [94].

Three update rules for forgetting factors are considered: constant, monotonically decreasing (cooling), and adaptive. In a conventional RLS filter, a constant forgetting factor rule is often used so that it has moderate tracking performance, but the error at convergence can be large. In order to achieve better learning performance, it is desirable to have a large λ during initial learning to increase convergence speed and a small λ near convergence to minimize error. To this end, a cooling forgetting factor may lead to satisfactory performance. We use the rule defined in [123] as $\lambda_n = \frac{\lambda_0}{n^\gamma}$, where λ_0 is the initial forgetting factor and γ determines its decay rates. The convergence rate of the cooling scheme is analyzed in [5]. Despite the good convergence, the annealed λ cannot follow non-stationarity in the data. Clearly, there is a need for an adaptive forgetting factor which increases during the presence of non-stationarity and otherwise decreases for fast and stable convergence.

To realize an adaptive forgetting factor, we need to quantify the non-stationarity in the data. Since ORICA is derived from the fixed point solution of minimizing $<\mathbf{I}-\mathbf{y}\mathbf{f}^T>$, the non-stationarity can be assessed by the error of nonlinear decorrelation, *i.e.* $\mathbf{I}-\mathbf{y}_n\mathbf{f}_n^T$. To increase the robustness of the estimates, a leaky average with weight δ is used:

$$\mathbf{R}_{n+1} = (1-\delta)\mathbf{R}_n + \delta(\mathbf{I}-\mathbf{y}_n\mathbf{f}_n^T) \tag{6.3}$$

The final non-stationarity index z_n is a scalar indicator of convergence for all sources which is defined as:

$$z_n = ||\mathbf{R}_n||_F \tag{6.4}$$

Motivated by Murata *et al*.[86], we use the following adaptive forgetting factor:

$$\lambda_{n+1} = \lambda_n - \alpha\lambda_n^2 + \beta G(z_{n+1})\lambda_n \tag{6.5}$$

where α controls the decay rate of λ in the learning stage (when data are relatively stationary) and β sets an upper bound on how fast λ can increase during the tracking stage (when data are non-stationary). In [86], the $\beta G(z)$ term in Equation (6.5) is defined as z_n/z_{max}. However, the linear mapping between z_n and λ_n does not work effectively when data dimension increases due to a large initial value of z.

To address the issue, inspired by the synaptic update from [23], we design $G(z)$ as a modified tanh function which maps the non-stationarity index z_n to a value from 0 to 1 indicating the level of λ adjustment:

$$G(z) = \frac{1}{2}\left\{1 + \tanh\left[\frac{1}{b}\left(\frac{z}{\max(z_{min}, \varepsilon)} - c\right)\right]\right\} \tag{6.6}$$

where b and c determine the bandwidth and the center of the tanh function. The term z_{min} is an indicator for initial convergence and ε controls the switching from learning stage to tracking stage. In the learning stage, we want $G(z)$ to be small so that λ can decrease to a small value to ensure convergence. In the tracking stage, we want $G(z)$ to be large when non-stationarity index increases.

Although many parameters are introduced in Equation 6.5, Murata *et al*.[86] have discussed how to choose α, β and δ. As for b and c in Equation 6.6, they can be designed according to the mean and variance of the non-stationarity index. This study follows the principles to search

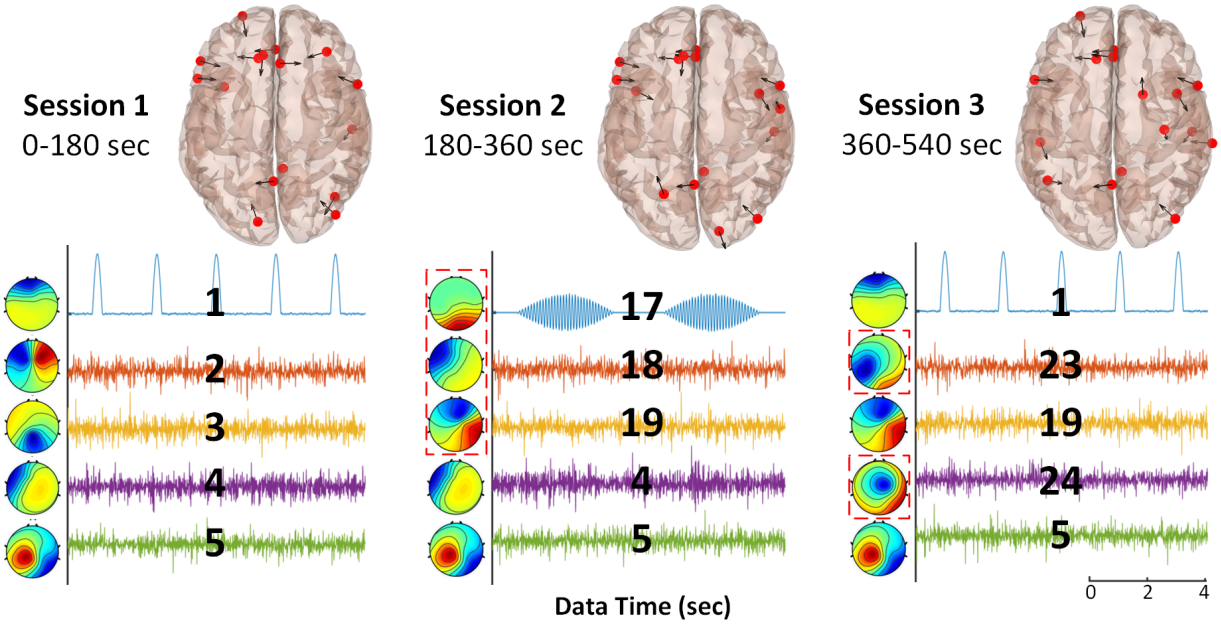

Figure 6.1: Source dipole locations and orientations (red dots and black arrows respectively in cortical head models), 10 second examples of source time-series, and component maps of select components in three sessions for the simulated 16-ch non-stationary EEG data. The red dashed boxes indicate which sources changed between sessions. Each source is labeled for comparisons with Figure 6.3.

for optimal parameters.

6.3 Experiments

6.3.1 Simulated EEG data

We simulate a realistic form of non-stationarity in the EEG data where the underlying sources in the brain are activating and deactivating alternately, using the EEG simulator in Source Information Flow Toolbox (SIFT) [29]. We first generate source time-series using a random-coefficient stationary autoregressive models with super-gaussian noises (128Hz sampling rate, nine minutes). In addition, two sources simulating eye blinks and alpha burst activities are manually designed (source #1 and #17 in Figure 6.1). Each source is assigned to the center of a randomly selected cortical region from MNI brain atlas, with its dipole orienting normal to the

cortical surface. Cortical regions for eye-blink and alpha components are manually selected to match reality (see head models in Figure 6.1).

The nine minutes of source data are partitioned into three 3-min sessions, each with a different set of 16 selected (active) sources. Between each session, six sources are changed, including alternating between the eye-blink and alpha components, while ten sources remain the same. The active sources in each session are designed such that all simulated sources (27 in total) are active in at least one session. Finally, the source-level activities are projected through a zero-noise 3-layer Boundary Element Method (BEM) forward model (MNI Colin27), yielding 16-channel EEG data.

6.3.2 Actual EEG data

Three sessions of EEG data are collected from a 26 year-old male subject using a 14-channel wearable wireless Emotiv headset. The first and third session are 3-min eye-closed music listening, and the second session is 3-min eye-open typing, with ten second transitions between sessions. Artifact Subspace Reconstruction (ASR) [83], an online-capable method to detect and repair temporal burst artifacts based on signals' covariance, is applied prior to online ICA pipeline. Our goal is to see the switching of active sources in each session.

6.3.3 Data processing

For running adaptive ORICA, both simulated and actual EEG data are processed in a simulated online environment using BCILAB [63]. The online ICA pipeline consists of an IIR high-pass filter (cutoff at 1Hz), an online RLS whitening filter, and an ORICA filter. As a comparison, we also apply a standard batch mode ICA (Infomax) available in EEGLAB [29] to both simulated and actual EEG datasets. Note that the Infomax ICA is performed on whole datasets instead of on data in each session.

6.4 Results

6.4.1 Simulated 16-ch non-stationary EEG data

Figure 6.2 demonstrates the learning and tracking performance of ORICA with three types of forgetting factors. The performance is quantified as the cross-talk error in Figure 6.2c, which is a measurement of ICA decomposition given the ground truth as defined in Equation (5) in [45]. The error zero implies perfect source separation, except for permutation and scaling ambiguities.

In the initial learning stage (0 to 3 minute), adaptive and cooling forgetting factors can achieve smaller errors comparing to the constant λ. In the tracking stage (3 to 9 minute), the non-stationarity index can successfully detect when the underlying sources change between sessions (Figure 6.2a), which leads to an increase of the adaptive λs (Figure 6.2b). This enables ORICA to relearn a new set of sources. The convergence speed after such a change is faster than that of the initial learning and reaches a comparable error level to previous sessions. Constant λ allows tracking to some extent; and while bigger λ values increase the convergence speed, they limit the quality of convergence. Cooling λs decline to values so small that tracking becomes impossible.

Figure 6.3 shows that adaptive ORICA can separate all sources in each session while Infomax ICA can only extract a subset of the sources, as it is inherently an over-complete problem: the number of distinct sources (27) across three sessions is larger than the number of channels (16). Infomax ICA successfully decomposes the simulated eye-blink component (#1) which activates in sessions 1 and 3 while losing one dimension to account for a new source in session 2. Similarly, Infomax ICA finds the occipital alpha component (#17) which only activates in session 2, at the expense of missing sources in sessions 1 and 3. Other sources activating during different sessions are presented. The errors of Infomax ICA decompositions are near -10dB, as opposed to -30dB of adaptive ORICA.

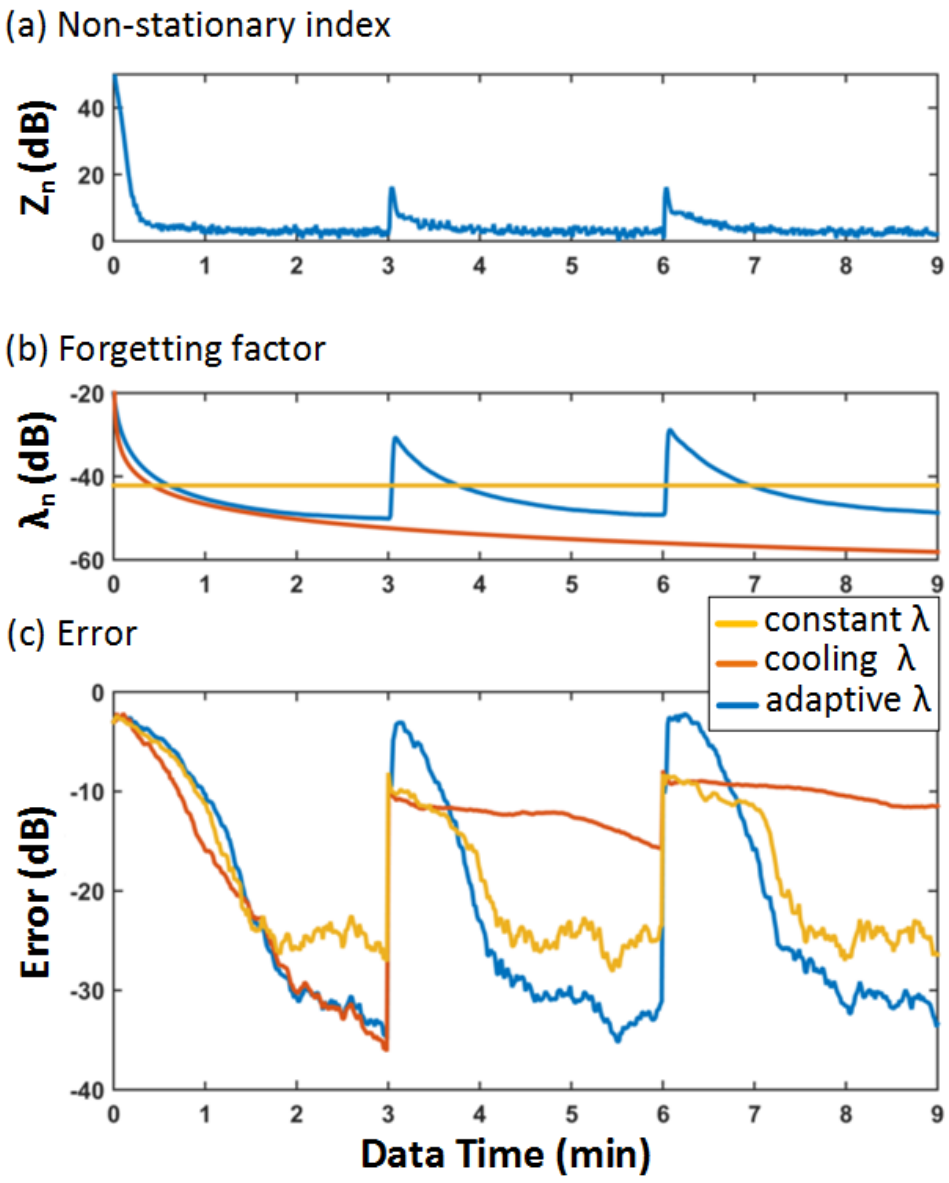

Figure 6.2: Time evolution of (a) the non-stationarity index z_n, (b) the forgetting factor λ_n, and (c) cross-talk error of ORICA decompositions with adaptive, cooling, and constant forgetting factor profiles applied to the simulated EEG data. Parameters are defined belows: for constant case, $\lambda = 0.0078$; for cooling case, $\lambda_n = 0.995/n^{0.6}$, for adaptive case, $\lambda_0 = 0.1, \alpha = 0.03, \beta = 0.012, \delta = 0.05, b = 1.5, c = 5$.

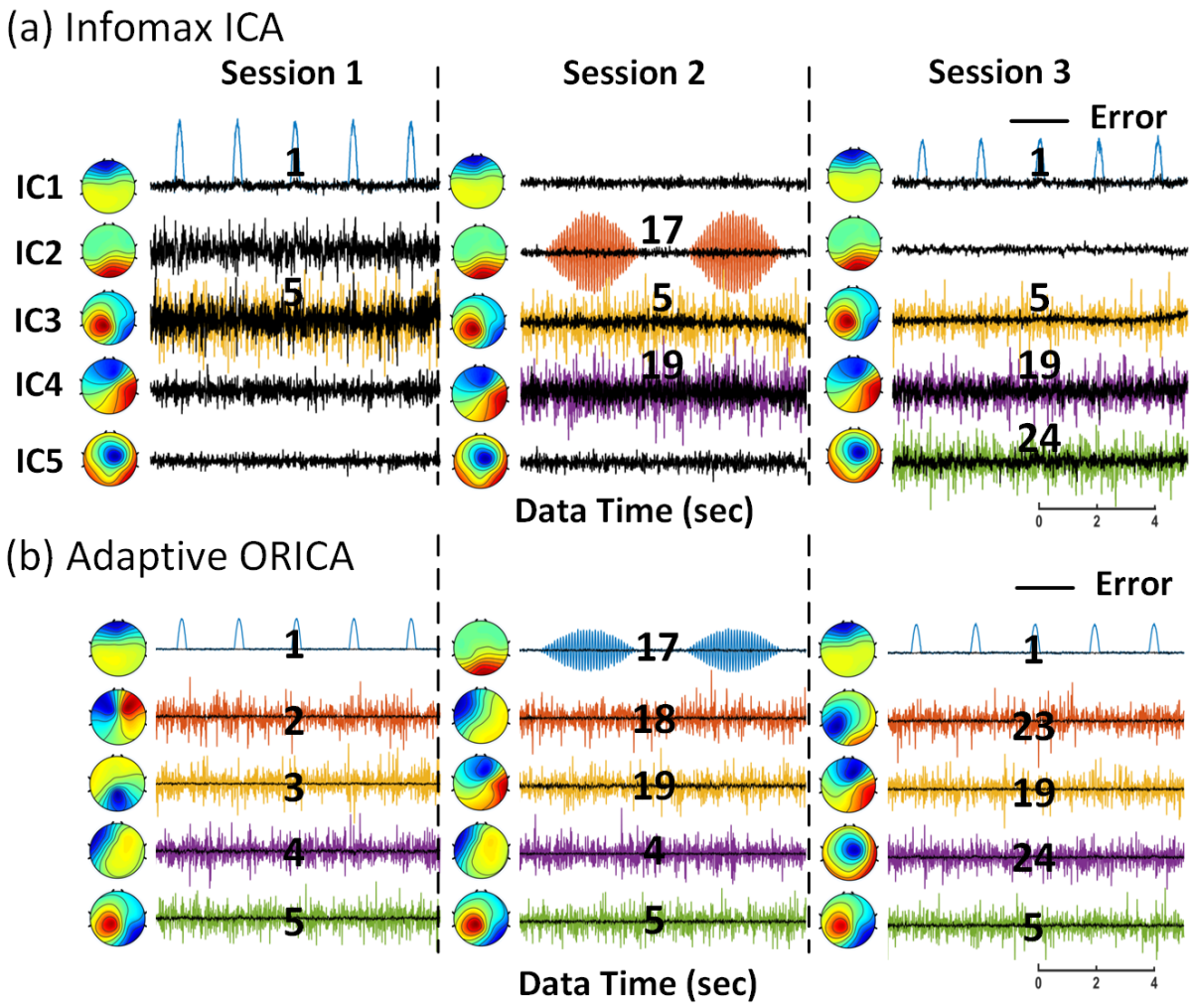

Figure 6.3: Component maps of select sources corresponding to those in Figure 6.1 (matched by correlation of component maps) decomposed and reconstructed by (a) Infomax ICA and (b) ORICA with an adaptive forgetting factor for the three sessions of simulated EEG data. 10 second windows of their reconstructed source time-series from the end of each session are presented. See Figure 6.1 for an explanation of source numbering.

6.4.2 Actual 14-ch EEG data from the task-switching experiment

Figure 6.4 demonstrates that adaptive ORICA can learn, forget, and relearn the task-specific sources in the task-switching experiment. In the first session (eye-closed), adaptive ORICA can quickly learn the alpha IC but not the eye-movement IC. In the second session (eye-open), the adaptive λ responds to the change of EEG sources at the transition between sessions (at 3 min 10 sec) and thus enables ORICA to learn the new eye-movement IC while gradually forgetting the alpha IC as it is not active anymore. Switching back to the eye-closed third session, adaptive ORICA can relearn the fading alpha IC and gradually lose the eye-movement IC. This demonstrates the applicability of adaptive ORICA to actual non-stationary EEG data. Note that some within-session non-stationarity is also present. Auditory and motor components are not clearly observed due to the limited scalp coverage of the Emotiv headset.

6.5 Conclusions

This study proposes and validates an adaptive forgetting factor update rule, which allows ORICA to adapt to changes in the underlying structure of the data. A realistically simulated EEG dataset mimicking brain sources activating and deactivating tests the performance of adaptive ORICA. Empirical results show that the adaptive rule maintains the fast and stable convergence of a cooling forgetting factor while presenting the ability to track changes when non-stationarity is detected. Standard offline ICA cannot solve the over-complete problem, while adaptive ORICA can learn the switching sources. Applied to actual EEG recorded from a task-switching experiment, we demonstrate adaptive ORICA can learn, forget, and relearn the task-related components.

Future works include (a) the development of an automatic, data-driven approach for selecting parameter values in the adaptive rule and (b) the integration of an online artifact removal method to stabilize ORICA convergence. The ability to track the changes in EEG source activities

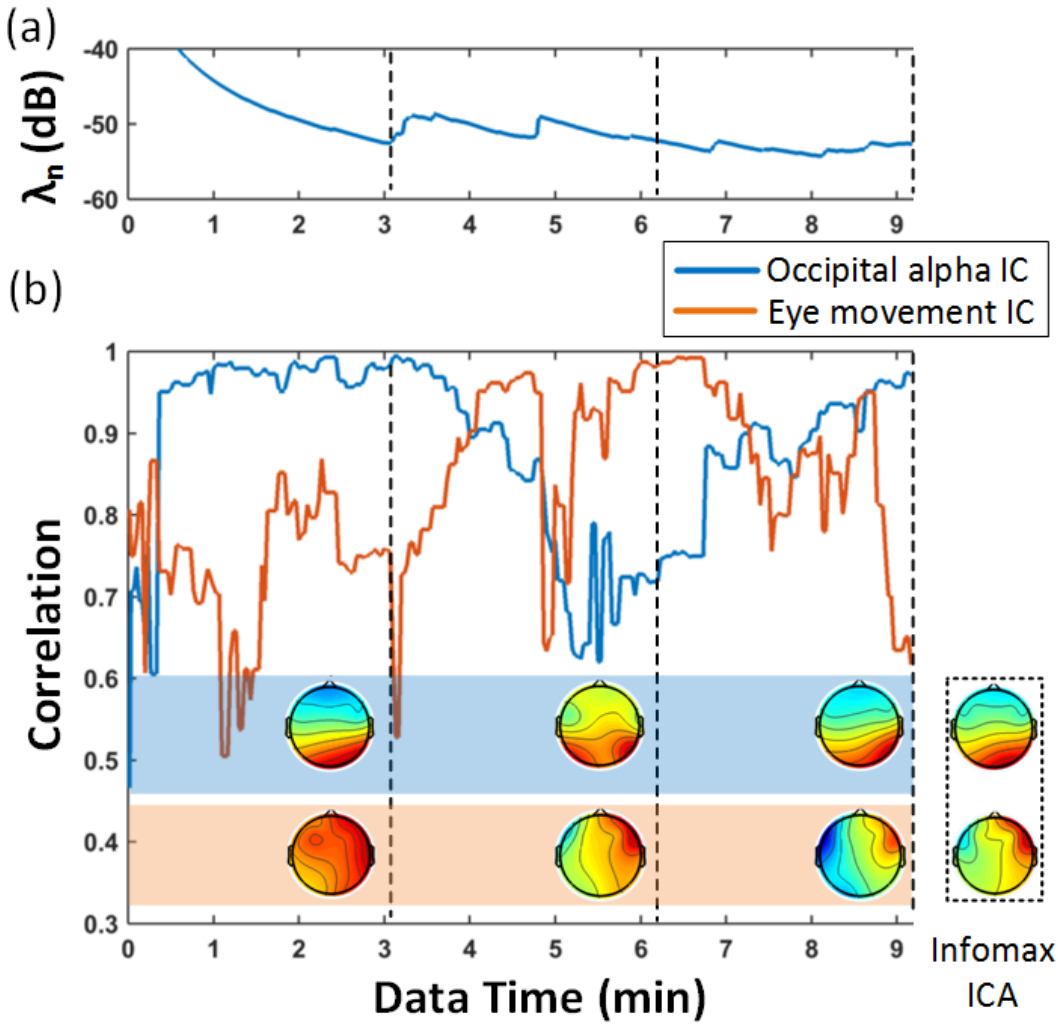

Figure 6.4: Adaptive ORICA applied to actual 14-ch EEG data from the task-switching experiment. Time evolution of (a) the adaptive forgetting factor and (b) the similarity (correlation) between two components decomposed by adaptive ORICA and Infomax ICA. The two components correspond to occipital alpha and horizontal eye movement activities. Dashed lines indicate the end of each session. The component maps solved by adaptive ORICA at the end of each session are shown to the left of the dashed lines with the component maps solved by Infomax ICA applied to the whole dataset are shown in a dashed box. Parameters used for adaptive ORICA: $\lambda_0 = 0.1, \alpha = 0.03, \beta = 0.005, \delta = 0.05, b = 0.5, c = 3$.

by the proposed method might lead to many real-world applications where modeling and tracking source dynamics are imperative.

6.6 Acknowledgments

Chapter 6, in part, is a reprint of the material as it appears in Hsu, S. H., Pion-Tonachini, L., Jung, T. P., and Cauwenberghs, G. (2015, August). "Tracking non-stationary EEG sources using adaptive online recursive independent component analysis." In *37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, (pp. 4106-4109). © 2015 IEEE. The dissertation author was the primary investigator and co-author of this paper. Hsu, S. H. and Pion-Tonachini, L. contributed equally to this work. This study was sponsored by the Army Research Laboratory under Cooperative Agreement Number W911NF-10-2-0022, by NSF EFRI-M3C 1137279, and by NSF Graduate Research Fellowship under Grant No. DGE-1144086.

Chapter 7

Real-time EEG Source-Mapping Toolbox (REST): Online ICA and Source Localization

In this chapter, we introduce the Real-Time Source Mapping Toolbox (REST), an extension to the EEGLAB environment. REST allows blind source separation of EEG data in real time using Online Recursive ICA (ORICA), plus real-time localization of separated sources. Two source localization methods are available to fit equivalent current dipoles or estimate spatial source distributions of selected sources. Two experiments demonstrate the accuracy and functionality of REST. We discuss some relevant applications.

7.1 Background

EEG source analysis combining ICA and source localization has generally been solved offline because of its computational cost. With faster processors and algorithmic advances, near real-time online applications are becoming even more viable. Bringing these analysis methods to

the domain of real-time processing would allow for the use of more specific neurophysiological information in closed-loop brain-computer interfaces (BCI) and neurofeedback paradigms, and could also provide experimenters online feedback useful for data quality control.

Analyzing EEG data at the level of cortical source dynamics is a complicated problem, but allows for much more biologically plausible, physiologically meaningful, and functionally significant results than treating scalp data channels as if they indexed single brain sources. A source-resolved imaging approach models the collected EEG as the sum of electric fields produced by many small patches of cortex whose local field activities are fully or partially synchronous, each such patch thus functioning as an effective EEG source with a scalp projection identical to that of a single equivalent current dipole (ECD). Source localization requires solutions to both the forward and inverse imaging problems: the forward problem (FP) determining the scalp projection patterns of the possible brain sources based on accurate modeling of head tissue geometries and conductivities, and the inverse problem (IP) estimating the locations and orientations or cortical surface distributions of one or more source projection patterns.

Many existing EEG processing toolboxes attempt to solve these problems, including core EEGLAB [28], BCILAB [63], LORETA-KEY [99], and Fieldtrip [92]. They all operate off-line or attempt to solve the IP by directly operating on, *e.g.* response-averaged EEG channel data. Approaching the IP directly from the EEG channel data complicates the problem by requiring determination of the number of sources to localize, a problem whose computational cost and number of false local minima increase dramatically with the number of sources being estimated. Other approaches simply attempt a low-resolution joint spatial estimate of all the active sources. Blind source separation can be used as an initial unmixing step to simplify an inverse problem by separating it into much simpler problems of finding the locations of the individual effective sources [74][75][79].

ICA has been shown to work exceedingly well when applied to EEG [30] as EEG data and ICA share many important assumptions. ICA assumes that input data is the result of a linear

mixing of spatially stationary independent time series or independent components (ICs). Here, we present the Real-Time EEG Source-Mapping Toolbox (REST), a collection of automated EEG analysis methods accessible through a graphic user interface (GUI). By applying Online Recursive ICA (ORICA) [4], we can estimate a solution to the source separation problem in near real-time, allowing low-latency access to source information, making possible innovations in experimental designs including a wide variety of clinical and non-clinical BCI paradigms. REST also allows the user to estimate the brain locations of the estimated sources using either LORETA or minimum-variance ECD fitting.

REST provides estimates of source activations and their current PSD, plus source scalp maps (source scalp projection patterns) and cortical source locations. Below, we show the layout of the REST GUI and detail the measures used in its analysis pipeline. We then test its accuracy and efficacy by applying it to simulated EEG data with known source locations and activations. Finally, we demonstrate the real-world utility of REST and its ease of use by applying it in a common BCI paradigm recording session.

7.2 REST: Architecture

REST is coded in MATLAB [39] using the EEGLAB environment. It uses a processing pipeline, shown in Figure 7.1A, designed to run from beginning to end with minimal user input. Preprocessing and source separation are implemented as a BCILAB pipeline followed by source localization implemented in part using routines in MoBILAB [88].

7.2.1 Preprocessing

The toolbox pulls EEG data from a data stream received through the Lab Streaming Layer framework [62]. The data are first preprocessed by IIR high-pass filtering. Artifact Subspace Reconstruction (ASR) [83] may be introduced as an additional preprocessing step to remove large

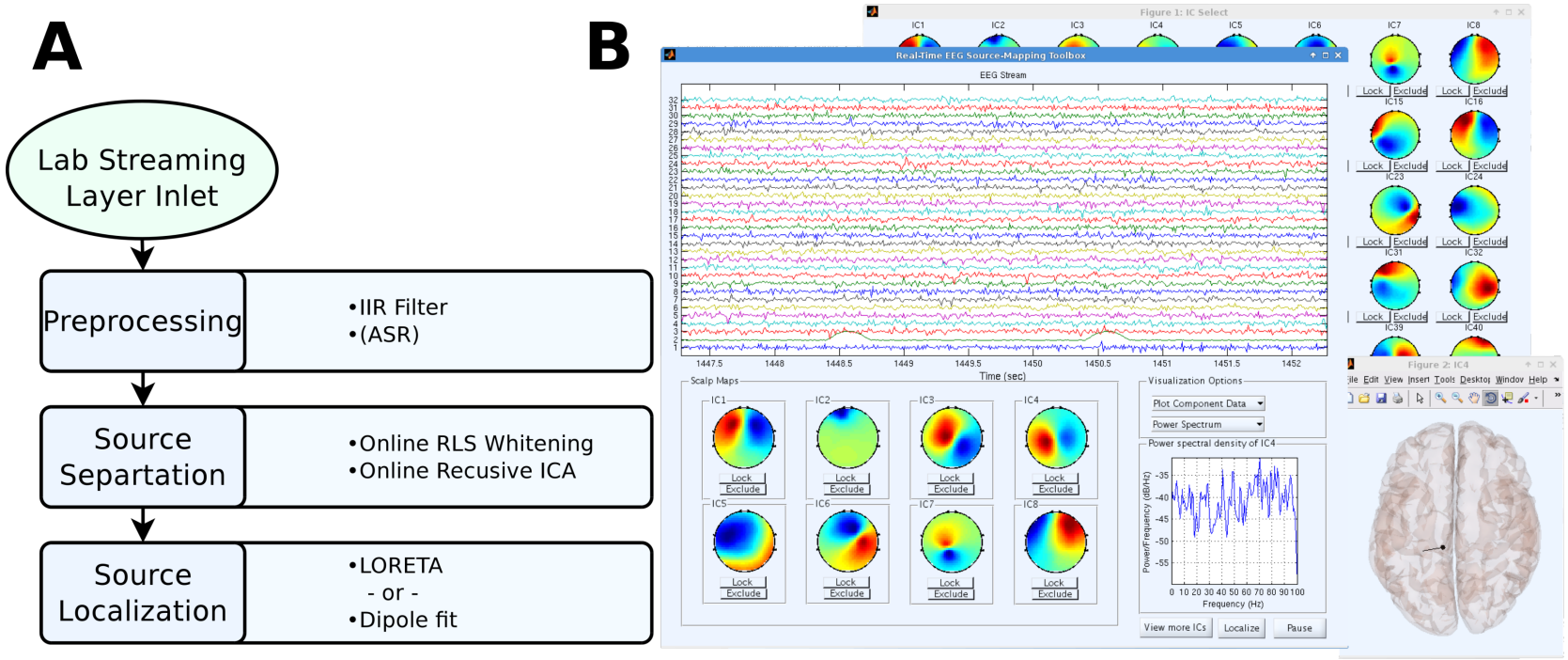

Figure 7.1: A. The pipeline used in the Real-Time EEG Source-Mapping Toolbox (REST). B. The toolbox GUI. The main window (left) shows the scrolling EEG channel or independent component (IC) activation data plus 8 (constantly updated) IC scalp maps. A source location estimate for IC4 is shown (lower right). Behind this (upper right), another REST window shows all the estimated IC scalp maps.

movement-based artifacts.

7.2.2 Source separation

Next, the EEG data are whitened using an online RLS whitening algorithm (to improve convergence) and then linearly unmixed using ORICA. ORICA is a centerpiece of this toolbox. It is, so far as we know, the only ICA implementation that is real-time capable with acceptable convergence rates for relatively large numbers of channels [45]. The output of ORICA is a set of linear IC filters that are used to separate the IC activation time courses and scalp maps from the EEG channel data. When the data sources are spatially and statistically stationary, the ICs that ORICA provide asymptotically approach those that (offline) Infomax ICA [66] returns. Unlike Infomax ICA (though compare AMICA [97]), ORICA can also adapt to source non-stationarities [4][45].

7.2.3 Source localization

Estimated IC source locations are calculated using one of two cortically-constrained source model (either distributed or ECD). Both the ECD and distributed source methods require a MoBILAB head model object to be created in advance, which can be computed easily using the included helper function. The head model uses spatial meshes representing the geometry of the cortex, scalp and one or more intervening head tissue types (*e.g.* skull, CSF, white matter). A lead field matrix (LFM) is calculated (automatically) using OpenMEEG [38], as well as a surface Laplacian operator for the cortical mesh. By default, the included helper function creates a 3-layer (scalp, skull, cortex) boundary element method (BEM) head model based on the MNI Colin 27 brain. The primary input to the source localization methods is an estimated ICA scalp map for the source being localized. Distributed source location model estimates are calculated using cortically-constrained LORETA with Bayesian hyper-parameter estimation [15] from MoBILAB,

while the ECD model estimates are computed using minimum residual variance fitting.

7.2.4 Graphical user interface

The REST main window, on the left of Figure 7.1B, displays either raw EEG or estimated IC activations. It also shows estimated scalp maps and power spectra for the estimated ICs, as well as convergence statistics. All the visualized information updates in near real-time. The (partially occluded) window in the top rights of Figure 7.1B provides an easy way to select which ICs are displayed on the main window. On the bottom right of Figure 7.1B is the source localization window which shows the current estimated source location for an IC as either an ECD or a distributed source.

7.3 Experiments

To show the utility of REST, we designed two experiments. One, using simulated source-resolved EEG data, for which we know the ground truth, tested the integrity of the REST pipeline. The other used actual EEG collected during a steady-state visually evoked potential (SSVEP) BCI paradigm to test the utility of the toolbox in interactive paradigms.

7.3.1 Simulated EEG Data

For the simulation, we used the default head model with ECD sources constrained to be normal to the cortical surface. We simulated 10 minutes of 64 channel EEG using SIFT [29] by placing ECDs at various vertices of the cortical mesh and generated source activation time-series for each. Two sources were handcrafted to imitate eye-blinks and occipital alpha while the rest were VAR processes driven by super-Gaussian noise. These are then mixed together using the LFM associated with the head model. As this was a test for accuracy rather than speed of convergence, we evaluated the accuracy of the ORICA decompositions and resulting source

location estimates at the end of the simulated data collection. For information on the convergence properties of ORICA, see [45].

7.3.2 Actual EEG Data

To collect the actual EEG we used a low-cost 14 channel Emotiv headset. This setup wirelessly streams data to a computer via Bluetooth. The streaming data were transferred to an LSL stream for REST. During the experiment, 2 minutes of eyes-closed resting allowed ORICA to identify relevant ICs. This was followed by 2 trials in which the subject (age 26) looked at flashing phone-pad style digits on a tablet. The subject first focused on the symbol "1" and then afterwards at "#" which were flashing at 9 Hz and 11.75 Hz respectively. This tested the adaptivity of the pipeline, since going from eyes-closed rest to viewing flashing stimuli could be expected to produce a noticeable change in brain sources and source activities.

7.4 Results

7.4.1 Synthetic 64-Channel Stationary EEG

As shown in Figures 7.2 and 7.3, ORICA quite accurately identifies the simulated source processes. Figure 7.2 visualizes the full REST pipeline applied to 3 of the 64 simulated sources and illustrates the accuracy of the reconstructed ORICA scalp maps as well as of their estimated dipole positions and orientations which were generally correct within 3 cm and 20 degrees respectively (Figure 7.3). The majority of the errors in dipole position were related to the depth of the dipole as the true source positions tended to be closer to the scalp than their estimates. Here we used the same simulated FP head model to solve the IP, something not possible in actual use where the true FP head model can only be estimated. Nevertheless, these results indicate that REST can generate accurate source locations and activations provided a minimum level of data

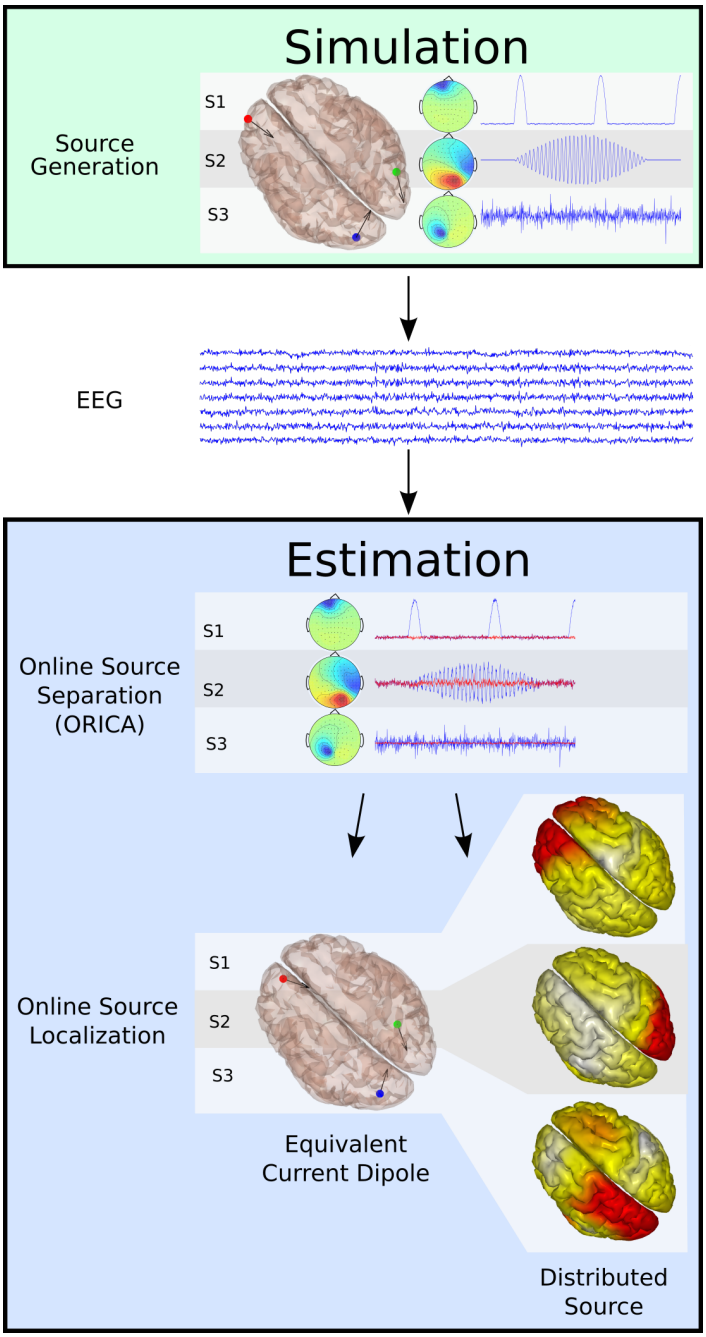

Figure 7.2: The simulated data experiment: data simulation to source estimation. (Green box) The simulated data source activations are mixed. (Blue box) The simulated EEG data are first decomposed within REST into estimated independent components (ICs) using ORICA. Then the source location of each IC is estimated as either an equivalent current dipole (ECD, left) or as a low-resolution cortical distribution (right).

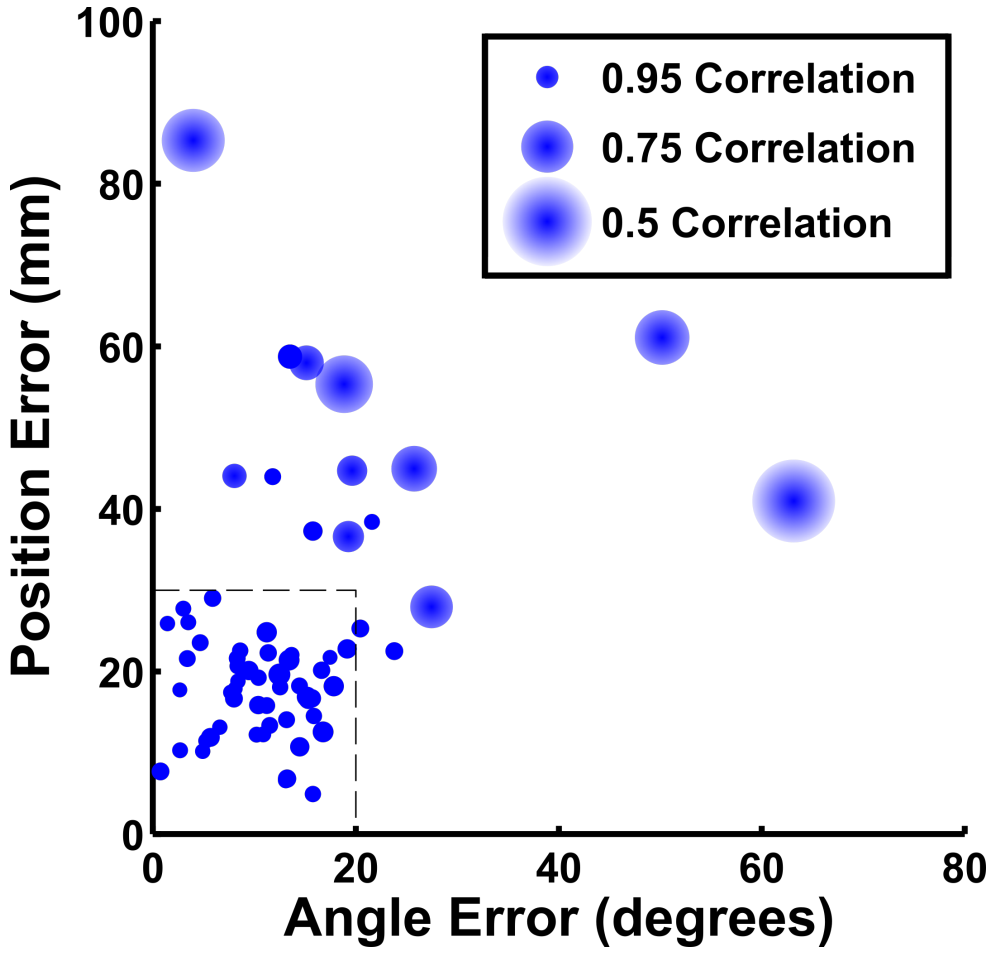

Figure 7.3: Source localization accuracy in the simulated data experiment using an equivalent current dipole (ECD) model for each estimated independent component (IC). Each disk represents an IC. Disk size shows how well the recovered IC scalp map correlated with the simulated source scalp map. For 48 of the 64 recovered sources, map correlations were above 0.95 with ECD model errors less than 3 cm and 20 degrees (lower left).

quantity and quality and sufficient head model accuracy.

7.4.2 Real 14-Channel EEG during SSVEP

The application of REST to data collected in an SSVEP paradigm showed that ORICA can converge to useful source solutions in real-life applications. Figure 7.4 compares REST outputs during eyes-closed rest and attention to 9 Hz and 11.75 Hz flashing stimuli. Clearly, ORICA extracted an occipital IC, first during rest with a weak 10 Hz peak (top panel), and then during attention to 9-Hz (middle panel) and 11.75-Hz (lower panel) flashing stimuli. The ECD during the latter condition (not shown) changed in orientation as indicated by the change in its scalp map.

7.5 Conclusion and future works

We have shown that REST can be accurate when applied to simulated data, and potentially usable in practice. There are many possible applications for real-time monitoring of sources of interest during an EEG experiment. The REST toolbox design allows possible extensions to implement near real-time computation, visualization, and application of other source-resolved EEG measures. REST could aid online data quality analysis, as when collecting EEG from particular sources if of specific importance. Additionally, the ORICA implementation in REST might be used to make a wide range of BCI paradigms more robust [76]. We plan to add more flexible and detailed data preprocessing, since ICA can be highly influenced by large amplitude artifacts, and also automated IC classification. In theory, the ORICA decomposition and, with some modifications, the source localization methods in the REST pipeline should be as applicable to MEG as to EEG data. Finally, this work follows in spirit, and some details, previous work [83] demonstrating a real-time application of the BCILAB [63] and SIFT [29] toolboxes, into which the source identification and localization methods in REST might easily be introduced.

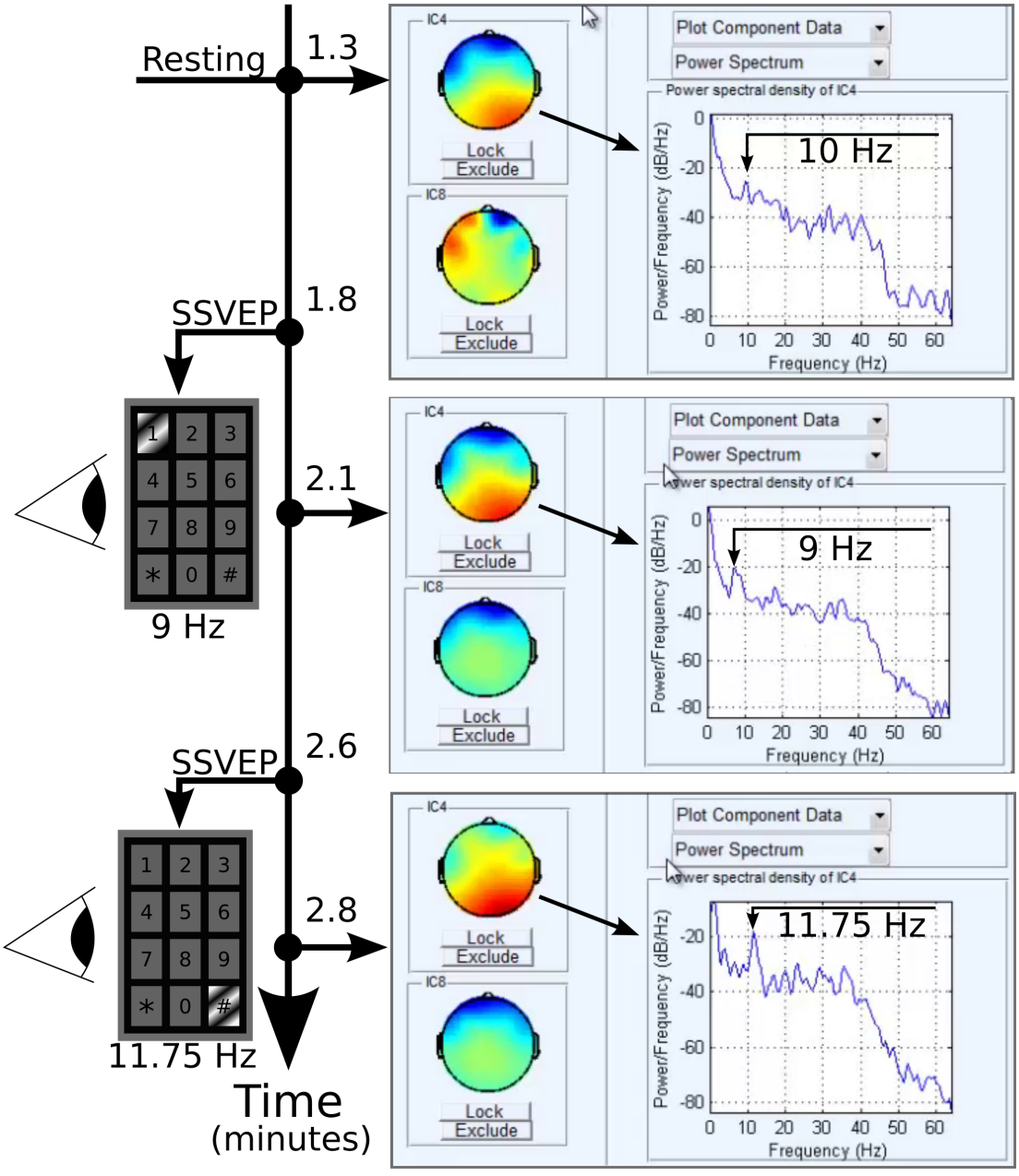

Figure 7.4: Screen captures of the REST GUI during an actual EEG experiment. At 1.3 min (top panel) during eyes-closed rest, the baseline PSD for IC had a peak at 10 Hz (alpha). At 2.1 min (middle panel), the subject attended the symbol "1" flashing at 9 Hz (note change in the IC4 spectrum). At 2.8 min (bottom panel), the subject attended the symbol "#" flashing at 11.75 Hz (note IC4 spectral shift and possible scalp map change).

7.6 Acknowledgments

Chapter 7, in part, is a reprint of the material as it appears in Pion-Tonachini, L., Hsu, S. H., Makeig, S., Jung, T. P., and Cauwenberghs, G. (2015, August). "Real-time EEG source-mapping toolbox (REST): Online ICA and source localization." In *37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, (pp. 4114-4117). © 2015 IEEE. The dissertation author was the secondary investigator and co-author of this paper. Pion-Tonachini, L. and Hsu, S. H. made equal contribution to this work. This study was sponsored by the Army Research Laboratory under Cooperative Agreement Number W911NF-10-2-0022, by NSF EFRI-M3C 1137279, and by NSF Graduate Research Fellowship under Grant No. DGE-1144086.

Chapter 8

Conclusion and Future Works

As a human brain performs various cognitive functions or undergoes different mental processes within ever-changing, real-world environments, states of the brain are inevitably nonstationary. While noninvasive, high temporal-resolution EEG recordings of brain activities have been available, there still lack computational data analysis tools for assessing the spontaneous, endogenous, nonstationary brain activities that reflect dynamics of brain states. Toward building an EEG-based brain-state monitoring system, specific challenges include the needs for: (1) computational models for effective and robust decoding of brain states from nonstationary EEG signals, (2) quantitative measures for assessing and tracking state changes, (3) data-driven approaches for addressing issues with unlabeled, continuous data, and (4) hypothesis-driven approaches for neurophysiological interpretations and scientific insights.

In this dissertation, the first need is addressed by the establishment and applications of three computational models (Table 1.1), namely the multi-model ICA (Chapter 2), the ICA mixture model (Chapter 3), and the adaptive ICA (Chapter 4), for modeling nonstationary, multichannel EEG signals at functional brain-sources level. Extensive studies of these models applied to both simulation datasets and experimental EEG datasets demonstrate successful detection of the nonstationarity and effective decoding of brain states, *e.g.* drowsiness, alertness, and sleep stages,

that are robust across subjects and experiments. To address the second need, this dissertation introduces the model-deviation index (Section 2.2.2 in Chapter 2) and non-stationarity index (Section 4.2.2 in Chapter 4) and explores the use of the model probability or data likelihood (Section 3.2.2 in Chapter 3) for quantitative assessment of brain state changes. As for the third challenge, the AMICA (Chapter 3) and ORICA (Chapter 4 and Chapter 6) approaches can be applied as data-mining tools to explore and reveal nonstationary brain dynamics from unlabeled data at milliseconds temporal resolution. Finally, these computational models inherent ICA's capability in separating non-brain artifact sources and identifying functional brain sources. With the source nonstationarity hypothesis – synchronized activities of functional brain sources under a brain state can be uniquely described by an ICA model, we demonstrate how these hypothesis-driven approaches can lead to neuroscientific insights in cognitive experiments.

The dissertation significantly contributes to the development of computational data analysis tools for decoding and monitoring rapidly-changing human cognitive and mental states, which may help advance the fields of cognitive sciences, clinical diagnosis and monitoring, neuroergonomics, and brain-computer interfaces. The ultimate goal is to make the EEG-based brain-state monitoring system capable of near real-time processing and applicable in addressing real-world problems. The Real-time EEG Source-mapping Toolbox (Chapter 7) shows our initial attempt at building such system and making it available to the scientific community and beyond. Toward this goal, remaining challenges include (real-time) separation of non-brain and brain signals in EEG to improve the robustness of brain-state decoding, automatic identification of functional sources in ICA models to enable neurophysiological interpretations and automatic labeling of brain states, and application of these tools to studying more variety of brain states (*e.g.* cognitive state changes during rest or meditation, mental strategy shifts during task performance, changes in fatigue and arousal, and changes in emotional states).

Bibliography

[1] Z. A. Acar, C. E. Acar, and S. Makeig. Simultaneous head tissue conductivity and eeg source location estimation. *NeuroImage*, 124:168–180, 2016.

[2] Z. A. Acar and S. Makeig. Neuroelectromagnetic forward head modeling toolbox. *Journal of neuroscience methods*, 190(2):258–270, 2010.

[3] H. Akaike. A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6):716–723, 1974.

[4] M. T. Akhtar, T.-P. Jung, S. Makeig, and G. Cauwenberghs. Recursive independent component analysis for online blind source separation. In *Proc. IEEE-ISCAS*, pages 2813–2816, 2012.

[5] S.-I. Amari. Natural gradient works efficiently in learning. *Neural computation*, 10(2):251–276, 1998.

[6] S.-I. Amari, A. Cichocki, H. H. Yang, et al. A new learning algorithm for blind signal separation. *Advances in neural information processing systems*, pages 757–763, 1996.

[7] J. S. Barlow. Methods of analysis of nonstationary eegs, with emphasis on segmentation techniques: a comparative review. *Journal of Clinical Neurophysiology*, 2(3):267–304, 1985.

[8] A. J. Bell and T. J. Sejnowski. An information-maximization approach to blind separation and blind deconvolution. *Neural computation*, 7(6):1129–1159, 1995.

[9] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, and E. Moulines. A blind source separation technique using second-order statistics. *IEEE Transactions on signal processing*, 45(2):434–444, 1997.

[10] R. F. Betzel, M. A. Erickson, M. Abell, B. F. O'Donnell, W. P. Hetrick, and O. Sporns. Synchronization dynamics and evidence for a repertoire of network states in resting eeg. *Frontiers in computational neuroscience*, 6, 2012.

[11] N. Bigdely-Shamlo, T. Mullen, C. Kothe, K.-M. Su, and K. A. Robbins. The prep pipeline: standardized preprocessing for large-scale eeg analysis. *Frontiers in neuroinformatics*, 9, 2015.

[12] J.-F. Cardoso and B. H. Laheld. Equivariant adaptive source separation. *IEEE Trans. Signal Processing*, 44(12):3017–3030, 1996.

[13] G. Carpaneto and P. Toth. Algorithm 548: Solution of the assignment problem [h]. *ACM Transactions on Mathematical Software (TOMS)*, 6(1):104–111, 1980.

[14] M. A. Carskadon, W. C. Dement, et al. Normal human sleep: an overview. *Principles and practice of sleep medicine*, 4:13–23, 2005.

[15] K. Chan, T.-W. Lee, and T. J. Sejnowski. Variational learning of clusters of undercomplete nonsymmetric independent components. *Journal of Machine Learning Research*, 3(Aug):99–114, 2002.

[16] S. Charbonnier, R. N. Roy, S. Bonnet, and A. Campagne. Eeg index for control operators mental fatigue monitoring using interactions between brain regions. *Expert Systems with Applications*, 52:91–98, 2016.

[17] J. Chen and X. Huo. Theoretical results on sparse representations of multiple-measurement vectors. *IEEE Trans. on Signal Processing*, 54(12):4634–4643, 2006.

[18] C. J. Chu, M. A. Kramer, J. Pathmanathan, M. T. Bianchi, M. B. Westover, L. Wizon, and S. S. Cash. Emergence of stable functional networks in long-term human electroencephalography. *Journal of Neuroscience*, 32(8):2703–2713, 2012.

[19] C.-H. Chuang, L.-W. Ko, Y.-P. Lin, T.-P. Jung, and C.-T. Lin. Independent component ensemble of eeg for brain–computer interface. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 22(2):230–238, 2014.

[20] A. Cichocki and S.-i. Amari. *Adaptive blind signal and image processing: learning algorithms and applications*, volume 1. John Wiley & Sons, 2002.

[21] A. Cichocki, S.-i. Amari, K. Siwek, T. Tanaka, A. H. Phan, R. Zdunek, S. Cruces, P. Georgiev, Y. Washizawa, Z. Leonowicz, et al. Icalab toolboxes, 2002.

[22] A. Cichocki and R. Thawonmas. On-line algorithm for blind signal extraction of arbitrarily distributed, but temporally correlated sources using second order statistics. *Neural Processing Letters*, 12(1):91–98, 2000.

[23] A. Cichocki and R. Unbehauen. Robust neural networks with on-line learning for blind identification and blind separation of sources. *IEEE Trans. on Circuits and Systems I: Fundamental Theory and Applications,*, 43(11):894–906, 1996.

[24] M. Congedo, C. Gouy-Pailler, and C. Jutten. On the blind source separation of human electroencephalogram by approximate joint diagonalization of second order statistics. *Clinical Neurophysiology*, 119(12):2677–2686, 2008.

[25] S. Cruces and A. Cichocki. Combining blind source extraction with joint approximate diagonalization: Thin algorithms for ica. In *Proceedings of 4th International Symposium on Independent Component Analysis and Blind Signal Separation (ICA2003)*, pages 463–468, 2003.

[26] S. Cruces-Alvarez, A. Cichocki, and L. Castedo-Ribas. An iterative inversion approach to blind source separation. *IEEE Trans. Neural Networks*, 11(6):1423–1437, 2000.

[27] L. De Lathauwer, B. De Moor, and J. Vandewalle. Fetal electrocardiogram extraction by blind source subspace separation. *IEEE Trans. Biomedical Engineering*, 47(5):567–572, 2000.

[28] A. Delorme and S. Makeig. Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. *Journal of neuroscience methods*, 134(1):9–21, 2004.

[29] A. Delorme, T. Mullen, C. Kothe, Z. A. Acar, N. Bigdely-Shamlo, A. Vankov, and S. Makeig. Eeglab, sift, nft, bcilab, and erica: new tools for advanced eeg processing. *Computational intelligence and neuroscience*, 2011:10, 2011.

[30] A. Delorme, J. Palmer, J. Onton, R. Oostenveld, and S. Makeig. Independent eeg sources are dipolar. *PloS one*, 7(2):e30135, 2012.

[31] W. Dement and N. Kleitman. Cyclic variations in eeg during sleep and their relation to eye movements, body motility, and dreaming. *Electroencephalography and clinical neurophysiology*, 9(4):673–690, 1957.

[32] S. C. Douglas. Blind signal separation and blind deconvolution. *Handbook of neural network signal processing*, 6:1–28, 2001.

[33] M. Falkenstein, J. Hoormann, S. Christ, and J. Hohnsbein. Erp components on reaction errors and their functional significance: a tutorial. *Biological psychology*, 51(2):87–107, 2000.

[34] F. Freyer, K. Aquino, P. A. Robinson, P. Ritter, and M. Breakspear. Bistability and non-gaussian fluctuations in spontaneous cortical activity. *Journal of Neuroscience*, 29(26):8512–8524, 2009.

[35] A. Gevins and M. E. Smith. Electroencephalography (eeg) in neuroergonomics. *Neuroergonomics: The brain at work*, pages 15–31, 2006.

[36] X. Giannakopoulos, J. Karhunen, and E. Oja. An experimental comparison of neural algorithms for independent component analysis and blind separation. *International Journal of Neural Systems*, 9(02):99–114, 1999.

[37] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley. Physiobank, physiotoolkit, and physionet. *Circulation*, 101(23), 2000.

[38] A. Gramfort, T. Papadopoulo, E. Olivi, M. Clerc, et al. Openmeeg: opensource software for quasistatic bioelectromagnetics. *Biomedical engineering online*, 9(1):45, 2010.

[39] M. U. Guide. The mathworks. *Inc., Natick, MA*, 5:333, 1998.

[40] J. T. Gwin and D. P. Ferris. An eeg-based study of discrete isometric and isotonic human lower limb muscle contractions. *Journal of neuroengineering and rehabilitation*, 9(1):35, 2012.

[41] S. Ha, C. Kim, Y. Chi, A. Akinin, C. Maier, A. Ueno, and G. Cauwenberghs. Integrated circuits and electrode interfaces for noninvasive physiological monitoring. *IEEE Trans. Biomedical Engineering*, 61(5):1522–1537, 2014.

[42] S. Haufe, R. Tomioka, G. Nolte, K.-R. Muller, and M. Kawanabe. Modeling sparse connectivity between underlying brain sources for eeg/meg. *IEEE Trans. Biomedical Engineering*, 57(8):1954–1963, Aug 2010.

[43] S.-H. Hsu and T.-P. Jung. Modeling and tracking brain nonstationarity in a sustained attention task. In *International Conference on Augmented Cognition*, pages 209–217. Springer, 2016.

[44] S.-H. Hsu and T.-P. Jung. Monitoring alert and drowsy states by modeling eeg source nonstationarity. *Journal of Neural Engineering*, 14(5):056012, 2017.

[45] S.-H. Hsu, T. Mullen, T.-P. Jung, and G. Cauwenberghs. Online recursive independent component analysis for real-time source separation of high-density eeg. In *Proc. IEEE-EMBC*, pages 3845–3848, 2014.

[46] S.-H. Hsu, T. Mullen, T.-P. Jung, and G. Cauwenberghs. Validating online recursive independent component analysis on eeg data. In *Proc. IEEE EMBS Conf. Neural Engineering*, pages 918–921, 2015.

[47] S.-H. Hsu, T. R. Mullen, T.-P. Jung, and G. Cauwenberghs. Real-time adaptive eeg source separation using online recursive independent component analysis. *IEEE transactions on neural systems and rehabilitation engineering*, 24(3):309–319, 2016.

[48] S.-H. Hsu, L. Pion-Tonachini, T.-P. Jung, and G. Cauwenberghs. Tracking non-stationary eeg sources using adaptive online recursive independent component analysis. In *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*, pages 4106–4109. IEEE, 2015.

[49] R.-S. Huang, T.-P. Jung, and S. Makeig. Tonic changes in eeg power spectra during simulated driving. *Foundations of augmented cognition. neuroergonomics and operational neuroscience*, pages 394–403, 2009.

[50] A. Hyvärinen, J. Karhunen, and E. Oja. *Independent component analysis*, volume 46. John Wiley & Sons, 2004.

[51] A. Hyvärinen and E. Oja. Independent component analysis: algorithms and applications. *Neural networks*, 13(4):411–430, 2000.

[52] J. R. Iversen and S. Makeig. Meg/eeg data analysis using eeglab. In *Magnetoencephalography*, pages 199–212. Springer, 2014.

[53] K. Jones and Y. Harrison. Frontal lobe function, sleep loss and fragmented sleep. *Sleep medicine reviews*, 5(6):463–475, 2001.

[54] T.-P. Jung, S. Makeig, C. Humphries, T.-W. Lee, M. J. Mckeown, V. Iragui, and T. J. Sejnowski. Removing electroencephalographic artifacts by blind source separation. *Psychophysiology*, 37(2):163–178, 2000.

[55] T.-P. Jung, S. Makeig, T.-W. Lee, M. J. McKeown, G. Brown, A. J. Bell, and T. J. Sejnowski. Independent component analysis of biomedical signals. In *Proc. Int. Workshop on Independent Component Analysis and Signal Separation*, pages 633–644. Citeseer, 2000.

[56] T.-P. Jung, S. Makeig, M. Stensmo, and T. J. Sejnowski. Estimating alertness from the eeg power spectrum. *IEEE transactions on biomedical engineering*, 44(1):60–69, 1997.

[57] C. Jutten and J. Herault. Blind separation of sources, part i: An adaptive algorithm based on neuromimetic architecture. *Signal processing*, 24(1):1–10, 1991.

[58] A. Kaplan, J. Röschke, B. Darkhovsky, and J. Fell. Macrostructural eeg characterization based on nonparametric change point segmentation: application to sleep analysis. *Journal of neuroscience methods*, 106(1):81–90, 2001.

[59] A. Y. Kaplan, A. A. Fingelkurts, A. A. Fingelkurts, S. V. Borisov, and B. S. Darkhovsky. Nonstationary nature of the brain activity as revealed by eeg/meg: methodological, practical and conceptual challenges. *Signal processing*, 85(11):2190–2212, 2005.

[60] J. Karhunen and P. Pajunen. Blind source separation and tracking using nonlinear pca criterion: A least-squares approach. In *Proc. IEEE International Conf. on Neural Networks*, volume 4, pages 2147–2152, 1997.

[61] A. Khanna, A. Pascual-Leone, C. M. Michel, and F. Farzan. Microstates in resting-state eeg: current status and future directions. *Neuroscience & Biobehavioral Reviews*, 49:105–113, 2015.

[62] C. Kothe. Lab streaming layer (lsl), 2013.

[63] C. A. Kothe and S. Makeig. Bcilab: a platform for brain–computer interface development. *Journal of neural engineering*, 10(5):056014, 2013.

[64] S. Kullback and R. A. Leibler. On information and sufficiency. *Ann. Math. Statist.*, 22(1):79–86, 03 1951.

[65] S. K. Lal and A. Craig. Driver fatigue: electroencephalography and psychological assessment. *Psychophysiology*, 39(3):313–321, 2002.

[66] T.-W. Lee, M. Girolami, and T. J. Sejnowski. Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. *Neural computation*, 11(2):417–441, 1999.

[67] T.-W. Lee, M. Lewicki, and T. Sejnowski. Ica mixture models for unsupervised classification of non-gaussian classes and automatic context switching in blind signal separation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(10):1078–1089, 2000.

[68] T.-W. Lee, M. S. Lewicki, and T. J. Sejnowski. Ica mixture models for unsupervised classification of non-gaussian classes and automatic context switching in blind signal separation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(10):1078–1089, 2000.

[69] D. Lehmann, P. Grass, and B. Meier. Spontaneous conscious covert cognition states and brain electric spectral states in canonical correlations. *International Journal of Psychophysiology*, 19(1):41–52, 1995.

[70] D. Lehmann, H. Ozaki, and I. Pal. Eeg alpha map series: brain micro-states by space-oriented adaptive segmentation. *Electroencephalography and clinical neurophysiology*, 67(3):271–288, 1987.

[71] C.-T. Lin, W.-C. Cheng, and S.-F. Liang. An on-line ica-mixture-model-based self-constructing fuzzy neural network. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 52(1):207–221, 2005.

[72] C.-T. Lin, K.-C. Huang, C.-F. Chao, J.-A. Chen, T.-W. Chiu, L.-W. Ko, and T.-P. Jung. Tonic and phasic eeg and behavioral changes induced by arousing feedback. *NeuroImage*, 52(2):633–642, 2010.

[73] C.-T. Lin, R.-C. Wu, S.-F. Liang, W.-H. Chao, Y.-J. Chen, and T.-P. Jung. Eeg-based drowsiness estimation for safety driving using independent component analysis. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 52(12):2726–2738, 2005.

[74] S. Makeig, A. J. Bell, T.-P. Jung, T. J. Sejnowski, et al. Independent component analysis of electroencephalographic data. *Advances in neural information processing systems*, pages 145–151, 1996.

[75] S. Makeig, S. Debener, J. Onton, and A. Delorme. Mining event-related brain dynamics. *Trends in cognitive sciences*, 8(5):204–210, 2004.

[76] S. Makeig, S. Enghoff, T.-P. Jung, and T. J. Sejnowski. A natural basis for efficient brain-actuated control. *Rehabilitation Engineering, IEEE Transactions on*, 8(2):208–211, 2000.

[77] S. Makeig and M. Inlow. Lapse in alertness: coherence of fluctuations in performance and eeg spectrum. *Electroencephalography and clinical neurophysiology*, 86(1):23–35, 1993.

[78] S. Makeig, M. Westerfield, T.-P. Jung, S. Enghoff, J. Townsend, E. Courchesne, and T. Sejnowski. Dynamic brain sources of visual evoked responses. *Science*, 295(5555):690–694, 2002.

[79] J. Marco-Pallares, C. Grau, and G. Ruffini. Combined ica-loreta analysis of mismatch negativity. *Neuroimage*, 25(2):471–477, 2005.

[80] M. McKeown, C. Humphries, P. Achermann, A. Borbély, and T. Sejnowsk. A new method for detecting state changes in the eeg: exploratory application to sleep data. *Journal of sleep research*, 7(S1):48–56, 1998.

[81] G. McLoughlin, B. Albrecht, T. Banaschewski, A. Rothenberger, D. Brandeis, P. Asherson, and J. Kuntsi. Performance monitoring is altered in adult adhd: a familial event-related potential investigation. *Neuropsychologia*, 47(14):3134–3142, 2009.

[82] T. Mullen, C. Kothe, M. Chi, A. Ojeda, T. Kerth, S. Makeig, T.-P. Jung, and G. Cauwenberghs. Real-time neuroimaging and cognitive monitoring using wearable dry eeg. *IEEE Transactions on Biomedical Engineering*, 62(11):1–1, 2015.

[83] T. Mullen, C. Kothe, Y. M. Chi, A. Ojeda, T. Kerth, S. Makeig, G. Cauwenberghs, and T.-P. Jung. Real-time modeling and 3d visualization of source dynamics and connectivity using wearable eeg. In *Proc. IEEE-EMBC*, pages 2184–2187, 2013.

[84] T. Mullen, G. Worrell, and S. Makeig. Multivariate principal oscillation pattern analysis of ica sources during seizure. In *Proc. IEEE-EMBC*, pages 2921–2924, 2012.

[85] T. R. Mullen. *The Dynamic Brain: Modeling Neural Dynamics and Interactions From Human Electrophysiological Recordings*. University of California, San Diego, 2014.

[86] N. Murata, M. Kawanabe, A. Ziehe, K.-R. Müller, and S.-I. Amari. On-line learning in changing environments with applications in supervised and unsupervised learning. *Neural Networks*, 15(4):743–760, 2002.

[87] K. Natarajan, R. Acharya, F. Alias, T. Tiboleng, and S. K. Puthusserypady. Nonlinear analysis of eeg signals at different mental states. *BioMedical Engineering OnLine*, 3(1):7, 2004.

[88] A. Ojeda, N. Bigdely-Shamlo, and S. Makeig. Mobilab: an open source toolbox for analysis and visualization of mobile brain/body imaging data. *Frontiers in human neuroscience*, 8, 2014.

[89] J. Onton, A. Delorme, and S. Makeig. Frontal midline eeg dynamics during working memory. *Neuroimage*, 27(2):341–356, 2005.

[90] J. Onton and S. Makeig. Information-based modeling of event-related brain dynamics. *Progress in brain research*, 159:99–120, 2006.

[91] J. Onton, M. Westerfield, J. Townsend, and S. Makeig. Imaging human eeg dynamics using independent component analysis. *Neuroscience & Biobehavioral Reviews*, 30(6):808–822, 2006.

[92] R. Oostenveld, P. Fries, E. Maris, and J.-M. Schoffelen. Fieldtrip: open source software for advanced analysis of meg, eeg, and invasive electrophysiological data. *Computational intelligence and neuroscience*, 2011, 2010.

[93] N. R. Pal, C.-Y. Chuang, L.-W. Ko, C.-F. Chao, T.-P. Jung, S.-F. Liang, and C.-T. Lin. Eeg-based subject-and session-independent drowsiness detection: an unsupervised approach. *EURASIP Journal on Advances in Signal Processing*, 2008:192, 2008.

[94] C. Paleologu, J. Benesty, and S. Ciochina. A robust variable forgetting factor recursive least-squares algorithm for system identification. *IEEE Signal Processing Letters*, 15:597–600, 2008.

[95] J. A. Palmer, K. Kreutz-Delgado, S. Makeig, et al. Super-gaussian mixture source model for ica. In *ICA*, pages 854–861. Springer, 2006.

[96] J. A. Palmer, K. Kreutz-Delgado, B. D. Rao, and S. Makeig. Modeling and estimation of dependent subspaces with non-radially symmetric and skewed densities. In *Independent Component Analysis and Signal Separation*, pages 97–104. Springer, 2007.

[97] J. A. Palmer, S. Makeig, K. Kreutz-Delgado, and B. D. Rao. Newton method for the ica mixture model. In *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, pages 1805–1808. IEEE, 2008.

[98] P. Parikh and E. Micheli-Tzanakou. Detecting drowsiness while driving using wavelet transform. In *Bioengineering Conference, 2004. Proceedings of the IEEE 30th Annual Northeast*, pages 79–80. IEEE, 2004.

[99] R. D. Pascual-Marqui, D. Lehmann, T. Koenig, K. Kochi, M. C. Merlo, D. Hell, and M. Koukkou. Low resolution brain electromagnetic tomography (loreta) functional imaging in acute, neuroleptic-naive, first-episode, productive schizophrenia. *Psychiatry Research: Neuroimaging*, 90(3):169–179, 1999.

[100] M. T. Peiris, R. D. Jones, P. R. Davidson, G. J. Carroll, and P. J. Bones. Frequent lapses of responsiveness during an extended visuomotor tracking task in non-sleep-deprived subjects. *Journal of sleep research*, 15(3):291–300, 2006.

[101] L. Pion-Tonachini, S.-H. Hsu, S. Makeig, T.-P. Jung, and G. Cauwenberghs. Real-time eeg source-mapping toolbox (rest): Online ica and source localization. In *Proc. IEEE-EMBC*, pages 4114–4117, 2015.

[102] S. J. Roberts and W. D. Penny. Mixtures of independent component analysers. In *International Conference on Artificial Neural Networks*, pages 527–534. Springer, 2001.

[103] R. N. Roy, S. Bonnet, S. Charbonnier, and A. Campagne. Mental fatigue and working memory load estimation: interaction and implications for eeg-based passive bci. In *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*, pages 6607–6610. IEEE, 2013.

[104] R. N. Roy, S. Charbonnier, A. Campagne, and S. Bonnet. Efficient mental workload estimation using task-independent eeg features. *Journal of neural engineering*, 13(2):026019, 2016.

[105] G. Safont, A. Salazar, L. Vergara, E. Gomez, and V. Villanueva. Probabilistic distance for mixtures of independent component analyzers. *IEEE Transactions on Neural Networks and Learning Systems*, pages 1–13, 2017.

[106] A. Salazar, L. Vergara, and R. Miralles. On including sequential dependence in ica mixture models. *Signal Processing*, 90(7):2314–2318, jul 2010.

[107] A. Salazar, L. Vergara, A. Serrano, and J. Igual. A general procedure for learning mixtures of independent component analyzers. *Pattern Recognition*, 43(1):69–85, jan 2010.

[108] J. Santamaria and K. H. Chiappa. The eeg of drowsiness in normal adults. *Journal of clinical Neurophysiology*, 4(4):327–382, 1987.

[109] M. Sarter, B. Givens, and J. P. Bruno. The cognitive neuroscience of sustained attention: where top-down meets bottom-up. *Brain research reviews*, 35(2):146–160, 2001.

[110] G. Schwarz et al. Estimating the dimension of a model. *The annals of statistics*, 6(2):461–464, 1978.

[111] L.-C. Shi and B.-L. Lu. Eeg-based vigilance estimation using extreme learning machines. *Neurocomputing*, 102:135–143, 2013.

[112] M. G. Terzano, L. Parrino, A. Smerieri, R. Chervin, S. Chokroverty, C. Guilleminault, M. Hirshkowitz, M. Mahowald, H. Moldofsky, A. Rosa, R. Thomas, and A. Walters. Atlas, rules, and recording techniques for the scoring of cyclic alternating pattern (cap) in human sleep. *Sleep medicine*, 3(2):187–99, mar 2002.

[113] N. V. Thakor and S. Tong. Advances in quantitative electroencephalogram analysis methods. *Annu. Rev. Biomed. Eng.*, 6:453–495, 2004.

[114] N. J. Trujillo-Barreto, E. Aubert-Vázquez, and P. A. Valdés-Sosa. Bayesian model averaging in eeg/meg imaging. *NeuroImage*, 21(4):1300–1319, 2004.

[115] D. Van de Ville, J. Britz, and C. M. Michel. Eeg microstate sequences in healthy humans at rest reveal scale-free dynamics. *Proceedings of the National Academy of Sciences*, 107(42):18179–18184, 2010.

[116] F. C. Viola, J. Thorne, B. Edmonds, T. Schneider, T. Eichele, and S. Debener. Semi-automatic identification of independent components representing eeg artifact. *Clinical Neurophysiology*, 120(5):868–877, 2009.

[117] P. von Bunau, F. C. Meinecke, S. Scholler, and K. Muller. Finding stationary brain sources in eeg data. In *IEEE EMBS*, pages 2810–2813, 2010.

[118] Y. Wang and T.-P. Jung. Improving brain–computer interfaces using independent component analysis. In *Towards Practical Brain-Computer Interfaces*, pages 67–83. Springer, 2013.

[119] C.-S. Wei, Y.-P. Lin, Y.-T. Wang, C.-T. Lin, and T.-P. Jung. Transfer learning with large-scale data in brain-computer interfaces. In *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the*, pages 4666–4669. IEEE, 2016.

[120] D. Whitmer, C. De Solages, B. C. Hill, H. Yu, J. M. Henderson, and H. Bronte-Stewart. High frequency deep brain stimulation attenuates subthalamic and cortical rhythms in parkinson's disease. *Frontiers in human neuroscience*, 6:155, 2012.

[121] H. H. Yang and S.-I. Amari. Adaptive online learning algorithms for blind separation: maximum entropy and minimum mutual information. *Neural computation*, 9(7):1457–1482, 1997.

[122] M. V. Yeo, X. Li, K. Shen, and E. P. Wilder-Smith. Can svm be used for automatic eeg detection of drowsiness during car driving? *Safety Science*, 47(1):115–124, 2009.

[123] G. Yin. *Adaptive filtering with averaging*. Springer, 1995.

[124] T. O. Zander and C. Kothe. Towards passive brain–computer interfaces: applying brain–computer interface technology to human–machine systems in general. *Journal of neural engineering*, 8(2):025005, 2011.

[125] T. O. Zander, L. R. Krol, N. P. Birbaumer, and K. Gramann. Neuroadaptive technology enables implicit cursor control based on medial prefrontal cortex activity. *Proceedings of the National Academy of Sciences*, page 201605155, 2016.

[126] X. Zhu, X. Zhang, and J. Ye. Natural gradient-based recursive least-squares algorithm for adaptive blind source separation. *Science in China Series F: Information Sciences*, 47(1):55–65, 2004.

[127] X.-L. Zhu and X.-D. Zhang. Adaptive rls algorithm for blind source separation using a natural gradient. *IEEE Signal Processing Letters*, 9(12):432–435, 2002.