

UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Delay Differential Analysis of Neural Data

Permalink

<https://escholarship.org/uc/item/91x6s3nc>

Author

Sampson, Aaron Lee Kasey

Publication Date

2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

Delay Differential Analysis of Neural Data

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

in

Neurosciences with a Specialization in Computational Neurosciences

by

Aaron Lee Kasey Sampson

Committee in charge:

Terrence J. Sejnowski, Chair
Henry D. I. Abarbanel
Eric Halgren
Claudia Lainscsek
Gregory A. Light
John Reynolds

2019

Copyright
Aaron Lee Kasey Sampson, 2019
All rights reserved.

The dissertation of Aaron Lee Kasey Sampson is approved,
and it is acceptable in quality and form for publication on
microfilm and electronically:

Chair

University of California San Diego

2019

DEDICATION

This dissertation is dedicated to my wife, Dr. Yvette Joon Ying Boon. She inspires and motivates me every day, and anything I have been able to accomplish is thanks to her constant love and support.

TABLE OF CONTENTS

Signature Page	iii
Dedication	iv
Table of Contents	v
List of Figures	viii
List of Tables	x
Acknowledgements	xi
Vita	xiv
Abstract of the Dissertation	xvii
Chapter 1	
Introduction	1
1.1 The Brain as a Nonlinear Dynamical System	1
1.2 Embeddings	2
1.2.1 A Simple Embedding Example	3
1.3 Delay Differential Analysis (DDA)	5
1.4 Estimating the Time Derivative	7
1.5 Delays and Sampling Rate	10
1.6 Structure Selection of the Model	11
1.7 DDA of Dynamical Systems	13
1.7.1 A Nine-Dimensional Lorenz System	14
1.7.2 The Lorenz-95 System	15
1.8 Cross-Dynamical Delay Differential Analysis (CD-DDA)	18
Chapter 2	
Nonlinear Dynamics Underlying Sensory Processing Dysfunction in Schizophrenia	24
2.1 Introduction	25
2.2 Delay Differential Analysis (DDA)	28
2.3 Results	30
2.3.1 Auditory Deviance Response Complex	30
2.3.2 Responses to Constituent Deviant and Standard Tones	30
2.4 Discussion	31
2.5 Materials and Methods	35
2.5.1 COGS-2 Data Collection	35
2.5.2 ERP Analysis	37
2.5.3 DDA of EEG data	37

Chapter 3	Delay Differential Analysis for Dynamical Sleep Spindle Detection	41
3.1	Introduction	42
3.1.1	Sleep Spindles	42
3.1.2	Delay Differential Analysis	43
3.2	Methods	46
3.2.1	Data	46
3.2.2	Spindle Marking	47
3.2.3	Supervised Structure Selection	48
3.2.4	Application to Full-Time Data	53
3.3	Results	54
3.3.1	Comparison with Established Methods	54
3.4	Discussion and Conclusions	64
Chapter 4	Cross-Dynamical Delay Differential Analysis Reveals Information Flow During Hippocampal Ripples	70
4.1	Hippocampal Ripples	71
4.2	Methods	72
4.2.1	Intracranial Recordings	72
4.2.2	Cross-Dynamical Delay Differential Analysis (CD-DDA)	72
4.2.3	Granger Causality	73
4.2.4	Convergent Cross Mapping	74
4.3	Results	75
4.3.1	CD-DDA During Ripples	75
4.3.2	Randomized Time Tests	78
4.4	Discussion	87
Chapter 5	Delay Differential Analysis of Seizures and Peri-Ictal States	89
5.1	Epilepsy Background	90
5.2	DDA of Epilepsy	91
5.3	Data Management	92
5.3.1	Data Collection	93
5.3.2	Processing Data Remotely	93
5.3.3	Data Processing and Analysis Code	93
5.4	Precise Seizure Timing with DDA features	96
5.5	Seizure Prediction	97
5.6	Discussion and Future Directions	101
Chapter 6	Conclusions	106
Appendix A	Supplementary Information for Chapter 2	109
A.1	Embeddings	109
A.2	DDA Details	113
A.2.1	Cross-Trial DDA	113

A.2.2	Structure Selection	115
A.3	COGS-2 Clinical and Demographical Characteristics	119
A.4	Statistical Details of t-test Results	122
Bibliography	125

LIST OF FIGURES

Figure 1.1:	A simple circular attractor with delay and differential embeddings	4
Figure 1.2:	Effect of the number of points m used in computing the numerical derivative	8
Figure 1.3:	Effect of the number of points m used in computing the numerical derivative and window length	9
Figure 1.4:	Classification of data from a 9-dimensional Lorenz system using Delay Differential Analysis (DDA)	16
Figure 1.5:	Behavior of a 40-dimensional Lorenz-95 system at nine timepoints following perturbation of one site	17
Figure 1.6:	Classification of data from the Lorenz-95 system using DDA	18
Figure 1.7:	Cross-dynamical DDA (CD-DDA) reveals causal interactions between two systems.	21
Figure 1.8:	Results from Granger Causality (GC), Convergent Cross-Mapping (CCM), and CD-DDA on data from interacting networks of Izhikevich neurons. . .	23
Figure 2.1:	Auditory Deviance Response (ADR) complex, deviant tone response minus standard tone response, as seen in DDA features and traditional event-related potential (ERP) analysis	29
Figure 2.2:	ADR components evident in responses to deviant tones using both DDA features and traditional ERP analysis	32
Figure 2.3:	Responses to standard tones as seen in DDA features and traditional ERP analysis	33
Figure 2.4:	Schizophrenia patients (SZ) and Nonpsychiatric Comparison Subjects (NCS) were presented with auditory stimuli consisting of 50 ms standard tones with randomly interspersed 100 ms deviant tones, and electroencephalogram (EEG) data were analyzed using both DDA and a traditional ERP paradigm	36
Figure 2.5:	Estimation of the DDA features from EEG data	38
Figure 3.1:	DDA is a powerful tool for time-domain data classification based on embed- ding theory from nonlinear dynamics.	45
Figure 3.2:	Peak spindle frequencies for each of the intracranial recordings	50
Figure 3.3:	Spindles as recorded in the EEG compared with a spectrogram and the output of the DDA spindle detector	56
Figure 3.4:	Comparison on DDA with several other methods for spindle detection based on F_1 score and CPU time	58
Figure 3.5:	Spindle detection performance of combined measures from multiple methods	60
Figure 3.6:	Spindle detection performance on DREAMS data	63
Figure 3.7:	Detecting content in specific frequency bands using DDA	69
Figure 4.1:	CD-DDA of hippocampus and cortical sites during hippocampal ripples . .	76
Figure 4.2:	CD-DDA of cortex-to-cortex connections during hippocampal ripples . . .	77

Figure 4.3:	CD-DDA of hippocampus-to-cortex connections for recordings from different patients break down into two groups	78
Figure 4.4:	GC, CCM, and CD-DDA when pairing hippocampal data with simultaneously recorded cortical data, data from random times in the cortex, and scrambled data	79
Figure 4.5:	CD-DDA results when pairing hippocampal ripples with data from random times in the cortex	81
Figure 4.6:	Difference in CD-DDA measure for time-locked pairing and random-time pairings	82
Figure 4.7:	GC results when pairing hippocampal ripples with data from random times in the cortex	83
Figure 4.8:	Difference in GC measure for time-locked pairing and random-time pairings	84
Figure 4.9:	CCM results when pairing hippocampal ripples with data from random times in the cortex	85
Figure 4.10:	Difference in CCM measure for time-locked pairing and random-time pairings	86
Figure 4.11:	Difference in CD-DDA measure for time-locked pairing and random-time pairings plotted across cortical recording sites	87
Figure 5.1:	Example DDA features from all channels for one seizure	97
Figure 5.2:	Differences between clinical onset time and visual marking based on DDA features	98
Figure 5.3:	DDA features show clear pre-ictal states in one example patient	100
Figure 5.4:	DDA features for seizure prediction	104
Figure 5.5:	Seizure prediction across long recording periods	105
Figure A.1:	Numerically integrated data x and y and the corresponding x - y phase space plots	111
Figure A.2:	Derivative and delay embeddings for the numerically integrated time series for $x(t)$	112
Figure A.3:	Cross-trial DDA	114
Figure A.4:	Cross-trial DDA coverage of the embedding space	114
Figure A.5:	Random subsampling cross-validation and construction of the weight matrix	116
Figure A.6:	Models selected for each of the 74 time windows	117
Figure A.7:	Delay pairs selected for each of the 74 time windows using supervised structure selection	118
Figure A.8:	Delay pairs selected for each of the 74 time windows using unsupervised structure selection	119
Figure A.9:	Statistical details of the t-tests on ERP and DDA	123
Figure A.10:	t-statistics values from unpaired t-tests performed for ERP and DDA grouped by the four age bins	124
Figure A.11:	Effect sizes (Cohen's d) values for ERP and DDA shown for each of the four age bins	124

LIST OF TABLES

Table 3.1:	Human-marked spindle properties for the fifteen recordings.	49
Table 3.2:	DDA spindle detection performance on all recordings	55
Table 3.3:	Comparison of detection methods for all data	59
Table 3.4:	Combining detection measures from the various methods	61
Table 3.5:	Selected delays (τ_1, τ_2) for specified bands, units of $\delta t = 1/f_s$	68
Table 5.1:	DDA seizure prediction performance from the best-predicting channel	101
Table A.1:	Clinical and demographic characteristics of the COGS-2 participants	120
Table A.2:	Exclusion criteria for the COGS-2 testing	121
Table A.3:	Demographic characteristics of the four age subgroups	122

ACKNOWLEDGEMENTS

I would like to acknowledge all of the many people who have guided me through my doctoral work. First, thank you to my advisor, Terrence Sejnowski, who has provided constant guidance and support. The entire Computational Neurobiology Laboratory (CNL), which assembles many of the smartest people I have ever met and provides the best environment I could imagine for groundbreaking neuroscience, is a testament to Professor Sejnowski's singular ability to inspire people and bring them together to advance science.

I would also like to give special thanks to Claudia Lainscsek. Her inexhaustible passion for science and discovery and her unparalleled ability to generate new ideas has been a constant source of inspiration and encouragement to me, and her deep knowledge of nonlinear dynamics has been essential for all of these projects. It would be impossible to overstate how much I have learned from her, or how fortunate I have been to work with her.

I am also grateful for the invaluable advice given by all of the other members of my thesis committee: Henry Abarbanel, Eric Halgren, Gregory A. Light, and John Reynolds. Each of them has provided numerous critical insights throughout my doctoral work, and I have been truly fortunate to have access to each of their perspectives. Thank you also to Sydney Cash for welcoming me to visit and work in his laboratory at the Massachusetts General Hospital and for his guidance of our epilepsy project.

I also acknowledge all of the many members of the CNL from whom I have learned so much during my graduate studies. In particular, I would like to thank everyone with whom I have had the privilege to work on the projects I've been involved in, including Lyle Muller, Anup Das, and all the other members of the Chaos Tower during my time, including Robert Kim, Christopher Gonzalez, Daniel Sexton, Jiayang Zhong, William McGrath, Nuttida Rungratsameetaweemana, Jonathan Weyhenmeyer, Erin Brown, Tiger Lin, Karen Man, Xenia Lainscsek, Margot Wagner, Burke Rosen, Charlie Dickey, Debha Amatya, Jack Knickrehm, Ahyeon Hwang, Alex Kuo, and Yinghao Li. Thank you also to Jorge Aldana, Brad Brockmeyer, Ramona Marchand, and Mary

Ellen Perry for all of the invaluable help they have given me during my time in the lab.

Thank you also to the Neurosciences Graduate Program, and particularly to Erin Gilbert and Linh Vandermar for all of the help they have given me, and all work they do on behalf of all of us in the program.

I am also grateful to my family for all of the support that they have given me during my graduate studies and throughout my life. Thank you to my father, Kenneth Sampson, for always providing me with reading material and innumerable interesting discussions, and to my sister, Sarah Sampson, for always sharing her wisdom and good judgment. And thank you to my mother, Zora Sampson, especially for her careful reading and comments on this dissertation, but also for all of her other support over the years. And thank you to my wife, Yvette Joon Ying Boon, who has made my life better every day since I met her, and who continues to inspire me every day.

There are also many people at UC San Diego who have vastly improved my time here. Thank you to my friend and longtime roommate Nandikishore Prakash, to all of my classmates and friends in the Neurosciences Graduate Program, to all the members of the UCSD Taekwondo Club, and to everyone from the American Indian Graduate Student Association, the Native American Student Alliance, and the Intertribal Resource Center.

I would also like to acknowledge that UC San Diego was built upon the unceded territory of the Kumeyaay Nation. The Kumeyaay continue to maintain both their political sovereignty and cultural traditions and are vital members of our community.

I would like to specifically acknowledge Claudia Lainscsek, Robert Kim, and Christopher Gonzalez for their help with preparing Chapter 1 of this dissertation, especially the analysis of interacting networks of Izhikevich neurons.

Chapter 2, in full, is a reprint of the material as it appears in the Proceedings of the National Academy of Sciences 2019. Claudia Lainscsek, Aaron L. Sampson, Robert Kim, Michael L. Thomas, Karen Man, Xenia Lainscsek, The COGS Investigators, Neal R. Swerdlow, David L. Braff, Terrence J. Sejnowski, and Gregory A. Light. The dissertation author was a

primary investigator and co-author of this paper.

Chapter 3, in full, is a reprint of the material as it appears in the Journal of Neuroscience Methods 2019. Aaron L. Sampson*, Claudia Lainscsek*, Christopher E. Gonzalez, István Ulbert, Orrin Devinsky, Dániel Fabó, Joseph R. Madsen, Eric Halgren, Sydney S. Cash, and Terrence J. Sejnowski. The dissertation author was the primary investigator and author of this paper.

Chapter 4 is coauthored with Claudia Lainscsek, Christopher E. Gonzalez, Xi Jiang, Jorge González-Martínez, Eric Halgren, and Terrence J. Sejnowski. The dissertation author was the primary author of this chapter.

Chapter 5 is coauthored with Claudia Lainscsek, Jiayang Zhong, William McGrath, Daniel Sexton, Jonathan Weyhenmeyer, Diana Ponce-Morado, Sydney S. Cash, and Terrence J. Sejnowski. The dissertation author was the primary author of this chapter.

VITA

2010	S. B. in Physics, Massachusetts Institute of Technology, Cambridge, MA
2010-2013	Research Technician, Massachusetts General Hospital, Boston, MA
2019	Ph. D. in Neurosciences, University of California San Diego, La Jolla, CA

PUBLICATIONS

Aaron L. Sampson*, Claudia Lainscsek*, Christopher E. Gonzalez, István Ulbert, Orrin Devinsky, Dániel Fabó, Joseph R. Madsen, Eric Halgren, Sydney S. Cash, and Terrence J. Sejnowski, “Delay Differential Analysis for Dynamical Sleep Spindle Detection”, *Journal of Neuroscience Methods*, 316 (2019) : 12-21.

Claudia Lainscsek, Aaron L. Sampson, Robert Kim, Michael L. Thomas, Karen Man, Xenia Lainscsek, The COGS Investigators, Neal R. Swerdlow, David L. Braff, Terrence J. Sejnowski, and Gregory A. Light, “Nonlinear Dynamics Underlying Sensory Processing Dysfunction in Schizophrenia”, *Proceedings of the National Academy of Sciences*, 116, no. 9 (2019): 3847-3852; DOI: 10.1073/pnas.1810572116

Claudia Lainscsek, Lyle E. Muller, Aaron L. Sampson, and Terrence J. Sejnowski. “Analytical derivation of nonlinear spectral effects and 1/f scaling artifact in signal processing of real-world data.” *Neural Computation* 29, no. 7 (2017): 2004-2020.

Anup Das, Aaron L. Sampson, Claudia Lainscsek, Lyle E. Muller, Wutu Lin, John C. Doyle, Sydney S. Cash, Eric Halgren, and Terrence J. Sejnowski. “Interpretation of the precision matrix and its application in estimating sparse brain connectivity during sleep spindles from human electrocorticography recordings.” *Neural Computation*, 29, no. 3 (2017): 603-642.

Kara J. Pavone, Oluwaseun Akeju, Aaron L. Sampson, Kelly Ling, Patrick L. Purdon, and Emery N. Brown. “Nitrous oxide-induced slow and delta oscillations.” *Clinical Neurophysiology* 127, no. 1 (2016): 556-564.

Patrick L. Purdon, Aaron L. Sampson, Kara J. Pavone, and Emery N. Brown. “Clinical electroencephalography for anesthesiologists part I: background and basic signatures.” *Anesthesiology: The Journal of the American Society of Anesthesiologists* 123, no. 4 (2015): 937-960.

Patrick L. Purdon, Kara J. Pavone, Oluwaseun Akeju, Anne C. Smith, Aaron L. Sampson, J. Lee, David W. Zhou, Ken Solt, and Emery N. Brown. “The ageing brain: age-dependent changes in the electroencephalogram during propofol and sevoflurane general anaesthesia.” *British Journal of Anaesthesia* 115, no. suppl. 1 (2015): i46-i57.

Oluwaseun Akeju, Kara J. Pavone, M. Brandon Westover, Rafael Vazquez, Michael J. Prerau, Priscilla G. Harrell, Katharine E. Hartnack, James Rhee, Aaron L. Sampson, Kathleen Habeeb, Lei Gao, Eric T. Pierce, John L. Walsh, Emery N. Brown, and Patrick L. Purdon. “A comparison of propofol-and dexmedetomidine-induced electroencephalogram dynamics using spectral and coherence analysis.” *Anesthesiology: The Journal of the American Society of Anesthesiologists* 121, no. 5 (2014): 978-989.

Oluwaseun Akeju, M. Brandon Westover, Kara J. Pavone, Aaron L. Sampson, Katharine E. Hartnack, Emery N. Brown, and Patrick L. Purdon. “Effects of sevoflurane and propofol on frontal electroencephalogram power and coherence.” *Anesthesiology: The Journal of the American Society of Anesthesiologists* 121, no. 5 (2014): 990-998.

Michael J. Prerau, Katherine E. Hartnack, Gabriel Obregon-Henao, Aaron L. Sampson, Margaret Merlino, Karen Gannon, Matt T. Bianchi, Jeffrey M. Ellenbogen, and Patrick L. Purdon. “Tracking the sleep onset process: an empirical model of behavioral and physiological dynamics.” *PLoS Computational Biology* 10, no. 10 (2014): e1003866.

Gaetano Valenza, Oluwaseun Akeju, Kara J. Pavone, Luca Citi, Katharine E. Hartnack, Aaron L. Sampson, Patrick L. Purdon, Emery N. Brown, and Riccardo Barbieri. “Instantaneous monitoring of heart beat dynamics during anesthesia and sedation.” *Journal of Computational Surgery* 1, no. 1 (2014): 13.

Patrick L. Purdon, Eric T. Pierce, Eran A. Mukamel, Michael J. Prerau, John L. Walsh, Kin Foon K. Wong, Andrés F. Salazar-Gomez, Priscilla G. Harrell, Aaron L. Sampson, Aylin Cimenser, ShiNung Ching, Nancy J. Kopell, Casie L. Tavares-Stoeckel, Kathleen Habeeb, Rebecca Merhar, and Emery N. Brown. “Electroencephalogram signatures of loss and recovery of consciousness from propofol.” *Proceedings of the National Academy of Sciences* 110, no. 12 (2013): E1142-E1151.

Aaron L. Sampson, Behtash Babadi, Michael J. Prerau, Eran A. Mukamel, Emery N. Brown, and Patrick L. Purdon. “Beamforming approach to phase-amplitude modulation analysis of multi-channel EEG.” In *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 6731-6734. IEEE, 2012.

Kin Foon Kevin Wong, Eran A. Mukamel, Andrés F. Salazar-Gomez, Eric T. Pierce, P. Grace Harrell, John L. Walsh, Aaron L. Sampson, Emery N. Brown, and Patrick L. Purdon. “Robust time-varying multivariate coherence estimation: Application to electroencephalogram recordings during general anesthesia.” In *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 4725-4728. IEEE, 2011.

Kin Foon Kevin Wong, Anne C. Smith, Eric T. Pierce, P. Grace Harrell, John L. Walsh, Andrés F. Salazar-Gomez, Casie L. Tavares, Aylin Cimenser, Michael J. Prerau, Eran A. Mukamel, Aaron L. Sampson, Patrick L. Purdon, Emery N. Brown. “Bayesian analysis of trinomial data in behavioral experiments and its application to human studies of general anesthesia.” In *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 4705-4708. IEEE, 2011.

Aaron L. Sampson. “An analysis of noise in the CoRoT data.” MIT Institute Archives–Noncirculating Collection 3. Thesis Phy 2010 S.B.

Vidya Ganapati, Stephan Schoenfelder, Sergio Castellanos, Sebastian Oener, Ringo Koepge, Aaron L. Sampson, Matthew A. Marcus, Barry Lai, Humphrey Morhenn, Giso Hahn, Joerg Bagdahn, and Tonio Buonassisi. “Infrared birefringence imaging of residual stress and bulk defects in multicrystalline silicon.” *Journal of Applied Physics* 108, no. 6 (2010): 063528.

Stephan Schoenfelder, Aaron L. Sampson, Vidya Ganapati, Ringo Koepge, Joerg Bagdahn, and Tonio Buonassisi. “Quantitative stress measurements of bulk microdefects in multicrystalline silicon.” In *Proc. 24th European Photovoltaic Solar Energy Conference* (2009): 977-980.

ABSTRACT OF THE DISSERTATION

Delay Differential Analysis of Neural Data

by

Aaron Lee Kasey Sampson

Doctor of Philosophy in Neurosciences with a Specialization in Computational Neurosciences

University of California San Diego, 2019

Professor Terrence J. Sejnowski, Chair

The brain can be seen as an extremely high-dimensional dynamical system. Despite the great complexity of the brain and the diversity of its activity, only limited measures of brain function are in general possible to record—e.g. the electroencephalogram (EEG) or intracranial recordings. The problem of extracting relevant information from a limited measurement of a complex system, and thereby determining the invariant nonlinear properties of the underlying system, has been well-studied in the field of nonlinear dynamics. Delay Differential Analysis (DDA) is a powerful nonlinear tool for time-domain data classification. In DDA, a low-dimensional nonlinear functional embedding is built from the dynamical structure of the data; this serves as a basis onto which the data can be mapped. By constraining the models used to low dimensionality, we ensure

that DDA is largely immune to overfitting, insensitive to noise, and generalizes well to new data. DDA has been successfully applied to a range of EEG classification problems. Cross-Dynamical DDA (CD-DDA) is a formulation of DDA which includes terms from multiple locations to study information flow or causal connections. Here, DDA and CD-DDA are applied to several EEG and intracranial data sets. In Chapter 2, DDA is applied to EEG data from a large study of schizophrenia patients and nonpsychiatric comparison subjects. In Chapter 3, DDA is used to detect sleep spindles in human intracranial recordings. In Chapter 4, CD-DDA is applied to intracranial recordings from the hippocampus and remote cortical sites to study information flow. In Chapter 5, the system that has been developed for analysis of large volumes of intracranial data from epilepsy patients is described, along with preliminary results related to seizure timing and prediction. Each of these distinct applications illustrates the power of DDA to access the underlying dynamics of neural activity and provide new insights.

Chapter 1

Introduction

1.1 The Brain as a Nonlinear Dynamical System

Nature is nonlinear. Across a striking breadth of fields focused on diverse sets of natural phenomena from the weather to brain activity, complex nonlinear dynamics have been observed. The result of these dynamics is behavior that often appears chaotic, unpredictable or counterintuitive. Early studies of these types of behavior focused on relatively simple physical systems, which nonetheless provide a great richness of behavior to explore. A key feature of such nonlinear systems is chaotic behavior: for certain parameter regimes, their behavior is deterministic but nonperiodic [Lorenz, 1963].

Neuroscience in particular is concerned with behavior of some of the most complex systems that occur in nature. The brain exhibits an astonishing complexity of structure across a range of scales from the molecular to that of the entire organ. This structural complexity is matched with a corresponding complexity of activity and behavior, which in turn underlies a great diversity of function.

Connections between neuroscience and nonlinear dynamics have long been noted, particularly in the context of developing models for the behavior of the nervous system at a range of

scales [Breakspear, 2017, Rabinovich et al., 2006]. Here, the goal is *not* to model neural data, but rather to study how techniques inspired by nonlinear dynamics can help to refine the analysis of neural data with a particular focus on problems of data classification and signal detection.

A fundamental challenge in approaching these types of problems in neural data is the necessarily limited view of the underlying systems. For a human brain composed of billions of neurons interacting across trillions of synapses, in general we can record its activity only at a few hundred locations. At each of these locations, for typical electroencephalogram (EEG) or intracranial recording techniques, the data represent the combined activity of a large number of neurons [Burle et al., 2015, Parvizi and Kastner, 2018]. While new technologies will allow for recordings from ever-increasing numbers of locations with far greater spatial resolution, the picture provided will remain incomplete, and the fundamental difficulty of understanding the sophisticated behavior of the underlying system with a limited view will remain. For this reason, it is essential to develop tools that are designed to provide information about the state of a complex system from limited data. And it is here that the tools of nonlinear dynamics are most useful.

1.2 Embeddings

Critical to studying the properties of dynamical systems is the concept of an embedding. This is a technique to reveal the nonlinear invariant properties of an underlying dynamical system from limited information, i.e. a single variable, by converting a single time series into a multi-dimensional object: an embedding [Whitney, 1936, Packard et al., 1980, Sauer et al., 1991, Takens, 1993]. Here we will consider both delay embeddings, where a time-delayed time series and the original time series form the embedding, and differential embeddings, which use time derivatives together with the original time series to construct the embedding. Both types of embeddings preserve the nonlinear invariant properties of the original dynamical system. To illustrate the usefulness of embeddings, an example is given of a very simple system of differential equations.

1.2.1 A Simple Embedding Example

The second-order differential equation

$$\ddot{x} = \dot{y} = ax \quad (1.1)$$

can be written as a system of differential equations:

$$\dot{x} = y \quad (1.2)$$

$$\dot{y} = ax. \quad (1.3)$$

A special solution to Eq. (1.1) is:

$$x = \cos(\omega t) \quad (1.4)$$

with

$$\dot{x} = y = -\omega \sin(\omega t) \quad (1.5)$$

and

$$\ddot{x} = \dot{y} = -\omega^2 \cos(\omega t) \quad (1.6)$$

we get

$$a = -\omega^2. \quad (1.7)$$

The simplest linear delay differential equation (DDE) is

$$\dot{x} = ax_\tau \quad (1.8)$$

where $x_\tau = x(t - \tau)$. Eq. (1.8) has the same special solution as Eq. (1.1) [Falbo, 1995]. We

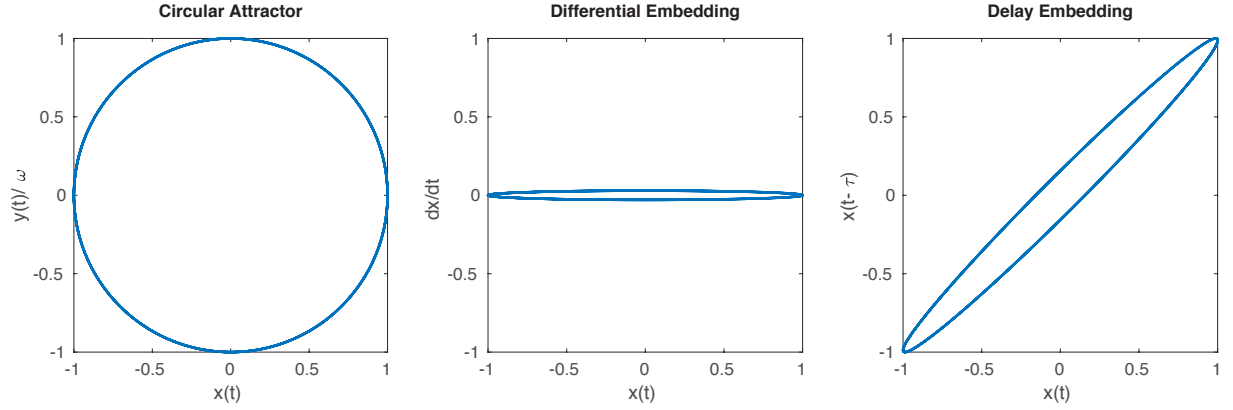


Figure 1.1: Eq. (1.8) corresponds to a system with an elliptical attractor in the x - y phase space. This is demonstrated by plotting $x(t)$ against $y(t)$ scaled by the angular frequency ω in the left panel. A differential embedding is shown in the center panel, and a delay embedding (with $\tau = 10\delta t$) is shown at right. Both embeddings are ellipses, which are topologically equivalent to the original attractor.

therefore get:

$$\begin{aligned}
 -\omega \sin(\omega t) &= b \cos(\omega t - \tau) \\
 b &= (-1)^n \omega \\
 \tau &= \frac{\pi(2n-1)}{2\omega}
 \end{aligned} \tag{1.9}$$

The behavior of this system in the phase space of x and y is constrained to an ellipse. Since this corresponds to the long-term behavior of the system, this can be said to be an elliptical *attractor* [Auslander et al., 1964]. This attractor is apparent from plotting x against y . This is shown in Fig. 1.1, where we plot x against y/ω to obtain a circular attractor, but it can also be revealed through embeddings. This allows us to see the underlying structure (a circle) even with access to only one variable. As an example of this, the middle and right panels in Fig. 1.1 respectively show differential and delay embeddings of $x(t)$.

This ability to recover the topology of the original phase space from a single time series through an embedding such as the one shown here forms the basis of delay differential analysis (DDA). It is also noteworthy that changing the single parameter a changes both the amplitude and frequency x and y , regardless of dt , but the embedding does not change.

1.3 Delay Differential Analysis (DDA)

Delay Differential Analysis (DDA) is a powerful tool for time-domain data classification, which is based on embeddings, as discussed in Section 1.2 [Kremliovsky and Kadtke, 1997, Lainscsek et al., 2013a]. Using embeddings, it is possible to reveal the nonlinear invariant properties of a complex dynamical system using one time series generated by that system. As discussed in Section 1.1, the brain can be thought of as a complex dynamical system, but we have only a very limited view of its activity. In DDA, we use both delay and differential embeddings, relating them in a function form. We also use a non-uniform embedding, which involves multiple delays that are allowed to vary freely, subject to constraints, rather than multiples of a single delay [Judd and Mees, 1995]. Thus, in DDA, we use a non-uniform functional embedding [Lainscsek et al., 2017]. (Note that the specific DDEs used in DDA will often be referred to as models, but it is not the aim in DDA to model the data or predict its future behavior, rather we use a sparse model for classification only.)

The embedding in DDA serves as a sparse, nonlinear functional basis onto which we map the data to be classified. By building this basis on the dynamical structure of the data itself, we can avoid extensive preprocessing that might be necessary for other approaches to classification. The models used in DDA are also constrained to be low-dimensional, and therefore yield a small number of features (typically around 4). This enforced sparsity helps to minimize the risk of overfitting and provides a framework for classification that is generally noise-insensitive. This approach allows sufficient flexibility to capture the relevant dynamical features of a wide range of systems, since we can, in each case, search over a space of the functional forms of the model, constrained by the number of delays, number of terms in the polynomial, and order of nonlinearity, then select the particular model form which provides the best performance for each task. For a given classification task, it is often possible to select a DDA model using a very limited subset of data (such a single channel of EEG data from one subject), and then apply it to a wide range of

recordings with different conditions (such as data from other subjects, other channels or locations, and other recording systems) and obtain good classification performance.

The general form of the models used in DDA, which combine both differential and delay embeddings using linear or nonlinear polynomial functions is:

$$\dot{x}(t) = \sum_{i=1}^I a_i \prod_{n=1}^N x_{\tau_n}^{m_{n,i}} \text{ for } \tau_n, m_{n,i} \in \mathbb{N}_0 \quad (1.10)$$

where I is the number of monomials in the model, N is the number of delays, $m_{n,i}$ is the order of the n^{th} term in the i^{th} monomial, and x_{τ_n} represents $x(t - \tau_n)$. In order to estimate the time derivative of the data, $\dot{x}(t)$, a weighted center derivative is used [Miletics and Molnárka, 2005]:

$$\dot{x}(t) = \frac{1}{2M} \sum_{m=1}^M \frac{x(t+m) - x(t-m)}{m} \quad (1.11)$$

where M is the number of points used. The choice of M could introduce error in certain cases, especially when using data with a low sampling rate or long windows. For most typical cases, the values of the computed coefficients will be quite robust to the choice of M . This will be discussed further in Section 1.4.

For a given model, we compute a small set of features. This low-dimensional feature space will be used to classify the data. These features are the estimated coefficients a_i in Eq. (1.10), which are computed using numerical singular value decomposition (SVD) [Press et al., 1990]. The least-square error is an additional feature, and is defined as:

$$\rho = \sqrt{\frac{1}{K} \sum_{k=1}^K \left(\dot{x}_{t_k} - \sum_{i=1}^I a_i \prod_{n=1}^N x_{\tau_n, t_k}^{m_{n,i}} \right)^2} \quad (1.12)$$

where K is the number of time points, and x_{τ_n, t_k} represents $x(t_k - \tau_n)$.

1.4 Estimating the Time Derivative

Following Miletics and Molnárka (2005), we approximate the derivative numerically by taking a Taylor expansion around $x(t)$ and evaluating at points $x(t \pm m)$:

$$x(t \pm m) = x(t) \pm m\dot{x}(t) + m^2 \frac{\ddot{x}(t)}{2} \dots \quad (1.13)$$

and then truncate the higher order terms and combine functions for $x(t + m)$ and $x(t - m)$ to solve for $\dot{x}(t)$:

$$x(t + m) - x(t - m) = 2m\dot{x}(t) \rightarrow \dot{x}(t) = \frac{x(t + m) - x(t - m)}{2m} \quad (1.14)$$

Applying this form for the derivative at each point m up to M , and computing the average over all values of m gives the first-order weighted center derivative in Eq. (1.11). Intuitively, the number of points m used in computing the numerical derivative should scale with sampling rate. The value of m will have some effect on the values of the features computed in DDA. To test the effect of m , we can again consider the simplest linear delay differential equation:

$$\dot{x} = ax_\tau \quad (1.15)$$

where x_τ is $x(t - \tau)$, and consider a case where we can determine the value of the coefficient a analytically:

$$\begin{aligned} x(t) &= \cos(\omega t), \quad \tau = \pi/2 \\ \dot{x}(t) &= -\omega \sin(\omega t), \quad x_\tau = \sin(\omega t) \\ \therefore a &= -\omega \end{aligned} \quad (1.16)$$

Using this equation, we can find the error in our estimate of a for a range of values of m and sampling rates f_s . Fig. 1.2 shows the results of this test for data generated by the simple DDE shown in Eq. (1.16) with added 10 dB noise. The estimate of a is in general quite robust to the number of points used in computing the derivative. Large errors, in fact, are only introduced when

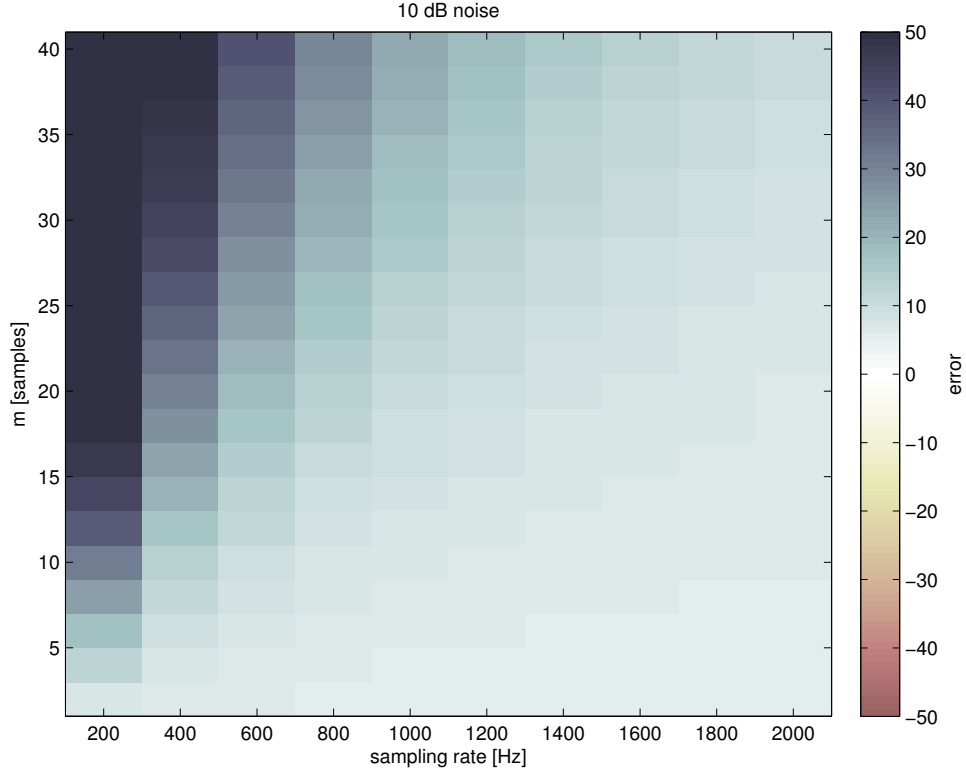


Figure 1.2: For sampling rates between 200 and 2000 Hz, data was generated from the simple DDE in Eq. (1.16) and 10 dB noise was added. The coefficient a was then estimated as in DDA, but using values of m for computing the numerical derivative (as in Eq. (1.11)) ranging between 2 and 40 points. The error in the estimate of the coefficient is shown with darker colors corresponding to greater error.

using high values of m with very low sampling rates. This is to be expected, since in this case, the amount of data used in computing the derivative would correspond to an appreciable portion of the period of the sine wave, and therefore to a segment of data over which the derivative is certainly not constant.

Additionally, we can test the effect of using data windows of different lengths. Fig. 1.3 shows the results of fixing the sampling rate to 1000 Hz and computing the error in the estimate of the coefficient a for window lengths ranging between 10 and 200 samples ($\delta t = 1/f_s$). For reasonably long window lengths (more than ≈ 0.06 s), again the estimate of the coefficient is quite robust to the number of points used, but large errors are introduced by shorter windows.

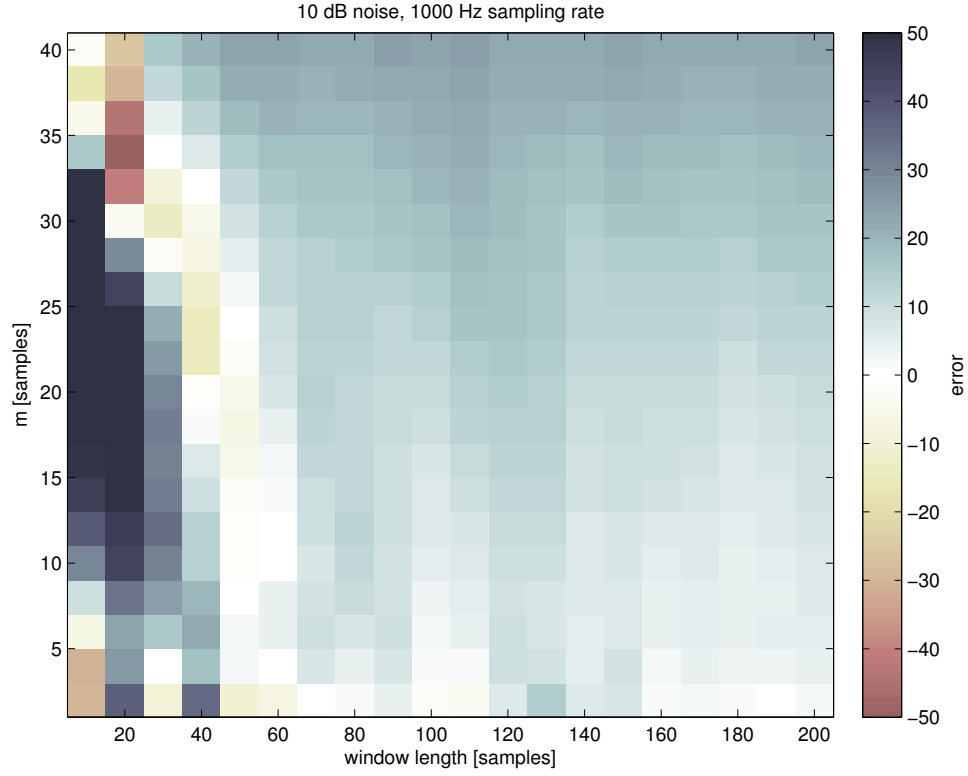


Figure 1.3: For 1000 Hz data generated from the simple DDE in Eq. (1.16) with 10 dB noise added, the coefficient a was estimated as in DDA, but using values of m for computing the numerical derivative (as in Eq. (1.11)) ranging between 2 and 40 points. The error in the estimate of the coefficient is shown with darker colors corresponding to greater error.

1.5 Delays and Sampling Rate

Since specific time delays are a fundamental aspect of the models used in DDA, it is necessary to make adjustments when using the same model with data of different sampling rates. Previously, the approach taken was to simply train a new model with new delays for each new sampling rate. E.g. to classify data sampled at 500 Hz and 1000 Hz, use recordings from each sampling rate as training data and perform structure selection separately. This approach presents significant drawbacks, however, since it is not possible to classify data with a different sampling rates without new corresponding training data, which significantly limits the otherwise broad applicability of the technique. A better approach is to simply scale the delays with the sampling rate and, critically, adjust the number of points used in estimating the derivative.

As an example, if a DDA model were selected based on data sampled at 500 Hz, it is possible to apply this same model to higher-sampled data, but the delays should be scaled. It is generally sufficient to scale them approximately (e.g. from 500 Hz to 1000 or 1024 Hz they should be doubled). In addition, it is necessary to adjust the points used in estimating the derivative. The total number of points should remain constant, but for higher-sampled data, they should sample more sparsely such that approximately the same amount of time is used (e.g. from 500 Hz to 1000 Hz, use every second point). To apply a DDA model to data with a lower sampling rate than the training data, the best approach is to upsample the new data to the training sampling rate.

It is also important to emphasize here that because of the nature of the models used, all terms are connected and both linear and nonlinear terms contain both linear and nonlinear information. This makes the interpretation of the delays somewhat more difficult. While intuitively, one might expect to be able to relate particular delays to corresponding frequencies, this is not actually the case. The presence of a specific delay in a nonlinear DDE does not denote the presence of a signal at the frequency f_s/τ [Lainscsek and Sejnowski, 2015].

1.6 Structure Selection of the Model

The particular sparse DDE model used for a given classification problem is selected based on a set of scored data for training purposes. This data in general relies upon expert scoring or clinical diagnosis to provide a “ground truth” for model selection. The set of models to be considered is first subjected to some reasonable constraints, informed by the types of models that have tended to be successful in past applications of DDA to diverse classification problems on different types of data. As discussed in Section 1.3, these constraints also ensure the sparsity of the model. From the general form of the DDA models shown in Eq. (1.10), it is necessary to set constraints on the number of delays N , the number of terms I , and the order of nonlinearity $\sum_n m_{n,i}$. Based on past applications of DDA, reasonable constraints would tend to be in the range of two delays ($N \leq 2$), three terms ($I = 3$), and up to third-order nonlinearities ($\sum_n m_{n,i} \leq 3$). Taking these constraints as an example, the resulting set of potential models would include 188 unique DDEs, a sufficiently low number to allow an exhaustive search of models. If using less strict constraints on the form of the models, it would be preferable to use a genetic algorithm or similar tool for model selection [Goldberg, 1998]. It is also necessary to constrain the delays to a reasonable timescale for the data, i.e. to detect short duration events, delays longer than the events should not be used. An important consideration is that among the set of unique DDE models, some will contain fewer delays than the upper limit (e.g. with a constraint of $N \leq 2$, some models will have only one delay present), and some models are symmetric with respect to the delays (the delays are interchangeable), reducing the total number of model-delay combinations to consider. Taking again the aforementioned constraints as an example, and considering delays τ_1 and τ_2 between 1 and 150 timepoints, we would consider, for non-symmetric models, 22350 delay pairs, for symmetric models, 11175 delay pairs, and for one-delay models, 150 delays.

In order to select a model which can generalize well to other data, we use repeated random subsampling cross-validation (CV) [Kohavi et al., 1995]. CV is a tool for estimating the accuracy

of a model's classification performance given limited data, preventing overfitting and improving generalizability. This method randomly divides the scored data into training and testing sets. This split is carried out with different random subdivisions a number of times. The breakdown between training and testing could be subject to variation, but in general 70% is assigned to training data, and 30% to testing. Note that throughout “training” and “testing” data will refer to these random splits of the scored data for model selection. Other data not included in the structure selection of the model will be referred to as “validation” data.

For each model, DDA is run on the training data, obtaining a set of features (coefficients $a_{k,i}$ and error ρ_k) for each subject, channel, trial, or epoch k in the data, depending on how the data are organized. We also use the “ground truth” labels l (from expert scoring, clinical diagnosis, etc.) for these data, and obtain a vector of weights W . These weights W which have been obtained from the training data can then be applied to the features computed from the testing data. This weighting of the features provides a one-dimensional distance D from an optimally separating hyperplane between the classes in features space. This distance from the hyperplane can then be compared to the labels l for the testing data. The correspondence between D and the scored labels is evaluated using the receiver operating characteristic (ROC). The ROC is constructed by setting various thresholds for D to separate the classes and then plotting the true positive rate and the false positive rate for each threshold. The area under the resulting ROC curve defined by these points, A' , is a threshold-independent measure of classification performance. For classification at the random chance level, A' should be around 0.5, and A' should approach 1 for perfect classification [Hand and Till, 2001]. Here, A' is computed using [Hand and Till, 2001] :

$$A' = \frac{S_0 - n_0(n_0 + 1)}{2n_0n_1} \quad (1.17)$$

where n_0 and n_1 represent the number of points in each of two classes, which are labeled 0 and 1. S_0 is obtained by ranking all points by their probability of being classified as 0, then taking

the sum of the ranks of the true class 0 points. CV is repeated a number of times for each model and delay pair, and the best-performing model and delays are selected to apply to validation data. The weighted combination of features D is often the optimal means of separating classes, but it is sometimes sufficient in practice to take a single feature for classification.

1.7 DDA of Dynamical Systems

A challenge in the application of any tool for classification of real-world data is finding “ground truth” for evaluating classification performance. In the brain, it is not possible in general to know the true labels for data coming from different disease states, states of consciousness, etc. Rather it is necessary to rely on the determination of human experts through clinical diagnosis or visual inspection of the EEG. Given this, it is useful to consider the performance of DDA on simulated data from well-studied dynamical systems. Lainscsek et al. used the Rössler system as a testing ground for DDA and found robust classification of data generated from systems with different parameters [Lainscsek et al., 2013d]. Here, we will consider two additional systems, a nine-dimensional Lorenz system [Reiterer et al., 1998], and the Lorenz-95 system, a model for placement of weather stations around the earth [Lorenz and Emanuel, 1998]. These systems are selected for their higher dimensionality. Using such high-dimensional systems may provide a closer approximation of data from natural systems, especially highly complex systems such as the brain. The Lorenz-95 system in particular is designed to allow extension to much higher dimensionality; here we will use a 40-dimensional system. While data generated from these systems is still very different from neural data, the robust performance of DDA on simulated data with such a range of dimensionality illustrates its generally broad applicability.

1.7.1 A Nine-Dimensional Lorenz System

Reiterer et al. introduced a nine-dimensional Lorenz system governed by a system of nine nonlinear ordinary differential equations (ODEs) for nine variables C_i :

$$\begin{aligned}
\dot{C}_1 &= -\sigma b_1 C_1 - C_2 C_4 + b_4 C_4^2 + b_3 C_3 C_5 - \sigma b_2 C_7 \\
\dot{C}_2 &= -\sigma C_2 + C_1 C_4 - C_2 C_5 + C_4 C_5 - \sigma C_9 / 2 \\
\dot{C}_3 &= -\sigma b_1 C_3 + C_2 C_4 - b_4 C_2^2 - b_3 C_1 C_5 + \sigma b_2 C_8 \\
\dot{C}_4 &= -\sigma C_4 - C_2 C_3 - C_2 C_5 + C_4 C_5 + \sigma C_9 / 2 \\
\dot{C}_5 &= -\sigma b_5 C_5 + C_2^2 / 2 - C_4^2 / 2 \\
\dot{C}_6 &= -b_6 C_6 + C_2 C_9 - C_4 C_9 \\
\dot{C}_7 &= -b_1 C_7 - r C_1 + 2 C_5 C_8 - C_4 C_9 \\
\dot{C}_8 &= -b_1 C_8 + r C_3 - 2 C_5 C_7 + C_2 C_9 \\
\dot{C}_9 &= -C_9 - r C_2 + r C_4 - 2 C_2 C_6 + 2 C_4 C_6 + C_4 C_7 - C_2 C_8
\end{aligned} \tag{1.18}$$

where r is the reduced Rayleigh number, and the parameters b_i are defined as:

$$\begin{aligned}
b_1 &= \frac{4((1+a^2))}{(1+2a^2)} \\
b_2 &= \frac{(1+2a^2)}{(2(1+a^2))} \\
b_3 &= \frac{2((1-a^2))}{(1+a^2)} \\
b_4 &= \frac{a^2}{(1+a^2)} \\
b_5 &= \frac{8a^2}{(1+2a^2)} \\
b_6 &= \frac{4}{(1+2a^2)}
\end{aligned} \tag{1.19}$$

leaving the system with three parameters, a , σ , and r . Following Reiterer et al., we set $\sigma = 0.5$ and $a = 0.5$ and vary the reduced Rayleigh number r .

In order to apply DDA to data generated from this system, data were generated with random initial conditions for a number of trials for a range of values of the parameter r . We select

two overlapping ranges to simulate two realistically difficult-to-classify groups (these could be compared to different patient populations, or data from the brain in different states, etc.). In each range, 10 values of r were selected, and 50 trials of data with different initial conditions were generated for each. 5 dB noise was added to the generated data. As described in Sections 1.3 and 1.6, these data were dividing into training and testing sets, DDA features were computed for a range of model forms and delay pairs, and repeated random subsampling cross-validation was used to select the best-performing model. For these data, the best model was found to be:

$$\dot{x} = a_1 x_{\tau_1} + a_2 x_{\tau_2} + a_3 x_{\tau_1}^2 \quad (1.20)$$

with $\tau_1 = 2\delta t$ and $\tau_2 = 44\delta t$; x_{τ} represents $x(t - \tau)$. Following the selection of this model, new data were generated for validation. These data were generated using new values of r selected from the same ranges for the two classes. Again, 10 values of r were selected from each range, 50 trials of data were generated using random initial conditions, and 5 dB noise was added to the data.

Fig. 1.4 shows the results of applying the selected DDA model to the validation data. Applying the optimally separating weights chosen after CV transforms the four features into a single distance from the hyperplane D which clearly separates the two classes. Performance is evaluated by computing the area under the ROC, A' . Classification performance is high for this system: across all validation data, A' is 0.792.

1.7.2 The Lorenz-95 System

Lorenz introduced a highly simplified model for observations at weather stations spaced around a single latitude circle on the earth [Lorenz, 1996, Lorenz and Emanuel, 1998]. The

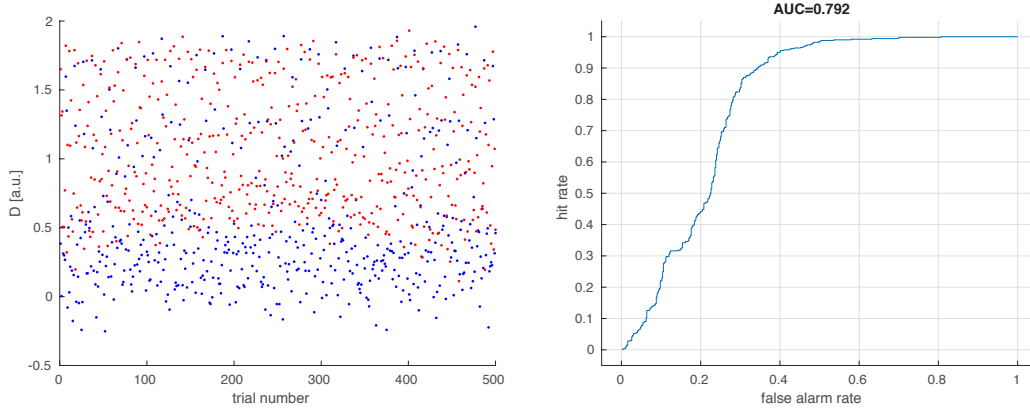


Figure 1.4: DDA was run on simulated data from the 9-dimensional Lorenz system described by Reiterer et al. Data were generated for 10 different values of the reduced Rayleigh number r in each of two overlapping ranges. Data from these two classes were labeled as belonging to two distinct classes. 5 dB noise was added to the data before running DDA. The DDA model and weights were selected on different data from the same system. At left, the distance from the hyperplane D , a weighted combination of the DDA features is plotted for class 1 in blue and class 2 in red for each of 50 trials for the 10 values of r in each class. At right, the ROC curve is plotted.

governing equations for this model are:

$$\dot{X}_j = (X_{j+1} - X_{j-2})X_{j-1} - X_j + F \quad (1.21)$$

where $j = 1, \dots, J$ for J weather stations, and J can be set to the desired dimension of the system. F is a parameter to be varied. X_{-1} is defined as X_{J-1} , X_0 as X_J , and X_{J+1} as X_1 , creating a cyclic chain. This system provides a straightforward way to generate data from a highly-dimensional system, perhaps with a dimensionality approaching that of the natural systems we wish to study (e.g. the brain).

Fig. 1.5 shows the evolution of a 40-dimensional Lorenz-95 system with $F = 8$ and an initial condition of $X_{20} = F + 0.008$ and all other $X_j = F$, following Fig. 1 in Lorenz & Emanuel 1998.

This system provides an ideal testing ground for DDA on data from a much higher-

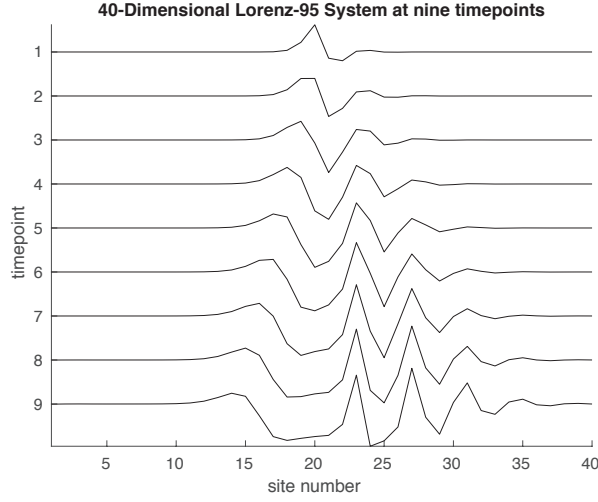


Figure 1.5: A 40-dimensional Lorenz-95 system with $F = 8$ and an initial condition of $X_{20} = F + 0.008$ and all other $X_j = F$, following Fig. 1 in Lorenz & Emanuel 1998. The horizontal portion of each curve is equal to F , and y-axis ticks correspond to 10 units.

dimensional system. Data were generated for a 40-dimensional system (akin to 40 weather stations) with differing values of the parameter F . Randomly selected values of F between 6 and 7 were used to generate data for class 1, and randomly selected values between 7 and 8 for class 2. A single variable, X_{10} , corresponding to a single weather station, was selected to simulate the type of limited data that might be available when measuring a natural system. These data were then used for DDA model selection. For these data, although a range of models and delays provide good classification performance, the best-performing DDA model is:

$$\dot{x} = a_1 x_{\tau_1} + a_2 x_{\tau_2} + a_3 x_{\tau_1}^2 \quad (1.22)$$

with $\tau_1 = 1 \delta t$ and $\tau_2 = 3 \delta t$; x_{τ} represents $x(t - \tau)$. Following structure selection, new data were generated for validation. These data were generated using new values of F selected from the same ranges for the two classes, and using the same variable X_{10} .

Fig. 1.6 shows the results of applying the selected DDA model to the validation data. The distance from the separating hyperplane D is plotted for both classes with blue points for class 1

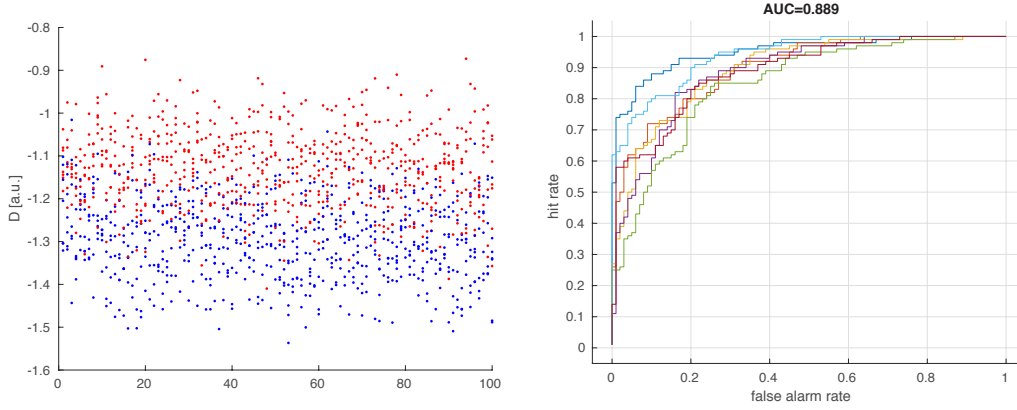


Figure 1.6: DDA was run on simulated data from the Lorenz-95 system. Data were generated for 10 different values of the parameter F in each of two adjacent ranges. Data from these two parameter ranges were labeled as belonging to two distinct classes. The DDA model and weights were selected on different data from the same system. At left, the distance from the hyperplane D , a weighted combination of the DDA features is plotted for class 1 in blue and class 2 in red. At right, the ROC curve is plotted for classification of seven subsets of the validation data.

and red points for class 2. Performance was evaluated for each of seven subsets of the validation data by computing the area under the ROC, A' . The average A' across these data sets is 0.901.

1.8 Cross-Dynamical Delay Differential Analysis (CD-DDA)

Despite the temptation to equate the two concepts, correlation does not imply causation, a fact that was noted by Yule in 1926. Yule further described the effect of time delays on our ability to find causal relationships between time series [Yule, 1926]. Drawing conclusions about causality with a firmer foundation than simple correlation has remained a longstanding challenge ever since. In 1969, Granger introduced a causality measure widely used in modern signal processing [Granger, 1969]. Despite its wide usefulness, however, Granger causality (GC) does have certain limitations, as it uses linear autoregressive (AR) models and treats data as stochastic. A new, nonlinear technique for studying causality, convergent cross mapping (CCM), was introduced by Sugihara et al. in 2012 and “specifically aimed at a class of system not

covered by GC” [Sugihara et al., 2012]. This includes nonlinear systems that can produce chaotic (deterministic but unpredictable) behavior [Lorenz, 1963, Lorenz, 1984].

Since it is a nonlinear technique which is fundamentally reliant on time delays appropriate to the data, DDA seems a natural candidate for extension to the study of causality, and perhaps as a means of further expanding the class of systems that can be interrogated for causal interactions. Such a formulation of DDA should be applicable to a wide array of natural, nonlinear systems, since DDA fundamentally relies on nonlinearities in the data. In order to consider the causal interactions between two systems, Eq. (1.10) can be modified to a cross-dynamical formulation of DDA (CD-DDA):

$$\begin{aligned}
\mathcal{F}_x &= \sum_{i=1}^I a_i \prod_{n=1}^N x(t - \tau_n)^{m_{n,i}} \quad \text{with } \tau_n, m_{n,i} \in \mathbb{N}_0 \\
\mathcal{F}_y &= \sum_{i=1}^I b_i \prod_{n=1}^N y(t - \tau_n)^{m_{n,i}} \quad \text{with } \tau_n, m_{n,i} \in \mathbb{N}_0 \\
\dot{x}(t) &= \mathcal{F}_x + \mathcal{F}_y + \rho_{x \leftarrow y} \\
\dot{y}(t) &= \mathcal{F}_x + \mathcal{F}_y + \rho_{y \leftarrow x}
\end{aligned} \tag{1.23}$$

where, as before, I is the number of monomials in the model, N is the number of delays, $m_{n,i}$ is the order of the n^{th} term in the i^{th} monomial, and x_{τ_n} represents $x(t - \tau_n)$. In CD-DDA, we also have $x(t)$ and $y(t)$ representing the data from two systems (e.g. two EEG or intracranial electrodes). Additionally, the x and y terms are the same (hence the number of monomials being equal to I for both x and y terms in the above equations). This constraint allows for the use of models found to describe the system well using the classical DDA approach detailed previously. It could, however, also be useful to allow these terms to differ and select a cross-dynamical model independently from classical DDA. This could be a promising subject of future research.

In order to obtain an interpretable measure of causal interactions between the systems generating $x(t)$ and $y(t)$, we compare the least squares error ρ from the cross-dynamical models in Eq. (1.23) to the one-system model, as in Eq. (1.10). For example, if we consider one specific

DDA model in its classical and cross-dynamical formulation:

$$\begin{aligned}\dot{x}(t) &= a_1 x_1 + a_2 x_2 + a_3 x_1^2 + \rho_x \\ \dot{x}(t) &= e_1 y_1 + e_2 y_2 + e_3 y_1^2 + e_1 x_1 + e_2 x_2 + e_3 x_1^2 + \rho_{x \leftarrow y}\end{aligned}\tag{1.24}$$

where $x_i = x(t - \tau_i)$ and $y_i = y(t - \tau_i)$. Then we define C_{xy} as our CD-DDA measure of the causal influence of $y(t)$ on $x(t)$:

$$C_{xy} = | \rho_x - \rho_{x \leftarrow y} | .\tag{1.25}$$

If, for example, $x(t)$ originates from a dynamical system A that responds to dynamical changes of another dynamical system B which is measured by $y(t)$, then $\dot{x} = f(x_{\tau_i}, y_{\tau_i})$ in Eq. (1.24) should reflect this causal link while $\dot{y} = f(x_{\tau_i})$ in Eq. (1.24) should be blind to such a configuration. Fig. 1.7 shows a schematic of such a pair of systems and the corresponding CD-DDA equations. Note that, in general, $\rho_{x \leftarrow y}$ will tend to be lower than ρ_x simply because more terms are present in the model. The comparison of the errors across time is a useful measure of how the interaction of the systems changes.

Assuming the dynamical systems A and B shown in Fig. 1.7 are two different brain regions, we can gain new insight into pathways of information flow in the brain by applying CD-DDA to recordings from these regions. CD-DDA has much in common with CCM as a nonlinear measure of causality, but the key difference between CD-DDA and CCM is the more general embedding used in CD-DDA. This presents an advantage because when using linear embeddings, causality will not necessarily result in a correlation. With a nonlinear functional embedding, such as in CD-DDA, the number of cases where an underlying causal link will result in observed correlation is expanded.

As a test of CD-DDA, and for comparison with both GC and CCM, simulated data were generated from two intermittently interacting networks of Izhikevich neurons. Data were generated for a scenario in which network A inhibits the activity of network B for one brief period

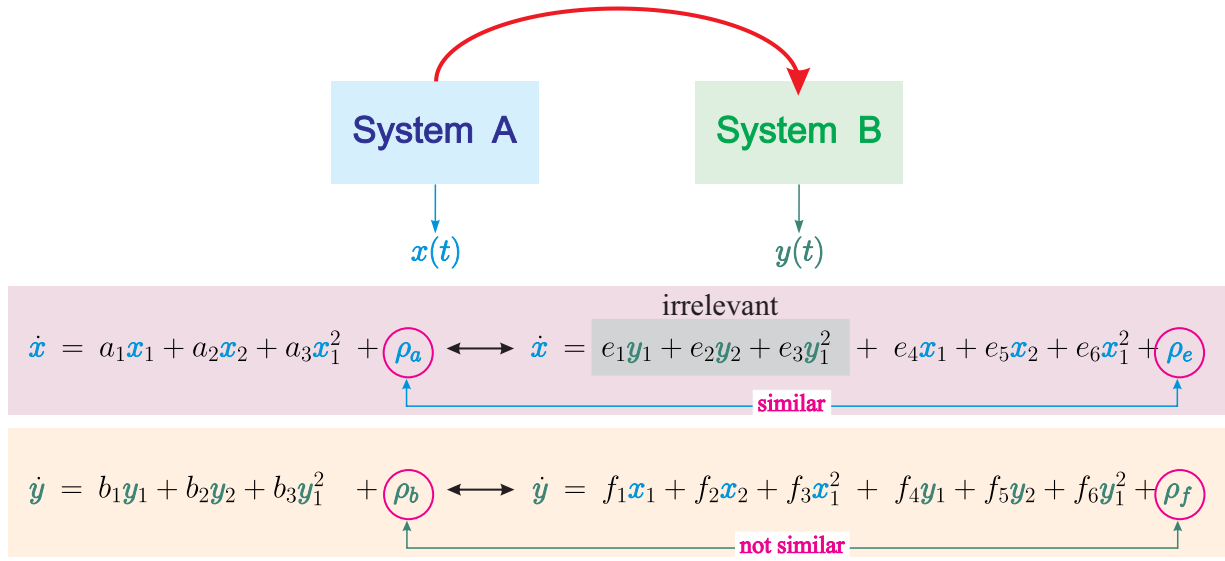


Figure 1.7: For two dynamical systems, System A which generates $x(t)$ and System B which generates $y(t)$, information is sent from System A that influences the behavior of System B, but there is no feedback from B to A. In this case, applying CD-DDA will show that there is a marked reduction in the error for a model for \dot{y} when x terms are introduced, but only a small reduction in the error when y terms are introduced into the DDE for \dot{x} , since these terms do not provide relevant information. ($|\rho_f - \rho_b| > |\rho_e - \rho_a|$)

and then network B inhibits the activity of network A for another brief period. GC, CCM, and CD-DDA were applied to these data, and the results are shown in Fig. 1.8. For these data, both GC and CD-DDA seem to recover the direction of inhibition at the correct times. Chapter 4 will detail the application of CD-DDA to intracranial recordings and its comparison with these other measures of causality.

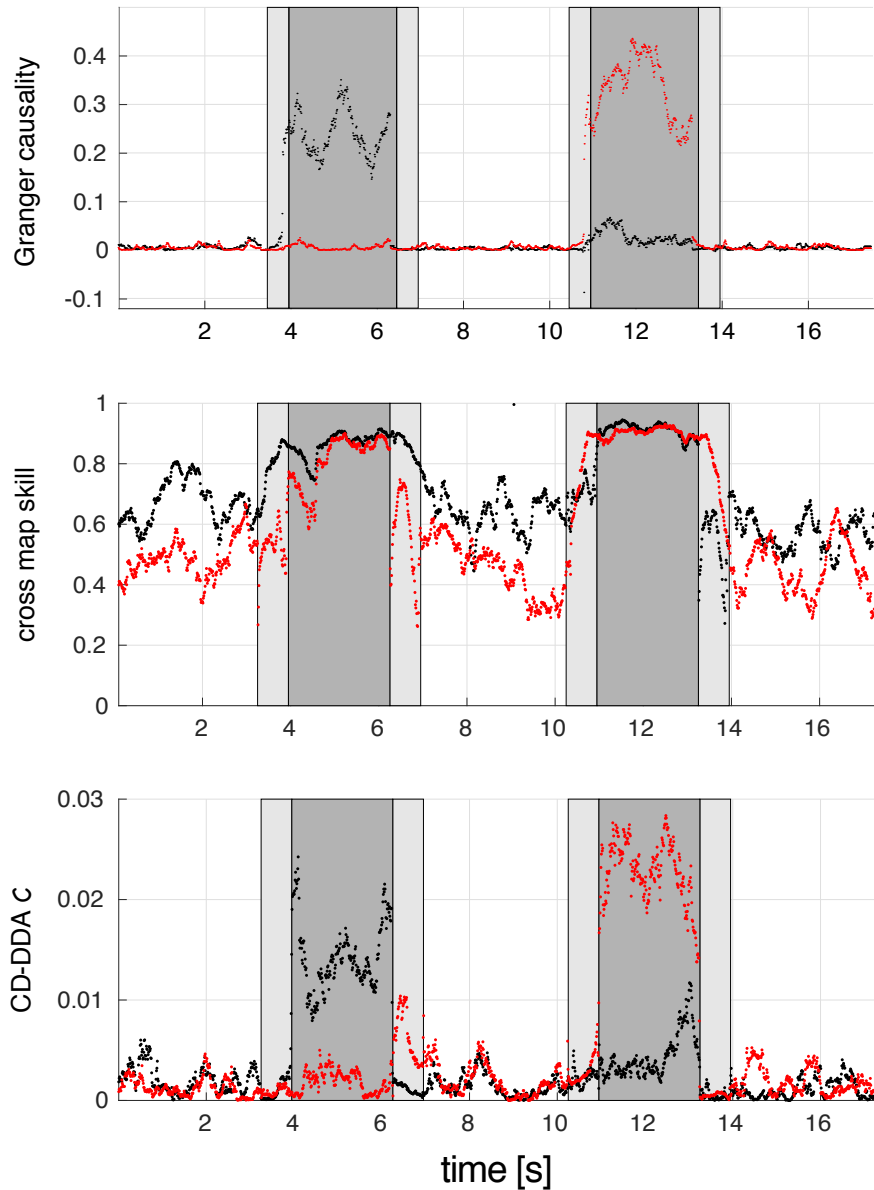


Figure 1.8: Simulated neural data were generated from two intermittently interacting networks of Izhikevich neurons. At time $t = 4$ s, network A inhibits the activity of network B for a period of two seconds. At time $t = 11$ s, network B inhibits the activity of network A, again for a period of two seconds. At all other times the networks do not interact. GC, CCM, and CD-DDA were all applied to these data and the results are plotted in the top, middle, and bottom panels. Black points represent measures of network A’s causal influence on network B, red points represent measures of network B’s causal influence on network A. Dark gray areas represent the periods of interaction, and light gray areas are the regions where the time windows used in computing each measure would include a mix of data from the interacting and non-interacting time periods.

Chapter 2

Nonlinear Dynamics Underlying Sensory Processing Dysfunction in Schizophrenia

Abstract

Natural systems, including the brain, often appear chaotic since they are typically driven by complex nonlinear dynamical processes. Disruption in the fluid coordination of multiple brain regions contributes to impairments in information processing and the constellation of symptoms observed in neuropsychiatric disorders. Schizophrenia, one of the most debilitating mental illnesses, is thought to arise, in part, from such a network dysfunction leading to impaired auditory information processing as well as cognitive and psychosocial deficits. Current approaches to neurophysiologic biomarker analyses predominantly rely upon linear methods and may therefore fail to capture the wealth of information contained in whole EEG signals, including nonlinear dynamics. In this study, delay differential analysis (DDA), a novel nonlinear method based on embedding theory from theoretical physics, was applied to EEG recordings from 877 schizophrenia (SZ) patients and 753 nonpsychiatric comparison subjects (NCS) who underwent mismatch negativity (MMN) testing via their participation in the Consortium on the Genetics

of Schizophrenia (COGS-2) study. DDA revealed significant nonlinear dynamical architecture related to auditory information processing in both groups. Importantly, significant DDA changes preceded those observed with traditional linear methods. Marked abnormalities in both linear and nonlinear features were detected in SZ patients. These results illustrate the benefits of nonlinear analysis of brain signals, and underscore the need for future studies to investigate the relationship between DDA features and pathophysiology of information processing.

2.1 Introduction

In neuropsychiatric disorders, subtle abnormalities in low-level processes contribute to the complex constellation of symptoms. Schizophrenia is among the most intractable and disabling illnesses with impairments in multiple domains of cognitive and psychosocial functioning [Thomas et al., 2017]. In this study, a novel nonlinear analysis technique for characterizing large-scale systems from a theoretical physics perspective was applied to leading candidate biomarkers in healthy nonpsychiatric subjects and schizophrenia patients.

Among many domains of clinically relevant dysfunctions, altered early auditory information processing (EAIP), as measured by event-related potentials (ERPs), is a fundamental feature of schizophrenia [Del Re et al., 2014, Turetsky et al., 2015a]. Mismatch negativity (MMN) is an ERP biomarker of early auditory information processing with promise for improving our understanding and treatment of schizophrenia. MMN has been shown to reliably correlate with cognition and psychosocial functioning [Thomas et al., 2017, Wynn et al., 2010]. MMN can be used for predicting the development of psychosis among individuals at high clinical risk and is sensitive to interventions that target cognition [Swerdlow et al., 2016, Perez et al., 2017]. MMN is automatically elicited via a passive auditory oddball paradigm in response to infrequent deviant sounds randomly interspersed in a sequence of frequently presented standard sounds. MMN is followed by a positive component also reflective of early auditory information processing (EAIP)

that peaks at 250-300 ms called P3a. Though less studied, the P3a is followed by a negative wave called the re-orienting negativity (RON) [Rissling et al., 2012]. This MMN-P3a-RON response has been referred to as the auditory deviance response (ADR) complex [Rissling et al., 2014]. Many studies have found significant ADR component reductions in schizophrenia which correlate with impairments across multiple domains of higher-order cognitive and psychosocial functioning [Todd et al., 2012, Light et al., 2015]. Recently, Thomas et al. proposed and validated a hierarchical information processing cascade model where impairments in these early auditory sensory processing measures propagate and substantially contribute to deficits in cognition, clinical symptoms, and real-world functional disability in schizophrenia [Thomas et al., 2017]. Thus, small changes in early auditory processing measures contribute to deficits in higher-order functions and macroscopic manifestations of schizophrenia.

The vast majority of ERP studies in neuropsychiatry have employed linear analysis of electroencephalogram (EEG) signals. These linear methods include ERP peaks and latencies [Oribe et al., 2015, Turetsky et al., 2015b], frequency [Knott et al., 2001] or time-frequency analyses [Roach and Mathalon, 2008, Light et al., 2017], and cross-frequency coupling [Moran and Hong, 2011, Kirihaara et al., 2012]. Such conventional linear methods, while highly informative, focus on *a priori* determined time-locked events or frequency ranges (i.e. delta, theta, gamma bands) and therefore fail to capture important underlying nonlinear system dynamics. Thus, characterizing general dynamical components of broadband data without such restrictions may reveal novel information about altered systems-level states associated with schizophrenia.

In order to characterize the large-scale, neural system-level dynamics present in brain electrical activity in patients with schizophrenia, computational models based on nonlinear dynamics and systems theory recently emerged as a promising tool in neuroscience [Breakspear, 2017, Murray et al., 2014]. In contrast to our prevailing focus on microscale neuronal activity, methods that characterize large-scale system level dynamics may help us understand emergent phenomena of behavior and cognition [Paulus and Braff, 2003, Breakspear, 2017]. This perspective also

underlies past work linking nonlinear analysis of the EEG to the “disconnection hypothesis” of schizophrenia [Breakspear et al., 2003, Stephan et al., 2006]. Breakspear et al. used a technique for estimating the nonlinear dynamical interdependence of several EEG channels across and within both hemispheres [Breakspear and Terry, 2002, Terry and Breakspear, 2003] and found significant differences in the topography of dynamical interdependence across the scalp between schizophrenia patients and matched healthy comparison subjects. Here, we apply a distinct but related technique to probe the nonlinear dynamics of a single EEG channel, which allows us to investigate the nonlinear properties of a broader range of data. Delay differential analysis (DDA) is designed to capture large-scale dynamics present in nonlinear time-series signals [Lainscsek and Sejnowski, 2015, Lainscsek et al., 2014, Lainscsek and Sejnowski, 2013b, Kremliovsky and Kadtke, 1997]. DDA operates in the time domain and maps unprocessed data onto a functional embedding basis. Embeddings are used in nonlinear dynamics to reveal invariant dynamical properties underlying a more extensive, largely unknown dynamical system (i.e. the brain) when only a single time series (e.g. EEG data) is available. Previous studies have shown that DDA can be used to extract disease-specific dynamical features: Parkinson’s movement data [Lainscsek et al., 2009, Lainscsek et al., 2012], electrocardiographic (ECG) recordings [Lainscsek and Sejnowski, 2013a, Lainscsek and Sejnowski, 2013b], sleep EEG [Lainscsek et al., 2014], classification of Parkinson’s disease EEG data [Lainscsek et al., 2013c, Lainscsek et al., 2013b], and electrocorticography (ECoG) data for epileptic seizure characterization [Lainscsek et al., 2017].

In this study, DDA was applied to the unprocessed EEG recordings from schizophrenia (SZ) and nonpsychiatric comparison subjects (NCS) who participated in the Consortium on the Genetics of Schizophrenia (COGS-2), a multi-center case-control study of schizophrenia and related endophenotypes [Light et al., 2015]. In contrast to our prior focus on linear ERP features (MMN and P3a amplitudes) [Light et al., 2015, Thomas et al., 2017], this study aimed to determine whether 1) nonlinear dynamics can be detected in raw EEG recordings, and if so, 2)

can they be used to differentiate SZ patients from NCS. By mapping the trial information onto the DDA feature space, we also aimed to 3) determine whether the MMN, P3a, and RON responses reflect nonlinear processes even after controlling for SZ-NCS amplitude differences, which would provide an important new perspective on EAIP deficits in schizophrenia. Such information may be important for guiding future studies of the genomic and neural substrates of schizophrenia and provide targets for new treatment developments.

2.2 Delay Differential Analysis (DDA)

We developed and applied a novel signal processing technique, DDA, based on embedding theory in nonlinear dynamics. DDA combines two types of embeddings via a delay differential equation: delay embedding and derivative embedding. The general nonlinear DDA model can be formulated as:

$$\dot{x} = \sum_{i=1}^I a_i \prod_{n=1}^N x_{\tau_n}^{m_{n,i}} \quad (2.1)$$

where $\tau_n, m_{n,i} \in \mathbb{N}_0$ with $x = x(t)$, $x_{\tau_n} = x(t - \tau_n)$, relating the signal derivative $\dot{x}(t)$ to the signal non-uniformly shifted in time. Most of the terms in this template are set to zero depending on the data type. The following DDA model has been shown to capture important dynamical information from EEG signals [Lainscsek et al., 2013b, Lainscsek et al., 2014, Sampson et al., 2015, Lainscsek et al., 2017]:

$$\dot{x} = a_1 x_1 + a_2 x_2 + a_3 x_1^2. \quad (2.2)$$

This model was chosen through a structure selection procedure, repeated random subsampling cross-validation (see [Kohavi et al., 1995] and supplementary information in the Appendix), using separate EEG and intracranial EEG datasets [Lainscsek et al., 2013b, Lainscsek et al., 2014, Sampson et al., 2015, Lainscsek et al., 2017]. The coefficients $a_{1,2,3}$ were used as features to distinguish dynamical differences in time series data. Together with these coefficients, the

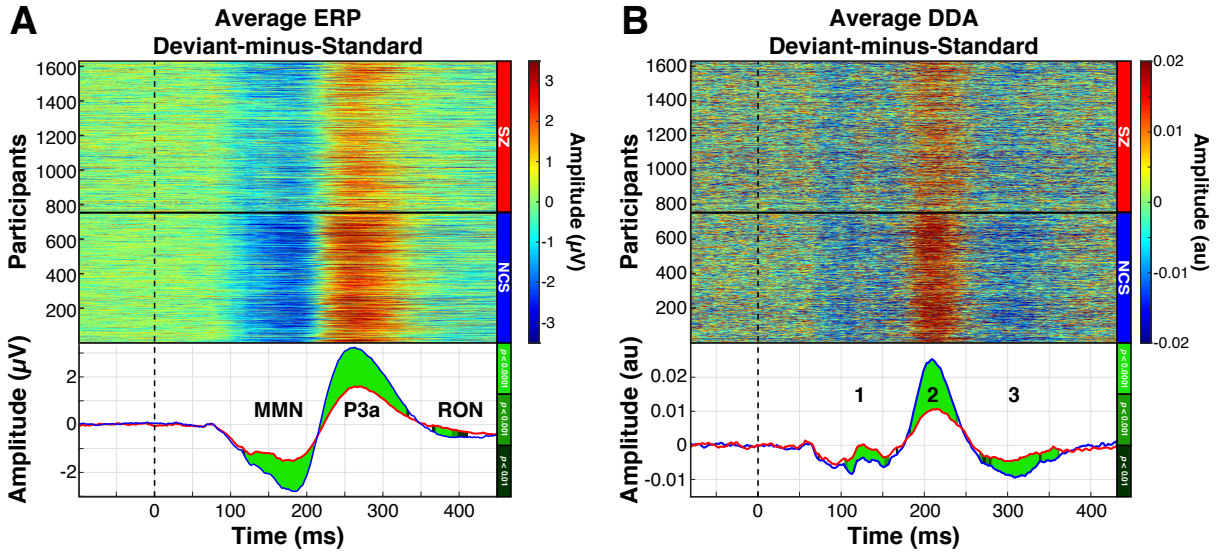


Figure 2.1: DDA identified dynamical state changes preceding each of the auditory deviance response complex components. **A.** NCS subjects demonstrated robust MMN, P3a, and RON components as shown in the heatmap of the individual subject difference (deviant tone ERP - standard tone ERP) average signals (top panel). All three components of the ADR (MMN, P3a, and RON) can be appreciated in the group-level average signals (bottom panel). **B.** DDA a_3 coefficient values averaged within each subject revealed significantly decreased a_3 in the SZ patients (top panel). As with the ERP results, the DDA group averages displayed three components with homologous waveform morphology and severity of deficits in schizophrenia, but the DDA components (numbered in the bottom panel) preceded their corresponding ERP peaks identified in **A** by 71, 54, and 82 ms respectively. The shaded regions in the group average signals represent group differences that are statistically significant after adjusting for multiple comparisons (FDR). Cohen's d , t -values, and degrees of freedom are shown in Supplementary Information Fig. A.9 in the Appendix.

least squares error ρ was also considered. Therefore, the full DDA feature set was composed of $\{a_1, a_2, a_3, \rho\}$. These previous studies have shown that the four DDA parameters of Eq. (2.2) reflect both linear and nonlinear features underlying nonlinear signals [Lainscsek and Sejnowski, 2015]. Since previous results converged on a_3 as most informative [Lainscsek and Sejnowski, 2015], this parameter was carried forward for detailed characterization in this dataset.

2.3 Results

2.3.1 Auditory Deviance Response Complex

Consistent with established methods [Light et al., 2015], individual subject deviant minus standard waveform averages derived from the preprocessed ERP signals showed significantly reduced mismatch negativity (MMN) and P3a positivity components for the SZ patients (Fig. 2.1A top panel). The group averaged signals (Fig. 2.1A bottom panel) also revealed significantly diminished auditory deviance response (ADR) complex, composed of MMN, P3a positivity, and reorienting negativity (RON) for the SZ group.

Performing the DDA analysis on the difference waveforms and computing individual subject a_3 averages showed significantly lower a_3 values across the SZ patients in the 180–240 ms time window (Fig. 2.1B top panel, mean t-value = 10.8, mean Cohen’s d = 0.55). The group averaged a_3 signals revealed three distinct areas corresponding to the ADR components (numbered areas in Fig. 2.1B bottom panel). Interestingly, the timing of the three DDA peaks occurred before their corresponding ADR components identified in the group averaged ERP signals. For instance, the peak group difference in area 1 in Fig. 2.1B occurred 70 ms before the peak group difference in the ERP MMN window in Fig. 2.1A. This suggests that DDA reflects dynamical state changes (as estimated by the amplitude changes in a_3) occurring immediately before each of the ERP ADR components. Thus, DDA not only captures the group differences, but also identifies the underlying dynamical state changes that contribute to the linear ERP changes of the ADR complex components.

2.3.2 Responses to Constituent Deviant and Standard Tones

To determine whether findings in the deviant-minus-standard difference waves were driven by effects to constituent tones, responses to deviant and standard tones were also examined separately. Consistent with the results presented above, similar waveform morphology and

magnitude of schizophrenia impairment were observed in both ERP and DDA results. Also consistent with the previous section, peaks observed with DDA, including those related to schizophrenia deficits, occurred earlier than those observed in the ERP analysis.

Responses to Deviant Tones

Averaging the ERP signals corresponding to only deviant tones revealed three ERP components (Fig. 2.2A) that were reduced for the SZ group consistent with other studies [Light et al., 2015]. DDA identified dynamical changes unique to each group occurring before each of the ERP components (Fig. 2.2B). In addition, the dynamical states of the NCS group were significantly different from those of the SZ patients during the 50 – 100 ms time window (black arrow in Fig. 2.2B, mean t -value = -2.6, mean Cohen's d = -0.13). Notably, this window corresponds to the earliest time window a stimulus can be identified as deviant, since the deviant auditory tones were 100 ms in duration (vs. 50 ms for the standard tones). No such difference was observed in the grand average ERP signals (Fig. 2.2A).

Responses to Standard Tones

Unlike the robust group differences observed in the ERP responses to the deviant tones, the group differences in the standard-tone ERP responses were much more modest as shown in Fig. 2.3A. DDA identified a similar group difference occurring between 90–130 ms (Fig. 2.3B). Consistent with ERP analysis of the standard tones, no significant dynamical changes were detected in the 50–100 ms time window.

2.4 Discussion

In this study, DDA, a novel technique based on dynamical systems theory, was used to determine whether nonlinear features significantly contribute to leading neurophysiologic

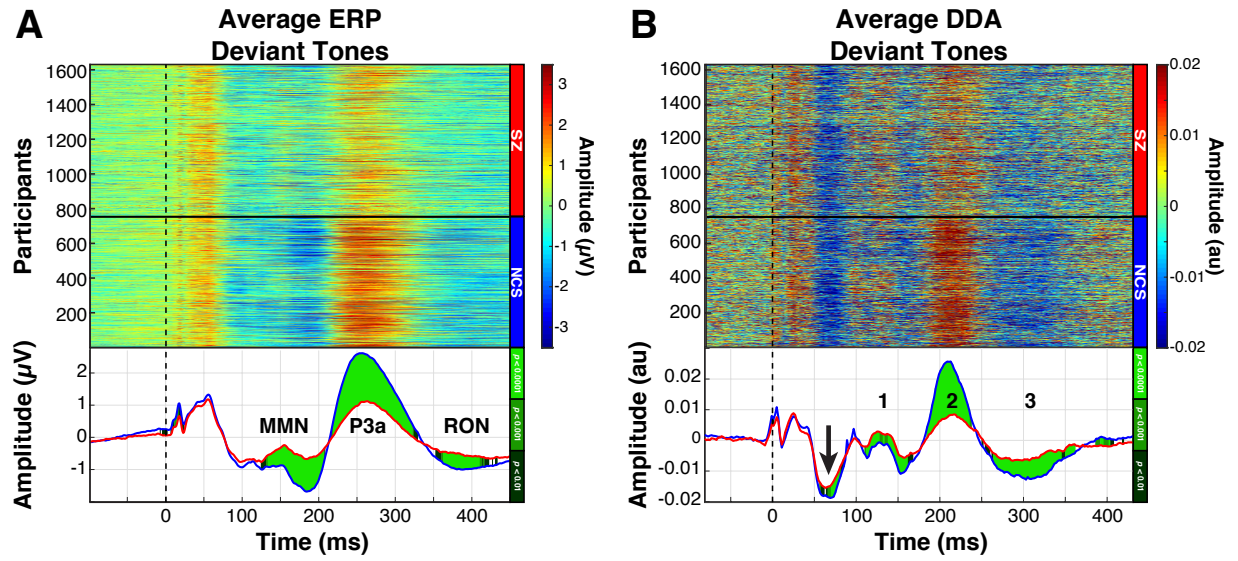


Figure 2.2: ADR components were evident from both ERP and DDA analyses of waveforms elicited by deviant auditory tones. **A.** Individual subject average signals from deviant tones revealed pronounced MMN and P3a components in the NCS subjects (top panel). The grand average signals from the two groups showed significantly different ADR component signals (bottom panel). **B.** DDA a_3 amplitude changes were observed before each of the ADR components (numbered in the bottom panel). Furthermore, statistically significant differences in a_3 amplitude between groups were detected during the earliest time window (50–100 ms) an auditory tone could be processed as deviant (black arrow, mean t-value = -2.6, mean Cohen's d = -0.13).

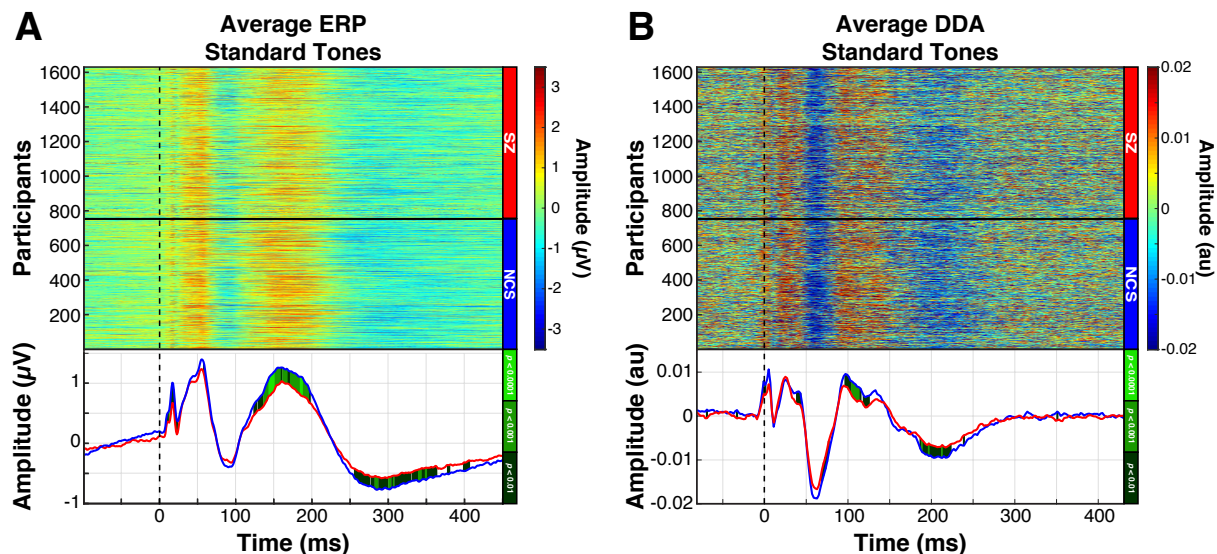


Figure 2.3: DDA of standard tones detected significant dynamical state alterations preceding the P150 window. **A.** ERP signals corresponding to the standard tones revealed a reduced P150 component in the SZ grand average signal (bottom panel). **B.** Dynamic changes preceding the P150 ERP changes were observed in the 90–130 ms window (bottom panel; mean t -value = 3.3, mean Cohen’s d = 0.17). The a_3 changes were reduced for the SZ group during this time window.

biomarker candidates of early auditory information processing. The concept that nonlinear, chaos theory-based perturbations occur early in information processing in SZ patients has been hypothesized for some time [Paulus and Braff, 2003]. In this context, these early dynamical state changes precede commonly studied ERP features in both groups. SZ-related impairments in DDA metrics may contribute to impairments in higher order neurocognitive and psychosocial functioning. A challenge has been with the identification and quantification of these early nonlinear abnormalities in the information processing cascade of function [Thomas et al., 2017]. DDA may provide a solution to this longstanding challenge.

DDA was employed on large cohorts of 877 SZ patients and 753 NCS using a well-validated paradigm from the Consortium on the Genetics of Schizophrenia multi-center study. DDA revealed nonlinear dynamical state changes, which preceded well-established auditory deviance response components of the ERP.

Importantly, the dynamical information extracted from EEG may also provide additional

information about the underlying nature of the MMN, P3a, and RON components of the ADR. Interestingly, nonlinear features are evident nearly instantaneously in response to standard and deviant auditory stimulus onset as well as in the deviant minus standard comparison that is commonly used for measuring the ADR components. DDA detected nonlinear activity > 50 ms prior to conventional ERP peaks. Since this early DDA effect does not appear to be an artifact of the analysis method, questions remain as to the functional significance of these early non-linear effects. It is possible that the nonlinear features captured via DDA provide a more proximal and direct readout of the neurophysiological substrate of ERP measures and the earliest detectable changes in the brain representing deviance detection.

DDA utilizes embeddings to reveal the underlying dynamical patterns of brain activity in the feature space. These patterns, which are otherwise not observable in the EEG using conventional linear ERP methods, are sensitive to even small perturbations in brain network dynamics—including those on a small, regional scale. The extracted features also provide a natural measure of the magnitude of dynamical changes, which may correspond to the severity of dysfunction in neuropsychiatric illness. In addition to revealing important information about nonlinear dynamics hidden in the EEG signals, DDA offers a number of important technical advantages over commonly used ERP measures: it is computationally fast, provides fine temporal resolution (10 ms data windows), and requires only minimal preprocessing (i.e., demeaning of the data) to avoid discarding meaningful contextual information. Since time-consuming preprocessing steps (e.g. blink correction and filtering) are not required for DDA, even very large datasets can be analyzed in minutes, rather than days, weeks, or even months.

Although DDA has great promise, several caveats should be noted. As a new measure, DDA findings do not neatly map onto the vast existing literature, which has nearly exclusively relied upon linear methods; DDA findings therefore may be challenging to interpret. Likewise, the DDA methods have not undergone the same degree of validation for use in clinical studies (e.g., reliability, suitability for use as a repeated measure, sensitivity to therapeutic interventions).

As with the original papers describing the underlying data used in this report, another caveat is that medications in the schizophrenia patients were not experimentally controlled. It is possible that some of the observed deficits in the DDA features from the SZ patients could be attributable to differences in their medication status. Likewise, other clinical factors could also influence DDA findings. It is noteworthy that the large COGS-2 dataset allows for additional tests to verify that the observed SZ-related effects are not driven by group differences in age or sex (see Supplementary Information Fig. A.10 and A.11 in the Appendix). Future studies are required to disentangle medication effects and interacting clinical variables from the current DDA findings in SZ [Thomas et al., 2017, Light et al., 2015]. On the other hand, if DDA is sensitive to medication or clinical changes, this rapid approach may be useful for biomarker-guided clinical trial approaches.

Despite the caveats of this study, these results represent a clear, novel, and powerful demonstration of the potential benefits of using nonlinear techniques to study important nonlinear aspects of natural phenomena, including neural function. By applying DDA to this large, rich data set, we were able to obtain linear and nonlinear features related to auditory information processing in healthy subjects and corresponding deficits observed in SZ patients. Elucidating the links among DDA features and brain network dynamics can lead to a better understanding of pathophysiology of cognitive impairment in SZ and neuropsychiatric disorders as well as contribute to novel biomarker-guided strategies to accelerate the development of treatments for CNS disorders.

2.5 Materials and Methods

2.5.1 COGS-2 Data Collection

Data were collected at five centers across the United States: the University of California San Diego, University of California Los Angeles, University of Washington, University of

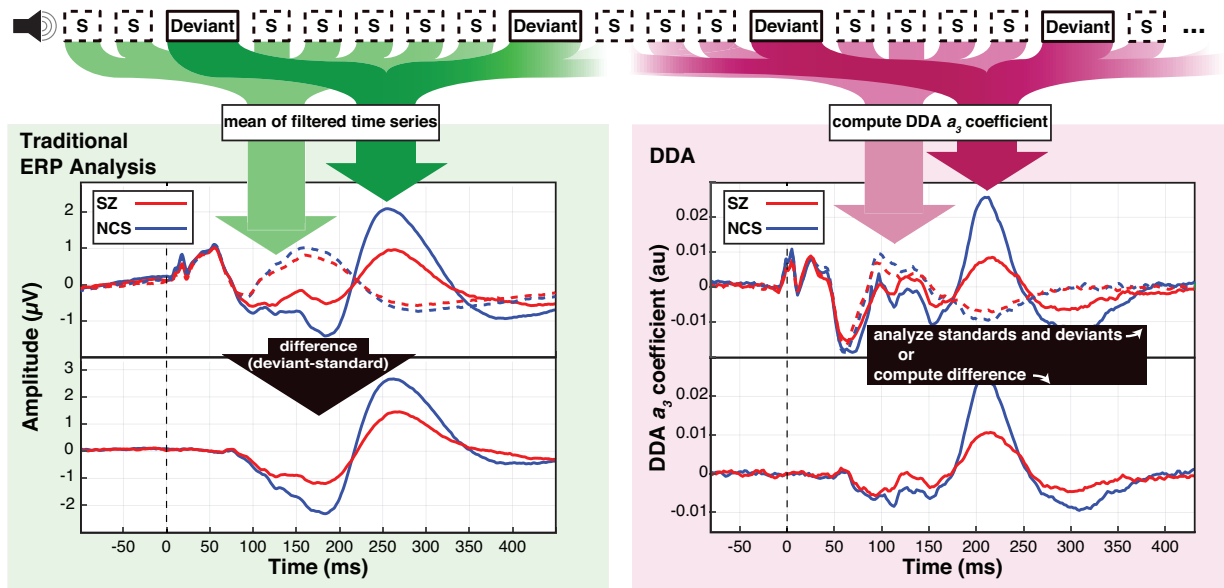


Figure 2.4: Subjects in both SZ and NCS groups were presented with auditory stimuli consisting of 50 ms standard tones with randomly interspersed 100 ms deviant tones. Data were analyzed according to a traditional ERP paradigm (left) and using DDA to assess nonlinear dynamics (right). In the standard ERP approach, EEG time series from standard (dashed lines) and deviant trials (solid lines) were averaged after preprocessing, and the difference waveform was computed from these averages. With DDA, a three-term delay differential equation was applied to the data, and the nonlinear coefficient a_3 was computed for both standard and deviant trials. Both ERP difference waveforms and the DDA a_3 coefficient timecourses are aligned to tone onset, marked with a vertical dashed line at 0 ms.

Pennsylvania, and Mount Sinai School of Medicine. Written consent was obtained from all the participants, and the study was approved by the local human research protection committees at each cite. Samples from a total of 1630 COGS-2 subjects were analyzed: 877 SZ patients and 753 NCS subjects. Trains of auditory stimuli presented to each subject consisted of 50 ms standard tones (90% of stimuli) and 100 ms deviant tones (10% of stimuli) as shown in Fig. 2.4. There were at minimum 1 standard and maximum 18 standard tones between deviants. EEG recordings were conducted at a sampling rate of 1 kHz from a single electrode at the CZ (central zero) position.

2.5.2 ERP Analysis

The COGS-2 data for each subject were segmented into trials with duration of 550 ms. Each trial contained 100 ms pre-tone and 450 ms post-tone ERP signals. Trials corresponding to deviant tones ($n = 150$) and standard tones ($n = 150$) were extracted from each subject. Eye movement artifact was removed using a second-order blind identification (SOBI) algorithm, and trials containing residual artifact (signal activity $\geq \pm 50 \mu V$) were discarded. Finally, the average standard tone responses were subtracted from the average deviant tone responses.

2.5.3 DDA of EEG data

DDA features are usually estimated from a data window of length L as shown in Fig. 2.5. The data for each such window are normalized to zero mean and unit variance to remove amplitude information. Sliding overlapping windows are then applied to the data. For the COGS-2 data, windows of 10 ms were used, and the corresponding points are plotted at the beginning of each time-window for all the figures. Fig. 2.5 shows the single trial version of DDA: The feature set $\{a_{1,2,3}, \rho\}$ is estimated from each trial and then the mean over several trials is taken as nonlinear counterpart to traditional ERP where the mean of the data (after pre-processing) is

$$\dot{x} = a_1 x_1 + a_2 x_2 + a_3 x_1^2$$

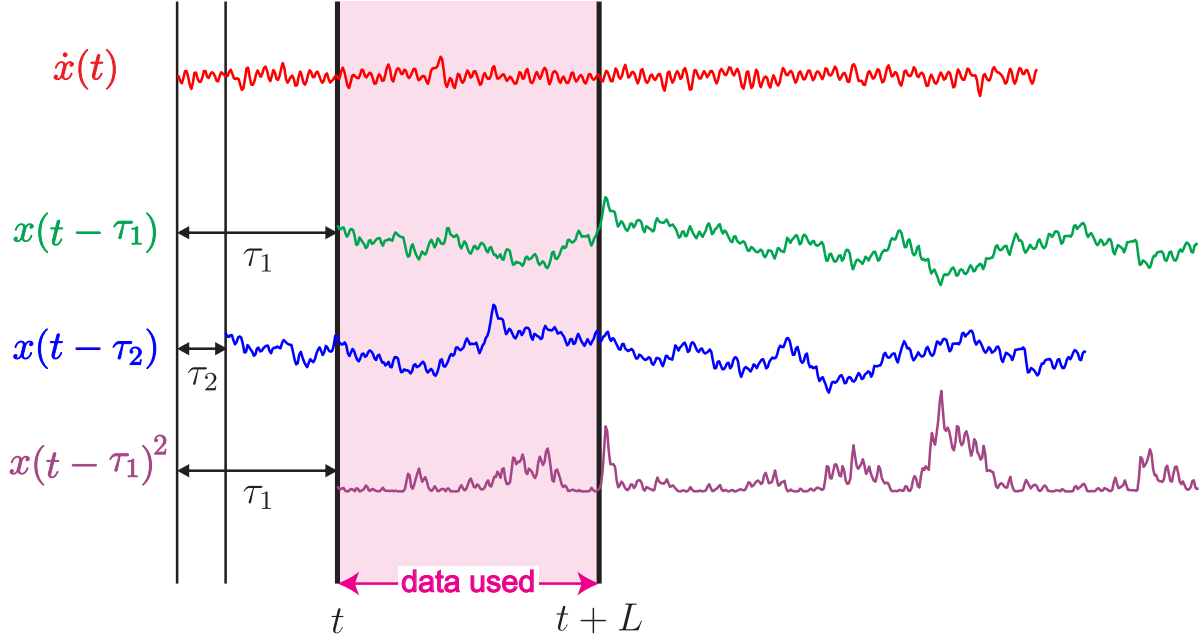


Figure 2.5: Estimation of the features $a_{1,2,3}$ for a data window of length L for Eq. (2.2).

used. The difference is that for DDA, the raw data are processed for each data window separately. To improve the performance and to reduce the number of trials needed, cross-trial DDA was used (see Supplementary Information Fig. A.3 in the Appendix). Using the ergodic hypothesis [Boltzmann, 1898], data windows of multiple trials were combined and features were computed across trials. The DDA coefficients were estimated simultaneously from 150 trial-windows (see supplementary figures). Supervised structure selection was used to identify the delays that led to optimal group classification (A') [Lainscsek et al., 2013c]. Note that these delays have no frequency correspondence due to the nonlinear nature of the DDA model [Lainscsek and Sejnowski, 2015]. Results from a_3 , the most salient feature, are presented. As a measure of classification performance, the area A' under the receiver operating characteristic (ROC) was used [Hand and Till, 2001]. An unpaired student's t-test was used to assess the significance of

differences between the SZ and NCS groups as observed in both traditional, trial-mean difference waveforms and the DDA a_3 coefficient. The false discovery rate was adjusted using the Benjamini-Hochberg procedure. The optimal delays for the COGS-2 dataset were $\tau = (3, 8)$ time points, corresponding to (3,8) ms.

Acknowledgements

This work was supported in part by the NIH grants R01-MH065571, R01-MH042228, R01-MH079777, K23-MH102420, R01-MH065554, R01-MH65707, R01-MH65578, and R01-MH65558, R01-MH059803, R01-MH094320. This work was supported by the Sidney R. Baer, Jr. Foundation, the Department of Veterans Affairs, the Medical Research Service of the Veterans Affairs San Diego Health Care System, the Department of Veterans Affairs DesertPacific Mental Illness Research, Education, and Clinical Center (MIRECC), the Brain and Behavior Research Foundation.

The COGS Investigators: Gregory A. Light^{a,b}, Neal R. Swerdlow^a, Monica E. Calkins^c, Michael F. Green^{d,e}, Tiffany A. Greenwood^a, Raquel E. Gur^c, Ruben C. Gur^c, Laura C. Lazzeroni^f, Keith H. Nuechterlein^d, Allen D. Radant^{g,h}, Larry J. Seidman^{i,j}, Larry J. Siever^{k,l}, Jeremy M. Silverman^{k,l}, Joyce Sprock^a, William S. Stone^{i,j}, Catherine A. Sugar^m, Debby W. Tsuang^{g,h}, Ming T. Tsuang^{a,n,o}, Bruce I. Turetsky^c, David L. Braff^{a,b}

^aDepartment of Psychiatry, University of California San Diego, La Jolla, CA ^bVISN-22 Mental Illness, Research, Education and Clinical Center (MIRECC), VA San Diego Healthcare System

^cDepartment of Psychiatry, University of Pennsylvania, Philadelphia, PA ^dDepartment of Psychiatry and Biobehavioral Sciences, University of California Los Angeles, Los Angeles, CA

^eVA Greater Los Angeles Healthcare System, Los Angeles, CA ^fDepartment of Psychiatry and Behavioral Sciences, Stanford University, Stanford, CA ^gDepartment of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA ^hVA Puget Sound Health Care System,

Seattle, WA ⁱDepartment of Psychiatry, Harvard Medical School, Boston, MA ^jMassachusetts Mental Health Center Public Psychiatry Division of the Beth Israel Deaconess Medical Center, Boston, MA ^kDepartment of Psychiatry, The Mount Sinai School of Medicine, New York, NY ^lJames J. Peters VA Medical Center, New York, NY ^mDepartment of Biostatistics, University of California Los Angeles School of Public Health, Los Angeles, CA ⁿCenter for Behavioral Genomics, and Institute for Genomic Medicine, University of California San Diego, La Jolla, CA ^oHarvard Institute of Psychiatric Epidemiology and Genetics, Boston, MA

Chapter 2, in full, is a reprint of the material as it appears in the Proceedings of the National Academy of Sciences 2019. Claudia Lainscsek, Aaron L. Sampson, Robert Kim, Michael L. Thomas, Karen Man, Xenia Lainscsek, The COGS Investigators, Neal R. Swerdlow, David L. Braff, Terrence J. Sejnowski, and Gregory A. Light. The dissertation author was a primary investigator and co-author of this paper.

Chapter 3

Delay Differential Analysis for Dynamical Sleep Spindle Detection

Abstract

Background

Sleep spindles are involved in memory consolidation and other cognitive functions. Numerous automated methods for detection of spindles have been proposed; most of these rely on spectral analysis in some form. However, none of these approaches are ideal, and novel approaches to the problem could provide additional insights.

New Method

Here, we apply delay differential analysis (DDA), a time-domain technique based on nonlinear dynamics to detect sleep spindles in human intracranial sleep data, including laminar electrode, stereoelectroencephalogram (sEEG), and electrocorticogram (ECoG) recordings.

Results

We show that this approach is computationally fast, generalizable, requires minimal preprocessing, and provides excellent agreement with human scoring.

Comparison with Existing Methods

We compared the method with established methods on a set of intracranial recordings and this method provided the highest agreement with human expert scoring when evaluated with F_1 score while being the second-fastest to run. We also compared the results on the DREAMS surface EEG data, where the method produced a higher average F_1 score than all other tested methods except the automated detections published with the DREAMS data. Further, in addition to being a fast and reliable method for spindle detection, DDA also provides a novel characterization of spindle activity based on nonlinear dynamical content of the data.

Conclusions

This additional, non-frequency-based perspective could prove particularly useful for certain atypical spindles, or identifying spindles of different types.

3.1 Introduction

3.1.1 Sleep Spindles

Sleep spindles are discrete events consisting of 11 to 16 Hz oscillations (the precise frequency range varies across subjects) recorded primarily in stage 2 non-REM sleep, and to a lesser extent in stage 3 non-REM sleep [Berry et al., 2012]. Spindles display a characteristic waxing and waning pattern in amplitude, and generally last between 0.3 and 3 seconds, recurring every 5 to 15 seconds [Bonjean et al., 2012, Leresche et al., 1991]. Sleep spindles arise from the

activity of thalamocortical circuitry. They have become a subject of study for their potential roles in memory consolidation and other cognitive functions [Sejnowski and Destexhe, 2000, Schabus et al., 2004, Fogel et al., 2007], as well as in psychiatric and neurological disorders [Ferrarelli et al., 2007, Petit et al., 2004, Ktonas et al., 2007].

Numerous methods for automated spindle detection have been proposed, most of which rely on spectral analysis in some form [Warby et al., 2014, O’Reilly and Nielsen, 2015]. Here, we propose an alternative approach using a nonlinear time-domain algorithm which is computationally fast and therefore capable of detecting spindles in real time.

3.1.2 Delay Differential Analysis

Delay differential analysis (DDA) is a time-domain classification framework based on embedding theory in nonlinear dynamics [Kremliovsky and Kadtke, 1997, Lainscsek et al., 2013a]. An embedding reveals the nonlinear invariant properties of an unknown dynamical system (here the brain) from a single time series (here intracranial recordings). The embedding in DDA serves then as a sparse nonlinear functional basis onto which the data are mapped. Since the basis is built on the dynamical structure of the data, preprocessing (such as filtering) is not necessary. DDA yields a small number of features (around 4), far fewer than traditional spectral techniques, which provide a power at each frequency (often 100-200 frequencies). In either case, the size of the feature set might vary depending on the parameters used. Also, either set of features can be combined or collapsed to yield a measure that can be thresholded. However, working with a constrained feature space is often desirable. This approach greatly reduces the risk of overfitting, and therefore helps to ensure that a model that was selected using a single EEG channel from one subject can be applied to a wide range of data from different subjects, channels, and recording systems.

One can also view DDA models as sparse Volterra series [Volterra, 1887, Volterra, 1959]. A general nonlinear real-valued function can be expressed as a Taylor series expansion of

functionals of increasing complexity around a fixed point. Rather than retain all low-order terms in the expansion, DDA imposes restricted complexity on the analysis by using a low-dimensional sparse delay differential equation (DDE) model. In a model of this type, linear and nonlinear components of the data are analyzed in an interconnected manner. This reduces the computational load, and further, by leaving some of the non-relevant dynamics unmodeled, it is possible to greatly reduce the effect of artifacts and other signals unrelated to the particular classification task of interest.

DDEs combine differential with delay embeddings as a functional embedding where (non-) linear polynomial functions of the delay terms are used [Lainscsek et al., 2017]. The general form of the DDEs is

$$\dot{x}(t) = \sum_{i=1}^I a_i \prod_{n=1}^N x_{\tau_n}^{m_{n,i}} \text{ for } \tau_n, m_{n,i} \in \mathbb{N}_0 \quad (3.1)$$

where I is the number of monomials in the model, N is the number of delays, $m_{n,i}$ is the order of the n^{th} term in the i^{th} monomial, and x_{τ_n} represents $x(t - \tau_n)$. The time derivative of the data, $\dot{x}(t)$, is estimated with a weighted center derivative [Miletics and Molnárka, 2005]:

$$\dot{x}(t) = \frac{1}{2M} \sum_{m=1}^M \frac{x(t+m) - x(t-m)}{m} \quad (3.2)$$

where M is the number of points used.

For a given model, we compute a small set of features, which are the estimated coefficients a_i in Eq. (3.1) as well as the least-squares error. The error is defined as:

$$\rho = \sqrt{\frac{1}{K} \sum_{k=1}^K \left(\dot{x}_{t_k} - \sum_{i=1}^I a_i \prod_{n=1}^N x_{\tau_n, t_k}^{m_{n,i}} \right)^2} \quad (3.3)$$

where K is the number of time points, and x_{τ_n, t_k} represents $x(t_k - \tau_n)$.

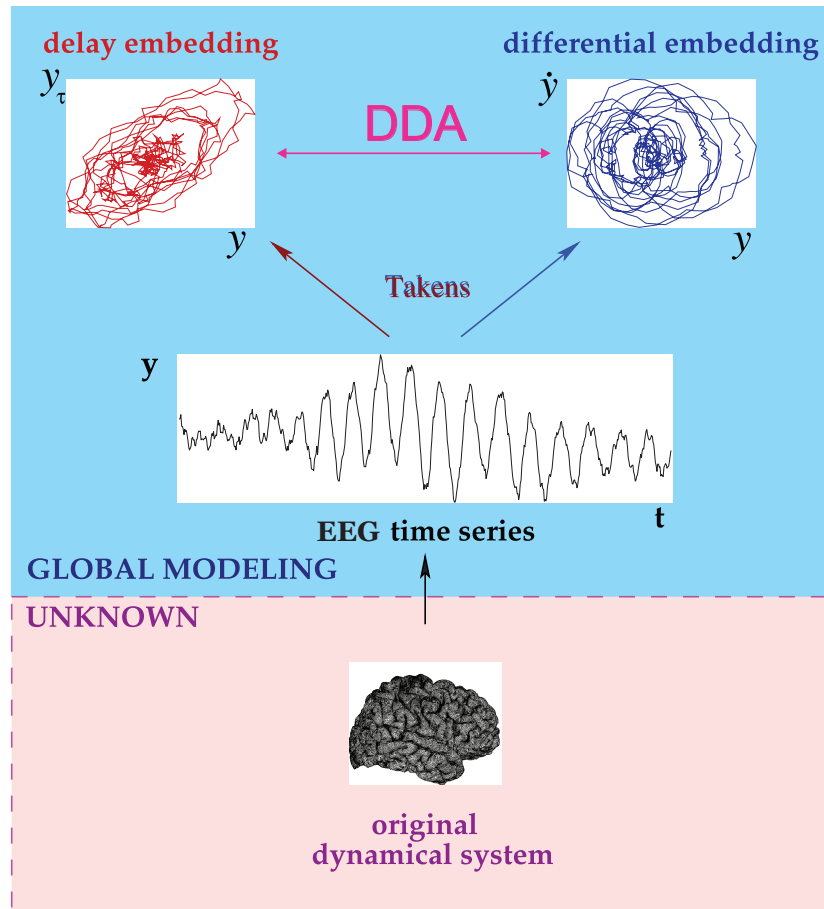


Figure 3.1: Delay Differential Analysis (DDA). For an unknown dynamical system (such as the brain) from which we can record a single variable over time (such as ECoG data), embedding theory states that we can recover the nonlinear invariant properties of the original system. DDA combines delay and differential embeddings in a functional form which allows time-domain classification of the data. For a given polynomial model form, we estimate the coefficients and least-squares error, which form a low-dimensional feature space. This figure is adapted from [Lainscsek and Sejnowski, 2015].

3.2 Methods

3.2.1 Data

DDA was applied to laminar, stereoelectroencephalogram (sEEG), and electrocorticogram (ECoG) recordings from patients with intractable epilepsy.

The laminar recordings studied here come from five patients, designated L1 to L5. Recordings and data were obtained under Institutional Review Board (IRB) approval with informed consent from participants in accordance with the Declaration of Helsinki.

The additional recordings used for this study consisted of sEEG (depth electrode) recordings from five patients, designated S1 to S5, and ECoG recordings from two patients, designated E1 and E2, with long-standing pharmaco-resistant complex partial seizures. These recordings used a standard clinical recording system (XLTEK, Natus Medical Inc., San Carlos, CA) with sampling rates of 500, 512, or 1024 Hz. The reference for the sEEG electrodes was an electrode placed over the C2 spinous process on the posterior neck. For the ECoG (cortical surface electrode) recordings, the reference channel was a strip of electrodes located outside the dura mater and facing the skull at a region remote from other grid and strip electrodes. Placement of the intraparenchymal (sEEG) electrodes and subdural electrode arrays was chosen to confirm the hypothesized seizure focus and locate epileptogenic tissue in relation to essential cortical areas, thus directing surgical treatment.

The decision to implant, as well as the electrode targets and the duration of implantation were entirely clinically based with no input from this research study. All data were handled following protocols approved by the IRB of the Massachusetts General Hospital according to National Institutes of Health guidelines.

sEEG data used for this study consist of three channels from subject S1, four channels from subject S2, one channel each from subjects S3 and S5, and two channels from subject S4. ECoG data used here consist of three channels from subject E1 and one channel from subject E2.

All data selected for use in this study were exclusively from stage two sleep, during time periods when no seizures were occurring.

3.2.2 Spindle Marking

Both the data used for developing the detector and those used for testing were drawn from human expert-scored intracranial recordings: 23-channel laminar electrodes in five subjects (L1-L5) and single-channel scored sEEG and ECoG recordings from subjects S1-S5 and E1-E2. In the laminar data set, the scorer marked a single time point for each identified spindle based on evaluation of all 23 channels (here designated type I scoring). In the sEEG and ECoG data, the beginning and end of all spindles were marked on the basis a single channel (type II scoring). In type II scoring, therefore, the beginnings of spindles are defined as the point where spindle oscillations become visually apparent to the scorer, and the end is defined as the point where these oscillations are no longer apparent. Also, in type II scoring, the scorer marked all potential spindles, regardless of clarity. By including both types of human scoring as well as a range of spindle quality, we aim to develop a robust detector that can function even with non-ideal data.

Since only a single time point was marked in type I scoring, a window of one second around each marker was taken as the spindle (that is, the beginning of each spindle was defined as 0.5 seconds before the mark and the end was defined as 0.5 seconds after the mark), and a wider window of one to three seconds around each marker was excluded from classification as non-spindle data (only data at least 1.5 seconds before or after a mark were considered non-spindle data). Table 3.1 summarizes the properties of the marked spindles in both data sets: the recording type (laminar electrodes, sEEG, or ECoG), the scoring type (I or II), the sampling rate f_s , the number of marked spindles, the mean spindle duration, and the mean peak frequency (between 11 and 17 Hz) for all spindles in each recording. Since type I scoring involved marking spindles on the basis of multiple channels, the peak frequencies are computed as the mean of the peak frequency across the five channels in which spindles are most visually apparent. The peak

frequencies for all channels for each subject are plotted in Fig. 3.2.

3.2.3 Supervised Structure Selection

Structure selection of the model ultimately relied on data from one channel from one subject. Since DDA uses specific time delays, adjustments need to be made for sampling rate, and to facilitate this, the model (polynomial form and delays) was selected using data with the lowest sampling rate in the available data set (this allows for easy adjustment to higher sampling rates). Here, we used an sEEG recordings sampled at 500 Hz. Data from these subjects and channels were divided into half-second epochs and marked as spindle or non-spindle based on how each epoch had been marked by a human expert in the manner described above. Among these 500 Hz recordings, the one for which spindle and non-spindle epochs proved most separable was used to select a model for use with new data.

In order to select the model from these training data, the set of models to be considered was first subjected to constraints based on model forms that had proven effective in previous applications of DDA, ensuring the sparsity of the model. The general form of the model shown in Eq. (3.1) was constrained to two delays ($N \leq 2$), three terms ($I = 3$), and up to third-order nonlinearities ($\sum_n m_{n,i} \leq 3$). This resulted in a total of 188 unique DDE model forms, upon which we performed an exhaustive search. The delays τ_1 and τ_2 were allowed to vary between approximately 1 and 80 ms at intervals of $1/f_s$.

We performed repeated random subsampling cross-validation [Kohavi et al., 1995] to evaluate the performance of each model. This method involves repeatedly dividing the data at random into training and testing sets. (Note that throughout we use the terms “training” and “testing” to refer to these repeated random splits of the data for cross-validation. New data, not used in the structure selection of a particular model, are referred to as “validation” data.) This prevents overfitting of the model and ensures generalizability. Here, the repeated random splits were carried out for the model selection data, assigning 70% of spindle and non-spindle epochs

Table 3.1: Human-marked spindle properties for the fifteen recordings.

Subject	Channel**	Type	Scoring	f_s [Hz]	Number	Mean duration [s]	Mean peak freq. [Hz]
L1	1-23, left frontal	Laminar	I	2000	144	1*	15.0580
L2	1-23, right frontal	Laminar	I	2000	48	1*	11.8063
L3	1-23, right frontal	Laminar	I	2000	137	1*	12.8836
L4	1-23, right frontal	Laminar	I	2000	50	1*	12.4320
L5	1-23, right temporal	Laminar	I	2000	72	1*	13.2750
S1	1 (RCIN3)	sEEG	II	500	57	0.84	12.5395
S1	2 (LCIN4)	sEEG	II	500	135	0.91	12.8115
S1	3 (LSF6)	sEEG	II	500	47	0.72	12.6363
S2	1 (LCIN3)	sEEG	II	500	213	1.79	12.7073
S2	2 (LSF3)	sEEG	II	500	218	1.42	13.1963
S2	3 (RCIN5)	sEEG	II	500	146	1.25	12.9723
S2	4 (LFR1)	sEEG	II	500	227	1.57	12.3713
S3	1 (OF7)	sEEG	II	500	138	0.87	12.7769
S4	1 (RPF5)	sEEG	II	512	152	1.15	12.7569
S4	2 (ROF4)	sEEG	II	512	81	0.98	13.9615
S5	1 (RAF6)	sEEG	II	512	124	0.96	13.0326
E1	1 (GR28)	ECoG	II	512	82	1.05	12.4093
E1	2 (GR53)	ECoG	II	512	13	1.36	11.7415
E1	3 (GR38)	ECoG	II	512	92	1.18	13.2799
E2	1 (AGR52)	ECoG	II	1024	47	0.71	12.1440

*The mean duration cannot be determined from Type I scoring because only a single time point was marked across all channels (1-23). One second of data is designated as spindle data for structure selection.

**RCIN–right cingulate, LCIN–left cingulate, LSF–left subfrontal, LFR–left frontal, OF–orbitofrontal, RPF–right posterior frontal, ROF–right orbitofrontal, RAF–right anterior frontal, GR–grid (subject E1 grid channels 28, 38, and 53 were all located over posterior frontal cortex with 28 the most inferior and 53 the most superior), AGR–anterior grid (subject E2 anterior grid channel 52 was located over middle posterior frontal cortex)

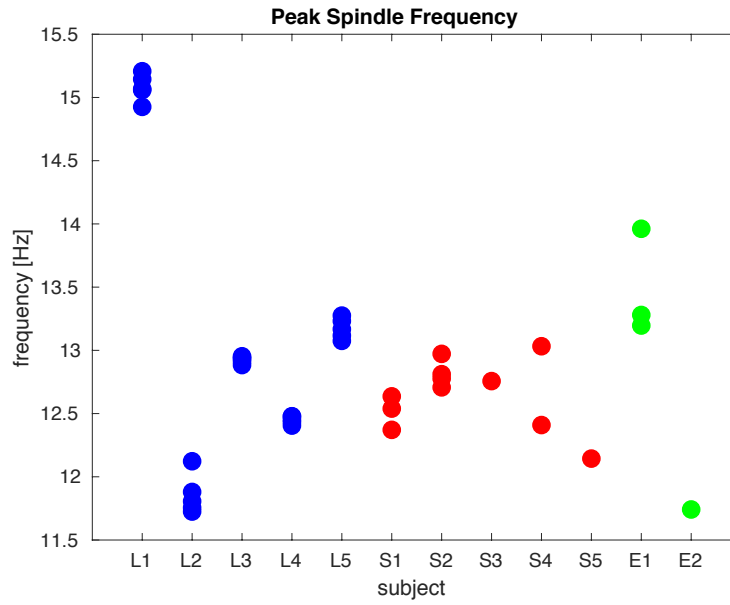


Figure 3.2: Spindle frequencies. For each of the five laminar, five sEEG, and two ECoG subjects, the peak frequency (between 11 and 17 Hz) was computed for all human-marked spindles and the mean peak frequency across all spindles is plotted as one point for each channel. For laminar subjects, five of the channels are plotted—spindles were marked based on evaluation of all channels. For the sEEG and ECoG subjects, spindles were marked on an individual-channel basis, and all scored channels are plotted. Color indicates the type of recording. Note that laminar recordings were collected from cortex identified as probably epileptogenic.

to the training set, and the remaining 30% to the testing set. Using the model coefficients $a_{k,i}$ and error ρ_k obtained from each epoch k of the training data, we used the human expert-scored labels l_k (i.e. 0 for non-spindle and 1 for spindle) to obtain a vector of weights W for the features by finding a least-squares solution to:

$$\begin{pmatrix} 1 & a_{1,1} & a_{1,2} & a_{1,3} & \rho_1 \\ 1 & a_{2,1} & a_{2,2} & a_{2,3} & \rho_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & a_{k,1} & a_{k,2} & a_{k,3} & \rho_k \end{pmatrix} W = \begin{pmatrix} l_1 \\ l_2 \\ \vdots \\ l_k \end{pmatrix}. \quad (3.4)$$

The additional constant term avoids constraining the separating hyperplane to pass through the origin in feature space. The weights W can be applied to the features computed from the testing data which provides a one-dimensional distance D from an optimal hyperplane of separation between spindle and non-spindle feature sets. We can evaluate how well this distance corresponds to the human expert-scored labels of the testing data by computing the area under the receiver operating characteristic (ROC) curve or F_1 score. The ROC is constructed by plotting the hit rate against the false alarm rate for various spindle detection thresholds for D . The area under the curve defined by the plotted points, A' , should be equal to 0.5 for random chance detection, and 1 for perfect separation of the groups [Hand and Till, 2001]. A' can be obtained by taking

$$A' = \frac{S_0 - n_0(n_0 + 1)}{2n_0n_1} \quad (3.5)$$

where n_0 and n_1 represent the number of points in each of two classes labeled 0 and 1 (here, non-spindle and spindle epochs), and S_0 is obtained by first ranking all points by their probability of being classified as 0, then summing the ranks of the true class 0 points. In practice, once a specific model form has been selected, it is often sufficient to use a single feature for classification.

While A' is useful for structure selection of the DDA model, we evaluate final performance

with another measure, the F_1 score, which is more widely used for evaluating spindle detection [Dice, 1945, Sørensen, 1948]. F_1 scores are computed from the confusion matrix according to:

$$F_1 = \frac{2TP}{FN + FP + 2TP} \quad (3.6)$$

where TP is the number of true positives, FN is the number of false negatives, and FP is the number of false positives. For this purpose, the human scoring is considered the “ground truth”. F_1 scores are used in Sec. 3.3.1 for comparison between the outputs of several spindle detection methods. As additional measures, we also compute the false discovery rate ($FDR = \frac{FP}{TP+FP}$) and false negative rate ($FNR = \frac{FN}{FN+TP}$).

The cross-validation was repeated 100 times and the maximal A' was used to select the optimal model form and values of the delays. Using this procedure, for spindle detection in the laminar, sEEG, and ECoG data at all sampling rates, an effective DDE model is:

$$\dot{x} = a_1 x_{\tau_1} + a_2 x_{\tau_2} + a_3 x_{\tau_1}^2 \quad (3.7)$$

with $\tau_1 = 16 \delta t = 32$ ms and $\tau_2 = 25 \delta t = 50$ ms for 500 Hz data. For spindle detection, we find that the single feature a_2 provides sufficient information for good detection performance. In general, the threshold for spindle detection is set to 1.2 standard deviations above the mean of a_2 . This threshold has been empirically determined to provide good agreement with human scoring and was fixed throughout.

Despite the fact that these data come from subjects with different types of electrodes and different sampling rates, it is possible to obtain spindle detection which agrees with human scoring across multiple recordings as well as multiple human scorers would tend to agree with each other [Basner et al., 2008]. Because we use nonlinear models, all terms are connected and linear as well as nonlinear terms contain both linear and nonlinear information. For this reason the delays do not correspond to particular frequencies as one might expect [Lainscsek and Sejnowski,

2015]. Adjustments need to be made for data with different sampling rates. In order to apply a selected DDA model to data with a higher sampling rate, we need to change the delays and derivatives in the following way: The delays can be just the approximate multiples (e.g. from 500 Hz to 1000 or 1024 Hz they would be doubled). For the derivatives we keep the number of total points constant but take for this example every second data point. For data with lower sampling rates (e.g. the DREAMS data in Sec. 3.3.1), results can only be obtained by upsampling the data to the minimum sampling frequency of 500 Hz before applying the model.

3.2.4 Application to Full-Time Data

Having selected a model form and delay pair according to the above procedures, we compute the corresponding a_2 coefficient in sliding time windows across the full length of all recordings. We use windows of length around 650 ms, shifted by around 200 ms per step. Since the number of spindle and non-spindle epochs in the training data are not equal, the optimal threshold for spindle detection may vary slightly between recordings. Nevertheless, for the sake of testing a fully automated method, we maintained the aforementioned 1.2 standard deviation above mean a_2 threshold for all results shown here. The beginning of each detected spindle is therefore defined as the point at which the normalized a_2 value increases this threshold, and the end is defined as the point at which it subsequently decreases below the threshold. (Note that threshold-setting does not affect A' , since this is a threshold-independent measure, but does determine the F_1 scores, which are computed from the confusion matrix for a particular threshold.) As a final step, any threshold crossings less than 300 ms in length are excluded and marked as non-spindle. The remaining threshold-crossings are the identified spindles. We evaluate detector performance by comparing these time points identified as spindle by the detector with those identified by the human expert.

3.3 Results

Applying the detector to laminar, sEEG, and ECoG data, we obtain a mean area under the ROC curve, A' , of 0.82 and a mean F_1 score of 0.50. For the laminar data, we take just one central channel from each electrode array for evaluating all methods. Since these data were scored based on all channels, but some superior and inferior channels lacked clearly visible spindles, one of the channels (channel 11) with apparent spindles was chosen for evaluating spindle detection performance. All available (individually scored) sEEG and ECoG channels were used. For comparison, DDA frequency-band detectors (discussed in Appendix 3.4) for 11-14 Hz and 11-17 Hz yield mean A' values of 0.72 and 0.77 and mean F_1 scores of 0.21 and 0.18 respectively. Such a difference in performance indicates that in addition to the frequency characteristics of spindles, nonlinear information might also be relevant. Fig. 3.3 shows the output the data-trained DDA spindle detector. Since the data-trained DDA detector shows higher agreement with human scoring than the frequency-based DDA detector, it is used exclusively for the remainder of the manuscript.

The A' values, F_1 scores, false discovery rates, and false negative rates for the DDA spindle detector on all subjects are listed in Table 3.2. Note that in Sec. 3.3.1, F_1 scores are used to compare methods. Where cross-recording averages are reported, two recordings are excluded since all automated detectors perform poorly, and these were originally selected as recordings that were difficult to score.

3.3.1 Comparison with Established Methods

Warby et al. presented a comparison of several automated methods for spindle detection with scoring by human experts and non-experts [Warby et al., 2014]. Here, we compare the DDA spindle detector to two of the automated methods considered there [Mölle et al., 2002, Martin et al., 2013] and a modified version [Andrillon et al., 2011] of a third [Ferrarelli et al., 2007], as

Table 3.2: DDA spindle detection performance on all recordings

Subject	Channel	A'	F_1	False discovery rate	False negative rate
L1	11	0.6023	0.2685	0.5323	0.8117
L2	11	0.6934	0.2991	0.7107	0.6903
L3	11	0.7423	0.2892	0.4701	0.8011
L4	11	0.7784	0.4948	0.5590	0.4365
L5	11	0.7529	0.3679	0.6682	0.5872
Laminar mean		0.7139	0.3439	0.5881	0.6654
S1	1 (RCIN3)	0.8785	0.5404	0.5924	0.1983
S1	2 (LCIN4)	0.9066	0.7685	0.2340	0.2290
S1	3 (LSF6)	0.8716	0.4345	0.6953	0.2428
S2	1 (LCIN3)	0.9120	0.3464	0.0380	0.7887
S2	2 (LSF3)	0.9170	0.5410	0.0265	0.6254
S2	3 (RCIN5)	0.8514	0.5601	0.1723	0.5768
S2	4 (LFR1)	0.9262	0.3970	0.0386	0.7499
S3	1 (OF7)	0.9062	0.8211	0.1718	0.1858
*S4	1 (RPF5)	0.4886	0.0749	0.8372	0.9514
S4	2 (ROF4)	0.8421	0.7201	0.1541	0.3731
S5	1 (RAF6)	0.8186	0.6290	0.3222	0.4133
sEEG mean		0.8830	0.5758	0.2445	0.4383
E1	1 (GR28)	0.8385	0.6081	0.3954	0.3884
*E1	2 (GR53)	0.6254	0.0462	0.9722	0.8636
E1	3 (GR38)	0.7726	0.5128	0.4000	0.5522
E2	1 (AGR52)	0.8112	0.3478	0.7692	0.2941
sEEG mean		0.8074	0.4896	0.5215	0.4116

*These recordings are excluded from the means and further analysis due to poor quality.

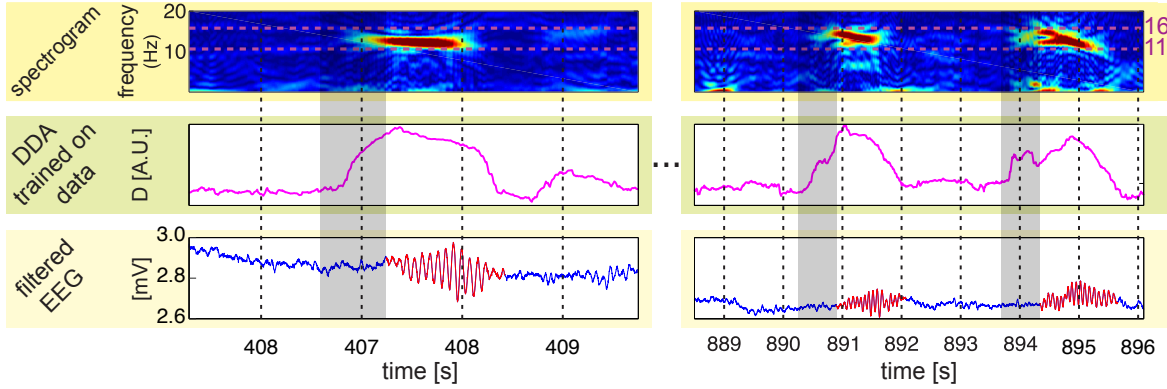


Figure 3.3: Spindle detection. The lowest row in the plot shows the data with the spindles in red, as marked by a human expert. In the middle row, a DDA spindle detection output (trained on one channel from a different subject) is shown. We also show the spectrograms (in the top row) for reference. The gray-shaded regions indicate the width of the time windows used for computing both the DDA features and the spectrogram (650 ms). Since we plot the time points on the x -axis for the start points of the sliding windows, all points within a shaded region use windows that include some amount of spindle data.

well as an additional method designed for intracranial data [Hagler et al., 2018]. Warby et al. used two additional detectors [Bódizs et al., 2009, Wendt et al., 2012] which are excluded here due to their reliance on the comparison of specific channels from a standard EEG montage, making them unsuitable for use with intracranial recordings from disparate locations.

It is important to note that for all of these methods, spindle detection performance may be lower here than with some other data, since no preprocessing or artifact removal steps have been applied here prior to the core processing steps for spindle detection intrinsic to each method. Further, these data present a mix of recordings of different quality and spindle clarity, as evaluated by human expert scoring.

Möller et al. used a 12-15 Hz bandpass finite impulse response (FIR) filter and subsequently computed a root mean square (RMS) signal with 50 ms time resolution and a 100 ms time window from the filtered data. Spindles were then detected using a thresholding procedure, with beginning and end threshold crossings between 0.4 and 1.3 s required for spindle detection. This threshold was set automatically by the algorithm for each subject as originally published, but was always greater than $5 \mu\text{V}$ [Möller et al., 2002].

The approach of Martin et al. was similar: data were first bandpass filtered from 11 to

15 Hz using an FIR filter applied both forward and reverse. The RMS of the signal was then computed using 0.25 s windows. The threshold for spindle detection was set at the 95th percentile and required two consecutive RMS time points (corresponding to 0.5 s) for a spindle [Martin et al., 2013].

We also use a slightly modified version of the detector of Andrillon et al., itself a modified version of the method of [Ferrarelli et al., 2007]. Putative spindles were identified by applying a zero-phase fourth-order Butterworth bandpass filter for 9 to 16 Hz. Instantaneous amplitude was computed using a Hilbert transform, and the threshold for detection was set at three standard deviations from the mean, with a threshold for the beginning and end of spindles set at one standard deviation. Only events with durations between 0.5 and 2 s were marked as spindles, and spindles separated by less than 1 s were merged.

Finally, we also apply a method developed for and previously applied to intracranial recordings of the type we consider here, which was developed by Hagler et al. This technique relies on an initial detection based on instantaneous power in the spindle band (11-17 Hz) using a smoothed wavelet convolution. Any initially identified spindles under 0.5 s in duration are excluded. Further, the ratio of Fourier power in the spindle band relative to power in the 4 to 9 Hz range is used to remove artifacts and weak spindles. [Hagler et al., 2016].

In order to compare these various techniques with differing methodologies, we convert the raw outputs of each technique to a binary index of spindle or non-spindle for each time point. These binary detection indices are then compared by computing the F_1 score of each method against the human expert-marked spindles. The mean across subjects of the number of spindles detected (expressed as a percentage of the number of spindles marked by the human expert), spindle length, F_1 score, and false positive and negative rates (relative to human expert scoring) for each of these methods are shown in Table 3.3. The F_1 scores as well as CPU time for all methods and recordings are shown in Fig. 3.4. DDA provides the highest average F_1 score and the second lowest average CPU time.

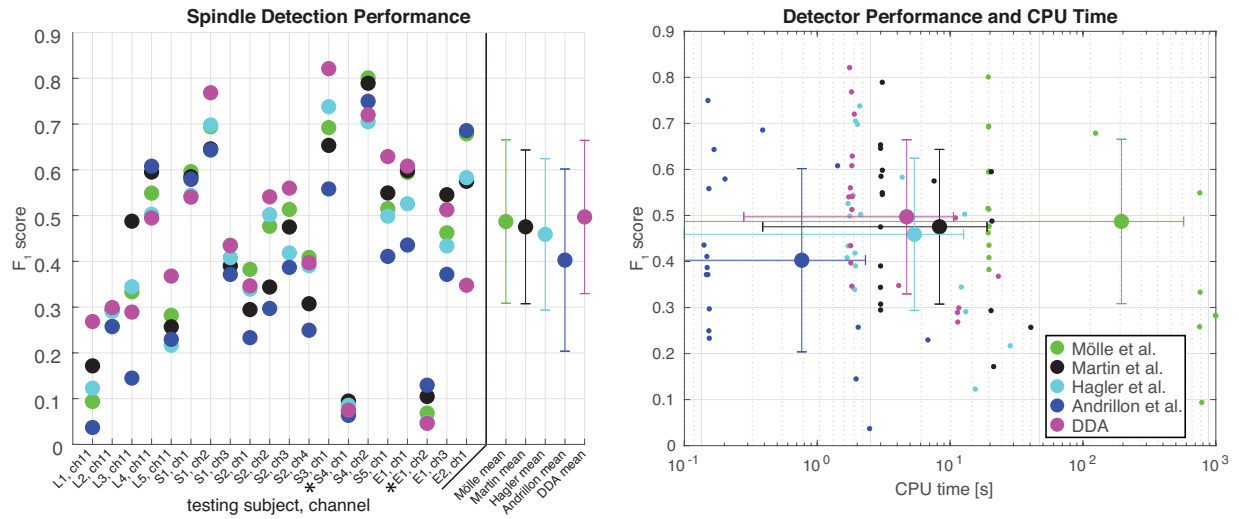


Figure 3.4: Detection methods comparison. In the left panel, F_1 score is plotted for a set of automated spindle detection methods and DDA for the various Laminar, sEEG, and ECoG recordings. The means (points) and standard deviations (bars) across all recordings for each detector are plotted at the far right—these exclude two recordings (denoted by *) of poor quality for which all methods yield low performance. These recordings are also omitted from the right panel. At right, the F_1 score for all recordings is plotted against CPU time for each detection method. Each detector was run on twenty intracranial recordings, the mean across all recordings (except the two noted exclusions) is plotted with a larger marker, standard deviations across all recordings are plotted as bars in both CPU time and F_1 score. Note that not all recordings are of equal length, so some variation in the CPU time is to be expected.

Table 3.3: Comparison of detection methods for all data

Method	Mean percentage of human-scored spindles	Mean length [s]	Mean F_1	False discovery rate	False negative rate	CPU* time [s] per recording
Möller	105.0457	0.4871	0.4871	0.2856	0.5994	30.5645
Martin	141.9600	0.4754	0.4754	0.3427	0.5441	2.5615
Andrillon	46.3362	0.4028	0.4028	0.2078	0.7022	0.3922
Hagler	116.2967	0.4591	0.4591	0.2963	0.6225	1.8177
DDA	89.8979	0.4970	0.4970	0.3861	0.4969	1.6389

*All methods were implemented in MATLAB 9.4 (R2018a) and tested on the same 12-core (Intel Xeon X5690 @ 3.47 GHz) system. The DDA detector calls an executable written in C for a key step in the procedure.

Notably, as shown in Fig. 3.2, one of the recordings (L1) had a higher mean peak spindle frequency than all others. That recording has a low F_1 score (see Fig. 3.4) for all comparison methods. DDA, in contrast, detected those spindles relatively well since the goal was to detect dynamical patterns in the data.

To assess the advantage provided by using DDA features in addition to spectral features, Fig. 3.5 and Table 3.4 show the mean F_1 scores for various combinations of the different detection methods. Of note is the fact that combining the DDA measure of spindle activity with other measures generally provides a better measure than combining two or more spectral methods, since it provides different information. Note that the F_1 scores for the DDA detector alone in Fig. 3.5 and Table 3.4 do not match exactly the scores in the earlier figures and tables. This is due to an additional step of averaging the DDA features across the overlapping windows at each time point. This provides a measure with time resolution equal to original data which can then be combined with other measures on a point-by-point basis.

Finally, for comparison, DDA and the other detection methods were applied to the DREAMS dataset, collected and made available by Université de Mons, TCTS Laboratory (Stéphanie Devuyst, Thierry Dutoit) and Université Libre de Bruxelles, CHU de Charleroi Sleep

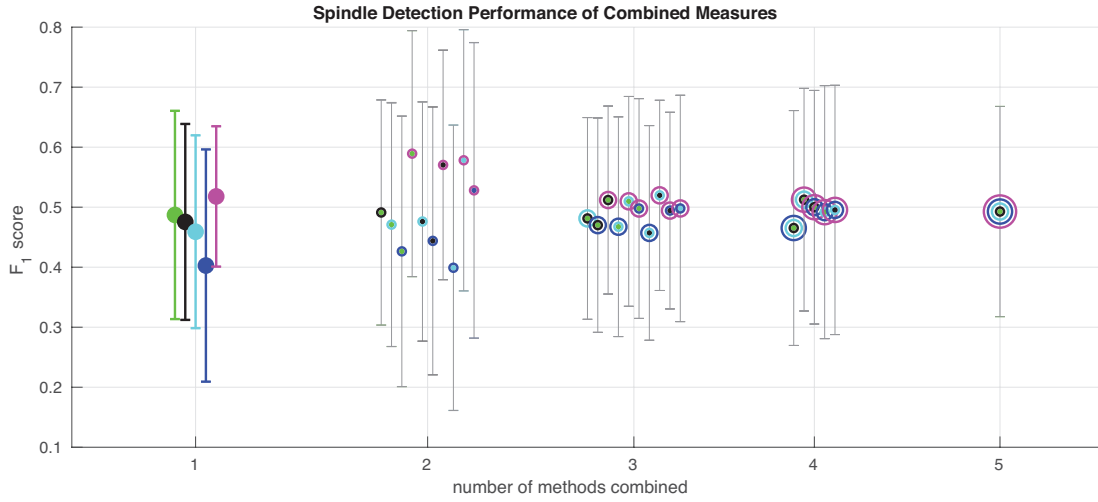


Figure 3.5: Combining features from the various methods. Spindle detection measures from the various tested methods were combined by taking a mean at each time point, and the agreement of these averaged measures with human scoring was evaluated via F_1 score. Two recordings with poor detector performance for all methods were omitted here. Colors correspond to the different methods, when methods are combined, concentric circles corresponding to the combined measures are plotted at one point. For all methods and combinations of methods, the mean across all recordings is shown. Error bars represent the standard deviation across recordings. Mean F_1 scores for these combinations of detectors are also shown in Table 3.4. It is noteworthy that there is a significant boost in detection performance only when combining DDA with any one of the spectral methods. No other combination of methods provides such a boost.

Table 3.4: Combining detection measures from the various methods. The highest-performing combinations of detectors are marked in red.

# combined	Möller et al.	Martin et al.	Hagler et al.	Andrillon et al.	DDA	F_1 score
1	0.4871	0.4754	0.4591	0.4028	0.5179	
2	X	X				0.4912
	X		X			0.4709
	X			X		0.4264
	X				X	0.5892
		X	X			0.4761
		X		X		0.4439
		X			X	0.5704
			X	X		0.3991
			X		X	0.5781
				X	X	0.5280
3	X	X	X			0.4813
	X	X		X		0.4701
	X	X			X	0.5119
	X		X	X		0.4674
	X		X		X	0.5098
	X			X	X	0.4978
		X	X	X		0.4571
		X	X		X	0.5197
		X		X	X	0.4943
			X	X	X	0.4979
4	X	X	X	X		0.4653
	X	X	X		X	0.5125
	X	X		X	X	0.5000
	X		X	X	X	0.4917
		X	X	X	X	0.4954
5	X	X	X	X	X	0.4927

Laboratory (Myriam Kerkhofs) [Devuyst et al., 2011]. The DREAMS data consist of surface EEG with spindles marked by two human experts. Using these data allow the above detection methods to be compared on surface EEG data, as well as compared to automated spindle detections from a method implemented by the original authors and made available with the data. This technique is based on bandpass filtering and applying a recording-specific threshold. While the DREAMS automated detections provide better agreement with the human scorers than the intracranial data-trained DDA detector or any of the other tested methods [Devuyst et al., 2011], we cannot compare directly with this method since only the data and automated detections are available, and not the code. We therefore cannot test the DREAMS method on our dataset. Further, as can be seen in Fig. 3.6, there is also a large discrepancy between the two human scorers, with one scorer also only having scored six of the eight subjects. Issues with the scoring of these data were previously noted by [O'Reilly and Nielsen, 2015]. Further, it is noteworthy that DDA still provides reasonable spindle detection after structure selection based solely on intracranial data. Most significantly, however, we also show the combinations of two detectors (as shown in Fig. 3.5). For these data, combining our DDA measure with the measure produced by the method of Martin et al. provides the highest average agreement with the two human scorers among all tested methods and combinations of methods.

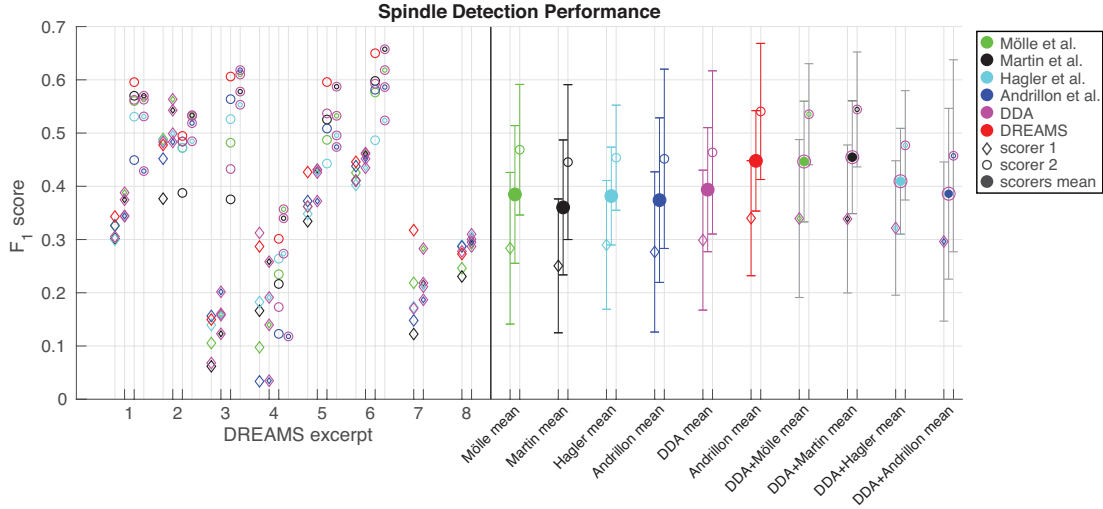


Figure 3.6: Spindle detection on DREAMS data. F_1 score is plotted for a set of automated spindle detection methods and DDA for the eight surface EEG excerpts included in the DREAMS data set. For six of the eight excerpts, two human experts scored the data. For these six recordings, F_1 scores based on the first expert’s markings are plotted as diamonds at left, and the scores based on the second expert’s markings are plotted as open circles at right. The means (diamonds and open circles) and standard deviations (bars) across all recordings for each detector’s agreement with both experts are plotted at the far right with the same markers, and the means of each method’s agreement with both scorers are plotted as larger filled circles. Combinations of the other measures with DDA, as shown in Fig. 3.5 are shown with the colors for each of the methods combined. In addition to the six methods shown previously, we also show here the F_1 scores of the automated spindle detections included with the DREAMS data with both human experts in red.

3.4 Discussion and Conclusions

DDA is a powerful novel tool for detecting sleep spindles in EEG and intracranial recordings. DDA requires minimal pre-processing of signals and can be rapidly applied to large datasets. When compared with several well-established and reliable frequency-based methods, DDA provides the highest level of agreement with human scoring (evaluated here with F_1 score). Further, DDA is the second fastest of the tested methods, where the only faster method produces the lowest F_1 scores. DDA therefore holds great promise for real-time applications. We also tested all methods on the publicly available DREAMS data, consisting of surface EEG recordings scored by two expert scorers. Again, DDA provides the highest F_1 score of the previously tested methods when taking the average across both scorers. The automated detections made available with the DREAMS data however, do provide better agreement with the human scorers. It should be noted that the DREAMS data is a small and heterogeneous data set, and therefore somewhat limited for evaluation purposes [O'Reilly and Nielsen, 2015].

An important caveat for the results from intracranial data presented here is that they are based on comparison with the spindle markings by a single human expert. Despite this, the fact that several automated methods produce similar detections indicates that the markings are reasonable. Further, similar results are achieved using the same approaches on an EEG data set scored by two experts. It is also important to note the classic bias that our implementation of other previously published detectors may not be as fully perfected as the novel method developed for this paper. Other implementations on other data and comparing to other human scoring might not produce the same relative performance numbers. However, this is only a concern when looking at each method separately. As shown in Figs. 3.5 and 3.6, combining our nonlinear time-domain method with any of the tested spectral-based methods, the performance is increased dramatically, beyond the relatively differences between individual methods. This indicates that spectral and nonlinear methods account for different information in the original signal: DDA

looks for dynamical differences while spectral methods look for content in a specific spindle frequency band.

Combining two spectral measures does not provide the same advantage as combining linear and nonlinear features. Additionally, we have demonstrated that DDA models built on the data show superior performance to those built to detect specific frequencies, which indicates that using the nonlinear signature of the spindle provides access to additional information. Accessing this type of information could prove especially useful in future studies focused on spindles of different types, or occurring in patients with neuropsychiatric disorders. Finally, it is worth emphasizing again the robustness of DDA measures in general to noise and artifacts due to the sparsity of the feature space. This is a significant advantage for many data sets.

A version of the DDA spindle detector for use on Linux systems using MATLAB has been made available at http://cnl.salk.edu/~asampson/detect_spindles_DDA.zip.

Acknowledgements

The authors would like to thank Dr. Werner Doyle for his assistance with the collection of these data. This work was supported by the Howard Hughes Medical Institute and the Crick-Jacobs Center for Theoretical and Computational Biology, the U.S. Office of Naval Research under Grants N00014-10-1-0072 and N00014-12-1-0299, and the NIH grants R01-EB009282 and R01-NS104368.

Appendix: Frequency-Based Spindle Detection

All spindle detection techniques DDA is compared to are based on decomposing the signal into oscillatory components, and therefore have very different assumptions: while DDA assumes nonlinearity of the (unknown) underlying dynamical system, spectral methods assume

linear superposition of stationary sinusoids. To interpret the differences in detector performance we need to answer the question of what is gained by using nonlinear instead of linear analysis.

In [Lainscsek and Sejnowski, 2015] a connection between DDA and spectral analysis was made: a one term linear DDE can be used for frequency detection while a one term nonlinear DDE can detect frequency/phase couplings in the time domain. A DDE with linear and nonlinear terms will have vanishing nonlinear coefficients for purely harmonic signals. For data that contain nonlinear couplings between frequencies or other nonlinear signal components, linear as well as nonlinear terms contain both linear and nonlinear information. Superposition does not work due to nonlinearities in the model. Therefore no connection between frequencies and delays can be made for real-world signals that are generally nonlinear.

Applying the same three-term, nonlinear DDE used for the spindle data to simulated data (noise-diluted sinusoids) can serve as a test of what can be gained by adding nonlinear information, and as a bridge between this technique and traditional wavelet or other spectral methods. The effectiveness of the frequency detector at detecting spindles is also informative as to how much of the relevant dynamical information is related to the dominant frequencies, which is of interest since many spindle detection techniques rely on spectral analysis [Warby et al., 2014].

The DDA frequency detector relies on the same structure selection framework as the data-trained spindle detector, but the DDE model form is fixed to match the model selected using the real data, and only the values of the delays are selected based on the simulated data. For the purposes of comparison with the data-trained detector, we select for frequency bands in the simulated data that correspond to sleep spindles in the EEG sigma band, defined alternately as 11-14 Hz or 11-17 Hz. By comparing the delays which are most successful at detecting these frequencies with those that are selected for the task of sleep spindle detection, we can gain insight into the information added by nonlinear analysis.

The simulated data is generated according to:

$$S_i = A_i \cos(\omega_i t + \varphi_i) + \varepsilon \quad (3.8)$$

with $\omega_i = 2\pi f_i$ for 9991 equally-spaced frequencies f_i between 0.1 and 100 Hz, equal amplitudes $A_i = 1$, random phases $0 < \varphi_i \leq 2\pi$, and added white noise ε with a signal-to-noise ratio of 5 dB. Starting from the full set of frequencies, we divide into nearly-equal groups for training and testing, with training data consisting of frequencies f_i from 0.1 to 100 Hz, and the testing data consisting of frequencies f_i from 0.11 to 99.99 Hz, both sets with 0.02 Hz frequency intervals. This ensures that we validate on slightly different frequencies from the training data, still in the desired range. For our simulated training data, we select data with frequencies f_i in the sigma band. As was the case for the data-driven detector, we train separately for each sampling rate, generating simulated data to match each of the sampling rates in the laminar, sEEG, and ECoG data. We then choose delays for each sampling rate f_s .

Selecting a model to provide sensitivity to specific frequency bands requires an additional step, in that we first select “high-pass delays” which are sensitive to frequencies above the lower bound we wish to set (here, 11 Hz), and then additional “low-pass delays” which are sensitive to frequencies below the upper bound (here, 14 or 17 Hz).

The delays chosen for each sampling rate for each definition of the sigma band (11-14 Hz or 11-17 Hz) are shown in Table 3.5. Note that in some cases, the same delays can be used in both the “high-pass DDE” and “low-pass DDE”, since different weights can be applied to the features to select for different frequency ranges.

As with the data-driven detector, we apply a vector of weights to the features for both the lower and upper bounds, in this case obtaining two values of D , which we call D_1 and D_2 . We

Table 3.5: Selected delays (τ_1, τ_2) for specified bands, units of $\delta t = 1/f_s$

f_s	delays [δt]			
	11-14 Hz		11- 17 Hz	
	> 11 Hz	< 14 Hz	> 11 Hz	< 17 Hz
2000	(8,105)	(8,105)	(8,69)	(7,39)
1024	(1,44)	(19,4)	(4,37)	(4,20)
512	(23,43)	(8,2)	(17,19)	(10,2)
500	(39,18)	(10,2)	(2,17)	(2,9)

combine them by summing their absolute values and applying the sign of the lesser of d_1 and d_2 :

$$D = \frac{\min(D_1, D_2)}{|\min(D_1, D_2)|} (|D_1| + |D_2|). \quad (3.9)$$

We will therefore obtain positive values only in the region where both are positive, which should correspond to the “DDA pass band”.

Fig. 3.7 shows the frequency response of the detector on simulated data. Given its strong selectivity for frequencies in the desired range, it was applied to the sleep spindle data as a means of detecting frequency content in the spindle band which uses the same methodology as the data-based DDA spindle detector. This allows for direct comparison between the frequency-based and data-based DDA approaches.

Chapter 3, in full, is a reprint of the material as it appears in the Journal of Neuroscience Methods 2019. Aaron L. Sampson*, Claudia Lainscsek*, Christopher E. Gonzalez, István Ulbert, Orrin Devinsky, Dániel Fabó, Joseph R. Madsen, Eric Halgren, Sydney S. Cash, and Terrence J. Sejnowski. The dissertation author was the primary investigator and author of this paper.

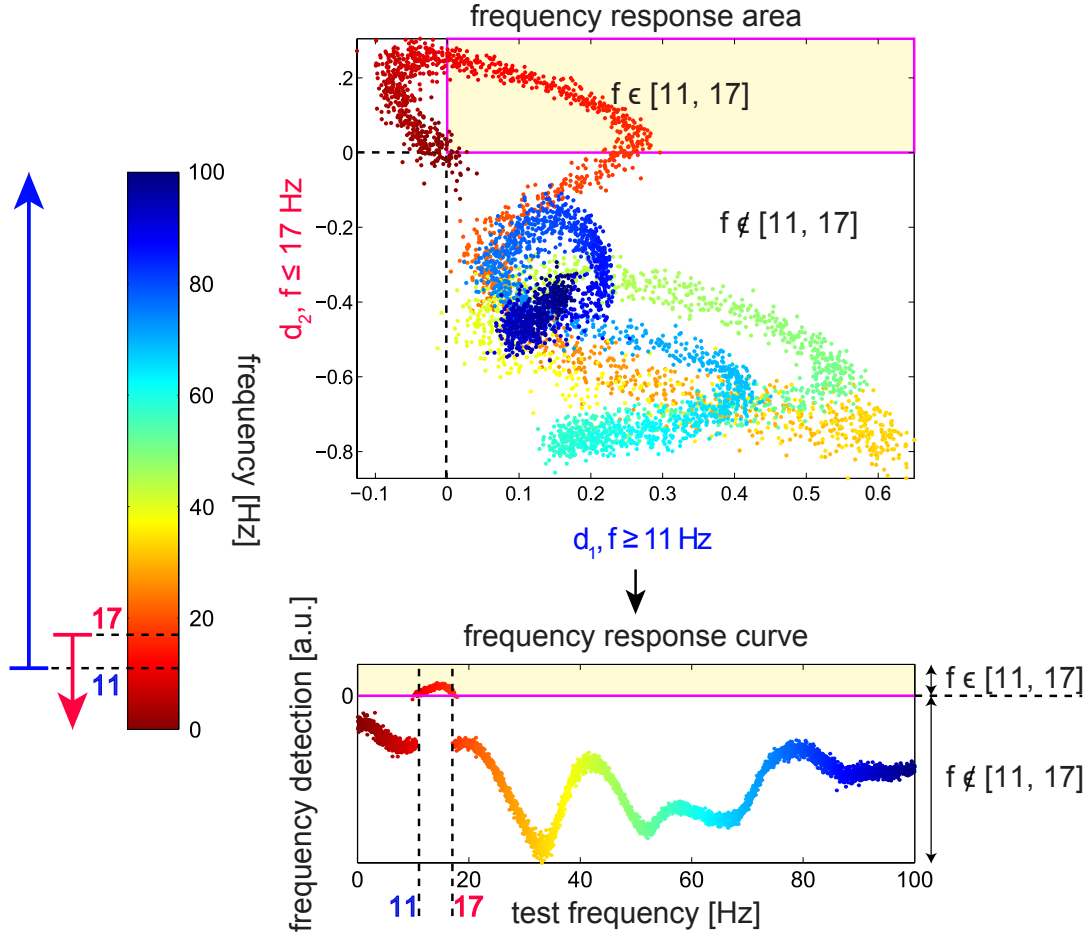


Figure 3.7: Frequency band detection. Applying the DDE model with two different delay pairs, one sensitive to frequencies above 11 Hz and one sensitive to frequencies below 17 Hz, we can obtain an output which is positive only in the desired band. In the top panel, the distance from the hyperplane values computed from both DDEs (d_1 and d_2) are plotted for test frequencies ranging from 0.1 to 100 Hz. The frequency of the test data is color-coded according to the color bar at left, from 0.1 (red) to 100 Hz (blue). Points falling into the upper right quadrant (shaded yellow) have positive values for both d_1 and d_2 , and we select delays such that only frequencies in the desired range (11-17 Hz) fall into this area. In the lower plot, d_1 and d_2 are combined according to Eq. (3.9) to obtain a one-dimensional index that is positive only for frequencies in the desired range. This procedure was also used to obtain delays and corresponding weights for frequency ranges 11-14 Hz and 12-15 Hz.

Chapter 4

Cross-Dynamical Delay Differential Analysis Reveals Information Flow During Hippocampal Ripples

Abstract

High-frequency hippocampal ripples mark the time when hippocampal cells replay sequences from waking during slow wave sleep. Evidence in rodents is consistent with ripples sending information to the cortex to permit memory traces to be transferred during consolidation. However, hippocampo-cortical interactions during ripples are poorly characterized in humans. Cross-Dynamical Delay Differential Analysis (CD-DDA) is a new tool to study causal connections between time series signals. Based on embedding theory from nonlinear dynamics, the classical formulation of Delay Differential Analysis (DDA) relates the differential and delay embeddings of a single time series in a functional form to uncover dynamical differences in the data. The features obtained from DDA provide a powerful basis for time-domain classification of data. In CD-DDA, we investigate causal interactions between two time series. Here, we apply

this technique to intracranial recordings from patients undergoing treatment for epilepsy. By applying CD-DDA to recordings from electrodes placed in both hippocampus and remote cortical areas, we seek to uncover directional information flow around the times when ripples occur. For comparison, we also apply Granger Causality (GC) and Convergent Cross Mapping (CCM) to the same data. A fundamental challenge for all approaches is achieving an event-related measure of causal interactions, as in recordings of highly distinct discrete events such as ripples. While this challenge limits the interpretability of these results, we do see some evidence of information flow between hippocampus and cortex. Further analysis could refine these results and solidify conclusions about the precise timing of these interactions, as well as characterize information flow in the brain more broadly.

4.1 Hippocampal Ripples

Perhaps the most striking feature of recordings of activity in the hippocampus is the sharp-wave ripple (SWR). These are short bursts of very high-frequency ($\approx 80\text{--}120$ Hz in humans) oscillations coupled with large deflections in amplitude [Vaz et al., 2019]. These events are believed to play a critical role in the memory consolidation process, likely by facilitating communication between hippocampus and cortex. SWRs have been associated with cortical down-to-up state transitions and cortical spindles [Rothschild et al., 2017, Isomura et al., 2006, Sirota et al., 2003]. This suggests that these specific cortical states could provide excitation to initiate SWRs in the hippocampus as well as spindles in the thalamus, providing a mechanism for coordinating these rhythms and thereby promoting the transfer of information from the hippocampus to the cortex as had been previously suggested [Diekelmann and Born, 2010].

4.2 Methods

4.2.1 Intracranial Recordings

Stereoelectroencephalogram (sEEG) recordings were obtained from patients undergoing treatment for intractable epilepsy. Procedures were approved by the Cleveland Clinic Institutional Review Board and all surgeries and placement of electrodes were part of standard patient care with no input from this research study.

Hippocampal ripples were identified in these data by visual inspection. Recordings were categorized based on the characteristics of the observed ripples, chiefly whether the ripple was co-occurring with a sharp wave, a particularly high-amplitude sharp-wave, or with another oscillation.

4.2.2 Cross-Dynamical Delay Differential Analysis (CD-DDA)

As discussed in Section 1.8, Cross-Dynamical Delay Differential Analysis (CD-DDA) is a tool for studying causal interactions between time series. CD-DDA is an extension of classical DDA which adds terms from one system to the delay differential equation (DDE) model for another system (here, another recording location). From the general form of such a cross-dynamical equation in Eq. (1.23), we select specific terms in the same way we choose a DDA model. For example, when considering recordings from two locations in the brain, we can choose a model which is effective at classifying data as coming from one location or the other in the classical DDA framework, then add the corresponding terms from the other location to consider the interactions between these locations. For concreteness, consider the model that will be used in the following sections, selected for its ability to separate hippocampal data from other cortical data:

$$\dot{x} = a_1 x_1 + a_2 x_2 + a_3 x_1^2 + \rho_x \quad (4.1)$$

where $x_i = x(t - \tau_i)$, $\tau_1 = 2\delta t$, and $\tau_2 = 8\delta t$ with $\delta t = 1/f_s$ and $f_s = 1000$ Hz. Note that this is again one of the “standard EEG models” with two linear terms and one quadratic term, which have been found in a number of applications to provide good classification performance on EEG and intracranial data, albeit with delays that are specific to a given classification task. For CD-DDA, we will call hippocampal data $x(t)$ and data from other sites in the cortex $y(t)$ and use the cross-dynamical models:

$$\begin{aligned}\dot{x} &= e_1 y_1 + e_2 y_2 + e_3 y_1^2 + a_1 x_1 + a_2 x_2 + a_3 x_1^2 + \rho_{x \leftarrow y} \\ \dot{y} &= f_1 x_1 + f_2 x_2 + f_3 x_1^2 + b_1 y_1 + b_2 y_2 + b_3 y_1^2 + \rho_{y \leftarrow x}\end{aligned}\tag{4.2}$$

with the same delays as in the classical DDA model in Eq. (4.1), but with terms from both locations included. We can then measure the causal influence of y on x or x on y by considering how the error in the fit to the data is reduced by the addition of these cross-dynamical terms. As in Eq. (1.25), we use

$$C_{xy} = |\rho_x - \rho_{x \leftarrow y}| \tag{4.3}$$

as a measure of this influence over time. We expect that when y is sending information to x that influences its behavior, the addition of y terms will significantly reduce $\rho_{x \leftarrow y}$ as compared to ρ_x , whereas if there is no such information flow, the reduction in error will be relatively small (though there will in general still be some reduction due to the additional terms in the model).

4.2.3 Granger Causality

Granger introduced his tool for studying causality in 1969 [Granger, 1969]. This was the first practical definition of causality based on prediction improvement by adding information from a separate time series [Runge, 2018]. GC is used to study causal interactions between systems, but it treats data as stochastic rather than deterministic and is best-suited to linear systems. In standard, bivariate GC, the idea is to determine whether there is unique information in the past

of a variable Y that can improve the predictability of the future behavior of a variable X . This is typically done using linear autoregressive models:

$$\begin{aligned} X(t) &= \sum_{\tau=1}^n A_{XX} X(t-\tau) + E_1(t) \\ X(t) &= \sum_{\tau=1}^n A_{XX} X(t-\tau) + \sum_{\tau=1}^n A_{XY} Y(t-\tau) + E_2(t) \end{aligned} \quad (4.4)$$

where n is the number of delayed time points used (the order of the model), and E is the model error. Causality is inferred based on comparing the model error before and after the inclusion of the Y terms. This is typically reported as the logarithm of the ratio of errors ($\ln(E_1/E_2)$). The results of applying GC to the three test cases for hippocampus-to-cortex connections are shown in Fig. 4.4A.

4.2.4 Convergent Cross Mapping

Sugihara et al. introduced CCM as a nonlinear dynamical tool capable of interrogating causality in a class of systems not covered by GC. CCM tests for causation by measuring the extent to which values of Y can estimate the state of X [Sugihara et al., 2012]. This is achieved by constructing multi-dimensional delay embeddings for both variables to construct what Sugihara refers to as shadow manifolds. In the case where Y does causally influence X , the times of nearby points on the Y manifold can be used to find nearby points on the X manifold.

Typically, CCM is used on longer time series such that the number of points can be varied to test for convergence. Here, because we wish to study causal interactions at the timescale of the ripple, we use CCM in short time windows (40 ms) and slide these windows across the data (with a 5 ms step size). The results should therefore be interpreted cautiously. We do however find a high cross-map skill for pairing of hippocampal and cortical recordings. The results of applying CCM in this manner to the three test cases for hippocampus-to-cortex connections are shown in Fig. 4.4.

4.3 Results

Studying information flow in the brain during discrete events (such as ripples) presents a unique challenge since the amount of data available will necessarily be limited, and since the characteristics of the time series around these events will necessarily be quite different from its longer-term behavior. For similar reasons, finding interpretable event-related measures of causality in general is a challenge. Despite these challenges, DDA is a tool that has been successfully applied to short time series, which might suggest that CD-DDA can provide useful information about causal connections between short time series as well. However, it must be applied cautiously. Here, we present preliminary results which suggest a distinctive pattern of information flow around the time of hippocampal ripples, but subsequently note limitations on the interpretability of these findings.

4.3.1 CD-DDA During Ripples

When applying CD-DDA to recordings from the hippocampus and other cortical sites, that there is a very consistent and distinctive pattern observed in \mathcal{C} that is highly characteristic of hippocampal ripples. Important limitations on the interpretability of this pattern will be discussed in Section 4.3.2, but nonetheless, the form of this ripple-associated pattern may hold some suggestions for when specific directed communication could occur.

Fig. 4.1A and B show the C_{xy} and C_{yx} for a hippocampal recording paired with one remote site in the cortex for 375 marked ripple events, and as a mean across ripples for one example patient (R1). At the time of the ripple, we see significant changes in both hippocampus-to-cortex and cortex-to-hippocampus information flow. In particular, we see a seeming biphasic pattern, where C_{xy} (cortex-to-hippocampus) rises first, at the beginning of the ripple, and C_{yx} (hippocampus-to-cortex) rises at the end of the ripple and for several hundred milliseconds afterwards. This pattern is apparent even on an individual ripple basis as shown in Fig. 4.1B.

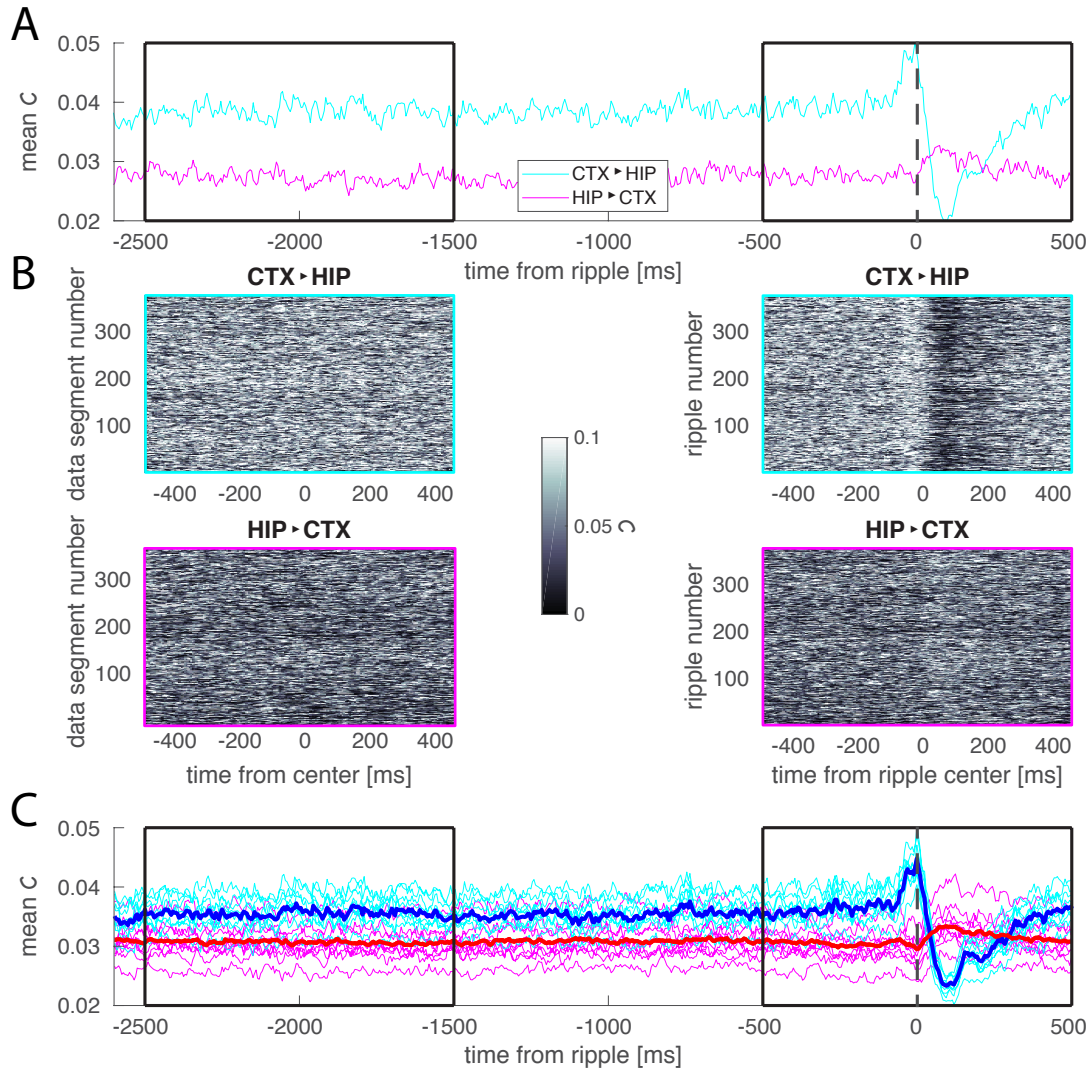


Figure 4.1: **A.** For recordings from the hippocampus and a single remote cortical site (in the posterior temporal lobe), CD-DDA was applied, and the resulting C is shown as a mean across 375 ripple events in cyan for cortex-to-hippocampus information flow (C_{xy}), and in magenta for hippocampus-to-cortex information flow (C_{yx}). **B.** For the same pairing of one hippocampal and one cortical location, each of the individual ripple events (at right), and a corresponding number of time segments without ripples (at left), C is shown as a heatmap. **C.** CD-DDA was applied to recordings from the hippocampus and 19 remote cortical sites and the resulting C_{xy} and C_{yx} are shown for each hippocampus-to-cortex pair as a mean across ripple events. As above, magenta corresponds to hippocampus-to-cortex and cyan to cortex-to-hippocampus. The mean across all hippocampus-to-cortex pairings is shown in red and the mean across all cortex-to-hippocampus pairings is shown in blue.

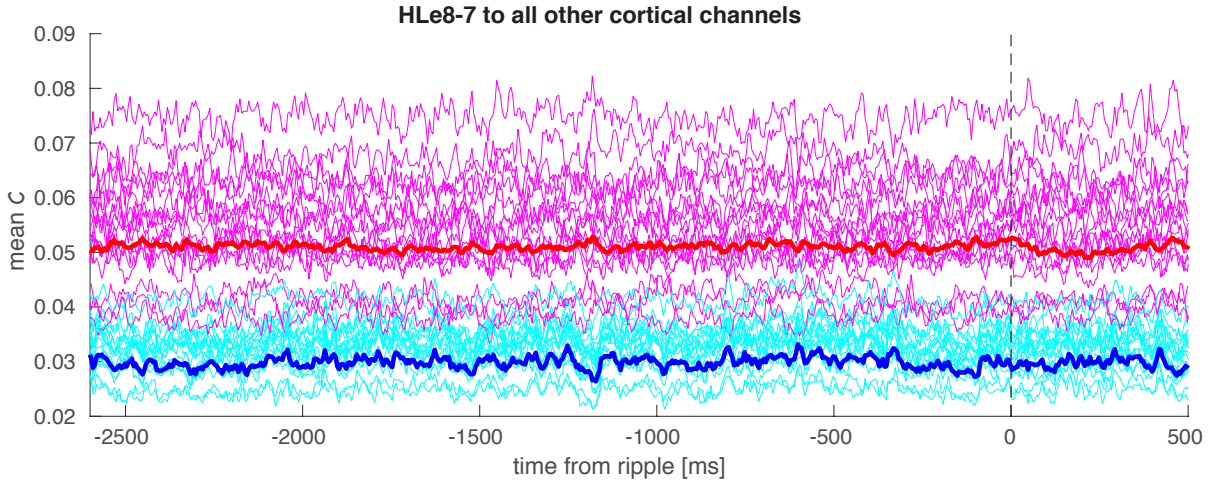


Figure 4.2: CD-DDA was applied to recordings from the cortex, ignoring the hippocampus. A single cortical electrode was taken as $x(t)$ here, with the CD-DDA error comparison measure for information flow from this channel to other channels shown in magenta (with the mean across channel pairs in red), and the measure for information flow from other channels to this channel shown in cyan (with the mean across channel pairs in blue).

These results are highly consistent when pairing the hippocampus with a wide range of cortical recording sites. Although the magnitude of C changes, the same biphasic pattern is apparent when considering hippocampo-cortical communication across a large area of the cortex. For comparison, the results of pairing the hippocampus with nineteen cortical sites in this manner are shown in Fig. 4.1C.

To further illustrate that this pattern is driven by the activity of the hippocampus, Fig. 4.2 shows the results of pairing all cortical sites with one another and ignoring the hippocampus. There is no marked pattern that deviates from the baseline communication at the times at which hippocampal ripples occur. This indicates that the distinct biphasic pattern apparent in Fig. 4.1 is unique to pairing of the hippocampus with remote cortical sites.

In Figs. 4.1 and 4.2 we have taken recordings from a single intracranial patient as an example. The same analysis was applied to recordings from an additional nineteen patients. The full set of patients in fact breaks down into two groups: one group shows the same biphasic pattern seen in Patient R1 while the other shows a more monophasic pattern. These seem to

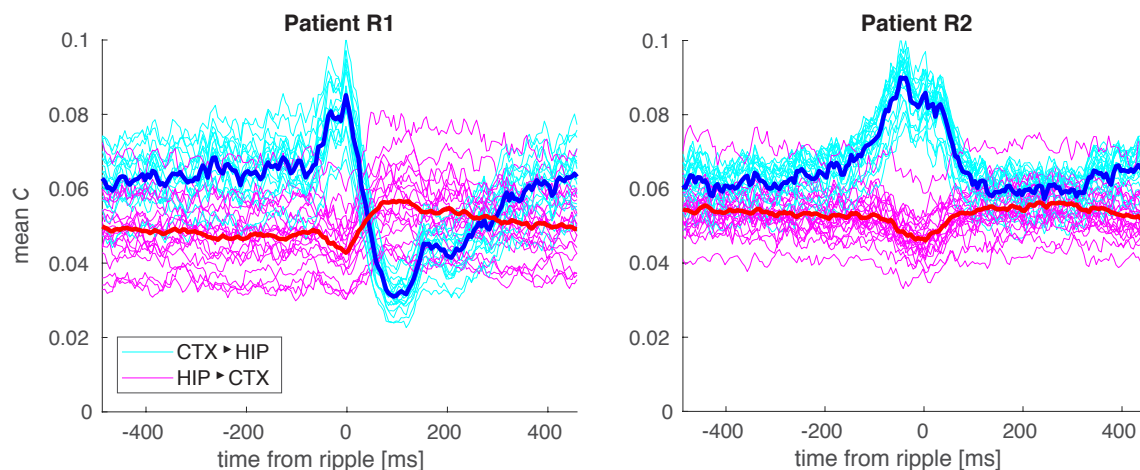


Figure 4.3: CD-DDA was applied to intracranial recordings from nineteen patients. The characteristics of the hippocampal ripples observed in each of these recordings were also noted. Recordings from Patient R1 showed ripples with particularly large sharp-wave deflections. Recordings from Patient R2 showed no clear sharp wave. These differences may be due to the precise recording location.

be determined by the presence or absence of sharp waves: for those recordings which were noted as having ripples without obvious sharp waves, we see only the monophasic response in C . Examples of both groups are shown in Fig. 4.3.

4.3.2 Randomized Time Tests

To test CD-DDA and two additional methods for inferring causality, GC and CCM, we use three pairings of time series as x and y : hippocampal data with simultaneously recorded cortical data, hippocampal data with cortical data from other randomly selected timepoints, and hippocampal data with temporally scrambled (order of data points randomized) cortical data. The results of the three methods on each are shown in Fig. 4.4.

As shown in Fig. 4.4C, CD-DDA, like GC and CCM, gives similar results for pairing hippocampal ripple data with simultaneously recorded cortical data, and cortical data from random timepoints. We do see however, that we find much more information is gained by adding terms

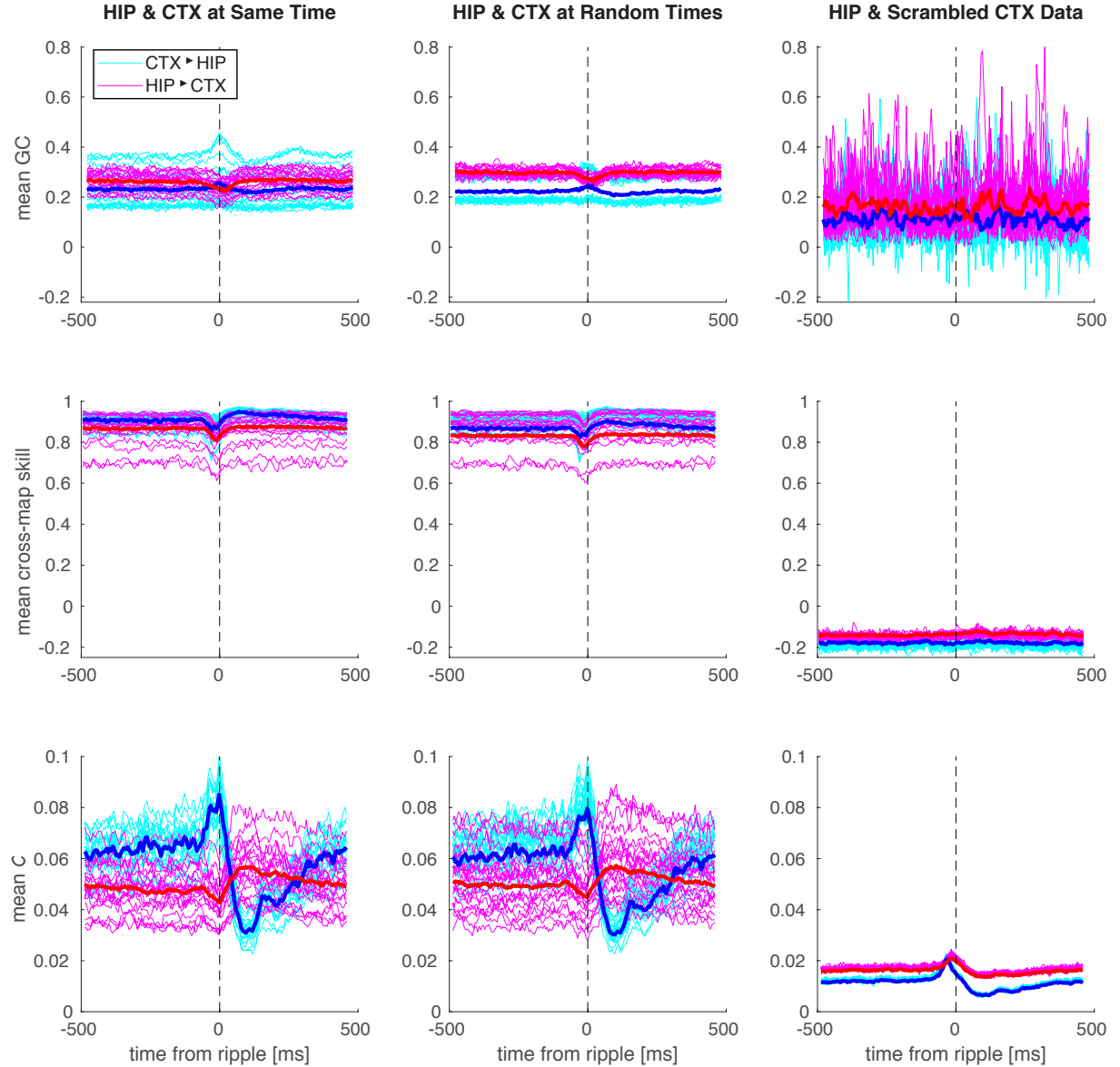


Figure 4.4: For GC (top row), CCM (middle row), and CD-DDA (bottom row), the data from one-second periods surrounding hippocampal ripples were paired with the data from a set of cortical sites at the same time (left column), random times (middle column), and temporally scrambled data from the cortex (right column). Magenta lines correspond to each of the individual hippocampus-to-cortex pairings, cyan lines to each individual cortex-to-hippocampus pairing, red lines to the average across cortical sites for hippocampus-to-cortex, and blue to the average across cortical sites for cortex-to-hippocampus.

with cortical data than temporally scrambled data with all relevant dynamics removed. This indicates that there is a high degree of dynamical similarity between the activity recorded in the hippocampus and the cortex, and perhaps that there is ongoing communication between the two. However, it is not possible to draw clear conclusions about the precise timing of this communication.

We do, however, find some evidence that there may be more information provided by pairing these recordings at the same time. To test this more thoroughly, in addition to pairing the data at random times as shown in Fig. 4.4, the data from these random timepoints were subjected to two additional shuffles: first, for each channel, the data from these random timepoints were shuffled, so that there was no time-locking across channels; second, the data from the random timepoints were also shuffled across channels. C_{xy} and C_{yx} for all of these pairings are shown in Fig. 4.5. To determine the degree to which the observed signal is unique to the true, time-locked pairing of the hippocampus and cortex at ripple times, the difference between the CD-DDA measure under those conditions and under the fully shuffled conditions is shown in Fig. 4.6. When compared to each of these shuffled data sets, the true pairing of simultaneously recorded data does stand out as having an error difference that is distinguishable, at least in the hippocampus-to-cortex direction. Figs. 4.7 and 4.8 show the same set of tests for GC. Similarly, we do see a small difference between the true, time-locked pairing and the shuffled conditions. Finally, Figs. 4.9 and 4.10 show the same set of tests for CCM. Here, the difference between the same-time pairing and shuffled pairings is less obvious, but there is an apparent signal that remains at the time of the ripple in the difference between the same-time cross-map skill and the cross map skill for the pairing of hippocampus with cortical data shuffled across both epochs and time.

Fig. 4.11 shows the same differences shown in Fig. 4.6, but plotted at the locations of the cortical recording sites. This illustrates the varying baseline levels of unique information obtained from matching terms at the same time (as opposed to random times), and also shows the few locations that do show time-varying signals in this view. The most notable signal associated with

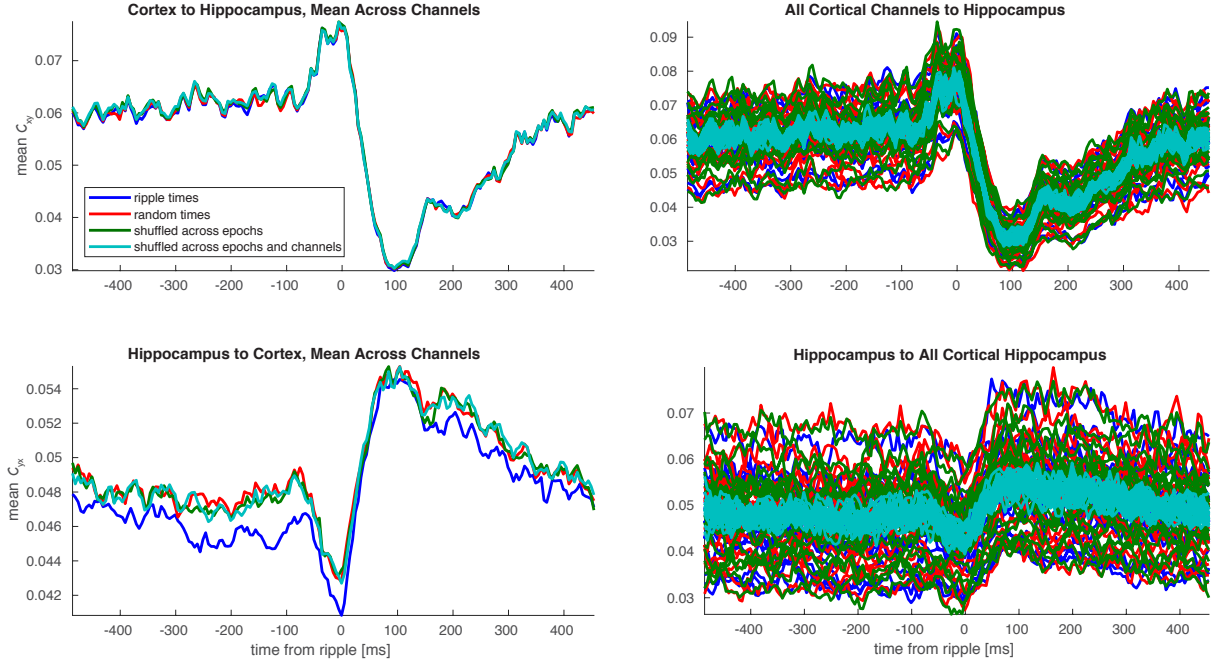


Figure 4.5: CD-DDA was applied to one second data from the hippocampus surrounding 375 ripple events. These hippocampal recordings were designated as $x(t)$ (as in Eq. (4.2)), and several different definitions of $y(t)$ were used. In blue, we plot the results of taking the simultaneously recorded cortical data as $y(t)$, in red, cortical data from random times, in green, cortical data from the same random times, but with the segments shuffled across epochs for each channel (i.e. paired with a different ripple events for different channels), and in turquoise, the same random segments shuffled across channels as well. C_{yx} (hippocampus to cortex) is shown in the top panels, and C_{xy} (cortex to hippocampus) in the bottom panels. The mean across channel pairs is shown at left and the individual channel pairs at right.

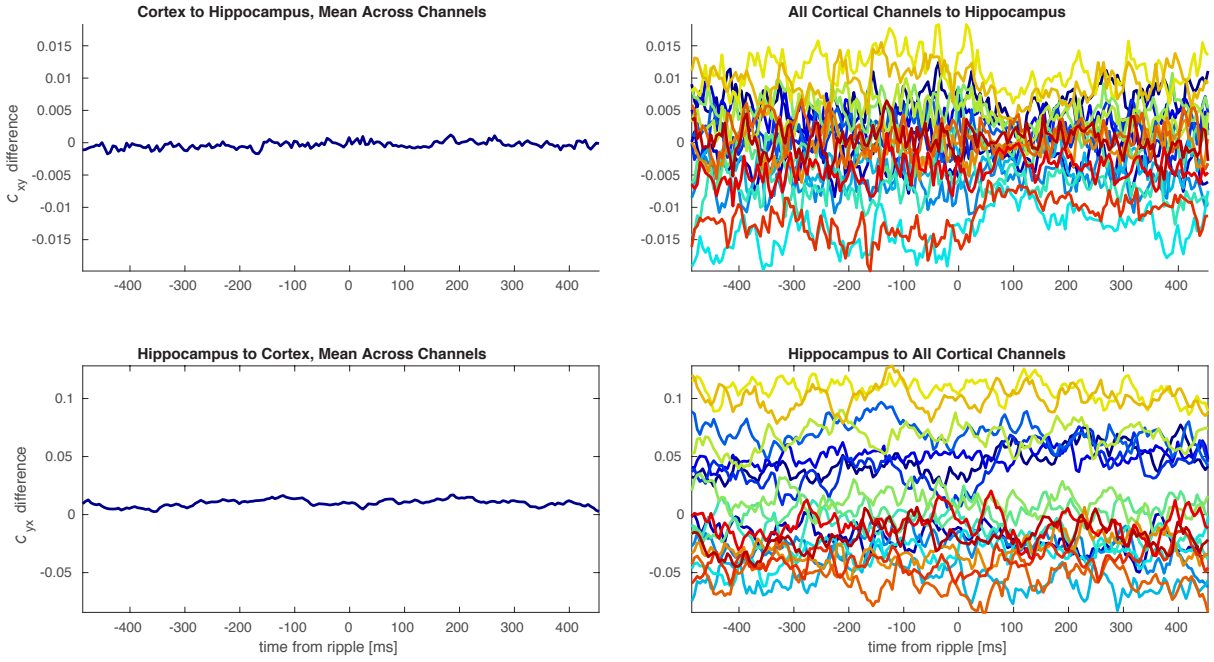


Figure 4.6: For the mean across channel pairs (at left) and for each individual channel pair (at right), the difference in C between the time-matched x and y pairings (plotted in blue in Fig. 4.5) and the pairing of x with y shuffled across both epochs and channels (plotted in turquoise in Fig. 4.5). The differences for C_{yx} (hippocampus to cortex) are shown in the top panels, and C_{xy} (cortex to hippocampus) in the bottom panels. The mean across channel pairs is shown at left and the individual channel pairs at right (with each channel pair plotted in a different color).

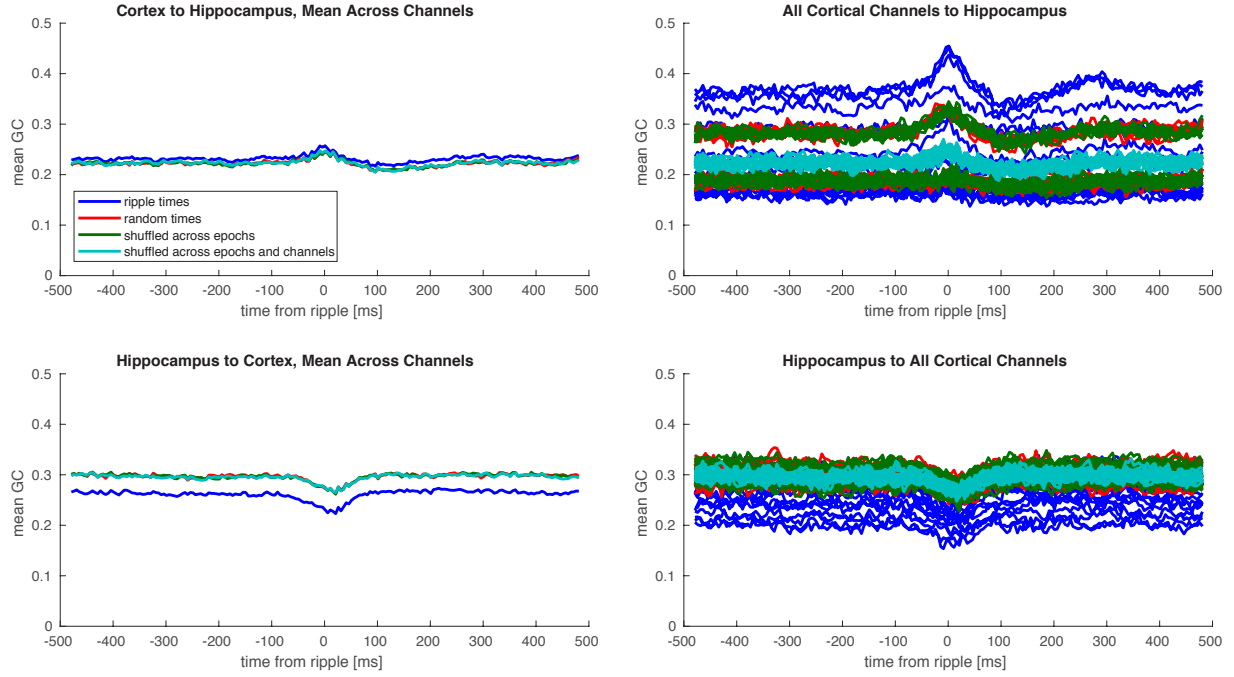


Figure 4.7: GC was applied to one second of data from the hippocampus surrounding 375 ripple events. These hippocampal recordings were designated as $X(t)$ (as in Eq. (4.4)), and several different definitions of $Y(t)$ were used. In blue, we plot the results of taking the simultaneously recorded cortical data as $Y(t)$, in red, cortical data from random times, in green, cortical data from the same random times, but with the segments shuffled across epochs for each channel (i.e. paired with a different ripple events for different channels), and in turquoise, the same random segments shuffled across channels as well. GC for hippocampus to cortex is shown in the top panels, and GC for cortex to hippocampus in the bottom panels. The mean across channels is shown at left and the differences for individual channels at right.

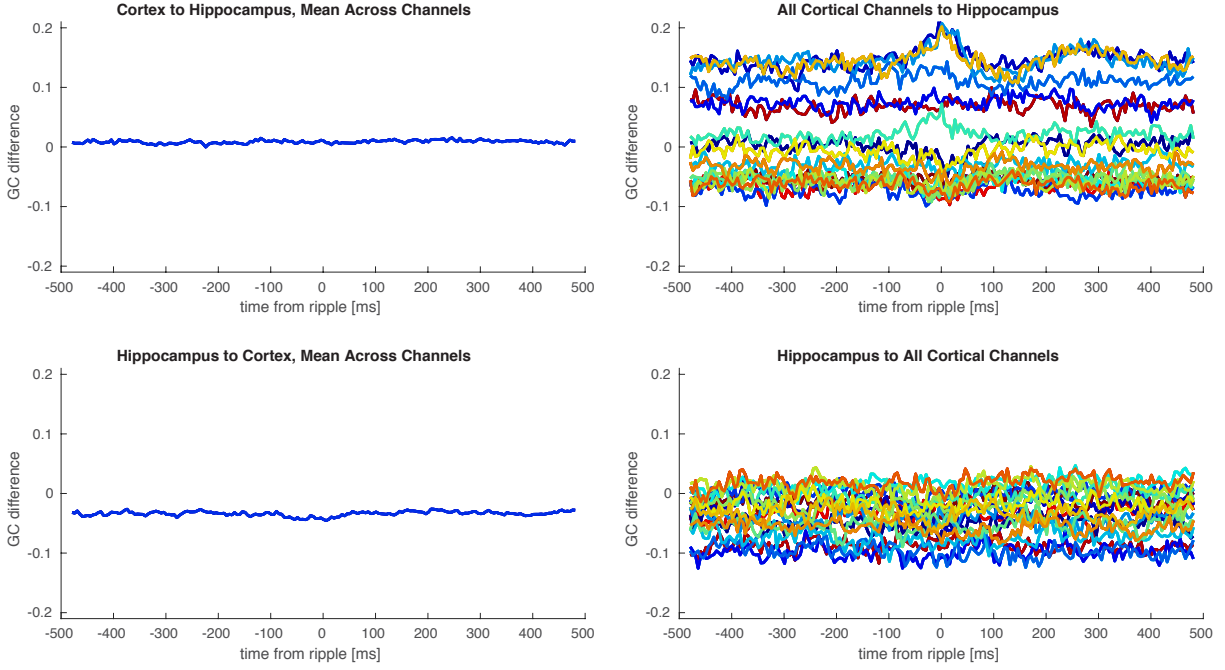


Figure 4.8: For the mean across channel pairs (at left) and for each individual channel pair (at right), the difference in the GC log ratio of errors between the time-matched X and Y pairings (plotted in blue in Fig. 4.7) and the pairing of X with Y shuffled across both epochs and channels (plotted in turquoise in Fig. 4.7). The GC differences for hippocampus to cortex are shown in the top panels, and for cortex to hippocampus in the bottom panels. The mean across channels is shown at left and the differences for individual channels at right (with each channel plotted in a different color).

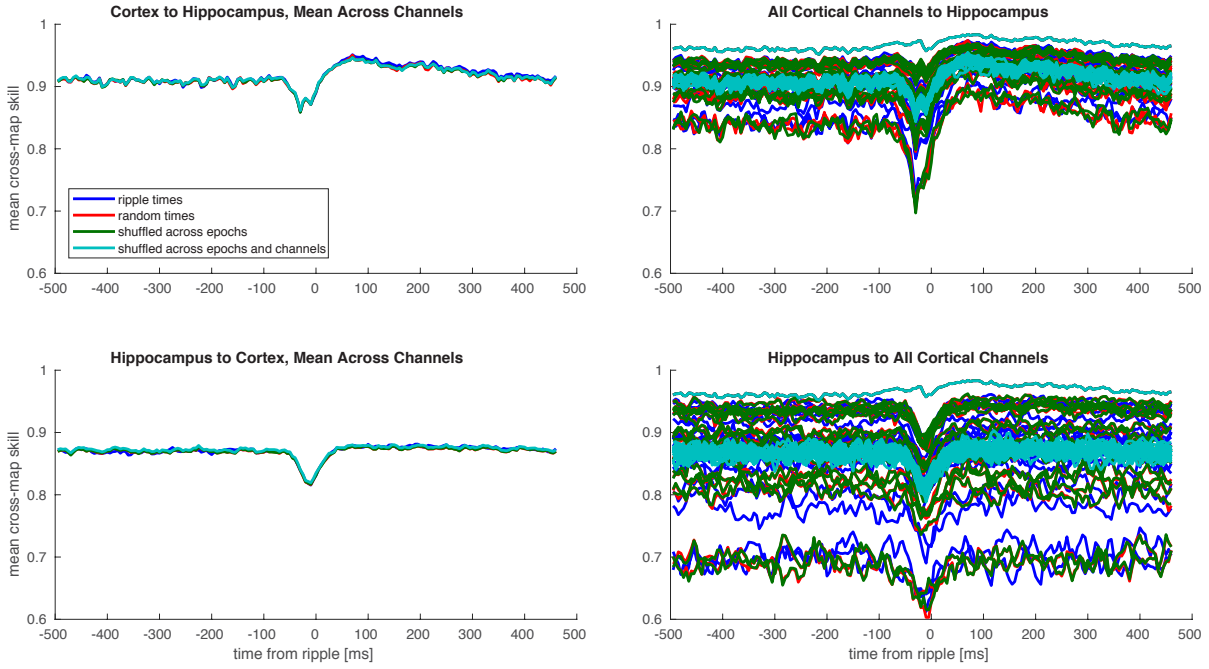


Figure 4.9: CCM was applied to one second of data from the hippocampus surrounding 375 ripple events. These hippocampal recordings were designated as X , and several different definitions of Y were used. In blue, we plot the results of taking the simultaneously recorded cortical data as Y , in red, cortical data from random times, in green, cortical data from the same random times, but with the segments shuffled across epochs for each channel (i.e. paired with a different ripple events for different channels), and in turquoise, the same random segments shuffled across channels as well. Cross-map skill for hippocampus to cortex is shown in the top panels, and cross-map skill for cortex to hippocampus in the bottom panels. The mean across channels is shown at left and the differences for individual channels at right.

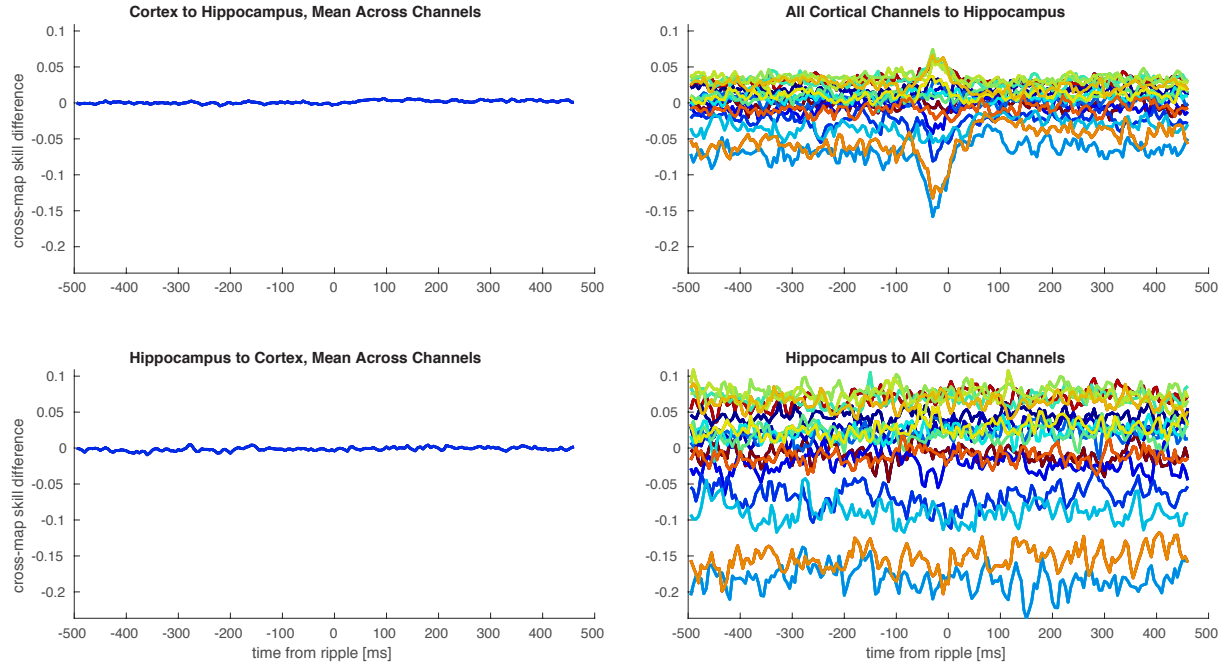


Figure 4.10: For the mean across channel pairs (at left) and for each individual channel pair (at right), the difference in the cross-map skill between the time-matched X and Y pairings (plotted in blue in Fig. 4.9) and the pairing of X with Y shuffled across both epochs and channels (plotted in turquoise in Fig. 4.9). The cross-map skill differences for hippocampus to cortex are shown in the top panels, and for cortex to hippocampus in the bottom panels. The mean across channels is shown at left and the differences for individual channels at right (with each channel plotted in a different color).

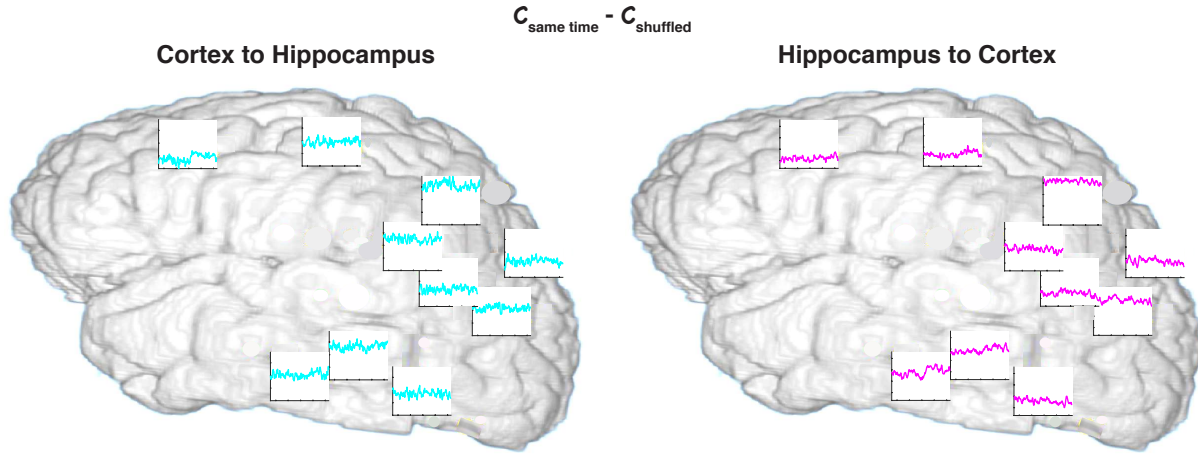


Figure 4.11: The difference (as in Fig. 4.6) in C between the time-matched hippocampus and cortex pairings and the pairing of hippocampus with cortical data shuffled across both epochs and channels are shown at each cortical recording site. The differences for C_{xy} (cortex to hippocampus) are shown at left in cyan, and C_{yx} (hippocampus to cortex) at right in magenta.

the ripple times in this view seems to be at a temporal location relatively near the hippocampus for the hippocampus-to-cortex direction. For the cortex-to-hippocampus direction, the most notable ripple-time-associated signal would seem to be in the frontal lobe.

4.4 Discussion

Despite the limitations of demonstrating information flow between interconnected brain regions in an event-related manner, these results do provide some tentative indications of how these tools could inform ideas about the memory consolidation process. The particular patterns of change in C that correspond to ripple events with sharp waves are highly evocative of a biphasic, back-and-forth communication pattern between hippocampus and cortex, which would correspond well to established ideas about the interactions of these regions [Rothschild et al., 2017]. No final conclusions about the precise timing of such communication can be drawn from these results, due to the lack of temporal specificity in the pairing of hippocampus with cortex,

but they strongly suggest the need for further study of these data and refinement of these tools. Also noteworthy is the relative speed of applying CD-DDA. As with classical DDA, once a model form and delays have been selected, the model can quickly be applied to new data. For a subset of data used for benchmarking consisting of 375 one-second epochs from one pair of channels, and using the same 40 ms windows for all methods, GC runs in 1.06 s, CCM in 21.97 s and CD-DDA in 0.01 s. All methods were implemented in MATLAB 9.5 (R2018b) on the same 12-core (Intel Xeon X5690 @ 3.47 GHz) system. The DDA detector calls an executable written in C for a key step in the procedure.

Chapter 4 is coauthored with Claudia Lainscek, Christopher E. Gonzalez, Xi Jiang, Jorge González-Martínez, Eric Halgren, and Terrence J. Sejnowski. The dissertation author was the primary author of this chapter.

Chapter 5

Delay Differential Analysis of Seizures and Peri-Ictal States

Abstract

Epilepsy is a common, debilitating neurological disorder. Reliable detection and prediction of seizures remain significant challenges. Intracranial recordings are a critical tool for clinicians who seek to identify the seizure onset zone in focal epilepsy cases, a target for surgical intervention. These recordings also provide an unparalleled opportunity to study in great detail the development and progression of seizures as well as both pre- and post-ictal states of the brain. Using both electrocorticogram (ECoG) and stereoelectroencephalogram (sEEG) recordings from human patients undergoing treatment for epilepsy, we can test new methods for detecting seizures and pre-ictal states. Reliable detection of pre-ictal states could form the basis of an early warning system for seizures, which could be highly beneficial to patients. Delay Differential Analysis (DDA) is among the most promising new techniques for detecting and predicting seizures. Past work applying DDA to epilepsy focused on data from 13 patients. Here the tools developed for systematically expanding this analysis to additional recordings are described, and preliminary

results from these new data are discussed. This includes the precise marking of seizure times and preliminary results of seizure prediction with DDA.

5.1 Epilepsy Background

Epilepsy is one of the most common neurological disorders, affecting more than 50 million people worldwide [WHO, 2019]. Epilepsy is a chronic neurological disorder characterized by seizures, interruptions of normal behavior due to abnormal synchronous neuronal activity [Fisher et al., 2005]. Given its broad and significant impact, treatments and strategies for ameliorating the symptoms and effects of epilepsy have been the subject of a great deal of effort. These remain a challenge however, given the extreme diversity of presentation of epilepsy, which is better characterized as a variety of related disorders than a singular disease [Fisher et al., 2005].

A major advance in the management and treatment of epilepsy could be achieved with reliable tools for the prediction of seizures. Early warnings of oncoming seizures could improve quality of life for patients and also provide a pathway to potential interventions to disrupt seizure onset. A great deal of research in recent years has focused on the challenge of seizure prediction, and this work has been aided by improvements in long-term electroencephalogram (EEG), electrocorticogram (ECoG), and stereoelectroencephalogram (sEEG) recordings and the establishment of large seizure databases. A number of techniques using different features for seizure prediction have been proposed. Using phase synchronization, Maiwald et al. achieved a sensitivity of 42% and a false prediction rate (FPR) of 0.15 seizures per hour [Maiwald et al., 2004]. This approach was improved upon by adding bivariate empirical mode decomposition to achieve a sensitivity over 70% [Zheng et al., 2014]. Another approach used spike rate difference in ictal and pre-ictal periods to achieve 75.8% sensitivity and an FPR of 0.09 seizures per hour [Li et al., 2013]. Bandarabadi et al. obtained similar performance using features from selected channel pairs and a support vector machine for classification [Bandarabadi et al., 2015].

Of particular interest here are approaches using tools and measures from the study of nonlinear dynamical systems. The rule-based seizure prediction method of Aarabi et al. uses correlation dimension, correlation entropy, noise level, Lempel-Ziv complexity and largest Lyapunov exponent. These features support a sensitivity for the method of 90.2% and an FPR of 0.11 seizures per hour [Aarabi and He, 2012].

Other exciting new approaches to seizure prediction utilize the tools of neural network-based machine learning. In particular, convolutional neural networks using the frequency content of scalp EEG or intracranial recordings have achieved a sensitivity of 81.4% and FPR of 0.06 seizures per hour [Truong et al., 2018].

5.2 DDA of Epilepsy

Lainscsek et al. previously applied DDA to intracranial recordings of seizures and found dynamics that could segment the data based on the occurrence of seizures and other events and provide insights into onset zone localization [Lainscsek et al., 2017]. In this study, in contrast to the structure selection approach described in Chapter 1 (an exhaustive search of models and delays subject to relatively tight constraints), structure selection was carried out using a genetic algorithm. This led to the selection of a single three-term DDA model:

$$\dot{x} = a_1x_1 + a_2x_2 + a_3x_1^4, \quad (5.1)$$

where $x_i = x(t - \tau_i)$, and a set of 8 delay pairs (see [Lainscsek et al., 2017]) which provided features with a clear delineation of seizures. DDA features obtained were further subjected to truncated higher-order singular value decomposition (THOSVD), using different unfoldings of three-dimensional DDA feature tensors across channels, delay pairs, and features (a_1 , a_2 , a_3 , and ρ) to obtain both global (with channel as a dimension) and local (without channel as a dimension) measures of the dynamics.

5.3 Data Management

The promising results of prior work applying DDA to intracranial recordings of seizures and peri-ictal data calls for similar analysis on additional subjects. The Cash laboratory at the Massachusetts General Hospital (MGH) has collected data from over 120 epilepsy patients over a number of years, which presents a singular opportunity to look for signatures of seizure activity and propagation as well as pre-ictal states that are common to multiple patients or patient populations. However, these data also present several challenges. First, the sheer volume of data makes handling them difficult and computationally memory-intensive. There are also a number of challenges introduced by the fact that these data were collected in a clinical environment for clinical purposes. The placement of electrodes is different for each patient depending on the clinical hypothesis of seizure origin, and often even subsequent recordings from a single patient will have channels sorted or labeled differently. There are also often limitations in the availability of additional relevant information about the recordings. We are able to get a great deal of useful information about the timing of seizures and inter-ictal activity from the clinical reports, but since these are written with clinical, rather than research needs in mind, they are typically not exhaustive. Finally, as these data are all exported from a clinical recording system, there are frequently inconsistencies in the clock times associated with the various files. Despite these challenges, these data are extremely valuable and provide an exciting opportunity to make significant progress in our understanding of epilepsy.

In order to overcome these noted challenges, we have established a set of tools for the processing and analysis of these intracranial recordings which place the data and relevant metadata into as uniform a format as possible. In the following section, these tools that have been developed for processing the data at MGH are described.

5.3.1 Data Collection

All data used for ongoing analysis of seizures and peri-ictal states with DDA are sEEG and ECoG recordings from patients with intractable epilepsy. These recordings used a standard clinical recording system (XLTEK, Natus Medical Inc., San Carlos, CA) with sampling rates of 500 Hz or higher. Placement of the intraparenchymal (sEEG) electrodes and subdural electrode arrays was chosen to confirm the hypothesized seizure focus and locate epileptogenic tissue in relation to essential cortical areas, thus directing surgical treatment. The decision to implant, as well as the electrode targets and the duration of implantation were entirely clinically based with no input from this research study. All data were handled following protocols approved by the MGH Institutional Review Board according to National Institutes of Health guidelines.

5.3.2 Processing Data Remotely

MGH maintains a Linux Cluster, ERISOne, which has proven to be a useful platform for processing and analyzing these data at MGH. It is also possible to access ERISOne from off-site using a virtual private network (VPN) or SSH key. There are a limited number of nodes on ERISOne designated as filemove nodes. These nodes allow for data to be directly read from and written to the archive server on which the data are stored. This allows for each step in the data processing pipeline to be run on ERISOne without having to copy large amounts of data back and forth from the archive server to another location for analysis.

5.3.3 Data Processing and Analysis Code

All code for the processing, formatting, and analysis of the intracranial epilepsy data is written in MATLAB, with the exception of the critical step of computing DDA features, which is done by an executable written in C. This executable requires input data to be saved in an ASCII format. This avoids some issues that arise when using MATLAB-formatted data files on different

systems and allows for the processing of large files.

The first steps in processing new data from new patients are carried out by the MATLAB function `run_all_prepare_files.m`. This function takes as its only input a list of the subject numbers to process. This function calls several other functions:

- `make_links.m` creates links to the original data files, exported from the clinical system to European Data Format (EDF) files
- `read_hdrs.m` saves informational MATLAB files with timing and other information from the EDF headers
- `check_times.m` checks for abnormalities in the clock times of the files (e.g. a file that starts before the previous file ends) and corrects these abnormalities if possible
- `check_channels.m` creates lists of the channel labels for all files, which must be checked manually for inconsistencies

All of the information files created by `run_all_prepare_files.m` are saved in a single location and organized by subject. Not all necessary information can be obtained from the EDF file headers, however. It is also necessary to consult a database or the relevant clinical reports for information on the seizures observed during the monitoring period. Specifically, we record the times of all seizures and whether they were electrographic or electroclinical, as well as the clinically-determined onset channels. This information, along with the full list of channel numbers and original data file numbers are added to a MATLAB script, `get_SUBJ_info.m`, which, when run, adds all of this information to the MATLAB workspace for further analysis. This is the most difficult and time-consuming step in the analysis process, since it cannot be automated and requires careful consideration of the information in the clinical report. In particular, the channels designated as onset channels are often not totally clearly delineated, therefore some degree of subjectivity in this designation is unavoidable. DDA features may be able to provide an additional

confirmation of electrodes as onset channels or not. It is also sometimes questionable whether the times as noted in the reports in fact align exactly with the clock times from the EDF headers, but usually any deviation from the apparent time of seizures is quite small. The full list of channels included in `get_SUBJ_info.m` includes only intracranial channels, but the scalp EEG channels, if available, can be added to another variable defined in the script.

To ensure consistency, all times (i.e. of seizures or of recording start and end times) are recorded in hours from midnight on the first day of recording for a given patient. The dates and clock times recorded in the reports are converted to these absolute times in hours using the MATLAB function `make_S_all.m`.

After these initial steps to obtain all necessary metadata, the intracranial recordings themselves must be converted to an ASCII format readable by the DDA executable. This can be done using the MATLAB function `save_ASCII_files.m`, which takes the subject number as its only input, which relies on a modified version of the `edf2ascii` utility [van Beelen, 2019].

After obtaining all necessary metadata and converting the intracranial recordings to ASCII format, DDA features can be computed using the MATLAB function `run_DDA_CLUSTER.m`, which calls the executable `run_DDA_EPILEPSY` to run one model with one delay pair. Depending on input options, this tool will compute single-channel and/or cross-channel features, and can use different specified window lengths and step sizes. Features are saved into specified locations, organized by patient and file number, as well as the parameters used in computing the features (single-channel vs. cross-channel, window length, etc.). Output files from the executable containing DDA features are also in ASCII format.

The ASCII files containing DDA features can be converted to MATLAB format using `convert_all_outputs_to_MAT.m`. This format is easier to work with for additional analysis and also uses less memory for long-term storage. These MATLAB files are also broken up into shorter time segments so that each individual file is smaller and easier to load as needed.

5.4 Precise Seizure Timing with DDA features

A key advantage of DDA features is that they can provide a more obvious or visually apparent readout of ictal activity than the raw EEG, especially with regard to the precise timing of both early ictal activity in the onset channels and generalization to other channels. This is due to the models used having been specifically selected for sensitivity to seizures. In order to indicate the potential utility of DDA features for seizure marking and timing, multiple non-epileptologists have visually inspected the DDA features obtained around clinically noted seizures, but without any indication of the precise clinically-determined onset times. Each individual marked the onset, generalization and end time of each seizure. This allows for the consistency of non-expert judgments of these times to be evaluated. These individual and group average markings of onset times can then be compared to the clinical onset times.

Fig. 5.1 shows an example of the plotted DDA a_1 coefficient in Eq. (5.1), which can be used for marking of seizure onset, generalization, and offset times. Coefficient a_1 is generally the feature that shows the seizure activity most clearly, although seizures are generally visible in the other features as well. Additionally, there is significant variation among patients and among seizures for the same patient, with some seizures being more apparent than others. In general, electroclinical seizures are quite apparent in the DDA features, but purely electrographic (i.e. subclinical) seizures are less so.

Fig. 5.2 shows the results of multiple non-epileptologists marking seizure onset times and the comparison of their marked times with the clinically designated onset time. Each individual viewed the a_1 coefficient for a series of seizures from several patients and marked the onset, generalization, and end times for each. The mean time across individuals was computed and the differences between each individual time and the mean time were then averaged across individuals and across seizures. These average differences are plotted as blue circles in Fig. 5.2; individual's average differences from the group mean are plotted as turquoise points. For each

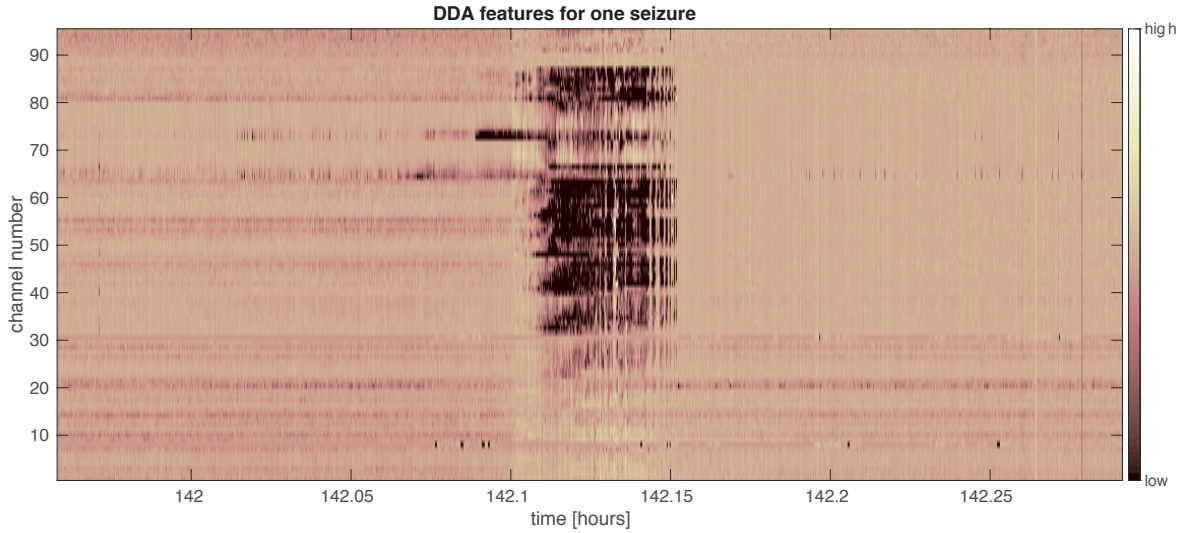


Figure 5.1: DDA a_1 coefficient plotted (as a heatmap) for all channels for one seizure. Viewing the seizure in this way makes the onset, generalization, and offset times visually apparent.

patient, the average across seizures of the difference between the clinical onset time and the average of the individuals' markings is plotted as a red cross. In general, the markings of multiple non-epileptologists are quite tightly clustered, with the average deviation from the average being lower than the difference between this group average time and the clinical onset time.

5.5 Seizure Prediction

DDA features also hold great promise for seizure prediction. A central challenge, however, is the high degree of heterogeneity of epilepsy. Indeed, Fisher et al. note that “epilepsy is not one condition, but is a diverse family of disorders, having in common an abnormally increased predisposition to seizures” [Fisher et al., 2005]. And even among the relatively restricted population of patients from whom these intracranial recordings are obtained (i.e. all have focal epilepsy which is pharmaco-resistant), there is great diversity among the types of seizures. Despite these challenges, it is possible in some patients to predict seizure onset using DDA features.

Fig. 5.3 shows an example of data from one patient (Patient EP4) in which pre-ictal

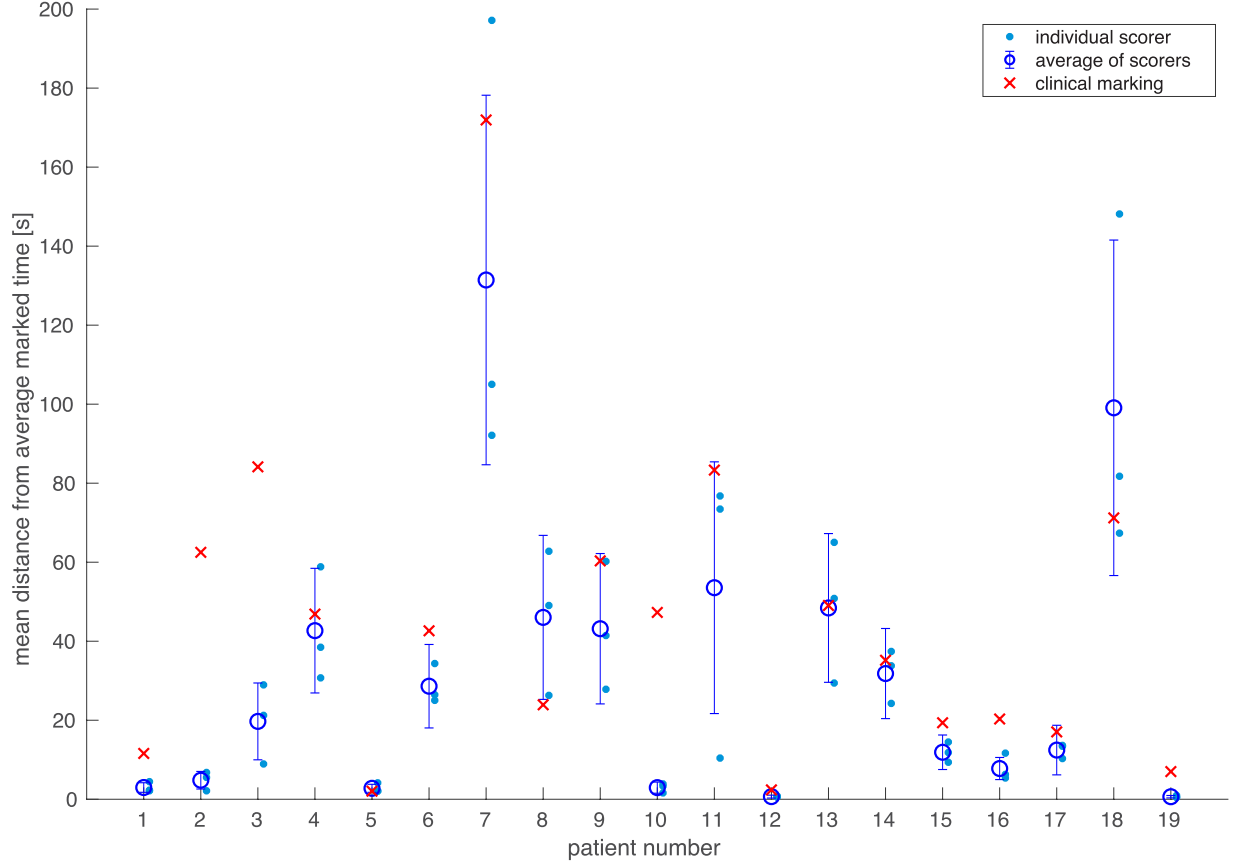


Figure 5.2: Three non-experts viewed all seizures for 19 patients and marked the onset time based on visual inspection of the DDA a_1 coefficient in Eq. (5.1) across all channels. The mean (across seizures) differences in seconds between the group average marked onset time and each individual's marked onset times are marked with blue circles above, with turquoise points showing each individual's deviation from the group average. The red crosses above show the mean (across seizures) differences between the clinically-determined onset times and the group average marked onset times.

states are highly distinguishable from the remainder of the recording. These states are marked by divergence between the cross-channel features obtained from onset channels and non-onset channels. The observed pre-ictal states for this patient initiate between half an hour and two hours prior to the corresponding seizure onset. These exceptionally clear pre-ictal states could serve as a highly reliable means of seizure prediction. It is not yet known whether Patient EP4 represents a particular subtype of epilepsy which could correspond to these predictive states.

More generally, seizure prediction using DDA can be framed as a classification problem similar to those discussed in Chapters 2 and 3 of this dissertation, and in other prior applications of DDA [Lainscsek and Sejnowski, 2013c, Lainscsek et al., 2013a, Sampson et al., 2019, Lainscsek et al., 2019]. The recordings from Patient EP4, with clearly observable pre-ictal states, provide ideal training data for the development of a DDA seizure predictor. For training and testing, data from the 30 minutes preceding seizure onset are categorized as pre-ictal and data from the period between 4 hours and 30 minutes preceding seizure onset are categorized as pre-pre-ictal. Data from the 10 minutes following clinical or subclinical seizure onset, and the 30 minutes before subclinical seizures are excluded. All other data are categorized as inter-ictal. One-second windows with no overlap are used to compute the DDA features.

Using data from Patient EP4 for structure selection, and following the procedures outlined in Chapter 1, a model form, delays, and a set of weights for optimally combining features was obtained to separate pre-ictal states, thereby providing seizure prediction. Fig. 5.4 shows the DDA features used for prediction and the corresponding area under the ROC (A') for classification of the three different designated states (pre-pre-ictal, pre-ictal, and inter-ictal) in two patients. When sorting channels by classification performance, the differences between the three states in the DDA features become visually apparent. The classification performance is also quite high for some channels. For Patient EP4, it is possible to obtain A' over 0.9 when classifying pre-ictal vs. inter-ictal one-second data segments in the onset channels.

Table 5.1 shows the results of this predictor on a set of recordings from 8 patients. While

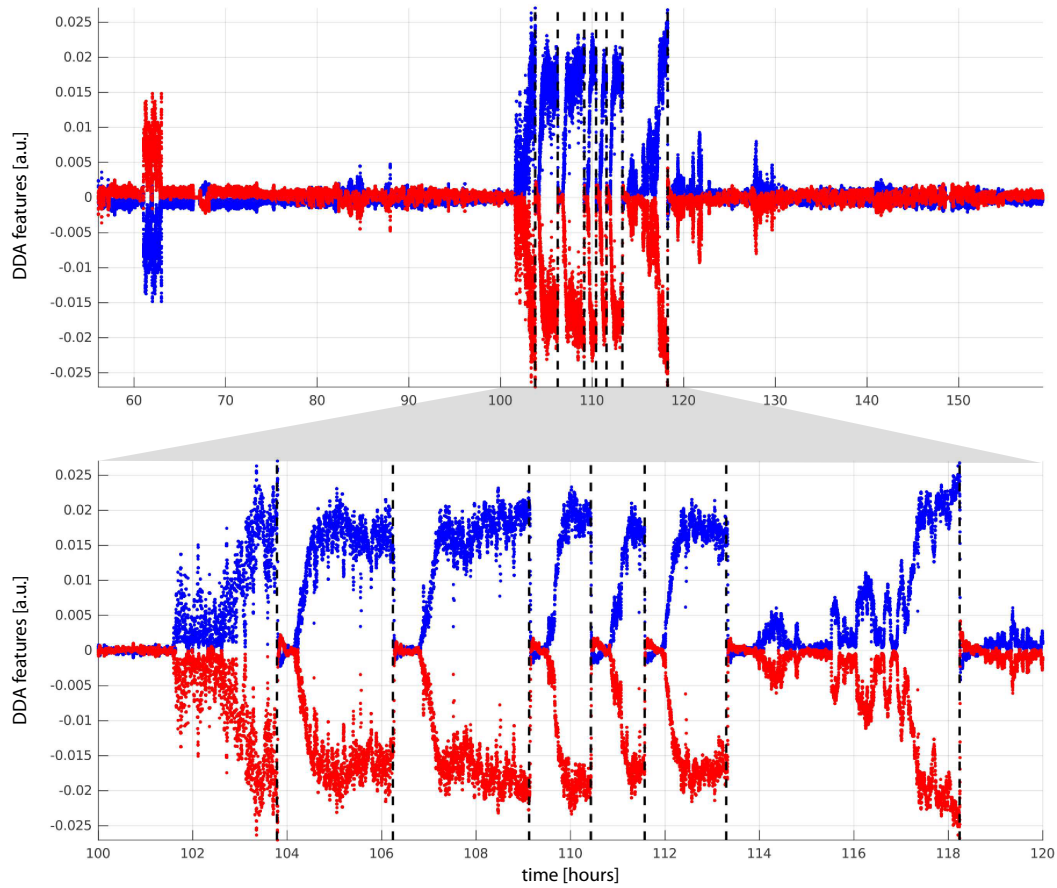


Figure 5.3: For Patient EP4, DDA features show clear pre-ictal states which begin between half and hour and two hours before seizure onset. In the above plots, onset channels and non-onset channels are plotted in blue and red respectively, with the mean of all channels subtracted from both to make visually apparent the divergence between these sets of channels in the pre-ictal states. The top panel shows these mean-subtracted features over a period of over four days, and the lower panel shows a period of 20 hours during which seven seizures occur. The clinically-marked onset time of each seizure is marked with a vertical dashed line in both panels. The deviations from the mean visible just after hour 60 in the top panel are the result of stimulation.

Table 5.1: DDA seizure prediction performance from the best-predicting channel in 8 patients

Patient Number	3-way classifier	pre-pre-ictal vs. inter-ictal	pre-ictal vs. inter-ictal	pre-pre-ictal vs. pre-ictal
EP1	0.6979	0.7252	0.7731	0.6209
EP3	0.6853	0.8042	0.8024	0.6627
EP4	0.8767	0.8937	0.9680	0.7894
EP5	0.6705	0.7211	0.7571	0.5915
EP8	0.7630	0.7766	0.8738	0.6804
EP10	0.6785	0.7222	0.8051	0.7511
EP20	0.7086	0.7538	0.8050	0.6453
EP21	0.8084	0.8977	0.9165	0.6420

there is a great deal of variation across subjects, pre-ictal vs. inter-ictal classification is quite robust in general. For the best-classifying channel from each patient, A' is always over 0.75. For pre-pre-ictal vs. inter-ictal classification in the same channels, A' is always over 0.7.

Finally, to illustrate how these models can be applied in a framework closer to real-time seizure prediction, the pre-ictal vs. inter-ictal and pre-pre-ictal vs. inter-ictal classifiers were applied in sliding 50-minute windows across the entire lengths of recordings. In Fig. 5.5 shows A' values for separating the features from these sliding windows and from all pre-ictal data from the same patient. (Pre-ictal recordings from a patient could equivalently be used to predict future seizures in that same patient.) Here, high values of A' indicate that no seizure is imminent, since the data at that timepoint are easily distinguishable from the pre-ictal data. Drops in A' indicate that the data at that time are more similar to pre-ictal data and that a seizure could therefore be imminent.

5.6 Discussion and Future Directions

A robust pipeline for the processing of large volumes of invaluable intracranial data has now been established, and the portion of the full data archive that has been fully processed has been greatly expanded and is still expanding. Further, by introducing a set of programs and

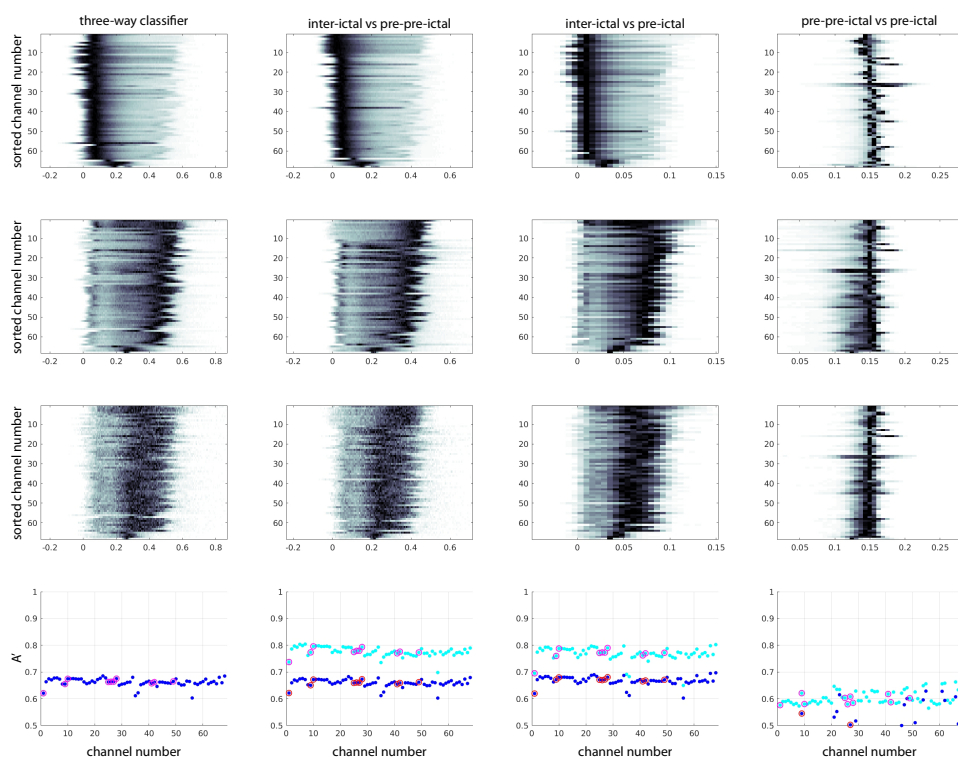
procedures designed for local processing of the data at MGH, we have also made easier the task of future data processing as new recordings are collected and archived. In addition, in the process of adapting this set of tools for use in a new computing environment, we have also identified limiting factors which can be minimized to accelerate the process of adapting similar tools for new applications in the future.

Significant work remains to be done in the realm of seizure prediction, as the results described in Section 5.5 are preliminary. Beyond establishing the optimal model form, delays, and weights for predicting the greatest possible number of seizures across patients, a key advance would be to categorize and subgroup patients or specific seizures depending on their characteristics—especially the characteristics of their respective pre-ictal states. Patient EP4 (shown in Fig. 5.3) currently stands out as having seizures with an exceptionally high level of predictability, but by training a seizure predictor on these data, it has already been possible to improve prediction for some other patients. It is likely that some other patients, perhaps representing a particular subtype of epilepsy, show similar pre-ictal states to this patient.

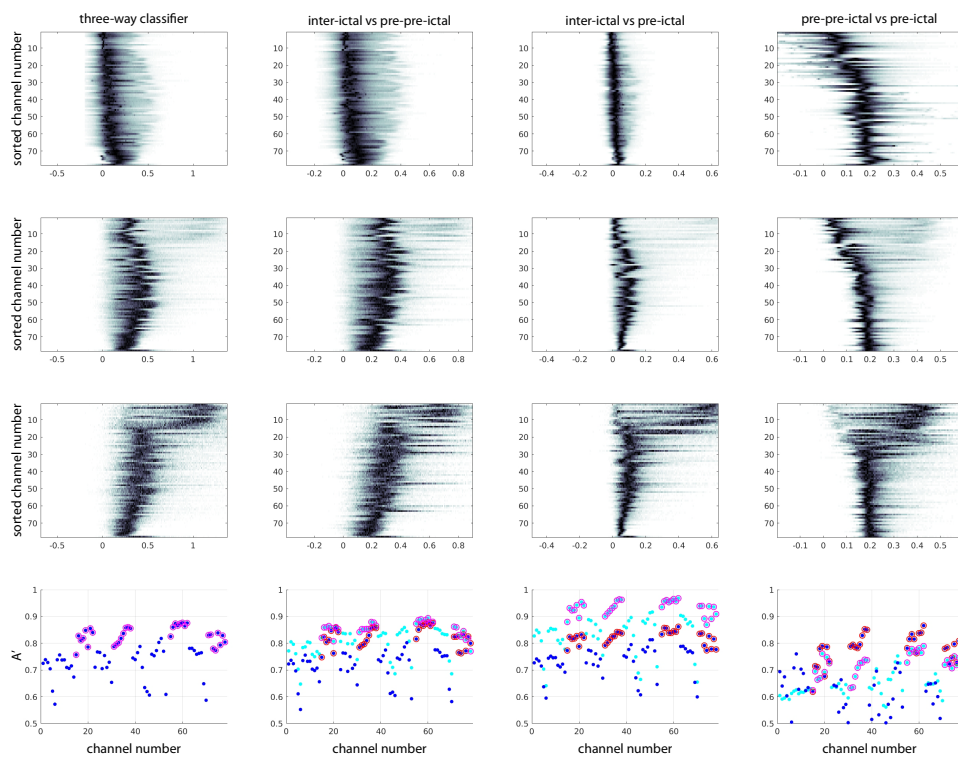
Chapter 5 is coauthored with Claudia Lainscsek, Jiayang Zhong, William McGrath, Daniel Sexton, Jonathan Weyhenmeyer, Diana Ponce-Morado, Sydney S. Cash, and Terrence J. Sejnowski. The dissertation author was the primary author of this chapter.

Figure 5.4: For Patient EP3 and Patient EP4, the top plots show heatmaps corresponding to histograms of DDA seizure prediction feature value across all channels. In these plots, the channels have been sorted by classification performance. The lowest row of plots for each patient shows the classification performance as measured by area under the ROC (A'). The leftmost column shows the results of a three-way (pre-pre-ictal vs. pre-ictal vs. inter-ictal) classifier, the second column shows the results of a pre-pre-ictal vs. inter-ictal classifier, the third column shows the results of a pre-ictal vs. inter-ictal classifier, and the rightmost column shows the results of a pre-pre-ictal vs. pre-ictal classifier. Across the lowest row of plots for each patient, A' for the three-way classifier is plotted in dark blue for comparison, and A' for the corresponding two-way classifier is plotted in cyan. Onset channels are circled in red or magenta.

Patient EP3



Patient EP4



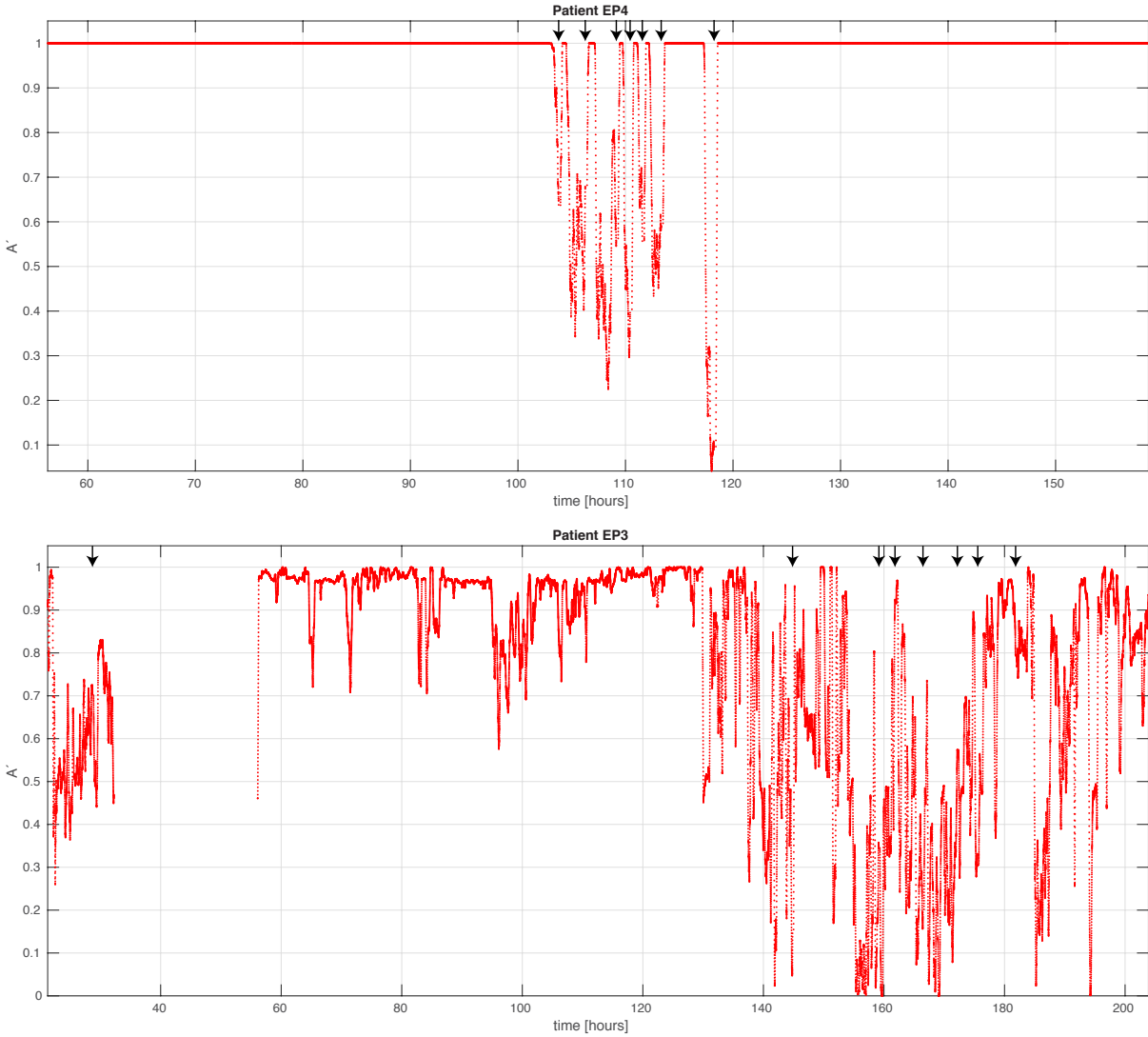


Figure 5.5: For Patient EP3 and Patient EP4, DDA features were computed in 50-minute sliding windows (with 20-second step size) across hours-long recordings using models selected to separate pre-ictal from inter-ictal data. A' is plotted for separating data in each window from all pre-ictal data (10 minutes or less before seizure onset) for the same patient. Each point is plotted at the end of the 50-minute time window used. High values of A' indicate that the data is highly separable from pre-ictal data, and therefore no seizure is imminent. Low values of A' indicate that the data is more similar to pre-ictal data and that a seizure could therefore be oncoming. Clinically marked seizure onset times are shown as arrows at the top of the plots.

Chapter 6

Conclusions

The focus of this dissertation work has been exploring the range of applications of delay differential analysis (DDA) to neural data. DDA has proven to be a robust and versatile tool for data classification in several contexts in neurosciences. This includes classifying data across large and diverse patient groups and detecting particular phenomena within a single recording.

In Chapter 1, the motivations for using nonlinear tools for neural data analysis were discussed, the methodological details of DDA were described, and details of tests related to the numerical derivative and windowing of the data were provided. Further, results of testing DDA on simulated data from dynamical systems were given as examples of the application of the technique, and to demonstrate its usefulness with high-dimensional systems. Finally, Cross-Dynamical DDA (CD-DDA) was introduced as a novel tool for information flow between systems.

In Chapter 2, DDA was applied to electroencephalogram (EEG) recordings from the Consortium on the Genetics of Schizophrenia (COGS-2). Using EEG corresponding to an auditory deviant stimulus paradigm, recordings from schizophrenia patients (SZ) and nonpsychiatric comparison subject (NCS) were separated, and the timing of each component of the auditory deviance response complex was investigated using DDA features and a traditional event-related

potential (ERP) analysis technique.

In Chapter 3, DDA was used to detect sleep spindles in human intracranial recordings. This novel approach to spindle detection was compared with several well-validated spindle detection techniques, and also applied to a publicly available scalp EEG data set. The best performance was achieved when combining a DDA measure with a spectral measure, gaining the benefit of unique perspectives from fundamentally different approaches.

In Chapter 4, a novel formulation of DDA, CD-DDA was used to investigate causal connections between brain regions in human intracranial recordings. Specifically, the communication between the hippocampus and various sites throughout the cortex was investigated during the observed occurrence of sharp wave ripple events in the hippocampus. Limitations in the interpretability of fluctuations in this measure at a fine timescale around discrete events were noted, but there remains some indication of detectable communication between these brain regions. This is a promising target for future study and refinement of the approach.

In Chapter 5, a set of tools developed for the analysis of a large archive of intracranial recordings from epilepsy patients was described. Preliminary results of the analysis of new data were also discussed, including the utility of DDA features for improving the measurement of the precise onset, generalization, and offset times of seizures, and using DDA features for seizure prediction.

The refinement of DDA as a tool for neural data analysis has provided a wealth of new results, and a set of tools that hold great potential for new analyses of new data. CD-DDA in particular is an exciting new technique, which is still in need of further study to ensure that the measures it provides can be interpreted with confidence, but it already shows great promise, despite the challenges related to high temporal resolution around short, discrete events in a time series. Like classical DDA, it can provide a valuable new tool to approach open questions from a new perspective. More broadly, with these nonlinear tools, we can begin to shine a light on aspects of the function of the nervous system that derive from its inherent nonlinearity and dynamical

complexity. These tools will have a critical role to play in the future, both in groundbreaking research in basic neuroscience and in developing new tools for the treatment of neurological and psychiatric disorders.

Appendix A

Supplementary Information for Chapter 2

A.1 Embeddings

Delay Differential Analysis (DDA) is a functional embedding technique that employs delay differential equations (DDEs) to provide a highly specific and low-dimensional representation of an observable time series. Even without direct access to all the system state variables, DDEs can provide a small subset of features that capture the relevant dynamics of a system [Kremliovsky and Kadtke, 1997, Kadtke and Kremliovsky, 1997, Lainscsek et al., 2012, Lainscsek and Sejnowski, 2013b, Lainscsek and Sejnowski, 2015]. DDEs combine two different types of embedding: a delay embedding and a derivative embedding. An embedding refers to the mapping of a single time series into a multidimensional object [Whitney, 1936, Packard et al., 1980, Takens, 1981, Sauer et al., 1991]. A delay embedding maps a single time series $x(t)$ into a multidimensional space spanned by $x(t)$ and its delayed versions $x(t - \tau_n)$, where the delays τ_n are positive integers. Similarly, a derivative embedding maps $x(t)$ into a higher dimensional space constructed by $x(t)$ and its n -order derivatives. Both types of embedding, delay and derivative embedding, are diffeomorph to the original unknown dynamical system. DDA combines the two types of embedding by relating them in a polynomial function to extract relevant dynamical

information from a single time series. For simplicity, Eq. (A.1) shows a DDE with a first-order derivative embedding and a delay embedding:

$$\dot{x} = \sum_{i=1}^I a_i \prod_{n=1}^N x_{\tau_n}^{m_{n,i}} \quad (\text{A.1})$$

Here, $\tau_n, m_{n,i}, I \in \mathbb{N}_0$, $x = x(t)$, and $x_{\tau_n} = x(t - \tau_n)$.

In order to illustrate the relationship between a dynamical system and its DDA features (the coefficients a_i and the delays τ_n) we discuss the following simple example. Consider a second-order linear differential equation defined by:

$$\ddot{x} = ax \quad (\text{A.2})$$

Eq. (A.2) is then equivalent to the following dynamical system:

$$\begin{aligned} \frac{dx}{dt} &= y \\ \frac{dy}{dt} &= ax \end{aligned} \quad (\text{A.3})$$

The above dynamical system has the special solution:

$$\begin{aligned} x &= \cos(\omega t) \\ a &= -\omega^2. \end{aligned} \quad (\text{A.4})$$

Therefore, it is easy to see that Eq. (A.3) is the functional form of the derivative embedding of $x(t) = \cos(\omega t)$. Both $x(t)$ and $y(t)$ can be solved by numerically integrating Eq. (A.3), and the phase space plot of x and $\frac{y}{\omega}$ can be plotted as shown in Supplementary Fig. A.1.

If only $x(t)$ is available, the phase space spanned by $x(t)$ and $y(t)$ can be still constructed using an embedding. In Supplementary Fig. A.2, two embeddings of the time series from Supplementary Fig. A.1 are shown—a delay embedding and a differential embedding. Since an

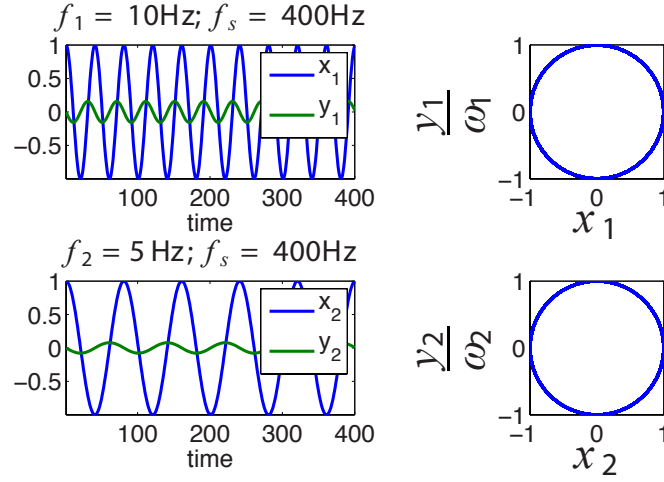


Figure A.1: The left plots show the numerically integrated $x(t)$ and $y(t)$ in Eq. (A.3). The same integration constant dt was used for both plots. Only the parameter a was different according to the sampling rates $f_1 = 10$ Hz for the upper plots and $f_2 = 5$ Hz for the lower plots. On the right plots the phase plots are shown. The y axes are scaled by ω to obtain a circle.

ellipse is topologically equivalent to a circle, both types of embeddings recover the structure of the original phase space (compare the right plots in Supplementary Fig. A.1 with Supplementary Fig. A.2).

For the above example, the simplest linear delay differential equation (DDE) is

$$\dot{x} = bx_\tau \quad (\text{A.5})$$

with $x_\tau = x(t - \tau)$. Again the special solution here is $x(t) = \cos(\omega t)$ [Falbo, 1995, Lainscsek et al., 2012]

$$\begin{aligned} -\omega \sin(\omega t) &= b \cos(\omega t - \tau) \\ b &= (-1)^n \omega \\ \tau &= \frac{\pi(2n-1)}{2\omega} \end{aligned} \quad (\text{A.6})$$

where b is proportional to the frequency ω , and τ is inversely proportional to the frequency ω .

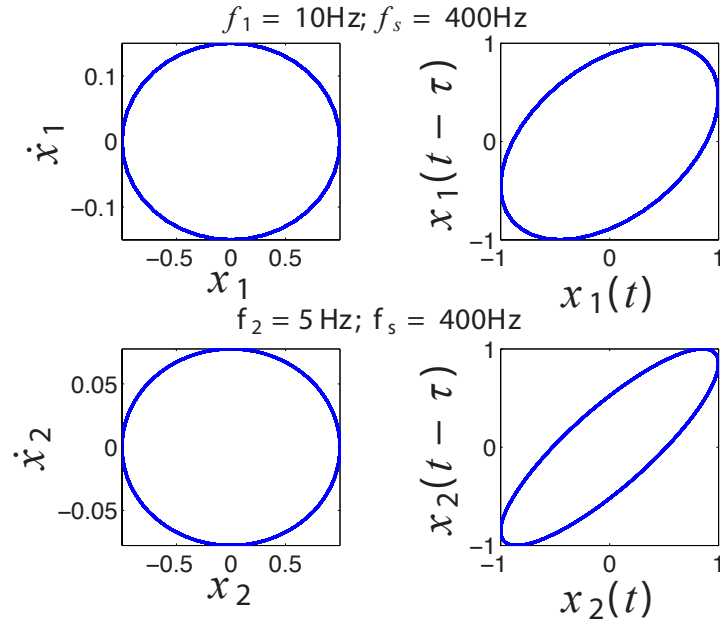


Figure A.2: Derivative (left) and delay (right) embeddings for the time series $x(t)$ of the numerically integrated Eq. (A.3).

DDA is therefore able to recover the systems parameters of the dynamical system. More examples including non-linear dynamical systems and the structure of the parameter space can be found in [Lainscsek, 2011, Lainscsek, 2012].

For a more complex dynamical system, relating the derivative and delay embeddings could provide additional information regarding the underlying dynamical system. This is the principal idea behind DDA: a derivative embedding is matched to a polynomial function of delay embeddings to uncover dynamical information.

A.2 DDA Details

For the derivative embedding, the first-order derivative is numerically estimated by using a weighted center derivative [Miletics and Molnárka, 2005]

$$\dot{x} = \frac{1}{N} \sum_{n=1}^N \frac{x(t-n) - x(t+n)}{2n}, \quad (\text{A.7})$$

where N is the number of forward and backwards steps in computing the derivative. Here we use $N = 10 \delta t$, where $\delta t = \frac{1}{f_s}$ and f_s is the sampling rate ($f_s = 1000$ Hz).

The least square mean error, used as one of the features in DDA, is computed according to:

$$\rho = \sqrt{\frac{1}{K} \sum_{k=1}^K \left(\dot{x}_{t_k} - \sum_{i=1}^I a_i \prod_{n=1}^N x_{t_k, \tau_n}^{m_{n,i}} \right)^2} \quad (\text{A.8})$$

for K data points, where $\dot{x}_{t_k} = \dot{x}(t_k)$ and $x_{t_k, \tau_n} = x(t_k - \tau_n)$.

A.2.1 Cross-Trial DDA

Cross-trial DDA is illustrated in Supplementary Fig. A.3. In this approach, data from equivalent time windows in multiple trials is combined to compute cross-trial DDA features. A key advantage of this approach is that it allows shorter time windows to be used. Additionally, it can serve as a test of dynamical coherence across trials.

In Supplementary Fig. A.4(a) the data show dynamical incoherence. This means each trial covers a small and different part of the embedding. In Supplementary Fig. A.4(b) the data are dynamically coherent and therefore the same part of the embedding is covered in each trial. There should be a much bigger difference between single-trial and cross-trial DDA in the case of dynamical incoherence (as in Supplementary Fig. A.4(a)). This is because in the case of dynamical incoherence, the short data windows do not sample enough of the embedding to efficiently estimate the coefficients and the variance of the coefficients for all trials will be large.

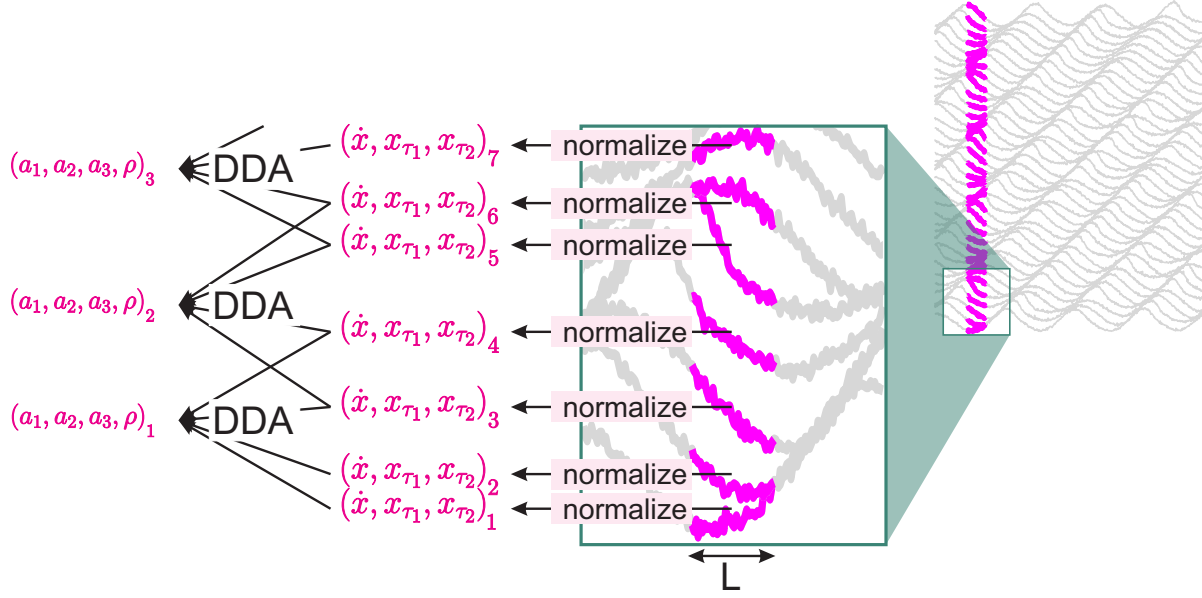


Figure A.3: Cross-trial DDA: In this example, data from sets of 4 trials are combined assuming ergodicity [Boltzmann, 1898]. L is the temporal window length (see also Fig. 2). The data from each time window and trial are normalized to zero mean and unit variance and then combined to compute the DDA features simultaneously.

Therefore, mean features computed across the individual channels will be very different from the cross-trial features, where despite the small data window a large part of the embedding is covered. For the case of dynamical coherence, both versions give similar results. In [Lainscsek and Sejnowski, 2015] the same concept was introduced for spectral analysis.

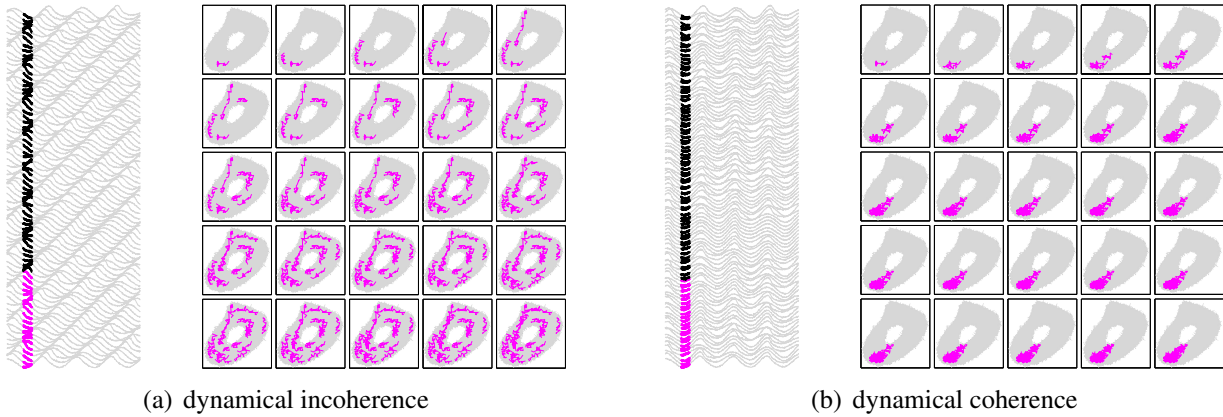


Figure A.4: Cross-trial DDA coverage of the embedding space.

A.2.2 Structure Selection

Supervised Structure Selection

Supplementary Fig. A.5 illustrates how the DDE model that best distinguishes NCS subjects from SZ patients was selected.

To obtain an equal amount of data for each subject, the first 150 deviant trials were used. The raw EEG data from these 150 trials were first normalized to zero mean and unit variance to disregard amplitude information. Then the DDA features were extracted from 50 ms time windows ($L = 50\delta t$) with a time shift of 5 ms (resulting in 74 temporal windows from trials of 450 ms) either single- or cross-trial. We combined 40 trials for the cross-trial DDA (in Supplementary Fig. A.3 the example of 4 trials is shown) and shifted the cross-trial windows by 10 trials (in Supplementary Fig. A.3 the example of 2 trials is shown). This resulted in 12 trial windows across the 150 trials for each temporal window (see Fig. A.3).

Next, repeated random subsampling cross-validation (CV) [Kohavi et al., 1995] was carried out: the data are divided into training data (5 from each group) and testing data (remaining 1620 subjects) 20000 times (see Supplementary Fig. A.5). This process is carried out for each of seven three-term DDEs that have been shown to be discriminative EEG models,

$$\begin{aligned}\text{Model \#1: } \dot{x} &= a_1x_1 + a_2x_2 + a_3x_1^2 \\ \text{Model \#2: } \dot{x} &= a_1x_1 + a_2x_2 + a_3x_1x_2 \\ \text{Model \#3: } \dot{x} &= a_1x_1 + a_2x_2 + a_3x_1^3 \\ \text{Model \#4: } \dot{x} &= a_1x_1 + a_2x_2 + a_3x_1^2x_2 \\ \text{Model \#5: } \dot{x} &= a_1x_1 + a_2x_2 + a_3x_1^4 \\ \text{Model \#6: } \dot{x} &= a_1x_1 + a_2x_2 + a_3x_1^3x_2 \\ \text{Model \#7: } \dot{x} &= a_1x_1 + a_2x_2 + a_3x_1^2x_2^2\end{aligned}\tag{A.9}$$

with $x_n = x(t - \tau_n)$. For each of these models all possible delay combinations for τ_1 and τ_2

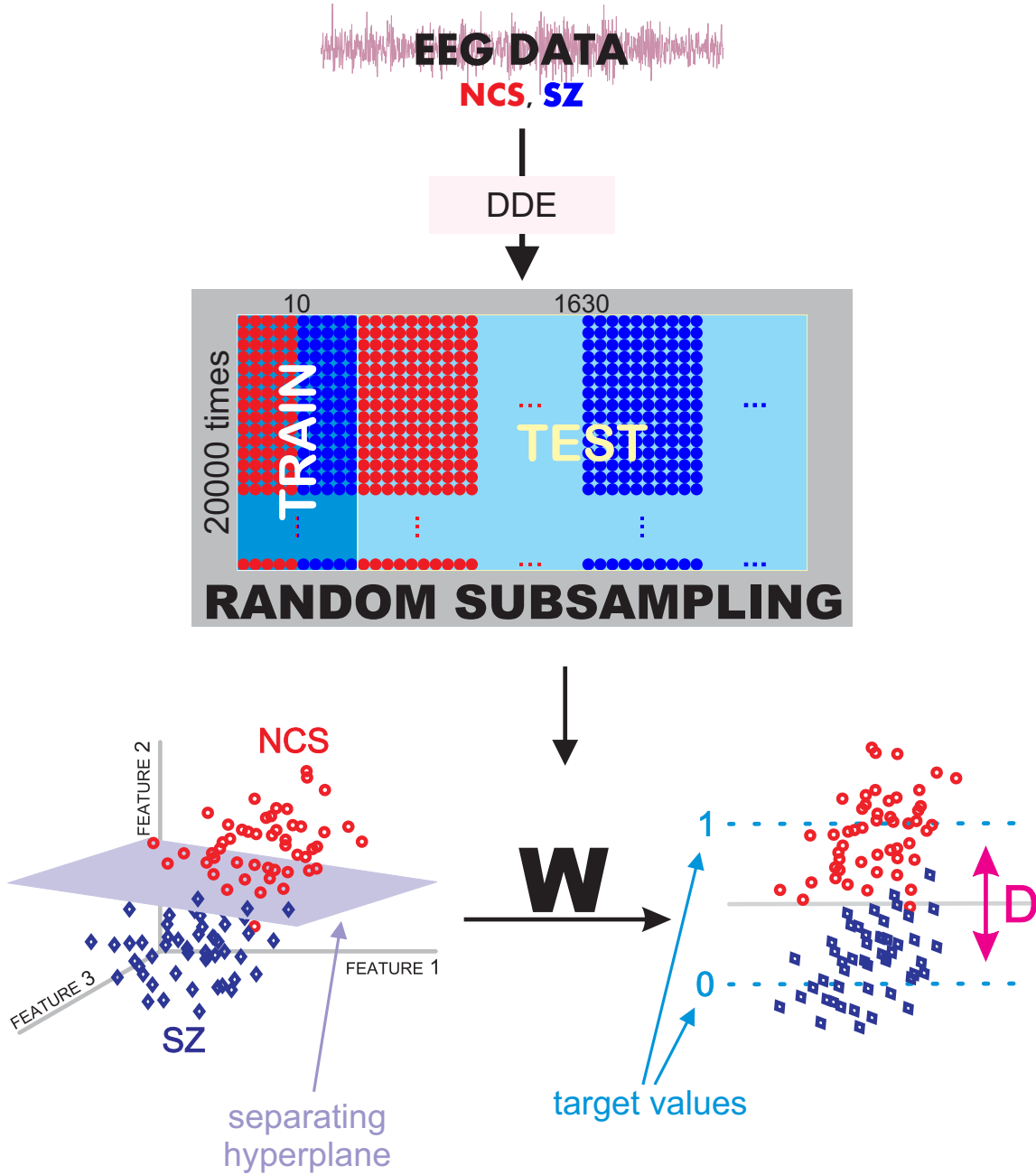


Figure A.5: Random subsampling cross-validation and construction of the weight matrix.

between $1\delta t$ and $15\delta t$ are run through the CV. This is done for each of the 74 temporal windows separately. We trained on only 5 subjects in each group to show that 1) if “typical” subjects are selected for training, a low number should be enough for training, 2) if “atypical” subjects are

selected for training, the classification/results can be opposite 3) the ratio of “typical” to “atypical” subjects sets the maximum performance possible by using the mean weights of the CV. Here we have a big data set of 1640 subjects. This is not the case for many studies. We therefore want to use the beforementioned properties of our CV to dynamically cluster the data and separate “typical” subjects from “atypical” ones. This dynamic clustering technique is a topic of its own and will be published separately.

The performance of the DDA classification was evaluated by computing the area under the ROC curve, A' . In this case, the ROC curve was a plot of the cumulative distribution function of the NCS subject group versus that of the SZ subject group. To calculate the area, the distances from the hyperplane D were ranked from the greatest positive value to the greatest negative value, and the A' approximated as follows [Hand and Till, 2001]:

$$A' = \frac{S_0 - \frac{n_0(n_0 + 1)}{2}}{n_0 n_1} \quad (\text{A.10})$$

where S_0 is the sum of the ranks of the NCS subject classifications, n_0 is the number of NCS subjects, and n_1 is the number of SZ subjects.

Supplementary Figs. A.6 and A.7 show the distributions of the best DDE model forms and the best delays for each of the 74 temporal windows.

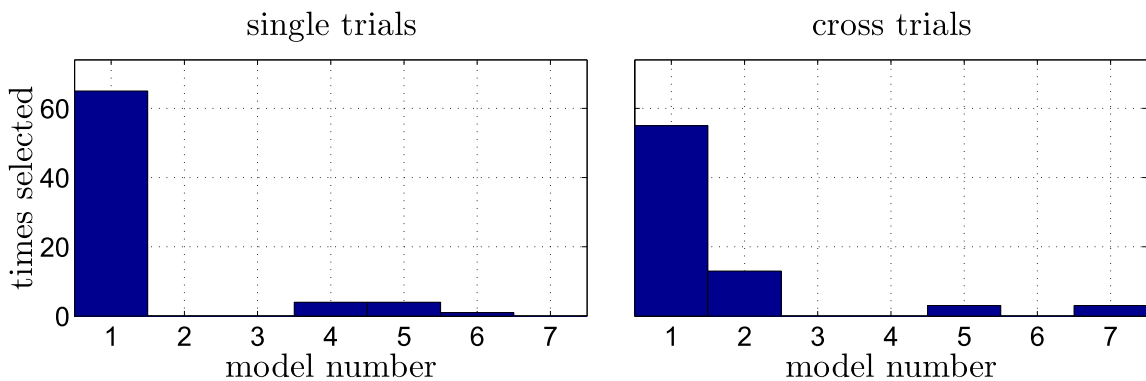


Figure A.6: Models selected for each of the 74 time windows.

Model #1 was the most often selected model for both approaches (Supplementary Fig. A.6). This makes sense since all the data came from the same dynamical system and therefore the model structure should be consistent across the data. This model was used for the analyses presented in the main paper. The delays selected across the time windows (Supplementary Fig. A.7) were

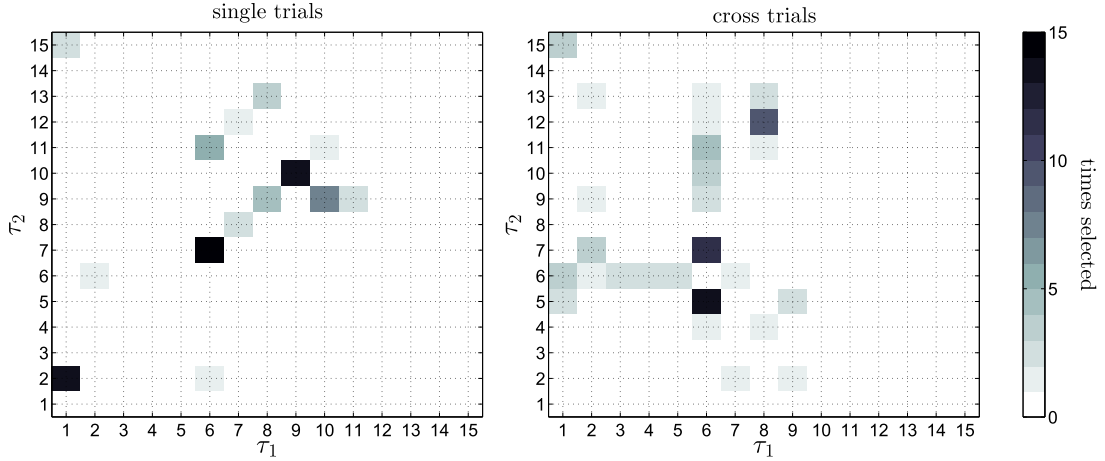


Figure A.7: Delay pairs selected for each of the 74 time windows using supervised structure selection for model #1 (A.11).

different for the single- and cross-trial approach. This could point to differences in the cross-trial dynamical coherence across the temporal windows and will be a topic of further investigation. In this paper, we chose the best performing delay pair for each temporal data window to account for the changes in the nonlinear cocktail of dominant time-scales throughout the trials.

Unsupervised Structure Selection

In an unsupervised approach, the model selected as best representing the dataset is the one which generated the minimum error ρ . The analysis was again done in both the single- and cross-trial approaches and for each temporal window separately. The only model used was

$$\dot{x} = a_1 x_1 + a_2 x_2 + a_3 x_1^2 \quad (\text{A.11})$$

and the best delays were very consistent across the trials. The best delays were $\tau = (9, 3) \delta t$ for the single-trial approach and $\tau = (3, 8) \delta t$ for the cross-trial approach for most of the windows (see Supplementary Fig. A.8).

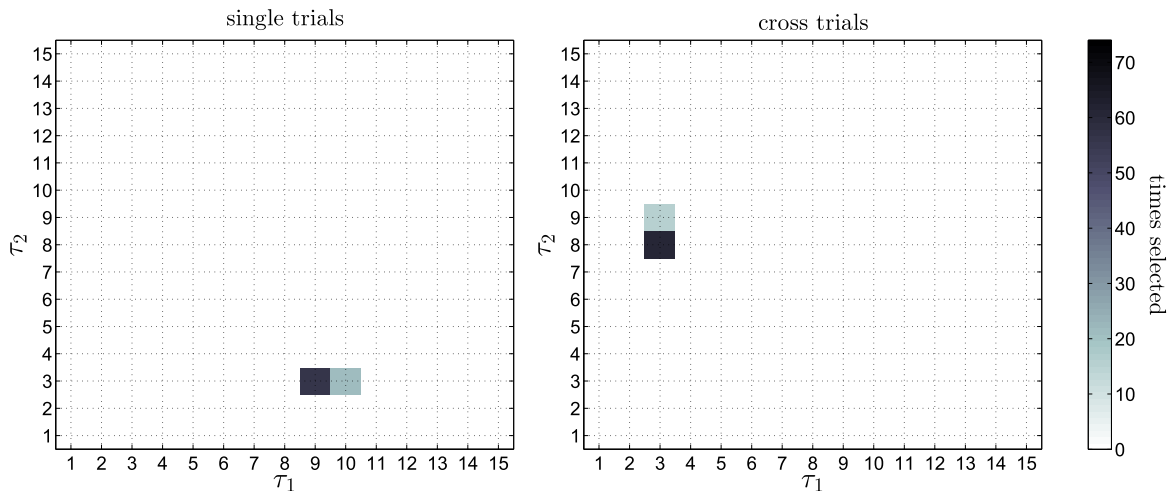


Figure A.8: Delay pairs selected for each of the 74 time windows using unsupervised structure selection for model #1 (A.11).

A.3 COGS-2 Clinical and Demographical Characteristics

COGS-2 included 1790 (966 SZ subjects and 824 NCS subjects) participants tested at the five COGS-2 test sites: University of California San Diego, University of California Los Angeles, University of Washington, University of Pennsylvania, and Mount Sinai School of Medicine. ERP data were collected from all 1790 participants, but 91% of the participants (SZ $n = 877$, NCS $n = 753$) had data with sufficient quality (as determined by signal-to-noise ratio) to permit further analysis. As noted in our previous publication ([Light et al., 2015]), the reason for data loss primarily included correctable errors in administration (e.g., failing to plug the electrodes into the correct port on the analog amplifiers, and failing to pause the recording due to poor signal quality, including excessive muscle artifact resulting in most trials being excluded using the automated artifact rejection procedures).

The clinical and demographic characteristics of the participants are shown in Supplementary Table A.1 (adapted from [Light et al., 2015]). The diagnostic criteria were established by a modified version of the Structural Clinical Interview for DSM-IV (SCID) [Swerdlow et al., 2015]. All SZ patients enrolled in COGS-2 met the DSM-IV diagnostic criteria for schizophrenia or schizoaffective disorder depressive type. Non-psychiatric comparison subjects were enrolled if they had:

- no current or past psychotic disorder
- a known biological family history
- no history of psychosis in a 1st degree relative
- no current Axis I mood disorder
- no Cluster-A Axis II disorder
- no current regular treatment with psychoactive medication

Table A.1: Clinical and demographic characteristics of the COGS-2 participants. Adapted from [Light et al., 2015]. MMSE = Mini-Mental State Exam; SAPS = Scale for the Assessment of Positive Symptoms; SANS = Scale for the Assessment of Negative Symptoms.

	Controls (NCS)	Patients (SZ)	<i>p</i>
Sample size	753	877	
Age Mean \pm SD	38.63 \pm 12.80	46.25 \pm 11.23	<0.001
Education Mean \pm SD	14.99 \pm 2.20	12.63 \pm 2.15	<0.001
Male	371 (49%)	616 (70%)	<0.001
Race	<i>n</i> (% full sample)	<i>n</i> (% full sample)	<0.001
Caucasian	438 (58%)	380 (43%)	
African American	159 (21%)	162 (18%)	
Other	156 (21%)	335 (38%)	
Smoker	84 (11%)	464 (53%)	<0.001
Age of Onset Mean \pm SD		22.52 \pm 7.23	
MMSE Mean \pm SD	33.61 \pm 1.69	31.12 \pm 3.31	<0.001
SAPS Mean \pm SD		6.88 \pm 4.09	
SANS Mean \pm SD		11.64 \pm 5.36	

The exclusion criteria (adapted from [Swerdlow et al., 2015]) for the COGS-2 study are shown in Supplementary Table A.2.

Table A.2: Exclusion criteria for the COGS-2 testing. Adapted from [Swerdlow et al., 2015].

Exclusion Criteria	SZ	NCS
Adopted or family history unknown		X
Outside study age range of 18-65	X	X
Unable to understand consent due to language or competency	X	X
Physically unable to participate in testing of at least two endophenotypes	X	X
Previous endophenotype testing in the last 1 month	X	X
Previous neuropsychological testing in the last 3 months	X	X
Positive illicit drug or alcohol screen at the time of testing	X	X
Severe systemic illness that interferes with ability to be endophenotyped	X	X
Electroconvulsive treatment in the last 6 months	X	X
Alcohol or substance abuse in the past 1 month	X	X
Alcohol or substance dependence not in remission for 6 months	X	X
Significant head injury (loss of consciousness > 15 min and/or neurological sequelae)	X	X
Neurological illness (e.g., seizures, stroke, Parkinson disease)	X	X
Less than one 1 month psychiatrically stable	X	
Estimated premorbid IQ < 70 per Wide Range Achievement Test-Third Edition	X	X
History of psychosis in themselves or a family member (1st degree)		X
Current Axis I mood disorder		X
Cluster A personality disorder		X
Current treatment with antipsychotic agents		X
Current treatment with any psychoactive medication		X
Participated in COGS-1 testing	X	X
First-degree relative who has already participated in this study	X	X

Table A.3: Demographic characteristics of the four age subgroups.

	18–30 years	31–40 years	41–50 years	51–65 years
Sample size	366	230	294	380
SZ	105 (29%)	120 (37.5%)	159 (54%)	260 (68%)
Male	183 (50%)	115 (50%)	147 (50%)	190 (50%)

A.4 Statistical Details of t-test Results

This section provides statistical details regarding the unpaired student's t-tests performed for Figs. 1–3 to identify time-windows where the group differences were statistically significant. Cohen's d effect sizes, t-statistic values, and degrees of freedom over all the time-windows are reported in Supplementary Figure A.9. For ERP signals (deviant-minus-standard, deviant, and standard ERPs), an unpaired t-test was performed at each time-point (i.e. every ms) to identify time-points with significant group differences. For DDA signals, an unpaired t-test was performed at each time-window (i.e. every 10 ms). The false discovery rate (FDR) was adjusted using the Benjamini-Hochberg procedure [Benjamini and Hochberg, 1995].

In order to ensure that the group differences observed in the ERP and DDA signals are not mainly driven by the effects of age and gender, we repeated the above statistical analyses in subgroups that were matched in age and gender. We first divided the data into four subgroups based on the following four age ranges: 18–30 years old, 31–40 years old, 41–50 years old, and 51–65 years old. Within each age subgroup, we then sampled randomly to match the gender distribution. Table A.3 summarizes the demographic characteristics of the four subgroups; t-statistic and Cohen's d values for the four age groups across all the time-windows are reported in Supplementary Figures A.10 and A.11, respectively.

Appendix A, in full, is a reprint of the supplementary information for Chapter 2. Chapter 2, in full, is a reprint of the material as it appears in the Proceedings of the National Academy of Sciences 2019. Claudia Lainscsek, Aaron L. Sampson, Robert Kim, Michael L. Thomas, Karen Man, Xenia Lainscsek, The COGS Investigators, Neal R. Swerdlow, David L. Braff, Terrence

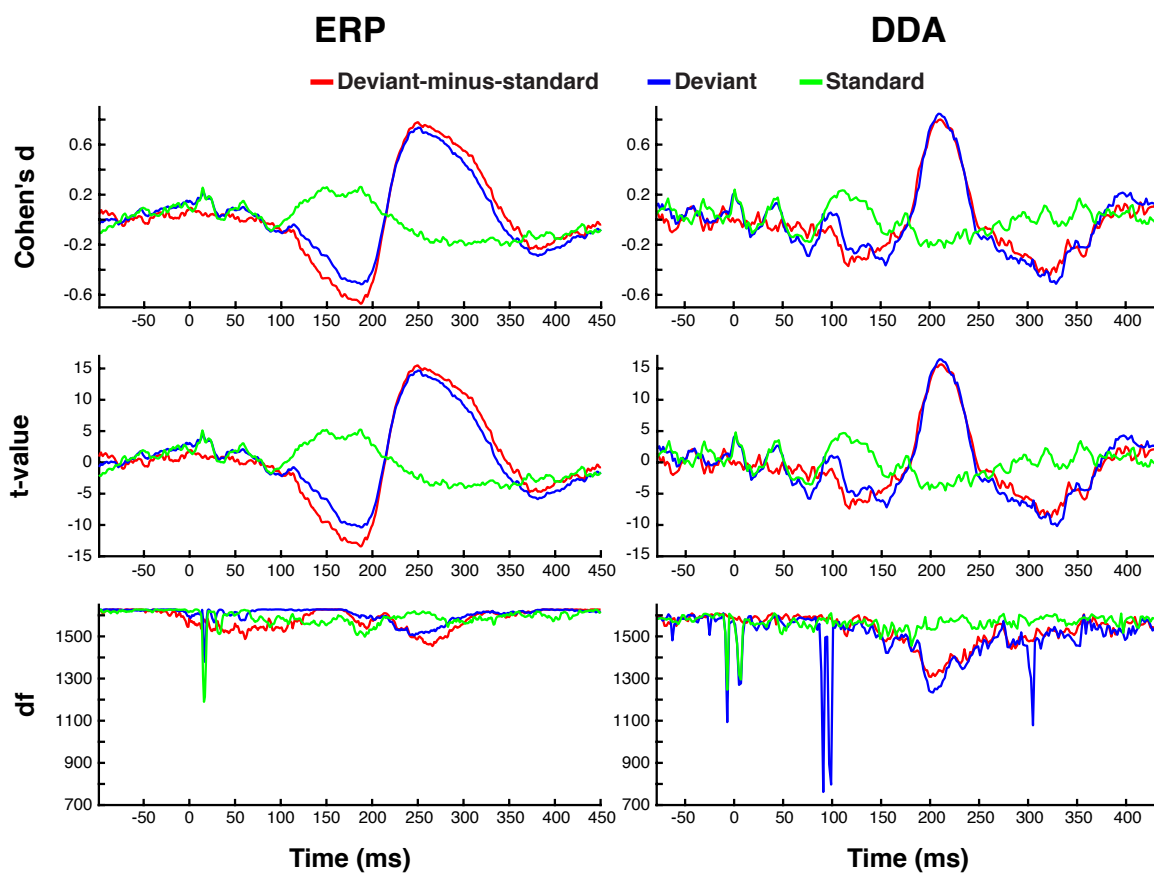


Figure A.9: Statistical details of the t-tests on ERP and DDA data. Effect sizes (Cohen's d), t-statistics, and degrees of freedom (df) for unpaired t-tests performed for ERP signals are shown in the left column. The same statistical measures are also plotted for the DDA time-series signals (right column).

J. Sejnowski, and Gregory A. Light. The dissertation author was a primary investigator and co-author of this paper.

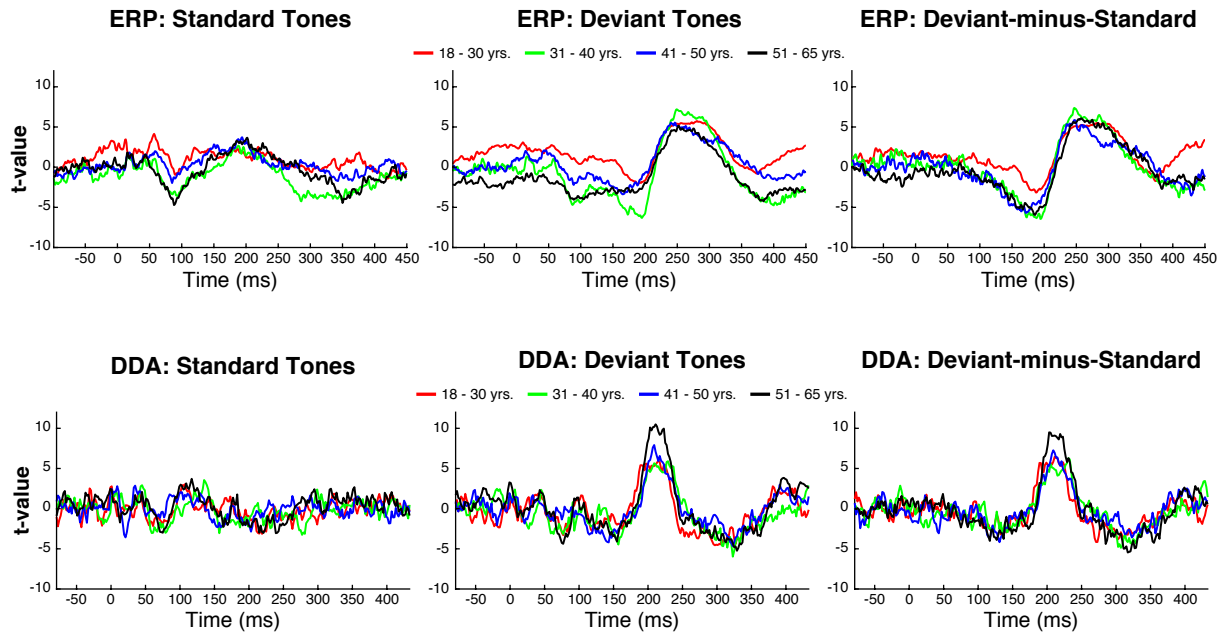


Figure A.10: t-statistics values from unpaired t-tests performed for ERP (top row) and DDA (bottom row) signals grouped by the four age bins are shown.

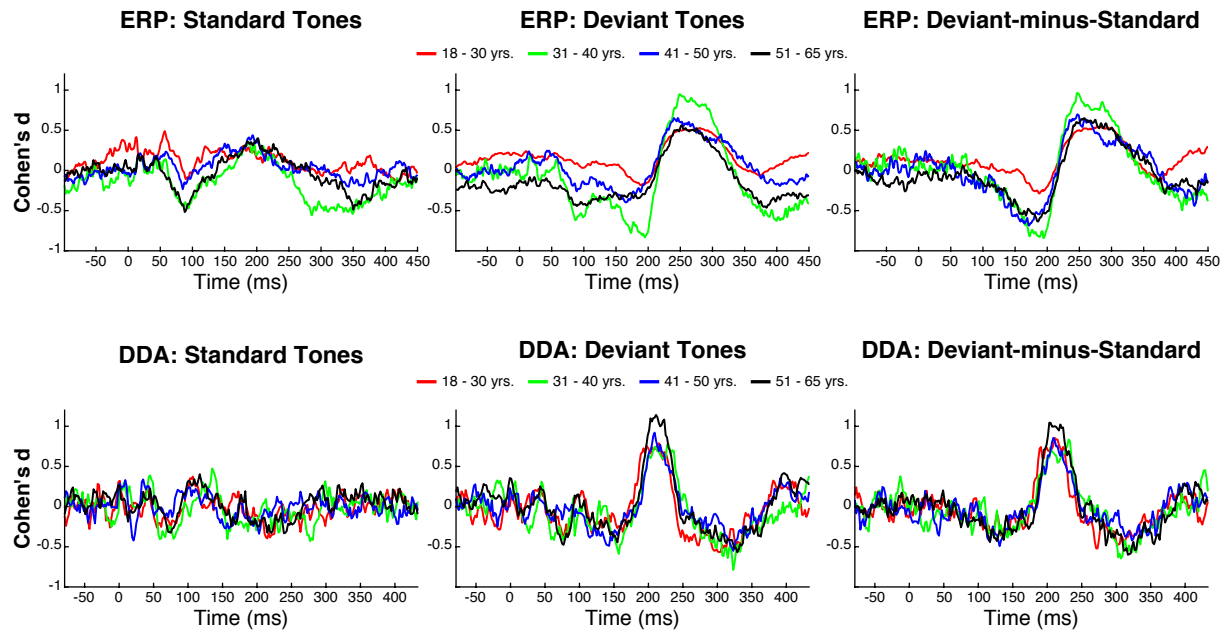


Figure A.11: Effect sizes (Cohen's d) values for ERP (top row) and DDA (bottom row) signals grouped by the four age bins are shown.

Bibliography

- [Aarabi and He, 2012] Aarabi, A. and He, B. (2012). A rule-based seizure prediction method for focal neocortical epilepsy. *Clinical Neurophysiology*, 123(6):1111–1122.
- [Andrillon et al., 2011] Andrillon, T., Nir, Y., Staba, R. J., Ferrarelli, F., Cirelli, C., Tononi, G., and Fried, I. (2011). Sleep spindles in humans: insights from intracranial EEG and unit recordings. *Journal of Neuroscience*, 31(49):17821–17834.
- [Auslander et al., 1964] Auslander, J., Bhatia, N., and Seibert, P. (1964). Attractors in dynamical systems.
- [Bandarabadi et al., 2015] Bandarabadi, M., Teixeira, C. A., Rasekhi, J., and Dourado, A. (2015). Epileptic seizure prediction using relative spectral power features. *Clinical Neurophysiology*, 126(2):237–248.
- [Basner et al., 2008] Basner, M., Griefahn, B., and Penzel, T. (2008). Inter-rater agreement in sleep stage classification between centers with different backgrounds. *Somnologie-Schlafforschung und Schlafmedizin*, 12(1):75–84.
- [Benjamini and Hochberg, 1995] Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the royal statistical society. Series B (Methodological)*, pages 289–300.
- [Berry et al., 2012] Berry, R. B., Brooks, R., Gamaldo, C. E., Harding, S. M., Marcus, C., Vaughn, B., et al. (2012). The AASM manual for the scoring of sleep and associated events. *Rules, Terminology and Technical Specifications, Darien, Illinois, American Academy of Sleep Medicine*.
- [Bódizs et al., 2009] Bódizs, R., Körmendi, J., Rigó, P., and Lázár, A. S. (2009). The individual adjustment method of sleep spindle analysis: methodological improvements and roots in the fingerprint paradigm. *Journal of neuroscience methods*, 178(1):205–213.
- [Boltzmann, 1898] Boltzmann, L. (1898). *Vorlesungen über Gastheorie*. Leipzig, J.A. Barth. (“Ergoden” in chapter III, §32).

- [Bonjean et al., 2012] Bonjean, M., Baker, T., Bazhenov, M., Cash, S., Halgren, E., and Sejnowski, T. (2012). Interactions between core and matrix thalamocortical projections in human sleep spindle synchronization. *The Journal of Neuroscience*, 32(15):5250–5263.
- [Breakspear, 2017] Breakspear, M. (2017). Dynamic models of large-scale brain activity. *Nature Neuroscience*, 20(3):340–352.
- [Breakspear and Terry, 2002] Breakspear, M. and Terry, J. (2002). Topographic organization of nonlinear interdependence in multichannel human eeg. *Neuroimage*, 16(3):822–835.
- [Breakspear et al., 2003] Breakspear, M., Terry, J., Friston, K., Harris, A., Williams, L., Brown, K., Brennan, J., and Gordon, E. (2003). A disturbance of nonlinear interdependence in scalp eeg of subjects with first episode schizophrenia. *Neuroimage*, 20(1):466–478.
- [Burle et al., 2015] Burle, B., Spieser, L., Roger, C., Casini, L., Hasbroucq, T., and Vidal, F. (2015). Spatial and temporal resolutions of eeg: Is it really black and white? a scalp current density view. *International Journal of Psychophysiology*, 97(3):210–220.
- [Del Re et al., 2014] Del Re, E., Bergen, S., Meshulam-Gately, R., Niznikiewicz, M., Goldstein, J., Woo, T., Shenton, M. E., Seidman, L., McCarley, R., and Petryshen, T. (2014). Analysis of schizophrenia-related genes and electrophysiological measures reveals ZNF804A association with amplitude of P300b elicited by novel sounds. *Translational psychiatry*, 4(1):e346.
- [Devuyst et al., 2011] Devuyst, S., Dutoit, T., Stenuit, P., and Kerkhofs, M. (2011). Automatic sleep spindles detection—overview and development of a standard proposal assessment method. In *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, pages 1713–1716. IEEE.
- [Dice, 1945] Dice, L. R. (1945). Measures of the amount of ecologic association between species. *Ecology*, 26(3):297–302.
- [Diekelmann and Born, 2010] Diekelmann, S. and Born, J. (2010). The memory function of sleep. *Nature Reviews Neuroscience*, 11(2):114.
- [Falbo, 1995] Falbo, C. E. (1995). Analytic and numerical solutions to the delay differential equation $y'(t) = \alpha y(t - \delta)$. In *Joint meeting of the Northern and Southern California sections of the MAA*, San Luis Obispo, CA.
- [Ferrarelli et al., 2007] Ferrarelli, F., Huber, R., Peterson, M. J., Massimini, M., Murphy, M., Riedner, B. A., Watson, A., Bria, P., and Tononi, G. (2007). Reduced sleep spindle activity in schizophrenia patients. *The American journal of psychiatry*, 164(3):483–492.
- [Fisher et al., 2005] Fisher, R. S., Boas, W. v. E., Blume, W., Elger, C., Genton, P., Lee, P., and Engel, J. (2005). Epileptic seizures and epilepsy: Definitions proposed by the international league against epilepsy (ilae) and the international bureau for epilepsy (ibe). *Epilepsia*, 46:470–472.

- [Fogel et al., 2007] Fogel, S., Nader, R., Cote, K., and Smith, C. (2007). Sleep spindles and learning potential. *Behavioral neuroscience*, 121(1):1–10.
- [Goldberg, 1998] Goldberg, D. (1998). *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley.
- [Granger, 1969] Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, pages 424–438.
- [Hagler et al., 2016] Hagler, D. J., Cash, S. S., and Halgren, E. (2016). Heterogeneous origins of human sleep spindles in different cortical layers.
- [Hagler et al., 2018] Hagler, D. J., Ulbert, I., Wittner, L., Erőss, L., Madsen, J. R., Devinsky, O., Doyle, W., Fabo, D., Cash, S. S., and Halgren, E. (2018). Heterogeneous origins of human sleep spindles in different cortical layers. *Journal of Neuroscience*, 38(12):3013–3025.
- [Hand and Till, 2001] Hand, D. and Till, R. (2001). A simple generation of the area under the ROC curve for multiple class classification problems. *Machine Learning*, (45):171–186.
- [Isomura et al., 2006] Isomura, Y., Sirota, A., Özen, S., Montgomery, S., Mizuseki, K., Henze, D. A., and Buzsáki, G. (2006). Integration and segregation of activity in entorhinal-hippocampal subregions by neocortical slow oscillations. *Neuron*, 52(5):871–882.
- [Judd and Mees, 1995] Judd, K. and Mees, A. (1995). On selecting models for nonlinear time series. *Physica D*, 82:426–444.
- [Kadtke and Kremliovsky, 1997] Kadtke, J. and Kremliovsky, M. (1997). Estimating statistics for detecting determinism using global dynamical models. *Phys. Lett. A*, 229(2):97.
- [Kirihaara et al., 2012] Kirihaara, K., Rissling, A. J., Swerdlow, N. R., Braff, D. L., and Light, G. A. (2012). Hierarchical organization of gamma and theta oscillatory dynamics in schizophrenia. *Biological psychiatry*, 71(10):873–880.
- [Knott et al., 2001] Knott, V., Mahoney, C., Kennedy, S., and Evans, K. (2001). EEG power, frequency, asymmetry and coherence in male depression. *Psychiatry Research: Neuroimaging*, 106(2):123 – 140.
- [Kohavi et al., 1995] Kohavi, R. et al. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Ijcai*, volume 14, pages 1137–1145.
- [Kremliovsky and Kadtke, 1997] Kremliovsky, M. N. and Kadtke, J. B. (1997). Using delay differential equations as dynamical classifiers. In *Applied nonlinear dynamics and stochastic systems near the millennium*, volume 411, pages 57–62. AIP Publishing.
- [Ktonas et al., 2007] Ktonas, P., Golemati, S., Tsekou, H., Paparrigopoulos, T., Soldatos, C., Xanthopoulos, P., Sakkalis, V., Zervakis, M., and Ortigueira, M. D. (2007). Potential dementia biomarkers based on the time-varying microstructure of sleep EEG spindles. In *Engineering*

in *Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pages 2464–2467. IEEE.

- [Lainscsek, 2011] Lainscsek, C. (2011). Nonuniqueness of global modeling and time scaling. *Phys. Rev. E*, 84:046205.
- [Lainscsek, 2012] Lainscsek, C. (2012). A class of Lorenz-like systems. *Chaos*, 22:013126.
- [Lainscsek et al., 2013a] Lainscsek, C., Hernandez, M. E., Weyhenmeyer, J., Sejnowski, T. J., and Poizner, H. (2013a). Non-linear dynamical analysis of EEG time series distinguishes patients with Parkinson’s disease from healthy individuals. *Frontiers in neurology*, 4.
- [Lainscsek et al., 2013b] Lainscsek, C., Hernandez, M. E., Weyhenmeyer, J., Sejnowski, T. J., and Poizner, H. (2013b). Non-linear dynamical analysis of EEG time series distinguishes patients with Parkinson’s disease from healthy individuals. *Frontiers in Neurology*, 4(200).
- [Lainscsek et al., 2014] Lainscsek, C., Messenger, V., Portman, A., Muir, J.-F., Sejnowski, T. J., and Letellier, C. (2014). Automatic sleep scoring from a single electrode using delay differential equations. In *Applied Non-Linear Dynamical Systems*, pages 371–382. Springer.
- [Lainscsek et al., 2012] Lainscsek, C., Rowat, P., Schettino, L., Lee, D., Song, D., Letellier, C., and Poizner, H. (2012). Finger tapping movements of Parkinson’s disease patients automatically rated using nonlinear delay differential equations. *Chaos*, 22:013119.
- [Lainscsek et al., 2019] Lainscsek, C., Sampson, A. L., Kim, R., Thomas, M. L., Man, K., Lainscsek, X., Swerdlow, N. R., Braff, D. L., Sejnowski, T. J., Light, G. A., et al. (2019). Nonlinear dynamics underlying sensory processing dysfunction in schizophrenia. *Proceedings of the National Academy of Sciences*, 116(9):3847–3852.
- [Lainscsek et al., 2009] Lainscsek, C., Schettino, L., Rowat, P., van Erp, E., Song, D., and Poizner, H. (2009). Nonlinear DDE analysis of repetitive hand movements in Parkinson’s disease. In In, V., Longhini, P., and Palacios, A., editors, *Applications of Nonlinear Dynamics, Understanding Complex Systems*, pages 421–427. Springer.
- [Lainscsek and Sejnowski, 2013a] Lainscsek, C. and Sejnowski, T. (2013a). Delay differential equation models of electrocardiograms. In *Proceedings of the International Conference on Theory and Applications in Nonlinear Dynamics; Seattle, 2012*.
- [Lainscsek and Sejnowski, 2013b] Lainscsek, C. and Sejnowski, T. (2013b). Electrocardiogram classification using delay differential equations. *Chaos*, 23(2):023132.
- [Lainscsek and Sejnowski, 2013c] Lainscsek, C. and Sejnowski, T. J. (2013c). Electrocardiogram classification using delay differential equations. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 23(2):023132.
- [Lainscsek and Sejnowski, 2015] Lainscsek, C. and Sejnowski, T. J. (2015). Delay differential analysis of time series. *Neural computation*.

- [Lainscsek et al., 2017] Lainscsek, C., Weyhenmeyer, J., Cash, S. S., and Sejnowski, T. J. (2017). Delay differential analysis of seizures in multichannel electrocorticography data. *Neural computation*, 29(12):3181–3218.
- [Lainscsek et al., 2013c] Lainscsek, C., Weyhenmeyer, J., Hernandez, M., Poizner, H., and Sejnowski, T. (2013c). Non-linear dynamical classification of short time series of the Rössler system in high noise regimes. *Frontiers in Neurology*, 4(182).
- [Lainscsek et al., 2013d] Lainscsek, C., Weyhenmeyer, J., Hernandez, M. E., Poizner, H., and Sejnowski, T. J. (2013d). Non-linear dynamical classification of short time series of the Rössler system in high noise regimes. *Frontiers in neurology*, 4:182.
- [Leresche et al., 1991] Leresche, N., Lightowler, S., Soltesz, I., Jassik-Gerschenfeld, D., and Crunelli, V. (1991). Low-frequency oscillatory activities intrinsic to rat and cat thalamocortical cells. *The Journal of Physiology*, 441(1):155–174.
- [Li et al., 2013] Li, S., Zhou, W., Yuan, Q., and Liu, Y. (2013). Seizure prediction using spike rate of intracranial eeg. *IEEE transactions on neural systems and rehabilitation engineering*, 21(6):880–886.
- [Light et al., 2015] Light, G. A., Swerdlow, N. R., Thomas, M. L., Calkins, M. E., Green, M. F., Greenwood, T. A., Gur, R. E., Gur, R. C., Lazzeroni, L. C., Nuechterlein, K. H., et al. (2015). Validation of mismatch negativity and P3a for use in multi-site studies of schizophrenia: characterization of demographic, clinical, cognitive, and functional correlates in COGS-2. *Schizophrenia research*, 163(1):63–72.
- [Light et al., 2017] Light, G. A., Zhang, W., Joshi, Y. B., Bhakta, S., Talledo, J. A., and Swerdlow, N. R. (2017). Single-dose memantine improves cortical oscillatory response dynamics in patients with schizophrenia. *Neuropsychopharmacology*, 42(13):2633.
- [Lorenz, 1963] Lorenz, E. N. (1963). Deterministic nonperiodic flow. *Journal of the atmospheric sciences*, 20(2):130–141.
- [Lorenz, 1984] Lorenz, E. N. (1984). The local structure of a chaotic attractor in four dimensions. *Physica D: Nonlinear Phenomena*, 13(1-2):90–104.
- [Lorenz, 1996] Lorenz, E. N. (1996). Predictability: A problem partly solved. In *Proc. Seminar on predictability*, volume 1.
- [Lorenz and Emanuel, 1998] Lorenz, E. N. and Emanuel, K. A. (1998). Optimal sites for supplementary weather observations: Simulation with a small model. *Journal of the Atmospheric Sciences*, 55(3):399–414.
- [Maiwald et al., 2004] Maiwald, T., Winterhalder, M., Aschenbrenner-Scheibe, R., Voss, H. U., Schulze-Bonhage, A., and Timmer, J. (2004). Comparison of three nonlinear seizure prediction methods by means of the seizure prediction characteristic. *Physica D: nonlinear phenomena*, 194(3-4):357–368.

- [Martin et al., 2013] Martin, N., Lafortune, M., Godbout, J., Barakat, M., Robillard, R., Poirier, G., Bastien, C., and Carrier, J. (2013). Topography of age-related changes in sleep spindles. *Neurobiology of aging*, 34(2):468–476.
- [Miletics and Molnárka, 2005] Miletics, E. and Molnárka, G. (2005). Implicit extension of Taylor series method with numerical derivatives for initial value problems. *Computers & Mathematics with Applications*, 50(7):1167–1177.
- [Mölle et al., 2002] Mölle, M., Marshall, L., Gais, S., and Born, J. (2002). Grouping of spindle activity during slow oscillations in human non-rapid eye movement sleep. *The Journal of neuroscience*, 22(24):10941–10947.
- [Moran and Hong, 2011] Moran, L. V. and Hong, L. E. (2011). High vs low frequency neural oscillations in schizophrenia. *Schizophrenia Bulletin*, 37(4):659–663.
- [Murray et al., 2014] Murray, J. D., Anticevic, A., Gancsos, M., Ichinose, M., Corlett, P. R., Krystal, J. H., and Wang, X.-J. (2014). Linking microcircuit dysfunction to cognitive impairment: effects of disinhibition associated with schizophrenia in a cortical working memory model. *Cerebral cortex*, 24(4):859–872.
- [O’Reilly and Nielsen, 2015] O’Reilly, C. and Nielsen, T. (2015). Automatic sleep spindle detection: benchmarking with fine temporal resolution using open science tools. *Frontiers in human neuroscience*, 9:353.
- [Oribe et al., 2015] Oribe, N., Hirano, Y., Kanba, S., del Re, E., Seidman, L., Meshulam-Gately, R., Goldstein, J. M., Shenton, M., Spencer, K. M., McCarley, R. W., and Niznikiewicz, M. (2015). Progressive reduction of visual P300 amplitude in patients with first-episode schizophrenia: An ERP study. *Schizophrenia Bulletin*, 41(2):460–470.
- [Packard et al., 1980] Packard, N. H., Crutchfield, J. P., Farmer, J. D., and Shaw, R. S. (1980). Geometry from a time series. *Phys. Rev. Lett.*, 45:712.
- [Parvizi and Kastner, 2018] Parvizi, J. and Kastner, S. (2018). Promises and limitations of human intracranial electroencephalography. *Nature neuroscience*, pages 474–483.
- [Paulus and Braff, 2003] Paulus, M. P. and Braff, D. L. (2003). Chaos and schizophrenia: does the method fit the madness? *Biological Psychiatry*, 53(1):3–11.
- [Perez et al., 2017] Perez, V. B., Tarasenko, M., Miyakoshi, M., Pianka, S. T., Makeig, S. D., Braff, D. L., Swerdlow, N. R., and Light, G. A. (2017). Mismatch negativity is a sensitive and predictive biomarker of perceptual learning during auditory cognitive training in schizophrenia. *Neuropsychopharmacology*, 42(11):2206.
- [Petit et al., 2004] Petit, D., Gagnon, J.-F., Fantini, M. L., Ferini-Strambi, L., and Montplaisir, J. (2004). Sleep and quantitative EEG in neurodegenerative disorders. *Journal of psychosomatic research*, 56(5):487–496.

- [Press et al., 1990] Press, W., Flannery, B., Teukolsky, S., and Vetterling, W. (1990). *Numerical Recipes in C*. Cambridge University Press, New York, NY, USA.
- [Rabinovich et al., 2006] Rabinovich, M. I., Varona, P., Selverston, A. I., and Abarbanel, H. D. (2006). Dynamical principles in neuroscience. *Reviews of modern physics*, 78(4):1213.
- [Reiterer et al., 1998] Reiterer, P., Lainscsek, C., Schürer, F., Letellier, C., and Maquet, J. (1998). A nine-dimensional Lorenz system to study high-dimensional chaos. *Journal of Physics A*, 31:7121–7139.
- [Risling et al., 2012] Risling, A. J., Braff, D. L., Swerdlow, N. R., Helleman, G., Rassovsky, Y., Sprock, J., Pela, M., and Light, G. A. (2012). Disentangling early sensory information processing deficits in schizophrenia. *Clinical Neurophysiology*, 123(10):1942 – 1949.
- [Risling et al., 2014] Risling, A. J., Miyakoshi, M., Sugar, C. A., Braff, D. L., Makeig, S., and Light, G. A. (2014). Cortical substrates and functional correlates of auditory deviance processing deficits in schizophrenia. *NeuroImage: Clinical*, 6:424 – 437.
- [Roach and Mathalon, 2008] Roach, B. J. and Mathalon, D. H. (2008). Event-related EEG time-frequency analysis: An overview of measures and an analysis of early gamma band phase locking in schizophrenia. *Schizophrenia Bulletin*, 34(5):907–926.
- [Rothschild et al., 2017] Rothschild, G., Eban, E., and Frank, L. M. (2017). A cortical–hippocampal–cortical loop of information processing during memory consolidation. *Nature neuroscience*, 20(2):251–259.
- [Runge, 2018] Runge, J. (2018). Causal network reconstruction from time series: From theoretical assumptions to practical estimation. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 28(7):075310.
- [Sampson et al., 2019] Sampson, A. L., Lainscsek, C., Gonzalez, C. E., Ulbert, I., Devinsky, O., Fabó, D., Madsen, J. R., Halgren, E., Cash, S. S., and Sejnowski, T. J. (2019). Delay differential analysis for dynamical sleep spindle detection. *Journal of neuroscience methods*, 316:12–21.
- [Sampson et al., 2015] Sampson, A. L., Lainscsek, C., Lin, T. W., Halgren, E., Cash, S. S., and Sejnowski, T. J. (2015). Nonlinear dynamical sleep spindle detection using delay differential analysis. presented at Society for Neuroscience Conference.
- [Sauer et al., 1991] Sauer, T., Yorke, J. A., and Casdagli, M. (1991). Embedology. *Journal of Statistical Physics*, 65:579.
- [Schabus et al., 2004] Schabus, M., Gruber, G., Parapatics, S., Sauter, C., Klosch, G., Anderer, P., Klimesch, W., Saletu, B., and Zeitlhofer, J. (2004). Sleep spindles and their significance for declarative memory consolidation. *Sleep*, 27(8):1479–1485.

- [Sejnowski and Destexhe, 2000] Sejnowski, T. J. and Destexhe, A. (2000). Why do we sleep? *Brain research*, 886(1):208–223.
- [Sirota et al., 2003] Sirota, A., Csicsvari, J., Buhl, D., and Buzsáki, G. (2003). Communication between neocortex and hippocampus during sleep in rodents. *Proceedings of the National Academy of Sciences*, 100(4):2065–2069.
- [Sørensen, 1948] Sørensen, T. (1948). A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons. *Biol. Skr.*, 5:1–34.
- [Stephan et al., 2006] Stephan, K. E., Baldeweg, T., and Friston, K. J. (2006). Synaptic plasticity and dysconnection in schizophrenia. *Biological psychiatry*, 59(10):929–939.
- [Sugihara et al., 2012] Sugihara, G., May, R., Ye, H., Hsieh, C., Deyle, E., Fogarty, M., and Munch, S. (2012). Detecting causality in complex ecosystems. *Science*, 338(6106):496–500.
- [Swerdlow et al., 2016] Swerdlow, N. R., Bhakta, S., Chou, H., Talledo, J. A., Balvaneda, B., and Light, G. A. (2016). Memantine effects on sensorimotor gating and mismatch negativity in patients with chronic psychosis. *Neuropsychopharmacology*, 41(2):419.
- [Swerdlow et al., 2015] Swerdlow, N. R., Gur, R. E., and Braff, D. L. (2015). Consortium on the Genetics of Schizophrenia (COGS) assessment of endophenotypes for schizophrenia: An introduction to this special issue of schizophrenia research. *Schizophrenia Research*, 163(1):9 – 16.
- [Takens, 1981] Takens, F. (1981). Detecting strange attractors in turbulence. In Rand, D. A. and Young, L.-S., editors, *Dynamical Systems and Turbulence, Warwick 1980*, volume 898 of *Lecture Notes in Mathematics*, pages 366–381. Springer Berlin/Heidelberg.
- [Takens, 1993] Takens, F. (1993). Detecting nonlinearities in stationary time series. *Int. J. Bifurcation and Chaos*, 3:241.
- [Terry and Breakspear, 2003] Terry, J. R. and Breakspear, M. (2003). An improved algorithm for the detection of dynamical interdependence in bivariate time-series. *Biological cybernetics*, 88(2):129–136.
- [Thomas et al., 2017] Thomas, M. L., Green, M. F., Helleman, G., Sugar, C. A., Tarasenko, M., Calkins, M. E., Greenwood, T. A., Gur, R. E., Gur, R. C., Lazzaroni, L. C., et al. (2017). Modeling deficits from early auditory information processing to psychosocial functioning in schizophrenia. *JAMA psychiatry*, 74(1):37–46.
- [Todd et al., 2012] Todd, J., Michie, P. T., Schall, U., Ward, P. B., and Catts, S. V. (2012). Mismatch negativity (MMN) reduction in schizophrenia—impaired prediction-error generation, estimation or salience? *International Journal of Psychophysiology*, pages 222–231.

- [Truong et al., 2018] Truong, N. D., Nguyen, A. D., Kuhlmann, L., Bonyadi, M. R., Yang, J., Ippolito, S., and Kavehei, O. (2018). Convolutional neural networks for seizure prediction using intracranial and scalp electroencephalogram. *Neural Networks*, 105:104–111.
- [Turetsky et al., 2015a] Turetsky, B. I., Dress, E. M., Braff, D. L., Calkins, M. E., Green, M. F., Greenwood, T. A., Gur, R. E., Gur, R. C., Lazzeroni, L. C., Nuechterlein, K. H., et al. (2015a). The utility of P300 as a schizophrenia endophenotype and predictive biomarker: clinical and socio-demographic modulators in COGS-2. *Schizophrenia Research*, 163(1):53–62.
- [Turetsky et al., 2015b] Turetsky, B. I., Dress, E. M., Braff, D. L., Calkins, M. E., Green, M. F., Greenwood, T. A., Gur, R. E., Gur, R. C., Lazzeroni, L. C., Nuechterlein, K. H., Radant, A. D., Seidman, L. J., Siever, L. J., Silverman, J. M., Sprock, J., Stone, W. S., Sugar, C. A., Swerdlow, N. R., Tsuang, D. W., Tsuang, M. T., and Light, G. (2015b). The utility of P300 as a schizophrenia endophenotype and predictive biomarker: Clinical and socio-demographic modulators in COGS-2. *Schizophrenia Research*, 163(1):53 – 62.
- [van Beelen, 2019] van Beelen, T. (2019). EDF/BDF to ASCII converter. <https://www.teuniz.net/edf2ascii/>.
- [Vaz et al., 2019] Vaz, A. P., Inati, S. K., Brunel, N., and Zaghloul, K. A. (2019). Coupled ripple oscillations between the medial temporal lobe and neocortex retrieve human memory. *Science*, 363(6430):975–978.
- [Volterra, 1887] Volterra, V. (1887). Sopra le funzioni che dipendono da altre funzioni. *Atti della Reale Accademia dei Lincei*, 3:97–105.
- [Volterra, 1959] Volterra, V. (1959). *Theory of Functionals of Integral and Integro-Differential Equations*. Dover Publ.
- [Warby et al., 2014] Warby, S. C., Wendt, S. L., Welinder, P., Munk, E. G., Carrillo, O., Sorensen, H. B., Jennum, P., Peppard, P. E., Perona, P., and Mignot, E. (2014). Sleep-spindle detection: crowdsourcing and evaluating performance of experts, non-experts and automated methods. *Nature methods*, 11(4):385–392.
- [Wendt et al., 2012] Wendt, S. L., Christensen, J. A., Kempfner, J., Leonthin, H. L., Jennum, P., and Sorensen, H. B. (2012). Validation of a novel automatic sleep spindle detector with high performance during sleep in middle aged subjects. In *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 4250–4253. IEEE.
- [Whitney, 1936] Whitney (1936). Differentiable manifolds. *Ann. Math.*, 37:645–680.
- [WHO, 2019] WHO (2019). Epilepsy. <https://www.who.int/en/news-room/fact-sheets/detail/epilepsy>. Accessed = 2019-4-3.
- [Wynn et al., 2010] Wynn, J. K., Sugar, C., Horan, W. P., Kern, R., and Green, M. F. (2010). Mismatch negativity, social cognition, and functioning in schizophrenia patients. *Biological Psychiatry*, 67(10):940 – 947.

- [Yule, 1926] Yule, G. U. (1926). Why do we sometimes get nonsense-correlations between time-series?—a study in sampling and the nature of time-series. *Journal of the royal statistical society*, 89(1):1–63.
- [Zheng et al., 2014] Zheng, Y., Wang, G., Li, K., Bao, G., and Wang, J. (2014). Epileptic seizure prediction using phase synchronization based on bivariate empirical mode decomposition. *Clinical Neurophysiology*, 125(6):1104–1111.