2015

Characterizing dynamically evolving functional networks in humans with application to speech

BOSTON UNIVERSITY

SCHOOL OF MEDICINE

Dissertation

CHARACTERIZING DYNAMICALLY EVOLVING FUNCTIONAL NETWORKS IN HUMANS WITH APPLICATION TO SPEECH

by

EMILY PATRICIA STEPHEN

Sc.B., Brown University, 2007

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

2015

Approved by

First Reader _____

Frank Guenther, Ph.D.
Professor, Department of Speech, Language and Hearing Sciences

Second Reader _____

Mark Kramer, Ph.D.
Associate Professor, Mathematical Neuroscience

Third Reader _____

Jason Ritt, Ph.D.
Assistant Professor, Biomedical Engineering

ACKNOWLEDGMENTS

The work that I describe in this dissertation was deeply collaborative, and I am sure I will only manage to name a subset of those who have played a role in its construction. Here I will do my best to highlight the major contributors as well as a few who have made a particular impact on my experience.

I have been extremely lucky to have several amazing mentors to help guide me through the process of building a project. My advisor Frank Guenther has always looked out for my best interests and reminded me to be practical. Mark Kramer, while not my formal advisor, has worked as closely with me as if he were. He always made time for my project, providing thoughtful advice and detailed feedback. Uri Eden, too, has been another close collaborator and mentor, teaching me how to think statistically and approach complex problems.

I also count among my mentors Kyle Lepage, who spent ridiculous amounts of time answering all of my questions about multitaper analysis, canonical coherence, and countless other topics.

Other collaborators include Jon Brumberg, who helped work out the scope and direction of the project early on, and Jason Tourville, who has been a great resource for speech theory and anatomy. The Wadsworth group, led by Gerwin Schalk and including Peter Brunner and Aysegul Gunduz, supplied the speech ECoG data and were generous with advice and very responsive to questions.

More thanks go to:

- Jason Ritt
- Alfonso Nieto-Castañón
- Ayoub Daliri
- Other members of my lab: Spencer Torene, Andrés Salazar-Gómez, Elisa Golfinopoulos, Nan Jia, Scott Kuzdeba, Dante Smith-Grasso
- The Cognitive Rhythms Collaborative and folks in the BU math department, especially Nancy Kopell, Xinyi Deng, Mikio Aoi
- Eric Kolaczyk and his group (especially Cedric Ginestet and Heather Shappell)
- Shelley Russek and Sandi Grasso
- My funding: NIH grants F31 DC011663 and R01 DC002852
- GPN, Compnet, NGSO, CNSO
- My wonderful family

Finally, I can't forget two people: Eric Denovellis and Caroline Moore-Kochlacs, who were always willing to read a first draft and offer moral support. You keep me sane!

CHARACTERIZING DYNAMICALLY EVOLVING

FUNCTIONAL NETWORKS IN HUMANS WITH APPLICATION TO SPEECH

EMILY PATRICIA STEPHEN

Boston University School of Medicine, 2015

Major Professor: Frank Guenther, Ph.D., Professor, Department of Speech, Language and Hearing Sciences

ABSTRACT

Understanding how communication between brain areas evolves to support dynamic function remains a fundamental challenge in neuroscience. One approach to this question is functional connectivity analysis, in which statistical coupling measures are employed to detect signatures of interactions between brain regions. Because the brain uses multiple communication mechanisms at different temporal and spatial scales, and because the neuronal signatures of communication are often weak, powerful connectivity inference methodologies require continued development specific to these challenges.

Here we address the challenge of inferring task-related functional connectivity in brain voltage recordings. We first develop a framework for detecting changes in statistical coupling that occur reliably in a task relative to a baseline period. The framework characterizes the dynamics of connectivity changes, allows inference on multiple spatial scales, and assesses statistical uncertainty. This

general framework is modular and applicable to a wide range of tasks and research questions.

We demonstrate the flexibility of the framework in the second part of this thesis, in which we refine the coupling statistics and hypothesis tests to improve statistical power and test different proposed connectivity mechanisms. In particular, we introduce frequency domain coupling measures and define test statistics that exploit theoretical properties and capture known sampling variability. The resulting test statistics use correlation, coherence, canonical correlation, and canonical coherence to infer task-related changes in coupling. Because canonical correlation and canonical coherence are not commonly used in functional connectivity analyses, we derive the theoretical values and statistical estimators for these measures.

In the third part of this thesis, we present a sample application of these techniques to electrocorticography data collected during an overt reading task. We discuss the challenges that arise with task-related human data, which is often noisy and underpowered, and present functional connectivity results in the context of traditional and contemporary within-electrode analytics. In two of nine subjects we observe time-domain and frequency-domain network changes that accord with theoretical models of information routing during motor processing.

Taken together, this work contributes a methodological framework for inferring task-related functional connectivity across spatial and temporal scales, and supports insight into the rapid, dynamic functional coupling of human speech.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ADHD ..... Attention deficit/hyperactivity disorder

BCa ..... Bias-corrected and accelerated (confidence intervals)

BOLD ..... Blood-oxygenation level dependent

CDF ..... Cumulative distribution function

DIVA ..... Directions into velocities of articulators (model)

ECoG ..... Electrocorticography

EEG ..... Electroencephalography

EMG ..... Electromyography

ERP ..... Event-related potential

FDR ..... False discovery rate

fMRI ..... Functional magnetic resonance imaging

LFP ..... Local field potential

M1 ..... Primary motor cortex

MEG ..... Magnetoencephalography

PCA ..... Principal component analysis

PET ..... Positron emission tomography

ROI ..... Region of interest

S1 ..... Primary somatosensory cortex

(a)SMG ..... (anterior) Supramarginal gyrus

SNR ..... Signal to noise ratio

(p)STG ..... (posterior) Superior temporal gyrus

CHAPTER I. INTRODUCTION

Some of the most powerful theories and discoveries in neuroscience have come from a model of brain activity in which different brain areas perform distinct, separable functions. This model assumes "functional segregation" between brain areas controlling sensory, motor, and cognitive sub-tasks, and explains complex behavior as arising from combinations of the sub-tasks. Assuming brain areas are functionally separate, however, leaves open the question of "functional integration", or how cooperation between areas occurs. Functional integration considers the ways that different brain areas can connect and communicate to generate complex processing. Describing functional integration in the brain has become an active area of research, both in terms of uncovering the mechanisms of functional integration and in terms of describing their properties during brain function in health and disease (Sporns et al., 2004; Fingelkurts et al., 2005; Friston, 2011).

This dissertation explores a set of problems that arise in inferring the dynamics of functional integration in the brain. We develop statistical methods to infer a type of functional integration over time during a task, and we apply those methods to an example electrocorticography (ECoG) dataset collected during an overt reading task. In this chapter, we review relevant literature and then summarize the contents of the dissertation.

1. Functional Connectivity

1.1 Approach

In the effort to describe functional integration between brain areas, researchers have taken two types of approaches. The first, effective connectivity, tries to detect

evidence of causal influences between brain areas. The second, functional connectivity, tries to quantify features of the signal that are signs of communication. The two approaches overlap, but in general effective connectivity makes stronger assumptions and involves using the data to fit generative models of the mechanisms of coupling, while functional connectivity makes weaker assumptions and reports descriptive statistics of the data that are thought to reflect coupling (Friston, 2011).

Here we focus on functional connectivity, which is especially useful in the absence of strong prior knowledge about the nature of the underlying coupling. Like other descriptive approaches, functional connectivity does not encompass an interpretation of the mechanisms generating the data. Instead, it uses statistics that have a direct, understandable relationship to the underlying data, such as means, correlations, and coherences, which can later be interpreted by the researcher. It therefore can be useful in the early stages of analysis before a theory of the generative processes underlying the data is available. This approach is standard for first-order descriptive statistics such as event related potentials and spectrograms, which are based on signal averages. Functional connectivity adds a set of second-order statistics to describe relationships between signals.

1.2 Resting state functional connectivity

One common example of functional connectivity analysis is resting state functional connectivity, which looks at correlations between activations in functional magnetic resonance imaging (fMRI) data when a subject is not performing any task. Using techniques that are sensitive to correlations between signals like principal

components analysis (PCA) and clustering algorithms, resting state functional connectivity groups brain areas based on patterns of coactivation, leading to "resting state networks" (reviews: Fox and Raichle, 2007; Van Den Heuvel and Pol, 2010). These networks are consistent across subjects (Yeo et al., 2011) and have been shown to be stable within subjects over time periods on the order of months (Shehzad et al., 2009). For example, even in the absence of a task, there is correlated activity between brain areas in the sensorimotor network (Biswal et al., 1995). Other resting state networks seem to be related to distributed control functions such as the ventral and dorsal attention networks (Fox et al., 2006). One network that was identified early on, the default mode network, is most active in the absence of a cognitive task (Raichle et al., 2001). The finding that resting state networks are robust across subjects and time has led to a surge in functional connectivity research into the causes of the correlated activity and how the networks change based on disease and other factors.

Importantly, Cordes et al. (2001) showed that resting state correlations are dominated by activity below 0.1 Hz, and most current studies therefore low-pass filter the signals below 0.08 or 0.1 Hz before performing a resting state analysis (Fox and Raichle, 2007). Hence, while resting state networks have been shown in some cases to persist during the performance of a task (Greicius and Menon, 2004; Esposito et al., 2006; Fransson, 2006; Calhoun et al., 2008; Smith et al., 2009), there are likely to be finer timescale changes in coupling that are more closely tied to task dynamics. Inferring task-related and dynamic functional connectivity can benefit from recording modalities with finer time resolution than fMRI, such as electroencephalography (EEG),

magnetoencephalography (MEG), and ECoG. The analyses developed in this study were designed for these finer-timescale signals, particularly ECoG (which will be described in more detail in Section 2).

1.3 Communication through coherence

Finer timescale network interactions can be inferred using correlation-based metrics just like resting state analysis, but there may be other signatures of communication between brain regions that become detectable with better time resolution. In particular, coherence in narrow frequency bands (that are too high frequency to be detected in fMRI data) has been proposed as a way that different brain regions can modulate their sensitivity to each other. The "communication through coherence" hypothesis (Fries, 2005; Fries et al., 2007; Fries, 2009) posits that information processing, and hence functional connectivity, between areas is gated by the coherence of rhythmic activity in those areas, particularly in the gamma frequency band (40-80 Hz). That is, when the population activity in two areas becomes synchronous (subject to a phase lag) in a given frequency band, it enables neurons to affect each other in one or both directions. Conversely, if the population activity is not coherent between the areas, action potentials from the sending area may reach the receiving area at a time when the receiving neurons are relatively insensitive to synaptic input (Fries, 2005; Börgers and Kopell, 2008). In this way, the sensitivity of one area to inputs from another can change according to mental state or task demands. A common way to look for communication through coherence is to place electrodes in two areas of the brain and measure the coherence of their dominant rhythms during a task (Buschman and Miller, 2007;

Womelsdorf et al., 2007). Ideally, functional connectivity using ECoG could be an extension of this approach, looking at coherence on all possible connections across a network of implanted electrodes (Canolty et al., 2010; Miller et al., 2012).

1.4 Graph theory and network inference

Before moving on to the details of ECoG in Section 2, it will be important to present some basic background in functional network inference. First, we introduce graph theory terminology and concepts. Then, we introduce the statistical problem of network inference and discuss the problem of dimensionality (for a thorough discussion of statistical network inference, see Kolaczyk (2009)).

1.4.1 Relevant concepts in graph theory. Functional networks are typically represented as "graphs", which are mathematical structures consisting of a set of elements, "nodes" or "vertices", and a set of connections between the nodes, "edges" or "links". Edges connect two nodes (all of the networks we consider here allow only pairwise connectivity), and the nodes connected by an edge are said to be "incident" to the edge. A "weighted" graph has a continuous value associated with each edge, while a "binary" graph stores only whether or not the edge exists. Here we consider only "undirected" graphs, where the ordering of the nodes does not matter for the edge definition.

A graph can be represented as a $N \times N$ matrix, where N is the number of nodes in the graph, and the $(i, j)^{th}$ element contains the edge weight between node i and node j. Edges that do not exist have a value of 0, and edges in a binary graph are given weights with a value of 1. Since we do not consider self-edges, the main diagonal will

contain all zeros. Also, since the edges are symmetric in undirected graphs, the matrix will be symmetric along the main diagonal.

1.4.2 Network inference. Network inference is the statistical process of inferring a graph from data. The statistical inference process defines a statistical model with a set of parameters related to the structure of the graph, makes some assumptions about the distributions of the parameters, fits them using the data, and makes statements about their uncertainty. For example, a typical functional connectivity analysis will define the nodes to be electrodes or voxels and the parameters to be the correlations between the neural signals on the nodes (for weighted networks) or a connected versus unconnected label (for binary networks). In both weighted and binary networks, there is assumed to exist a true underlying network with population values for the edge weights or edge/non-edge labels, and the statistical goal is to produce an estimate of this true network. The edges are typically assumed to be undirected (directed networks are usually relegated to effective connectivity analysis). Hence a full network model will involve at least $N(N-1)/2$ parameters, where N is the number of nodes: this is the total number of possible edges in the network.

Since the number of parameters scales with N^2, network inference can become extremely high-dimensional as the number of nodes increases: the number of parameters increases much faster than the amount of data, leading to the so-called “Curse of Dimensionality”. Dealing with this high dimensionality is a core problem in network inference, because statistical models with many parameters or degrees of freedom also

have high variance, which leads to high variance in the fitted result (i.e. the inferred network). Hence it is important to control the degrees of freedom in the statistical model in order to manage the performance of the inference. That is, it is important to constrain the model whenever possible, for example by incorporating prior knowledge, so that the inferred networks have a small enough variance to be meaningful.

This is related to statistical model selection and the "bias-variance tradeoff". The expected prediction error of a test statistic is the sum of (1) the variance of the measurement noise, (2) the squared bias of the statistic, which is the expected difference between the true value and the estimated value, and (3) the variance of the statistic, which relates to how sensitive it is to fluctuations in the data (Hastie et al., 2009). Since we often cannot control the variance of the noise in the data itself, when designing a statistical model of a system we aim to minimize the bias and the variance of the test statistic. In practice, for a given sample size, a statistic with a low bias will tend to have high variance and a similar statistic with low variance will tend to have higher bias; hence the "bias-variance tradeoff". A well-designed model having both low bias and low variance will have low mean squared error, meaning that the inferred parameter values are expected to be close to the true values. Including prior knowledge about the system tends to reduce variance, although it could bias the results if the prior knowledge is imperfect. Hence, in order to best use the given data, we try to select a statistical framework that has relatively few degrees of freedom (reducing variance) while avoiding imposing structure in the model that is not present in the system (reducing bias). Some popular ways to reduce the degrees of freedom in functional connectivity analysis include

grouping sensors by region of interest, focusing only on edges incident to a pre-selected "seed region", and eliminating nodes that are not expected to be involved in the network.

Another common way to reduce the degrees of freedom in functional connectivity analysis is to infer binary networks rather than weighted networks. The choice to infer binary networks reflects the underlying assumption that only a portion of the possible connections in the network is actually present at any given time. More strictly, binary networks assume that there is a classifiable distinction between edges that exist and edges that do not exist. While this is consistent with the communication through coherence hypothesis in which coherence is enforced mechanistically between specific communicating areas (and not others) during a task (Fries, 2005; Fries et al., 2007; Fries, 2009), there are other possible models of connectivity that may represent coupling on a continuum. For example, if the mechanism of functional connectivity acts at the neuron level, there may appear to be a continuum of coupling strengths at the population level based on the proportion of neurons that are coupled during the task. Thus, it may be that all brain areas are coupled to all other areas with differing strengths that transition smoothly between zero and larger values. Even if true functional coupling does occur between only a subset of possible pairs of areas, the data or the coupling statistic may be insufficient to distinguish these connections from noise. Imposing a binary characterization of networks when the underlying data do not support it can throw away information and obscure the patterns that do exist in the data (weighted networks can have problems, too; see Ginestet et al. (2011)), introducing bias. If the data support a

binary classification for a given coupling statistic, the test statistics for all edges will be a mixture of distributions that can be uncoupled empirically (Efron, 2010).

2. Electrocorticography

ECoG is an invasive procedure that is used clinically primarily in patients with intractable epilepsy. In these patients, electrodes are implanted above the cortical surface subdurally, and the recordings are monitored over the course of a few days or weeks to identify (1) the location of the seizure focus and (2) the spread of eloquent cortex, or cortex that is critical for sensory, motor, and language functioning. Electrodes are arranged in grids or strips, with electrodes typically spaced 1 cm or 4 mm apart, or closer for "micro-ECoG" recordings. After the seizure focus has been identified, the electrodes are removed and the seizure focus is resected surgically.

During the monitoring period, patients often agree to participate in research studies, providing an opportunity to study high-quality invasive recordings in human subjects. These studies sometimes relate to epilepsy, but they also often study more general topics under the assumption that epileptic brains behave like healthy brains under certain conditions. In particular, studies using data from epilepsy patients typically exclude from analysis time periods with seizure activity, as well as any electrodes over the epileptogenic zone or electrodes that have interictal epileptic activity. For the most part, results from ECoG studies on epileptic patients have been consistent with results in healthy subjects (often from other recording modalities). In one notable exception, Engel et al. (1982) performed a study of metabolism using positron computed tomography, and found that areas in the temporal lobe showed hypometabolism even in the absence of

interictal spikes in EEG. It is unclear how this hypometabolism may relate to neural activity, although Bettus et al. (2011) found evidence in both fMRI and ECoG that outside of seizures, epileptic cortex may have an influence on areas otherwise free of epileptic activity.

Because the electrodes do not penetrate the cortical surface, it is not possible to resolve individual action potentials from ECoG recordings. Instead, the recordings reflect local field potentials (LFPs) from the circle of cortex underlying the electrode (typically ~2.3 mm diameter) consisting of about a half a million neurons (Ojemann et al., 2013) and a billion synapses (Ritaccio et al., 2011). LFPs represent a spatial average of all of the ionic processes (action potentials, synaptic potentials, intrinsic currents, glial fluctuations, etc.) in the vicinity of the electrode. In practice, postsynaptic currents are thought to be the biggest contributor to the LFP because of their slow dynamics relative to action potentials, which allows them to overlap in time and hence summate spatially (Buzsaki et al., 2012). The spatial reach of an electrical source in cortex due to volume conduction depends on a number of factors such as neuron morphology, the distribution of synapses, and the amount of correlation in the synaptic activity (Lindén et al., 2011). In some situations electrical activity can spread over centimeters: for example, brain stem auditory evoked potentials can be observed using EEG electrodes on the scalp as much as 10 cm away (Jewett and Williston, 1971). However, invasive recording modalities such as ECoG are proposed to be relatively less sensitive to distant electrical sources in comparison to nearby sources (Zaveri et al., 2009).

LFPs recorded with ECoG have many of the same features observed in LFPs from intracortical electrode recordings, such task- and region-specific narrow-band rhythms in the delta (<4 Hz), theta (4-8 Hz), alpha (~10 Hz), beta (15-30 Hz), and gamma (30-80 Hz) frequency bands (Wang, 2010). In addition, there has been a recent focus on broadband power in the high gamma range (>100 Hz) (Crone et al., 2006). Experimental and theoretical work supports the idea that this broadband high gamma power does not arise from rhythms but from asynchronous population spiking and synaptic activity (Manning et al., 2009; Ray and Maunsell, 2011; Buzsaki et al., 2012; Miller et al., 2014). Modulation in the high gamma power is hence commonly interpreted as an indicator of changes in neuronal population activity under the electrode.

Broadband high gamma power is also related to the blood-oxygen-level-dependent (BOLD) response recorded by fMRI. A review of studies linking fMRI and LFP/ECoG recordings found broadband (30-250 Hz) power to be the electrophysiological feature with the strongest coupling to the BOLD signal (Ojemann et al., 2013). Furthermore, several studies have found that during the resting state, both the LFPs themselves and the envelope of high gamma instantaneous power are coherent at very low frequencies (<0.1 Hz) within the same networks that are observed in resting state fMRI analyses (He et al., 2008; Ko et al., 2011; Breshears et al., 2012; Hiltunen et al., 2014).

As discussed above, ECoG also holds promise for studying the dynamics of functional connectivity at higher temporal scales, with the goal of detecting connectivity at the pace of functional processing. For example, dynamic functional networks based on

ECoG data in epilepsy patients have been described over time during seizures (Kramer et al., 2010) and outside of seizures over the course of a full day (Kramer et al., 2011). Non-seizure networks are highly variable at fine temporal scales and start to show stability over time across cognitive states when computed over at least 100 seconds of data (Kramer et al., 2011). Dynamic networks can also be examined during tasks, using multiple trials to achieve the stability that was achieved above through temporal averaging (Greenblatt et al., 2012). While this dynamic approach is still unusual in the literature, there is a growing body of work assessing differences in ECoG coherence during a task relative to a baseline period (Marsden et al., 2000; Aoki et al., 2001; Ohara et al., 2001; Sehatpour et al., 2008). This dissertation will present a framework for inferring this style of functional connectivity, using ECoG data collected during overt reading as a sample application.

3. Speech and overt reading

The study of the neural basis of speech production poses a few challenges. First of all, speech production involves the use of many different brain areas spread over cortex and subcortical structures. Secondly, speech is highly dynamic, involving rapid movements of the tongue and other articulators. Finally, humans are the only animals capable of speech, so speech production must be studied in human subjects. This historically has restricted studies to noninvasive neural recordings such as fMRI, EEG, and MEG, although as discussed above ECoG is an invasive recording technique that is used in humans. Despite these challenges, there has been much progress towards

identifying areas involved in speech and their roles, as well as the physiological basis of learning, speech disorders, and feedback processing.

The DIVA (directions into velocities of articulators) model describes neural control of speech movements in a way that is consistent with current theory and fMRI results (Guenther et al., 2006). In the model, speech is controlled through the interaction of a feedforward subsystem consisting of premotor cortex, primary motor cortex, and the cerebellum, and a feedback subsystem controlling auditory and somatosensory feedback involving the use of superior temporal cortex (auditory) and inferior parietal cortex (somatosensory).

Overt reading involves the same areas as speech, with the addition of a few related to orthographic and semantic processing. Brain areas involved in mapping orthographic to phonological representations include the left ventral occipitotemporal cortex, the left posterior temporal cortex (middle and superior gyri), the inferior frontal gyrus, and the supramarginal gyrus (Price, 2012). These are thought to be organized into two pathways that distinguish (1) graphophonological processing in which sublexical and lexical features are mapped to phonology through involvement of areas related to short term memory and the articulatory loop, and (2) lexicosemantic processing in which the semantic content is accessed directly based on the visual form of the word, for example in reading irregularly spelled or very familiar words (Jobard et al., 2003; Share, 2008; Price, 2012).

Speech research groups have started to use ECoG to test dynamic predictions from these models. For example, DIVA predicts that primary motor cortex activation

should occur about 40 ms before speech onset (Guenther et al., 2006), a prediction that is difficult to test with low temporal resolution fMRI results. Recent work has begun to approach these kinds of questions with ECoG data (Edwards et al., 2010; Flinker et al., 2010; Leuthardt et al., 2012). One potential challenge is that ECoG electrode grids cover a limited portion of cortex, typically only on one hemisphere, and there is rarely simultaneous recording from the thalamus, basal ganglia, or cerebellum. On the other hand, ECoG grids almost always cover eloquent cortex, which includes key speech network areas such as the primary motor cortex, primary somatosensory cortex, and the superior temporal gyrus. Combining inferences across subjects when the ECoG grid placements differ is a difficult problem with no consensus solution (e.g., Miller et al., 2007; Leuthardt et al., 2012).

4. Summary of Dissertation: Motivation and Approach

The goal of this project was to develop a statistically principled framework for inferring dynamic task-related functional connectivity from human ECoG data. First, we present a general framework that targets dynamics, spatial scale, and uncertainty in task-related connectivity analysis. Second, we refine the methodology to incorporate prior knowledge about mechanisms of neural coupling and to reduce variance in the estimation of the coupling metrics. Finally, we apply the methods to speech ECoG recordings as a case study, discuss the challenges that arise with task-related human data, and describe the neural dynamics of speech production. Taken together, this work contributes a methodological framework for inferring task-related functional connectivity and a description of the fine timescale functional coupling of speech.

4.1. Chapter II: Assessing dynamics, spatial scale, and uncertainty in task-related brain network analyses

Chapter II describes work targeting three common statistical challenges in functional connectivity analysis (published in Stephen et al., 2014). First, temporally dynamic correlated activity between brain areas is best described by networks that change in time. Second, while electrophysiology data is collected at discrete sensor locations, researchers typically endeavor to describe connectivity across multiple spatial scales. Finally, it is important to describe the variability associated with individual edges in functional networks, as well as that of aggregate statistics such as network density. We developed a statistical approach to identify functional networks from task-related brain voltage data in light of these challenges.

The basic approach of functional connectivity is (1) to define nodes, (2) to calculate a test statistic between all possible pairs of nodes, and (3) optionally, to define binary edges between those pairs of nodes that exceed some threshold. For example, a functional network aiming to detect communication through coherence may use electrodes as the nodes, calculate coherence between all node pairs at a frequency of interest, and define p-values based on the Fisher transform of the coherences, which are normally distributed (Brillinger, 1981). This would represent a weighted network, which could then be thresholded using a multiple comparisons test to yield a binary network.

The methodology presented in Chapter II infers binary functional networks by comparing coupling statistics during a task to their empirical distributions during a baseline period. In order to analyze time-varying connectivity, we construct dynamic

networks using sliding time windows, boosting statistical power by incorporating data from many trials. To address multiple spatial scales, we construct networks (1) in sensor space using correlation values, and (2) in region-of-interest (ROI) space using canonical correlation. For all networks, we estimate the variability of individual edges and aggregate network statistics (e.g. density) using a bootstrap procedure.

Chapter II explores the properties of these methods in simulation and describes an illustrative example using ECoG data from one subject during a speech task. The results from real data highlight features of electrode-space and region-space networks. While electrode-space networks can be difficult to interpret visually because of the large numbers of edges, the aggregate network statistics can be less variable. In this case, the estimated density after speech onset is higher than before speech onset, and both are significantly nonzero. This is consistent with a description of speech processing in which brain areas generally increase coupling during both speech preparation and speech execution. Region-space networks have more variable density because of the fewer number of possible edges, but they are also easier to interpret visually. For example, in the data set analyzed here, the region corresponding to ventral primary motor cortex has 2 edges before speech onset and 6 edges after speech onset, showing local coupling with premotor and primary motor cortex during motor planning that expands during speech execution to somatosensory cortex and the supramarginal gyrus. These patterns are consistent with speech processing areas identified by fMRI (Hickok and Poeppel, 2007; Price, 2012).

The functional network inference methodology described in Chapter II is designed to be modular, so that each step of the process can be separately customized according to the requirements of the data and the assumptions of the experimenters. Using weighted or unweighted networks is one such customization. In Chapters III and IV we apply several versions of the inference methodology that are designed to target different aspects of the speech tasks and incorporate either strong or weak assumptions about the underlying mechanisms of functional connectivity. By analyzing the data in a few different ways, we aim to recognize where bias is adversely affecting our results and identify the best fitting models in order to refine our conclusions.

4.2. Chapter III: Pairwise coupling metrics for electrode-level and region-of-interest-level functional network analysis

As discussed above, the more we can incorporate appropriate prior knowledge into the network inference, the more we will be able to infer from the data. Chapter III expands on the test statistics developed in Chapter II by incorporating more prior knowledge about (1) theorized mechanisms of communication between brain areas, (2) the statistical behavior of the test statistics, and (3) the sources of variability in the experiment.

First, if the "communication through coherence" hypothesis (see Section 1.3) is correct and functional connectivity between brain areas is subserved by rhythmic activity, then using coherence rather than correlation to detect functional links may reduce bias and improve the power of the functional network analysis.

Furthermore, correlation, coherence, and their canonical counterparts are known to be normally distributed in the limit as the number of observations approaches infinity (Brillinger, 1981). While the convergence is slow for small samples, applying the Fisher transform can stabilize the variance and speed up the convergence to normality. Incorporating this knowledge about the behavior of the test statistics and using theoretical distributions rather than empirical distributions when appropriate can reduce sample variance and increase the power of statistical tests.

A third improvement to the methodology in Chapter II relates to the sources of variability in the experiment. The test statistics in Chapter II model the sampling distribution of the baseline period, but they do not model the sampling distribution during the test period, effectively assuming that the test statistic estimated from the test period has no sampling variability. The unmodeled variability in the test period propagates through the analysis, leading to excess variance in the p-values that are used to distinguish edges from non-edges in the inferred networks. In practice, this exclusion results in more false positives than are expected from the False Discovery Rate procedure. Chapter III eliminates this source of false positives by reformulating the test statistics in terms of the difference between the test and baseline period, and explicitly estimating the variance of the difference. The variance of the difference between the test and the baseline includes the variance of both the baseline period and the test period. This approach of formulating the hypothesis tests as two-sample tests for differences in means (Casella and Berger, 1990; Efron and Tibshirani, 1993) is analogous to the well-known

two-sample t-test. By accounting for as much of the experimental variability as possible, the resulting inferred networks have minimal false positives.

Chapter III presents methodology for inferring networks using four different test statistics that incorporate these improvements: correlation, coherence, canonical correlation, and canonical coherence. Since canonical correlation and canonical coherence are rare in the functional connectivity literature, we also include in an appendix a derivation of the solution to the optimization problem solved by canonical coherence, and a multitaper estimator of the sample canonical coherence[1]. Each test statistic is tailored to detect a different kind of coupling in the neural signals. Since the true forms of coupling between brain areas are still unknown, the hope is that inferring networks using multiple test statistics will help narrow down the options. That is, if one or more of the test statistics is a good match for the underlying coupling mechanism, then we will be more likely to reject the null hypothesis in a replicable way for that test statistic. This approach is demonstrated with simulations in Chapter III and applied to human ECoG recordings in Chapter IV.

4.3. Chapter IV: Characterizing the dynamically evolving functional networks of speech

Chapter IV considers ECoG recordings collected during overt reading, and analyzes the dynamic changes that occur in the neural activity around the time of speech onset. For completeness, the analysis covers both first order (within electrode) effects and second order (between electrode) effects. As a result, the analysis is grounded in classical

[1] Canonical correlation is a special case of canonical coherence, and we include a section describing how it relates.

results related to speech production (e.g. event-related potentials and sensorimotor rhythms) and more modern approaches to speech production research (e.g. high gamma power as an indicator of neural population activity).

Specifically, the event-related potentials (ERPs) over rolandic cortex are consistent with historical EEG observations of a readiness potential (Deecke et al., 1986; Wohlert, 1993), and we observe beta event-related desynchronization over rolandic cortex that matches the sensorimotor rhythm literature (Pfurtscheller and Lopes da Silva, 1999). Auditory cortex has a relatively weak ERP in the average, but individual electrodes have ERPs with varying timecourses, consistent with previous literature (Flinker et al., 2010). With regard to the more recent paradigm of tracking high gamma or broadband power as a correlate of population spiking activity, both rolandic and auditory cortex show increases in broadband power time-locked to the onset of speech. A latency analysis places the onset of the rolandic broadband response at an average of 52 ms prior to speech onset, and the auditory broadband response at an average of 92 ms after speech onset. In general, these broadband effects occurred after the beta desynchronization. This timing is consistent with a view of beta rhythms as reflecting spatially broad control mechanisms that gate local processing.

Within this context, the second-order functional connectivity results, while preliminary, are promising. We make an effort to manage the high dimensionality by (1) inferring networks only over electrodes in regions of interest (rolandic, auditory, and inferior parietal cortex), (2) controlling the false discovery rate of the networks, and (3) characterizing the uncertainty of region-level average coupling. Even so, the results

across subjects are inconsistent, likely due to a lack of statistical power. In two subjects with particularly large numbers of trials and particularly good spatial coverage of rolandic cortex, there appear to be dense intra-regional correlation networks that may be related to changes in coherence in the alpha and beta bands. These dense rhythmic networks are consistent with an interpretation of a common thalamic driver of alpha rhythms across rolandic cortex (Hughes, 2005), and beta traveling waves over motor cortex during movement preparation (Rubino et al., 2006).

In addition to these static descriptive analyses, we present an interactive visualization tool, SpectraVis, which facilitates exploration of the time-domain and frequency-domain networks with their corresponding spectrograms and coherograms. SpectraVis displays the network at a particular time and frequency alongside the spectrograms and coherogram for a selected edge, while the user can mouse-over the displays to change the selections. The binary and weighted networks can be shown in anatomical space and a force-layout format, and they can be played dynamically as movies across time. In such a high-dimensional exploratory analysis, an interactive visualization can make it possible to find strong network effects and make connections between different metrics without explicit a priori assumptions about the locations and timing of the effects. The results of these explorations are necessarily preliminary, however, because of the problem of multiple testing, and so will need to be replicated in future studies.

4.4. Chapter V: Conclusion

We conclude by bringing together the findings of the work, discussing its implications for the broader scientific community, and proposing future directions for research.

CHAPTER II. ASSESSING DYNAMICS, SPATIAL SCALE, AND UNCERTAINTY IN TASK-RELATED BRAIN NETWORK ANALYSES

1. Introduction

The recent neuroscience literature has seen a dramatic increase in the number of studies that investigate functional connectivity in brain networks. Roughly speaking, functional connectivity refers to coupling (i.e., systematic associations or relationships) between neural activities in different brain regions of interest (ROIs) or recording sites (Friston, 1994; Bullmore and Sporns, 2009). Functional connectivity can be estimated from a wide range of data types with varying degrees of temporal and spatial resolution, including data with high spatial resolution but poor temporal resolution collected with positron emission tomography (PET) and functional magnetic resonance imaging (fMRI), as well as data with high temporal resolution collected using electroencephalography (EEG), electrocorticography (ECoG), and magnetoencephalography (MEG). Here, we focus on functional connectivity estimated from brain voltage recordings, i.e. EEG and ECoG (for a review of similar considerations in the context of network inference for fMRI, see Hutchison et al., 2013). One of the foremost issues associated with functional connectivity analysis is the choice of coupling measure. Coupling measures include linear and nonlinear measures of statistical association, as well as information theoretic and model based measures, and can be chosen to highlight specific types of associations such as rhythmic or causal coupling (as reviewed in Pereda et al., 2005; Greenblatt et al., 2012).

In addition to the choice of coupling measure, a number of important statistical issues arise in the inference and analysis of functional brain networks. Here we highlight three such issues. First, researchers are often interested in detecting and characterizing dynamic transitions in functional connectivity structure. Such transitions may arise suddenly as a function of a specific stimulus or behavior, or may reflect gradual ongoing changes in connectivity through time. For example, in speech production – an example we will refer to throughout this paper in order to focus our thoughts – it has been shown that functional connectivity as measured with fMRI changes during voicing (Simonyan et al., 2009), and abnormal functional connectivity has been associated with disordered speech (e.g., Chang et al., 2011). During epileptic seizure, brain functional networks assessed from ECoG data exhibit characteristic dynamical patterns that include the aggregation and fragmentation of network structure (Schindler et al., 2007; Burns et al., 2012; Kramer and Cash, 2012).

A second statistical issue in the analysis of functional connectivity in the brain concerns multiple spatial scales. At the microscopic scale, associations occur between the activities of individual neurons (e.g., Cohen and Kohn, 2011), and between the aggregate activity of small neural populations (e.g., Schnitzler and Gross, 2005). At the macroscopic scale, associations emerge between entire brain areas, or previously identified ROIs (e.g., Golfinopoulos et al., 2011). Often, the spatial scale at which brain networks are analyzed is determined by the implicit scale of the method used to record neural activity. However, there has been increasing interest in understanding relations between functional connectivity structure at multiple spatial scales. This relationship can

be studied using either multiple simultaneous measures of neural activity or statistical methods that are capable of inferring associations across scales.

The final statistical issue we chose to highlight relates to estimating uncertainty in network level statistics. Typically, a selected coupling measure is estimated pairwise between all possible nodes, and the results are thresholded to produce a binary network (e.g., Stam, 2004; Micheloyannis et al., 2006; Ponten et al., 2007; Srinivas et al., 2007; Stam et al., 2007; Kramer et al., 2008; Supekar et al., 2008). Various network measures are then computed to summarize features of the resulting network topology (Kolaczyk, 2009; Rubinov and Sporns, 2010). However, the uncertainty pertaining to these network level statistics is typically not computed. This uncertainty will be related to the uncertainty associated with each of the pairwise connections, but propagating this edge uncertainty to network level uncertainty is nontrivial and remains an active area of research.

In this paper, we present a framework for inferring functional connectivity that addresses each of these issues. After introducing a statistical methodology appropriate for analyzing connectivity in a given time window with respect to a baseline condition (Sections 2.4 and 2.5), we demonstrate the use of this network methodology in the context of assessing dynamics, spatial scale, and uncertainty in networks. First, temporally dynamic changes in network structure are tracked (Section 3.1). Second, using a priori spatial information, functional connectivity is generalized from between-node connectivity to between-region connectivity. This procedure reduces the number of inferred connections relative to the number of measurements, and trades spatial resolution

for a more reliable estimate of functional connectivity. We apply both of these approaches to ECoG data collected during an overt speech task for one subject (Section 3.2), demonstrating their feasibility and highlighting some desirable features. Finally, we present several examples to illustrate advantages of the canonical correlation metric (Section 3.3). Throughout these discussions we quantify the uncertainty inherent in estimates of functional connectivity (Sections 3.1, 3.2, and 3.3). Together these approaches advance the burgeoning field of functional connectivity in important ways and provide a general tool for the construction of meaningful functional networks in task-related data with associated measures of confidence. The paper concludes with a discussion of limitations and of future research directions (Section 4).

2. Methods

Network analyses take a wide variety of forms; even for functional networks constructed using cross correlation (i.e., correlation-based networks) applied to brain voltage data, there are a number of design decisions that affect the interpretation of the resulting networks. Here we focus on how correlation-based networks differ during a task compared to a baseline period, and we note that the same framework developed here also applies to other choices of coupling measure, including coherence-based networks (see Discussion). We propose a multi-step analysis strategy that builds on previous work on edge-count uncertainty in binary networks (Kramer et al., 2009), utilizing the trial structure of the data to increase power, highlight differences between task and baseline, and assess uncertainty. In brief, the proposed method determines network edges (between individual sensors or groups of sensors collected across ROIs) using a statistical

hypothesis test, corrects for multiple comparisons, and computes confidence measures at the single edge and aggregate network measure levels. As described in detail below, nonparametric bootstrap estimation (Efron and Tibshirani, 1993) is used both for the statistical hypothesis test to detect individual edges, and for the determination of confidence intervals over edges and aggregate network measures. Because bootstrapping is nonparametric, it requires minimal assumptions about the data.

2.1 Motivation: speech task

Here we will consider networks constructed during speech production as compared to periods of silence. While this example will be used throughout the paper, any repeated task with a baseline period could be substituted. The production of speech involves a large network of brain regions that spans several lobes of the cerebral cortex along with numerous subcortical structures (e.g., Guenther et al., 2006). Fluent speech requires very rapid movements of the tongue and other articulators. For example, a typical speaker can easily produce 10 phonemes in 1 second; this involves the precise sequencing of individual articulatory gestures that each last approximately 100 ms. The most commonly used neuroimaging techniques for studying speech, PET and fMRI, have a temporal resolution on the order of 1 second, precluding them from measuring these rapid dynamical processes. Furthermore, speech articulation creates massive muscle-related artifacts in EEG and MEG, limiting the utility of these technologies for studying speech production. To overcome these problems, a number of neuroscientists have begun studying speech using ECoG recordings collected prior to epilepsy surgery (e.g., Korzeniewska et al., 2011; Pei et al., 2011; Leuthardt et al., 2012; Bouchard et al., 2013);

ECoG recordings have high temporal resolution (typically on the order of 1 ms) combined with reasonably high spatial resolution (on the order of 1–10 mm on the cortical surface). Thus, ECoG offers the possibility of studying dynamic changes of functional connectivity within the speech network, potentially providing a powerful tool for deciphering the neural dynamics underlying speech and other actions or cognitive tasks.

In general, quite complex topologies are derived using modern neuroimaging techniques. For example, a network consisting of 100 electrodes (typical in both noninvasive and invasive brain voltage recordings) may contain up to 4950 edges. Many measures exist for assessing the organization of these edges (Kolaczyk, 2009; Rubinov and Sporns, 2010). However, principled methods to determine confidence in these network measures are not well established. A notion of confidence is particularly important when assessing how networks change in time. For example, during a reading task, we may ask how the functional network changes during the different stages of the task, starting from visual and linguistic processing of the stimulus, proceeding to motor planning and execution, and finally involving the processing of auditory and somatosensory feedback during speech. In the simplest scenario that two functional networks are established – one preceding the onset of speech, and the other following – we might then ask how the density, a network measure defined as the ratio of the number of detected edges to the number of possible edges in the network, differs before and after speech onset. To answer this question in a meaningful way requires a measure of uncertainty in the density. In what follows, we propose a resampling procedure to

estimate the sampling variability of individual edges as well as to establish confidence intervals for density and other aggregate network measures.

2.2 Simulated Data

Simulated data were generated to mimic ECoG recordings in which a task is performed multiple times, and the resulting brain activity observed. For example, a speech task may involve showing the subject a specific word, which the subject would read multiple times during the experimental session. In general, we expect the pattern of correlations between electrodes to vary as a function of time with respect to task onset. If the task onset is defined as the start of the behavioral response (e.g., overt speech), there may be task-related activity and connectivity that precede the task onset, such as processing of the visual stimulus and motor preparation. Hence we consider the trials to start before task onset and last until some time after task onset (specifically, 500 ms before until 500 ms after task onset).

In the simulations performed here, we consider dynamic activity recorded from nine sensors. The synthetic data at each sensor consist of four dynamic components with a known pattern of correlations. The first component, pink noise, captures one feature of brain voltage activity, the "1/f" reduction in power as a function of frequency common in brain voltage activity (Miller et al., 2009a; He et al., 2010). The second component is white noise, meant to represent sensor noise. On top of these two uncorrelated components, two types of correlated signal are added: 2–50 Hz correlations that exist throughout the recording, and 8–25 Hz correlations that only come into effect during the trials. The first is meant to represent any persistent correlated structure existing in the

signals that does not change during the trials, such as correlations related to the recording apparatus. The trial correlations are meant to represent task-related changes in the correlation structure. For example, during a speech task there may be increased correlations between areas involved in speech, including high-level motor areas, primary motor cortex, and auditory and somatosensory cortices (Guenther et al., 2006). While in a true task this structure may change many times during the different stages of task execution, in the simulated data we introduce only two known network structures: one that gradually increases, peaks, and fades away in the 500 ms before task onset on every trial, and one that gradually increases, peaks, and fades away in the 500 ms after task onset on every trial. Details of these networks and how the simulated data were generated are described in Appendix 5.1 and summarized in Figure II.1.

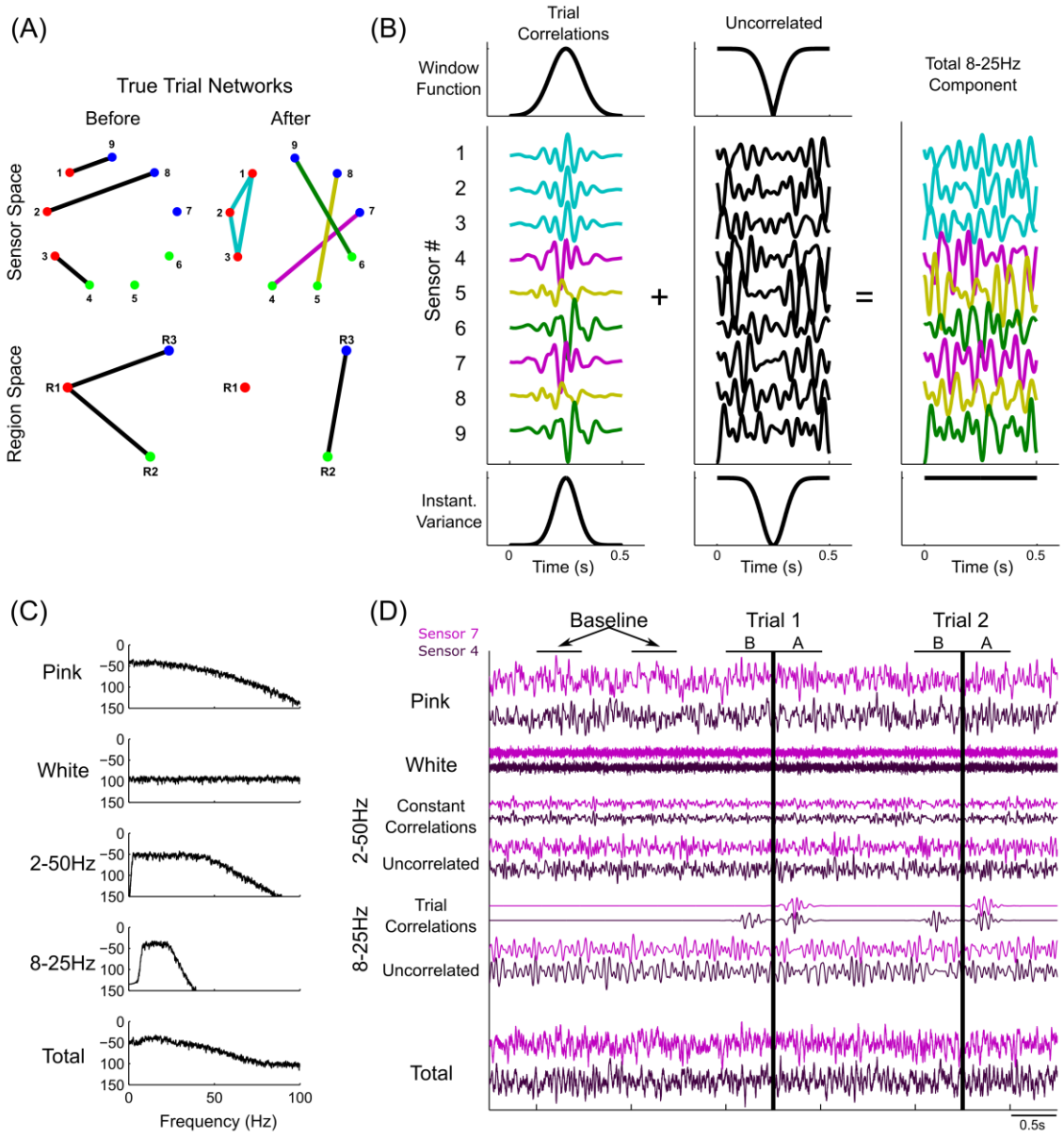


Figure II.1: Illustration of construction of the simulated data.

Predetermined networks (A) were chosen for periods before and after task onset. Sensor-space networks (A, top row) correspond to region-space networks (A, bottom row). Node colors correspond to region assignments. Panel B shows how the correlation structure defined in A was introduced into the simulated data, using the after network to illustrate. The correlation structure in the 500 ms of each trial following task onset consisted of 4 instantiations of 8–25 Hz noise (B, left), added to the sensors according to the defined network (colors in B correspond to the edge colors in the after network in A). The trial correlations were introduced into a background of uncorrelated 8–25 Hz noise by windowing (B, top row) in such a way that the variance remained constant in the total signal (B, bottom row). In addition to the 8–25 Hz component, 3 other signals were added: pink noise, white noise, and a 2–50 Hz signal containing a correlated and uncorrelated part (C: power spectra; D: time domain). Panel D shows sample data for 2.5 s of

baseline and 2 trials. Vertical black lines correspond to task onset: the before network is in effect for 500 ms before this time and the after network is in effect for 500 ms after this time. For more details regarding the construction of simulated data, see Section 2.2 and Appendix 5.1.

2.3 ECoG Data

To illustrate the use of the methods on experimental data, we analyze ECoG time series recorded during overt speech.

2.3.1 Experimental Protocol and ECoG recording. The neuronal recordings were collected from a single individual undergoing treatment for intractable epilepsy involving implantation of subdural electrocorticographic grids, used for localization of seizure foci for later resection of epileptic tissue. The electrodes consisted of two, 1-cm spaced grids positioned over the frontal and parietal lobes, and the temporal lobe, respectively, and a strip of electrodes over the occipital lobe. Signals were recorded using g.tec g.USBamp amplifiers (sampling rate 9600 Hz). Data acquisition and stimulus presentation were handled using BCI2000 software (Schalk et al., 2004).

During the task the subject read the Gettysburg Address (272 words) aloud from a video monitor, as the text scrolled from right to left. The full session lasted 295 seconds. There were no seizures during the experimental session, and electrodes near the putative seizure focus were excluded from the analysis. The subject gave informed consent to participate, and this research has been approved by the local institutional review boards.

2.3.2 Preprocessing of ECoG data. The analyses were based on recordings from 90 electrodes. The slow progression of the teleprompter forced the subject to pause periodically, and these pauses were used to define a trial structure. Specifically, trials

were defined as any time during the speech recording where 500 ms of silence was followed by at least 500 ms of speech, determined using an audio recording of the session. Hence, each trial lasted one second, and contained activity related to reading the word(s), preparing the motor command, and voicing the word or phrase aloud. Note that the specific words spoken during the trials varied. There were a total of 98 trials so defined. Furthermore, data from silences separated by more than 500 ms from the beginning or end of any utterance were extracted to be used in defining the baseline distribution, described below (Section 2.4.3).

2.3.3 ROI definitions. Regions of interest were defined anatomically for the subject based on manual inspection of structural MRI by an experienced anatomist. The 25 regions corresponded to anterior/middle dorsal premotor cortex, posterior middle frontal gyrus, inferior frontal gyrus, pars opercularis, posterior dorsal premotor cortex, middle premotor cortex, ventral premotor cortex, dorsal primary motor cortex, middle primary motor cortex, ventral primary motor cortex, dorsal somatosensory cortex, middle somatosensory cortex, ventral somatosensory cortex, superior parietal lobule, anterior supramarginal gyrus, posterior supramarginal gyrus, fronto-orbital cortex, temporal pole, anterior superior temporal gyrus, posterior superior temporal gyrus, anterior middle temporal gyrus, posterior middle temporal gyrus, anterior inferior temporal gyrus, posterior inferior temporal gyrus, ventral temporal cortex, and occipital cortex (Golfinopoulos et al., 2011). Electrode ROI assignments are indicated by node color in Figure II.9.

2.4 Construction of functional networks

The specific method for inferring functional network structure from brain voltage recordings used here contains a set of techniques for analyzing the dynamics, spatial scale, and uncertainty in network connectivity. While we discuss the behavior of this technique as a whole, the steps are modular and generally applicable: for example, the coupling statistic could be changed from correlation to coherence or phase locking value while keeping the other aspects of the analysis, like uncertainty estimation, intact. Alternatively, a different technique for correcting for multiple comparisons could be used to identify binary edge assignments. Below, we describe our choices for each module of the functional network analysis: preprocessing (including defining trial epochs), choice of test statistic, p-value calculation through comparison to a null distribution, multiple comparisons correction, and uncertainty assessment. This procedure is also outlined in Figure II.2. One additional module, the selection of network nodes, is implicit in this analysis. We use two different node definitions corresponding to the coupling statistics, defining each node as an individual sensor (for correlation) or all of the sensors in an ROI (for canonical correlation).

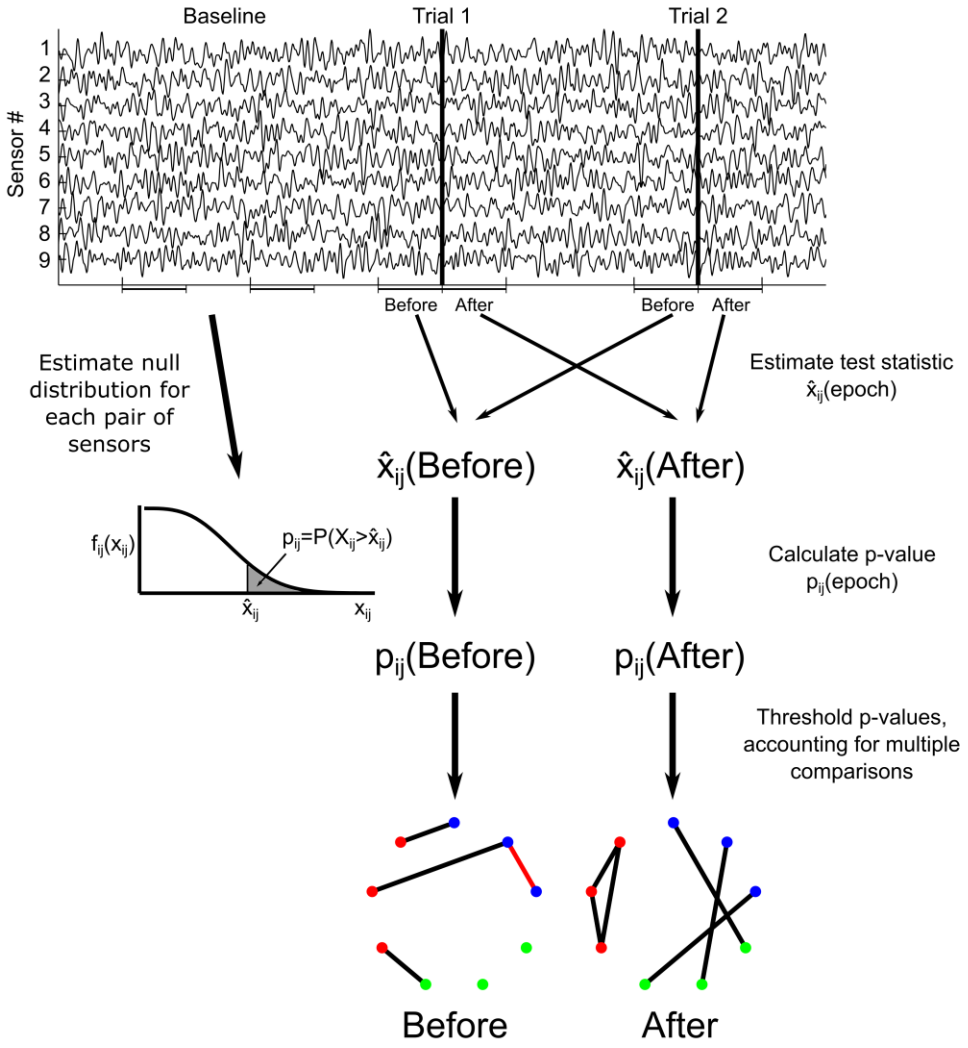


Figure II.2: Illustration of construction of functional networks from the time series data.

For each trial epoch of interest (here labeled “Before” and “After”), the test statistic $\hat{x}_{ij}$ for each potential network edge (between node i and node j) is calculated over all trials. The test statistic is compared to a null distribution estimated from baseline intervals, resulting in p-values for each potential edge, p_{ij} . A threshold is chosen for the p-values using a multiple comparisons correction, resulting in binary networks. Red edges in the networks indicate false positive edges, identified by comparison to the true networks. For more detail, see Section 2.4.

2.4.1. Preprocessing of the data. To infer functional networks from the synthetic time series, the data were first preprocessed by downsampling to 200 Hz (Matlab decimate

function, which in this case also low-pass filtered the signal below 100 Hz)[1]. The signals were then band-pass filtered between 0.1 and 30 Hz using a zero-phase 3rd order Butterworth filter (Matlab filtfilt function). This preprocessing step was performed to mimic the typical procedure of bandpass filtering observed brain voltage recordings to eliminate artifacts and focus on a particular frequency range. For the ECoG data, a common average reference was applied by subtracting the mean over all channels from each channel at each time step. This step was intended to reduce referencing effects from the experimental signals.

In order to analyze time-varying functional connectivity, the trial data were then divided into epochs of interest spanning the 1s-long trials. Two kinds of epoch were defined. The first were 500 ms non-overlapping epochs consisting of the first and second halves of the trials (500 ms each, with no overlap), corresponding to "before" and "after" task onset, defined as time zero. The second kind of epoch used a 200 ms sliding window (195 ms overlap) over trials, allowing for the construction of "dynamic" networks on overlapping time points starting from 400 ms before time zero until 400 ms after time zero (the 200 ms epochs were labeled according to their time midpoints).

The epochs of interest were collected from all trials resulting in a 4-dimensional matrix with dimensions ($E \times T \times N \times L$), where E is the number of epochs, T is the number of time points per epoch, N is the number of sensors, and L is the number of trials. For example, using before/after epochs in the simulations here, there were 2 epochs, each consisting of 100 time points (corresponding to 500 ms sampled at 200 Hz),

[1] Note that the ECoG data had a sampling rate of 9600 Hz while the simulated data had a sampling rate of 1200 Hz. After downsampling, they both had sampling rates of 200 Hz.

9 sensors, and 100 trials. Using dynamic epochs, there were 162 epochs, 40 time points per epoch (200 ms sampled at 200 Hz), 9 sensors, and 100 trials. The data from each sensor, trial, and epoch were normalized by subtracting out the mean and dividing by the standard deviation.

For each epoch type, a group of non-overlapping "baseline" intervals were also collected from the baseline period for characterization of the null distribution of the test statistic between each pair of sensors. We note that the baseline intervals contain no trial data and hence should not contain task-related correlated activity, yet may possess other correlated activity due to the persistent correlation signal common to all sensors. The intervals of baseline data were chosen to have the same duration as the trial epochs, and were stored in a 3-dimensional matrix with dimensions ($T \times N \times K$), where K is the number of baseline intervals. For example, for before/after epochs in the simulated data, 400 non-overlapping intervals of the (preprocessed) baseline data, 100 time points long, were collected from the baseline period. Note that K does not need to be the same as L: having more baseline intervals than trials will improve the estimation of the null distribution (described in Section 2.4.3). The baseline intervals are normalized in the same way as the trial intervals, by subtracting the mean and dividing by the standard deviation.

2.4.2 Computing the test statistic. Two measures of coupling (test statistics) were employed to establish the functional networks: correlation and canonical correlation. For a given epoch, the test statistics were calculated for each pair of nodes using all trials.

Hence calculating the test statistics resulted in a 3-dimensional matrix with dimensions ($E \times N \times N$) for sensor-space networks and ($E \times R \times R$) for region-space networks, where R is the number of regions.

We focus on correlations that become more extreme (more positive or more negative) during the task, so the absolute value of the raw correlation values was used as the sensor-level test statistic. The absolute value of the correlation was calculated as:

$$AbsCorr(x, y) = \left| \frac{\sum_{l=1}^{L} \sum_{t=1}^{T} x_l(t) y_l(t)}{\sqrt{\left(\sum_{l=1}^{L} \sum_{t=1}^{T} x_l(t)^2 \right) \left(\sum_{l=1}^{L} \sum_{t=1}^{T} y_l(t)^2 \right)}} \right|$$

where $x_l(t)$ and $y_l(t)$ are the voltage at time t of the given epoch of trial l for sensors x and y, respectively.

Canonical correlation was chosen as a region-level coupling measure because of its relationship to the sensor-level correlations. A common measure of network topology (used to describe networks that have already been constructed) involves the identification of groups of nodes (e.g., communities). However, when analyzing multivariate data recorded from the brain, there often exists an a priori natural grouping of nodes belonging to a region of interest (e.g., a brain region associated with a particular function). By using a classical statistical multivariate technique – canonical correlation – connectivity between groups of channels can be studied in a principled way. Canonical correlation finds the linear combinations of the group members such that the linear combinations are maximally correlated. Canonical correlation is reported in classical texts on multivariate statistics (e.g., Mardia et al., 1979; Anderson, 2003). When a priori knowledge of signal

presence on group member channels is not known, canonical correlation possesses increased statistical power and is a powerful method of dimensionality reduction in the context of network analysis. This is discussed further in Section 3.3.

Formally, the canonical correlation between two groups of signals $\vec{x} = (x_1, ..., x_n)^T$ and $\vec{y} = (y_1, ..., y_m)^T$ seeks to find linear combinations of $\vec{x}$ and $\vec{y}$ that are maximally correlated:

$$CanonCorr(\vec{x}, \vec{y}) = \max_{\vec{a},\vec{b}} Corr\left(\vec{a}^T\vec{x}, \vec{b}^T\vec{y}\right)$$

For convenience, we will switch to matrix notation: X and Y are ($n \times LT$) and ($m \times LT$) matrices, respectively, containing a row of appended trial voltages for each sensor in the region. The optimization problem for canonical correlation is solved using the singular value decompositions (Strang, 2003) of X and Y, $X = U_X \Sigma_X V_x^\dagger$ and $Y = U_Y \Sigma_Y V_Y^\dagger$. The canonical correlation is the first singular value of the matrix $Q_{XY} = U_X V_X^\dagger V_Y U_Y^\dagger$.

2.4.3 Assessing the significance of the test statistic. After computing the test statistic between each pair of nodes, a p-value is computed using the null hypothesis that the test statistic for that edge comes from the same distribution as the baseline period (which lacks task-related activity). The alternative hypothesis is that the test statistic is too large to be from the baseline distribution. For example, in our analysis of speech data we use a null distribution estimated from epochs of time without speech (i.e., silence). Note that for simplicity we have chosen a one-sided hypothesis test, focusing only on correlations or canonical correlations that are larger than baseline. All of the methods described here

could be developed for a two-sided hypothesis test, allowing for both higher and lower correlations and canonical correlations than baseline.

In order to estimate the null distribution, a bootstrap procedure is employed involving resampling of the baseline data. In particular, for each bootstrap sample, L baseline intervals (described in Section 2.4.1) are sampled with replacement from the K total baseline intervals. The test statistic for each potential edge is calculated from these baseline intervals just as it was for the set of trial intervals. Here, 1000 bootstrap samples were processed in this way, thus building an empirical distribution for the test statistic from the baseline period for each potential edge.

In order to determine a p-value associated with each edge, the test statistic is computed during the epoch of interest and compared to the null distribution. Specifically, the p-value associated with the edge between node i and node j is calculated as the number of samples in the bootstrapped empirical distribution for nodes i and j that fall above the observed test statistic between nodes i and j, divided by the total number of bootstrap samples (here we use 1000 samples for simulated data, and more for the real data: for discussion, see Appendix 5.2). When the test statistic is larger than any of the samples in the empirical distribution, this method assigns a p-value of zero, which is unrealistic. To account for this, we set all zero p-values to the smallest possible p-value for the number of bootstrap samples (e.g. 1/1000 for the simulated data).

2.4.4 Correcting for multiple comparisons. After a p-value has been computed for each potential edge, a correction for multiple comparisons is performed using the Benjamini-

Hochberg procedure to control for False Discovery Rate (hereafter referred to as the FDR procedure; Benjamini and Hochberg, 1995). This procedure sets a theoretical limit, q (here, 0.05), on the expected number of falsely detected edges as a proportion of the total number of detected edges. The procedure first sorts the p-values in increasing order and then chooses the highest integer k such that:

$$p_{(k)} \leq \frac{qk}{N_{MC}}$$

where N_{MC} is the total number of comparisons being performed, and $p_{(k)}$ is the k^{th} smallest p-value. The FDR procedure thus selects a set of edges to consider significant such that only 5% *of the detected edges* are expected to be false positives. In contrast, a typical p-value threshold of α = 0.05 implies that 5% *of all possible edges* will be expected to be false. The FDR procedure is thus expected to have many fewer false detections than would occur with a 5% p-value threshold. Note that when the null distribution is estimated from data, as it is here, more than the expected number of false positives may occur in the network due to sampling bias (Dudoit and Laan, 2008).

The resampling approach used here to estimate the baseline null distributions, combined with the FDR procedure for multiple comparisons correction, imposes a lower bound on the number of edges that can be detected in the networks. This lower bound scales with the square of the number of nodes in the networks, so it can become large for a network with many nodes. This phenomenon and some approaches to dealing with it are discussed more in the Appendix (Section 5.2).

After the FDR procedure has been applied, the resulting assessment of test statistics as significant or not significant is used to define a binary network for the epoch

of interest. This procedure is repeated for all epochs of interest (e.g., 500 ms before and after time zero), creating a dynamic series of networks. Note that the null distributions do not need to be estimated separately for each epoch. The null hypothesis is defined in terms of baseline distributions for each potential edge that are independent of trial epoch, so the same estimated distributions can be used for all epochs. The procedure for constructing a functional network is also illustrated in Figure II.2.

2.5 Assessing uncertainty

We focus on assessing the uncertainty of two network features: (1) edge existence, and (2) network density – the total number of edges in a network divided by the number of possible edges. To do so, we employ a classic nonparametric bootstrap procedure, making use of the trial-structure of the data (Efron and Tibshirani, 1993). Specifically, the time series data from the trials are resampled with replacement, the test statistic at each potential edge is calculated over the resampled trials, p-values are computed by comparison of the test statistic to the same baseline distribution used above, the p-values are thresholded using FDR to create a binary network, and an aggregate network measure (the density) is computed for the network. This process of resampling generates a population of surrogate networks (here we use 100 bootstrap samples) that can be used to approximate the true network population. For example, the probability of each edge can be estimated as the proportion of the surrogate networks that contain that edge. Statistically, this approach treats each potential edge as a Bernoulli random variable with parameter p, the probability of "success", or the probability of observing an edge.

The variance is therefore equal to $p(1-p)$. Hence, estimating the probability of an edge allows us to assess the variability of the network edges themselves.

For an aggregate network statistic such as the density, the distribution of the statistic calculated on the resampled networks is an estimate of the true sampling distribution of the statistic. Here, we estimate the standard error of the density as the standard deviation of the resampled densities, $\hat{s}$. Then, a 95% confidence interval around the observed network density d_{obs} is constructed using a Gaussian approximation as $\left[d_{obs} - 1.96\hat{s}, d_{obs} + 1.96\hat{s}\right]$.

The estimates of the sampling distributions of networks and network statistics constructed this way are biased because each surrogate network is constructed using fewer unique trials than the original network (constructed using all trials). This results in decreased degrees of freedom in the surrogate networks, which has the effect of increasing the occurrence of false positives. This is similar to training-set-size bias that occurs in prediction error estimation (Hastie et al., 2009). Increasing the number of false positives will tend to increase the probability of individual edges and the overall density of the surrogate networks relative to the original network. Hence, (1) the estimated probabilities of edges reported below will have a small upward bias, and (2) the full bootstrap distribution of densities estimated as described above will be biased upward. Note that we chose to construct confidence intervals of the density using only the standard error of the bootstrap distribution, in part to mitigate this issue. Alternative approaches exist, including the construction of confidence intervals using quantiles of the

bootstrap distribution, possibly correcting for the bias using a technique such as the "Bias-Corrected and Accelerated" method (BCa) (Efron and Tibshirani, 1993).

See Appendix 5.3 for the full algorithm.

3. Results

In this section we apply the network inference procedure described above to construct functional networks from synthetic data. We consider different simulation scenarios to illustrate the utility of the method in assessing the dynamic evolution of networks, in the principled spatial aggregation of network nodes using canonical correlation, and in the determination of uncertainty in network edges and measures.

3.1 Dynamics

3.1.1 Inference of functional networks before and after task onset. To begin we consider the scenario in which the only coupled activity between electrodes appears during trials. Outside of these time intervals, the simulated activity for each electrode consists only of uncorrelated noise; in these simulations, no correlated baseline activity exists between the electrodes (i.e., the constant correlations described in Sections 2.2 and 5.1.3 were set to zero). We perform these simulations for four different levels of signal-to-noise ratio (SNR). Functional networks were constructed using both the correlation and canonical correlation for these simulated data in two intervals: one interval immediately before the onset of the task (the first half of the trial), and another immediately after the onset of the task (the second half of the trial); both intervals are of duration 0.5 s. As expected, the network inference procedure identifies the appearance of the correct functional network before and after task onset when the SNR is sufficiently

large. This occurs for networks computed using both correlation (Figure II.3A) and canonical correlation (Figure II.3C). Note that for the correlation networks, edges appear between the (nine) individual nodes, while for the canonical correlation networks, edges appear between the (three) ROIs.

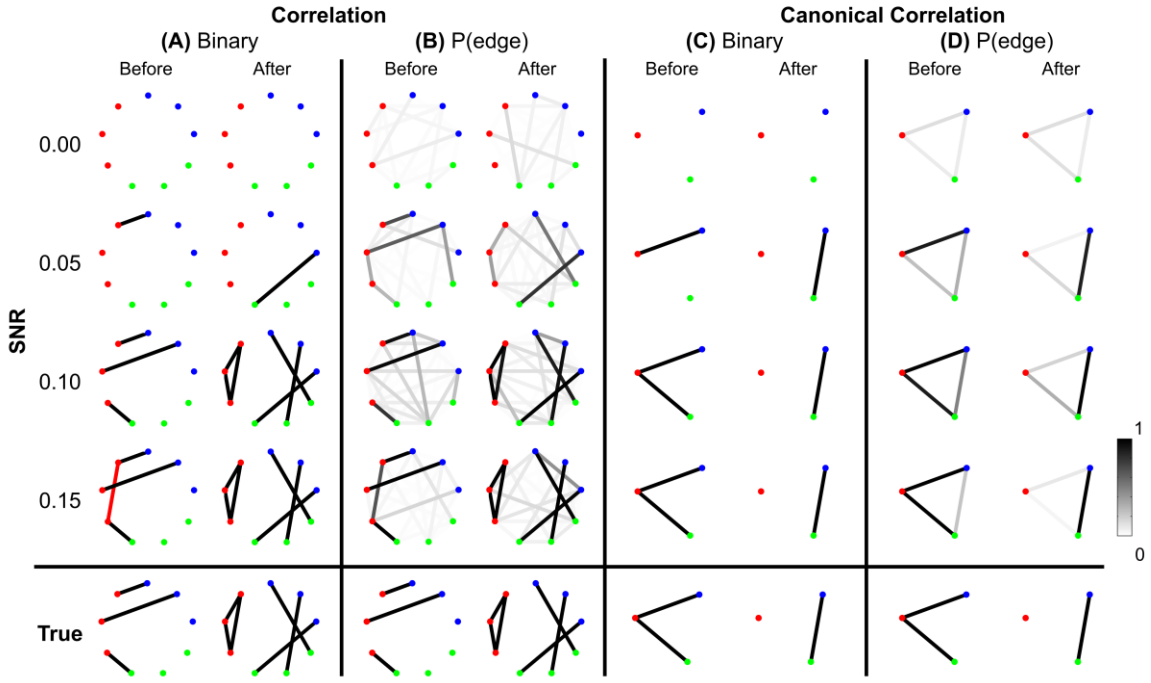


Figure II.3: Detected networks reflect the true correlation structure for sufficiently high SNR.

Correlation and canonical correlation networks are shown before and after task onset, for simulations with varying SNR. (A) Binary correlation networks for 500 ms before task onset (left) and the 500 ms after task onset (right). True positive edges are colored black, false positive edges are colored red. (B) Bootstrapped probability of each edge for the correlation before (left) and after (right) networks. Probabilities are indicated in grayscale from 0 (white) to 1 (black); scale bar at bottom right of the figure. (C) The same as (A), but for canonical correlations between regions. (D) The same as (B), but for canonical correlations between regions. Plots (A)-(D) are shown in the rows for four different levels of SNR: 0.00, 0.05, 0.10, and 0.15. True networks are shown in the bottom row. In all networks, nodes are color-coded according to region.

In addition to the binary network inference, the proposed method permits a characterization of the confidence in each edge. For both the correlation and canonical correlation networks, we associate with each edge a probability of appearance (determined through resampling, see Methods, Section 2.5). As expected, edges that appear in the binary networks possess a high probability of appearance in the correlation (Figure II.3B) and canonical correlation (Figure II.3D) networks. These results illustrate that the network inference procedure – and associated measures of edge uncertainty – accurately identify the underlying network when no coupled baseline activity exists between the nodes, even when the SNR is relatively small (0.1).

To summarize how the network topology changes from the interval before task onset to the interval after task onset, we compute a fundamental network measure: the density (see Methods, Section 2). For the true network, the density is 3/36=0.083 before task onset, and 6/36=0.167 after task onset. At each level of SNR, the density for the correlation networks is easily computed from the resulting binary networks in Figure II.3A (SNR=0.00, before=0, after=0; SNR=0.05, before=0.028, after=0.028; SNR=0.10, before=0.083, after=0.167; SNR=0.15, before=0.111, after=0.167). In addition to this single density value, we also determine the confidence interval for each density value (Figure II.4). To compute these confidence intervals, we use the resampling procedure described in Section 2.5. We find that, as the SNR increases, the density before and after task onset also increases. In addition, at larger SNRs (namely, SNR=0.10, and SNR=0.15), the confidence intervals for the densities of the inferred correlation networks include the known true network densities (Figure II.4). The confidence intervals at an

SNR of 0.05 do not enclose the true network density value after task onset, highlighting the fact that the estimated confidence intervals may themselves be biased, here towards zero, and are variable due to the limited sample of trials used to estimate them. In summary, this approach provides a technique to calculate not only the aggregate network measure (density), but also an associated level of confidence. The results for the canonical correlation networks are similar (in this case, the true network density is 2/3=0.66 before task onset and 1/3 = 0.33 after task onset) and are not included here.

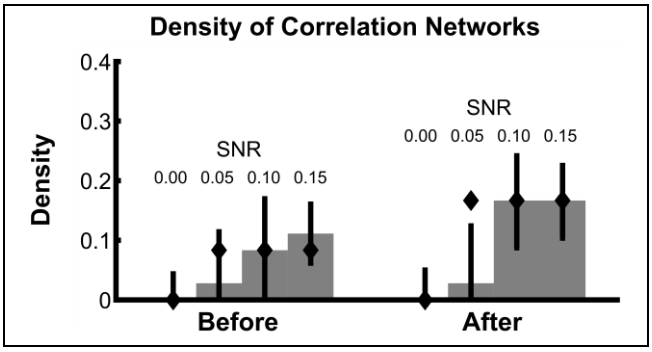


Figure II.4: Density estimates improve with increasing SNR.

Density of bootstrapped correlation networks in the before and after trial epochs are shown, for simulations with four different SNRs: 0.00, 0.05, 0.10, and 0.15. For details of the bootstrapping procedure, see Methods Section 2.5. Bar height indicates the observed density, and whiskers indicate 95% bootstrapped confidence intervals. The true values of the density in the before and after networks are indicated with diamonds.

Next we consider how the network inference procedure performs when additional coupling not related to the task exists in the data. This "constant" coupling corresponds to persistent correlations between the nodes that occur throughout the recording, i.e., both in the baseline periods and during the trials (described in Sections 2.2 and 5.1.3). The simulations varied by "correlation ratio", which we define to be the ratio of the variance

of the correlated trial component to the variance of the constant correlated component. Hence a small value of the correlation ratio represents weaker trial correlations relative to background (nuisance) correlations (see Appendix 5.1 for more detail). We apply the network inference procedure to the data, and find good performance (Figure II.5); the inferred binary networks agree with the true networks. Interestingly, spurious edges do not appear in the binary correlation networks with trial correlations equal to or lower than constant correlations (ratios 1.0 and 0.5, respectively), but do appear in the simulation in which the trial correlations were twice as strong as the constant correlations (red lines in Figure II.5A, for correlation ratio of 2.0). Computing the probability of each edge shows that the spurious edges possess lower probability of appearance, while the true network edges are highly probable (Figure II.5B). In this way, determination of the edge probability provides additional information to help identify spurious network edges. We note that, for the canonical correlation, these spurious edges are not present. Instead, the appropriate edges between the ROIs are correctly identified (Figure II.5C), with high probability (Figure II.5D). In this way the canonical correlation measure provides additional robustness to the appearance of spurious edges between individual node pairs.

Analysis of the correlation network density in the presence of constant correlations (Figure II.6) reveals that the 95% confidence intervals of the density contain the true density in all cases, even when the observed network contains false positives. These results show that even for weak values of trial-specific correlation (compared to constant correlation), the estimate of network density is accurate.

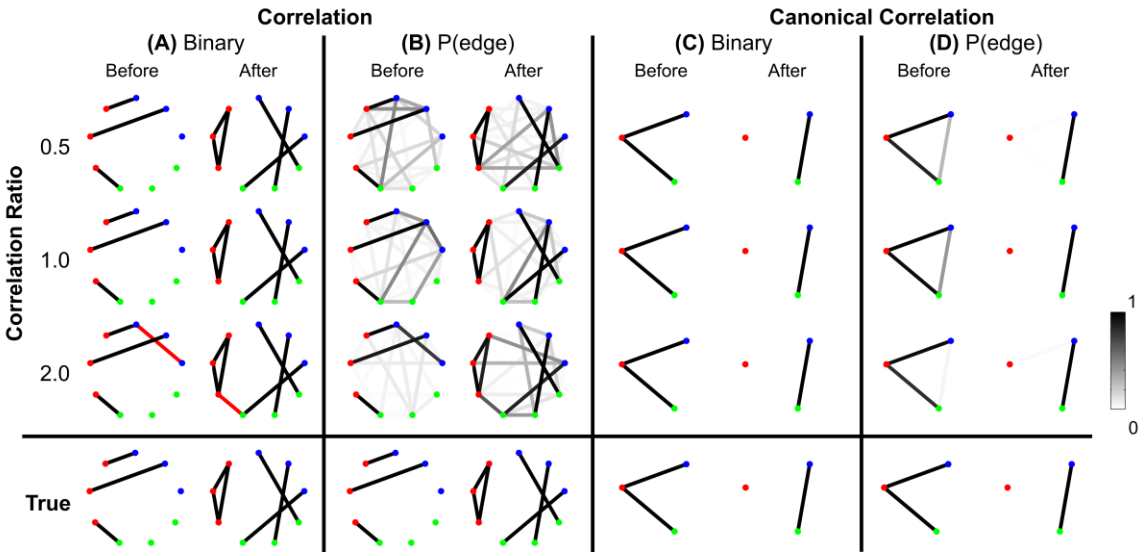


Figure II.5: True networks can be detected in the presence of background correlations.

Correlation and canonical correlation networks are shown before and after task onset, for simulations with varying correlation ratio. (A)–(D) as in Figure II.3, shown in the rows for three different levels of the ratio of trial correlation variance to constant correlation variance: 0.5, 1.0, and 2.0. All simulations have a fixed SNR of 0.11. False positives are indicated in (A) and (C) by red edges. True positives are displayed in black.

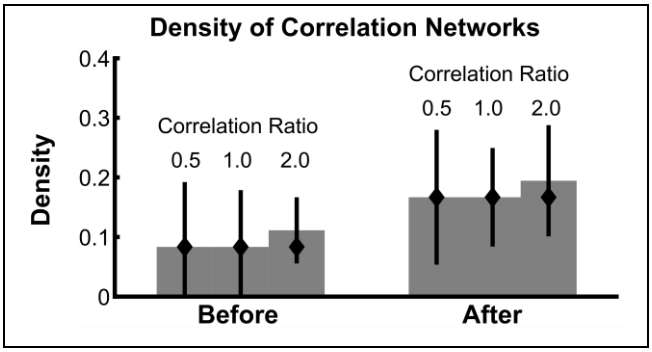


Figure II.6: Confidence intervals on network density include the true density, even in the presence of baseline correlations.

The density of bootstrapped correlation networks is plotted before and after task onset, for simulations with three different levels of the correlation ratio: 0.5, 1.0, and 2.0. Bars, confidence intervals, and true density as in Figure II.4.

We note that these simulation results correspond to a single instantiation of the synthetic time series data consisting of multiple trials. We focus only on a single instantiation to mimic the typical experimental paradigm, in which an individual subject performs the experiment once. Repeating the analysis for different instantiations (i.e., with different values of noise) we find similar qualitative results: namely, that the network inference procedure correctly identifies the true underlying network and changes in density (results not shown). What differs across these instantiations is both the number and location of false positive and false negative edges.

3.1.2 Inference of dynamic functional networks. In Section 3.1.1, we illustrated how the inferred functional networks changed from an interval immediately preceding task onset, to an interval immediately after task onset. We now further explore the applicability of the network inference procedure to dynamic (i.e., time-indexed) networks. To do so, we consider the same task-related data, but infer functional networks from overlapping windows of duration 0.2 s (0.195 s overlap) that begin 0.4 s before task onset, and end 0.4 s after task onset (networks are labeled by their midpoint, e.g., the network at time -0.4 s represents data from -0.5 s until -0.3 s). Task onset is defined to be time 0 s. Within each window, we apply the network inference procedure described above, and compute a binary network and the density with an associated confidence interval. We show the resulting dynamic correlation and canonical correlation networks for increasing SNR in Figure II.7. For comparison, we show both the dynamic density estimate with confidence intervals (sub-panels A) and example inferred functional networks (sub-panels B) for each value of SNR. In both cases, as the SNR increases, the

inferred functional networks more accurately capture the true networks. We note that, even for low SNR (=0.05), a significantly non-zero increase in density appears following task onset for the correlation networks. However, before task onset the confidence intervals rarely achieve significantly nonzero density. We note that the bootstrap density distributions tend to exceed the true density and possess long tails due to the occurrence of false positive edges (not shown). This leads to a large standard error that is relatively insensitive to SNR, and confidence intervals that include zero in this procedure. The same issue arises in the canonical correlation networks after task onset for high SNR. An alternative approach based on quantiles of the bootstrap distributions (Efron and Tibshirani, 1993) would account for this asymmetry in the sampling distribution, although these confidence intervals would be subject to bias (discussed in Section 2.5).

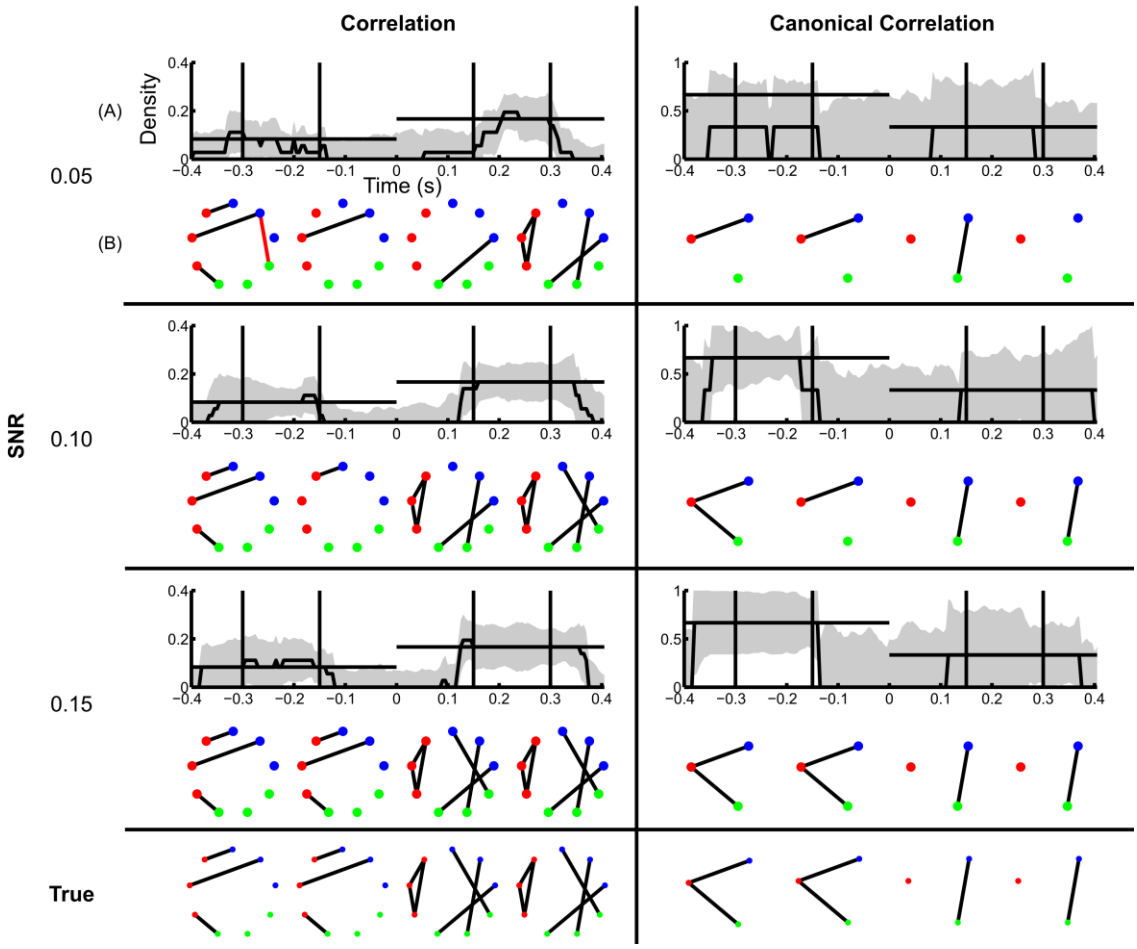


Figure II.7: Dynamic changes in network structure are better resolved at higher SNRs.

Correlation and canonical correlation networks are shown as a function of time with respect to task onset, for simulations with varying SNR. (A) Density of observed networks as a function of time (black), with 95% bootstrapped confidence intervals (shaded gray). The timing of four example binary networks (B) is indicated by black vertical lines, and the true densities are indicated by horizontal black lines. In **(B)**, network edges that exist in the true network are shown in black and false positive edges are shown in red. In all networks, nodes are color-coded according to region. (A)–(B) are shown for correlation (left) and canonical correlation (right) networks, for three different levels of SNR: 0.05 (top), 0.10 (middle), and 0.15 (bottom). True networks are shown in the bottom row.

When constant correlations are introduced, we find results similar to the SNR simulation (Figure II.8). Although spurious edges do appear in the correlation networks (see sub-panels B in Figure II.8), the inference procedure accurately identifies the true

networks both before and after task onset (examples in sub-panels B in Figure II.8). In addition, the confidence intervals for the density exclude zero, and include the known density value before and after task onset in some cases (sub-panels A in Figure II.8); for the reasons described above, the correlation networks before task onset and the canonical correlations after task onset often do not achieve significant nonzero density. In this way, the establishment of confidence intervals for the aggregate network measure (density) permits identification of dynamic changes in network structure. Surprisingly, in the presence of strong baseline correlations (ratio 0.5, when baseline correlations are twice as strong as trial correlations), the variability of the canonical correlation networks is very small after task onset: all of the bootstrapped networks had one edge, leading to an estimate of zero for the standard error. In this case, the presence of background correlations causes the baseline canonical correlations to be higher, making it more difficult for spurious correlations to appear as false positives in the network estimation. Hence, the presence of background correlations may in some cases make network inferences more robust to noise under this framework.

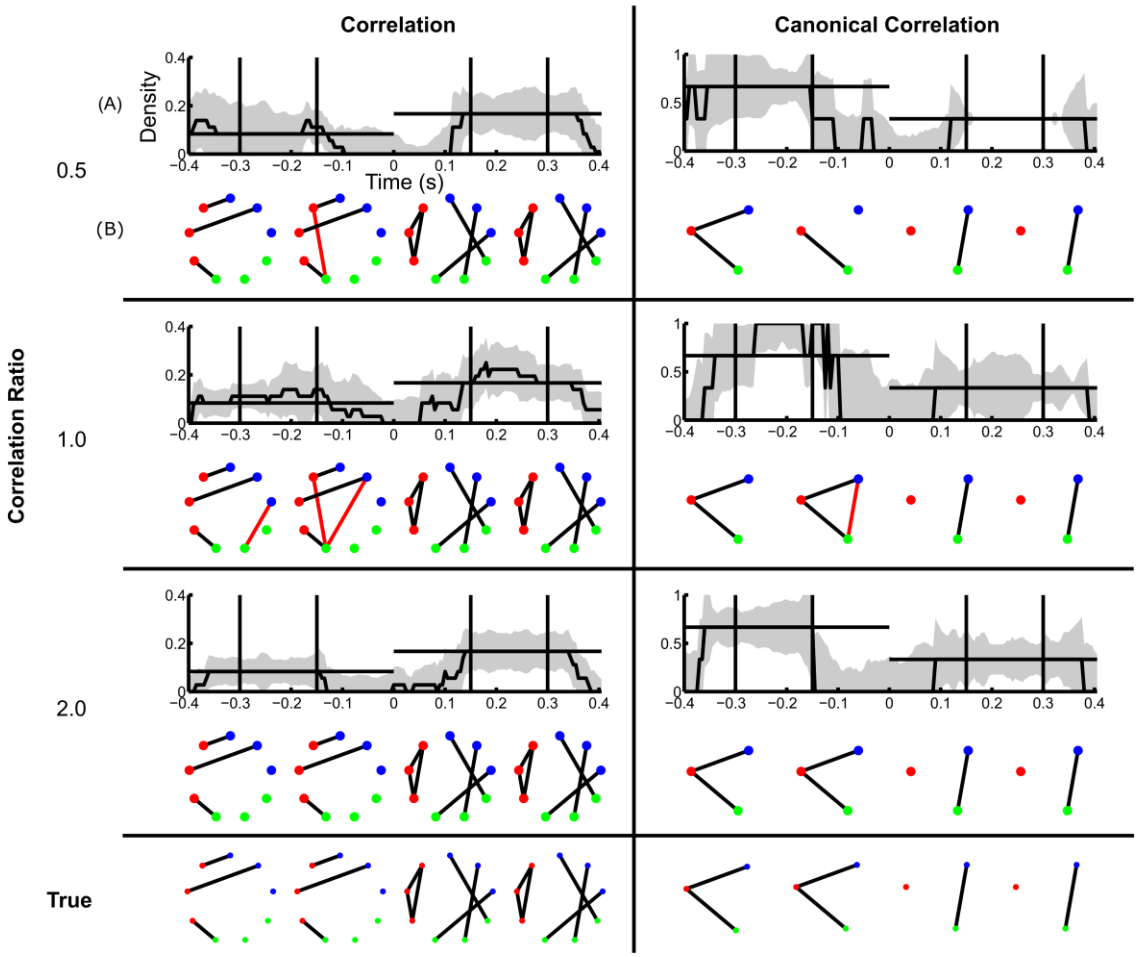


Figure II.8: Dynamic network structure can be detected in the presence of baseline correlations.

Correlation and canonical correlation networks are shown as a function of time with respect to task onset, for simulations with varying correlation ratio. (A)–(B) as in Figure II.7, shown for correlation (left) and canonical correlation (right) networks, for three different levels of correlation ratio: 0.5 (top), 1.0 (middle), and 2.0 (bottom).

3.2 ECoG Data

To demonstrate the use of the network inference procedure on real data, we analyze an ECoG dataset in which a subject read aloud from a teleprompter. Trials are defined by the 500 ms preceding and following speech onset (see Methods, Section 2.3),

which includes preparatory activity before speech onset in addition to activity related to motor execution and feedback after speech onset. Visual inspection of the binary correlation networks before and after speech onset is hindered by the large number of edges (Figure II.9A), although the network after speech onset appears more dense. This is confirmed by the density plot (Figure II.9B), in which the density after speech onset is higher than before, and significantly nonzero. In fact, the density before speech onset is also significantly nonzero, by a narrow margin. The trend of higher density after speech onset also appears in the canonical correlation networks (Figure II.9C), although neither achieves significant nonzero density (Figure II.9D). With fewer nodes and edges, the canonical correlation binary networks (Figure II.9C) are easier to interpret anatomically: for example, the region corresponding to ventral primary motor cortex (indicated with an asterisk (*)) has 2 edges before speech onset and 6 edges after speech onset, showing local coupling with premotor and primary motor cortex during motor planning that expands during speech execution to somatosensory cortex and the supramarginal gyrus. These patterns are consistent with speech processing areas identified by fMRI (Hickok and Poeppel, 2007; Price, 2012).

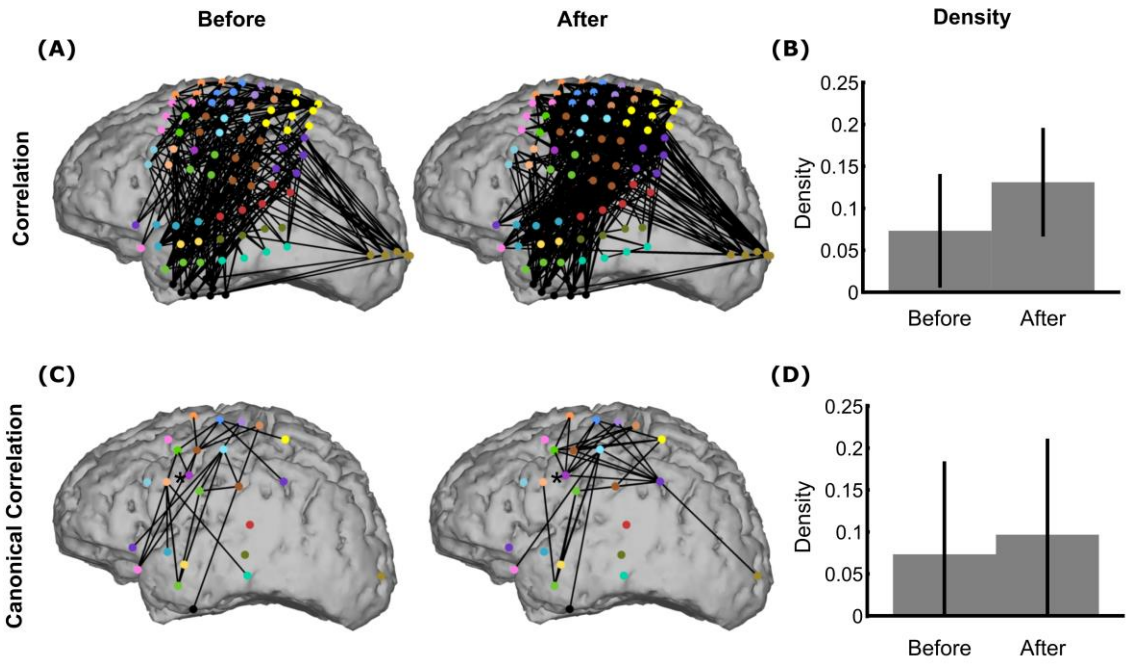


Figure II.9: The network inference procedure facilitates interpretations of real ECoG data.

Binary correlation **(A)** and canonical correlation **(C)** networks were inferred from speech data in the intervals 500 ms before (left) and 500 ms after (right) speech onset. The corresponding densities are significantly nonzero for both correlation networks **(B),** but not for the canonical correlation networks **(C)**. The node colors in **(A)** represent for each electrode the corresponding ROI, shown in **(C)**. Node locations in **(C)** are the average locations of the electrodes corresponding to the ROI. In **(C)**, the node indicated by an asterisk (*) is the ventral primary motor cortex ROI, which is discussed in the text (Section 3.2).

3.3 Advantages of the canonical correlation measure

The results in the previous section illustrate the utility of the network inference procedure for both the correlation and canonical correlation measures. We found that the canonical correlation provides additional robustness to the appearance of spurious edges between individual node pairs. We now consider two additional examples that illustrate the utility of the canonical correlation measure in functional network inference.

3.3.1 Canonical correlation improves detectability of weak inter-regional connections. One advantage of region-level analyses is that weak connectivity undetectable at the sensor level can be aggregated in such a way as to become detectable at the region level. To illustrate this idea, we constructed a simple scenario with 10 sensors comprising 2 regions of 5 sensors each (Figure II.10). During trials, every sensor from Region 1 (red nodes in Figure II.10) is weakly connected to every sensor from Region 2 (green nodes in Figure II.10): while the total SNR on each sensor is approximately 0.14, the correlated signals contributing to each true edge are weak, with SNR of about 0.03. Figure II.10 shows that, while the correlations are too weak to be detected in sensor-space networks, in the region-space networks the canonical correlation between the two regions is strong enough to reveal the true edge. There are two reasons for the improved performance of the region-space analysis. First, the weak connections between the sensors are combined in the canonical correlation calculation, increasing the effective SNR and statistical power. Second, the region-space network has many fewer edges (in this case only one), making the FDR correction for multiple comparisons much less severe and allowing higher p-values for the test statistics to be classified as edges.

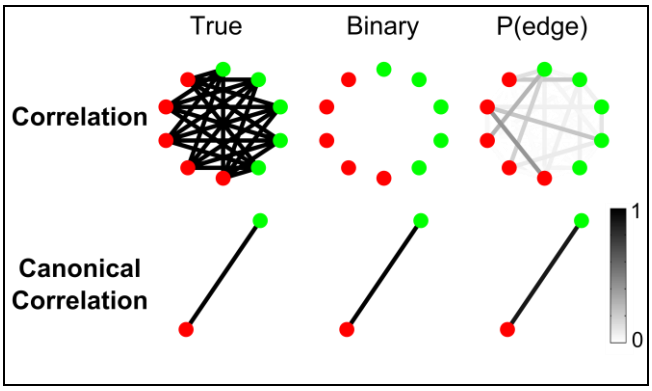


Figure II.10: Canonical correlation improves detectability of weak inter-regional connections.

While the total SNR on each electrode is about 0.14, the correlated signals contributing to each true edge (5 per node) are weak, with SNR of about 0.03. For more details about the simulation, see Section 3.3.1 ("Example 1" in Table 1). Shown are the true network (left), binary network (middle), and P(edge) (right) for correlation and canonical correlation networks. In all networks, nodes are color-coded according to region, and edges occur between regions.

3.3.2 Canonical correlation outperforms signal averaging within ROIs. In the fMRI literature, it is common to combine the signals from many image voxels into functional ROIs, each containing a large number of neurons believed to be performing a common function. In studies of speech production, for example, commonly used ROIs include the ventral premotor cortex, which contains neurons that represent syllabic motor programs, the supplementary motor area, which contains neurons that initiate the readout of speech motor programs, and the ventral primary motor cortex, which contains neurons involved in generating commands to the articulatory musculature (Tourville et al., 2008; Peeva et al., 2010; Golfinopoulos et al., 2011). ROI-based analyses can increase statistical power for activity contrasts and connectivity analyses by (1) reducing the number of statistical tests and corresponding correction for multiple comparisons, and (2) averaging many

noisy measurements together to form a single, more stable measure for each ROI (Nieto-Castanon et al., 2003).

In the case of fine time-resolution voltage recordings (e.g., scalp EEG or invasive ECoG), the fast dynamics of the signals within a region may not be phase-aligned; therefore, the averaged signals may interfere and cancel in the sum, or the SNR may be reduced. Canonical correlation does not require averaging, and can effectively detect and aggregate signals between sensors even when the sensor signals are not phase aligned. To demonstrate this, we constructed two example scenarios, each containing 9 nodes that are split into 3 regions, of which only 2 regions are connected by an edge (Figure II.11).

In the first scenario (Figure II.11A), one signal is shared by two sensors in the green region and two sensors in the blue region, but the sign of the signal is reversed on one sensor in each region such that the average signal within each region is zero. This creates a completely connected subnetwork, where the sensors within each region are perfectly negatively correlated. In this case, averaging the signals within each region cancels the correlated component, resulting in only the noise component. Hence, the correlation between the averaged region signals is zero and no edge is detected (middle row). The canonical correlation network, however, does detect the edge (bottom row), because a change of sign on the individual sensors can be compensated by the coefficients in the linear combination (vectors $\vec{a}$ and $\vec{b}$ in Section 2.4.2).

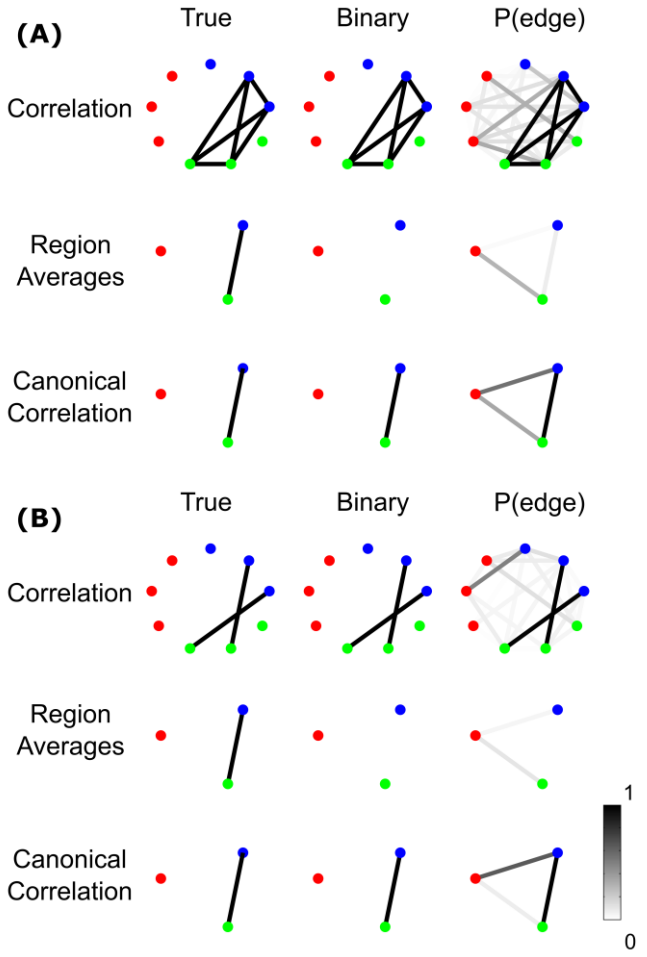


Figure II.11: Canonical correlation outperforms signal averaging within ROIs.

In (A), labeled “Example 2a” in Table 1, the signals on the electrodes within each region precisely cancel. In (B), labeled “Example 2b” in the Table 1, the signals on each edge are independent (so the signals on the electrodes do not precisely cancel in the region averages), but one edge represents a positive correlation and the other edge represents a negative correlation. For more details about the simulations, see Section 3.3.2. (A) and (B) show the true network (left), binary network (middle), and P(edge) (right) for correlation networks (top), networks resulting from averaging the signals on the electrodes within the regions and calculating the correlations between the resulting region-level signals (middle), and canonical correlation networks (bottom).

The second scenario (Figure II.11B) illustrates a related point: that averaging signals within regions assumes that the sign of the correlation will be the same for all connections between the regions. In this scenario, one sensor from the green region has a

positive correlation with one sensor from the blue region, while a second sensor in the green region has a negative correlation with a sensor from the blue region. At the sensor level, a negative correlation during trials represents an increase in the magnitude of the correlation, so the sensor-level networks detect both edges despite the fact that one reflects a positive correlation and the other reflects a negative correlation (top row). For region level networks, however, canonical correlation detects the true structure (bottom row) while averaging the signals within the region does not (middle row). This is because averaging the signals in the two regions combines the effects of the positively and negatively correlated edges, resulting in a combined signal that is neither positively nor negatively correlated. Hence the edge between the regions is not detected in the region-averaged network (middle row). Canonical correlation, however, is not sensitive to the sign of the correlations between sensors because of the freedom in the linear combination (Section 2.4.2), so it detects the edge in the region-level network (bottom row).

4. Discussion

Here we described a method to infer functional networks from time series data recorded from multiple sensors in a task-related paradigm. In doing so, we tracked the changes in dynamic network topology over time, and established measures of uncertainty for both individual edge and aggregate network measures. We also introduced an additional measure of coupling – canonical correlation – which provides a principled approach to aggregate sensor activity, and improves robust detection of network structure. We illustrated the performance of the network inference procedure applied to synthetic data in a variety of simulation scenarios, and verified its utility when used with

experimental ECoG data. In this section, we summarize the primary features of this research, mention associated limitations, and suggest avenues for future research.

4.1. Principled choices in functional network analysis

In this manuscript, we described a specific procedure for functional network inference. This procedure can be broken down into several modules that can be individually substituted with other choices without significant modification of the overall procedure. The modules include: (1) choice of trial epochs: here we used two kinds of epochs, 500 ms before and after task onset or 200 ms sliding windows over the duration of the trials; (2) selection of network nodes, either treating each sensor as a separate node or choosing groups of sensors according to, e.g., anatomical region; (3) choice of coupling measure: here we used correlation and canonical correlation; (4) choice of method for estimating the null distribution: here we used bootstrapping on baseline intervals; (5) choice of method for correcting for multiple comparisons, here FDR; and (6) choice of method for estimating uncertainty in edges and network measures: here we used bootstrapping over trials and standard error confidence intervals (see Methods, Section 2).

The methodology developed here could be extended to accommodate alternative choices. For example, the coupling measure used to define edges could be replaced by a frequency domain measure (e.g., coherence or canonical coherence) or a nonlinear coupling measure (e.g., synchronization likelihood: Stam and van Dijk, 2002; Pereda et al., 2005). In this case, the same procedures for estimating the null distribution, correcting for multiple comparisons, and estimating edge and network uncertainty can still be used.

In summary, the framework discussed here illustrates a specific implementation of a functional network inference procedure that researchers may adapt to suit their own needs.

4.2 Tracking of dynamic network topology over time

As brain activity changes dynamically to achieve specific functions (e.g., to respond to external stimulation), we expect that coupling between the activity from separate brain areas will also change. Because of this, the resulting functional networks – deduced for sensors observing these brain areas – will also change dynamically. To accurately track these dynamic changes requires that we choose analysis intervals of sufficiently short temporal duration. However, choosing a short temporal interval reduces the number of data points employed in the coupling analysis, and therefore reduces the statistical power of any coupling measure. The effect is mitigated for the task-related data considered here. In this case, the multiple repetitions of the task provide increased statistical power. In addition, the multiple trial structure permits a principled resampling procedure. Extending this approach to spontaneous data that lacks a task-related structure would require additional careful considerations, including whether resampling is appropriate.

4.3 Spatial Scale

One approach to dealing with issues of spatial scale is to group nodes into interconnected sub-groups after a full network has been constructed (Salvador et al., 2005; Ferrarini et al., 2009; Kolaczyk, 2009; Meunier et al., 2009; Rubinov and Sporns, 2010). These types of approaches can identify functional clusters in networks and were

not considered here. However, clustering on sensor-space networks could be used to define ROIs for subsequent ROI-based analyses, suggesting a multi-step procedure that utilizes fine spatial resolution networks to inform networks with coarser spatial resolution.

Note that in situations in which functional regions are known beforehand, using knowledge of ROIs can improve network inference. For example, in Section 3.3.1, we illustrated a situation in which region-level connectivity was too weak to be visible in sensor-space networks, and the greater power of region-level analysis was necessary to detect the connection.

A priori knowledge of ROIs can also be used to perform across-subject analyses in situations when it would otherwise be difficult. For example, in ECoG recordings, the electrode locations typically differ for each subject, but the spatial coverage of the electrodes often overlaps across subjects. Using the same ROIs to construct region-level networks for each subject could facilitate across-subject comparisons. In addition, the data for each ROI could be aggregated across subjects before network construction and aggregate across-subject networks could be estimated. These across-subject networks could be calculated as a function of time during a task, with the goal of detecting dynamic changes in connectivity that are not subject-specific.

The specific coupling measure used here for region-level network inference, canonical correlation, is a particularly robust tool for detecting connectivity between predefined regions. As described above, fMRI studies often average the signals of voxels within functional ROIs. We showed two examples of how averaging signals within an

ROI can mask inter-regional correlations (Section 3.3.2) that are detected using canonical correlation. In the first scenario, signals within an ROI are responsive to task onset but with opposite sign; in this case averaging results in a net signal for the ROI that is not responsive to task onset, which in turn results in weak or nonexistent correlations between the ROI and other ROIs involved in the task. This could occur in physiological data in which electrode location relative to an electrical source induces a change in sign of the recorded activity between neighboring electrodes (e.g., Wood et al., 1988). In the second example, two ROIs are both positively and negatively correlated with each other, through different pairs of electrodes. This situation could also occur in real data in which the location of neighboring electrodes may lead to a reversal in sign (Wood et al., 1988; Buzsaki et al., 2012). In this case the averaged signal for each ROI is still responsive to task onset, but averaging the signals masks the positive and negative correlations with other regions. Our simulation results demonstrate that canonical correlation successfully identifies inter-regional correlations in these situations.

One drawback of canonical correlation is that it can depend on the number of signals (electrodes) in the groups (ROIs) being compared: because more signals provide more degrees of freedom for the linear combinations involved in the calculation, canonical correlation values tend to be higher when calculated between ROIs with more component electrodes. Here, the raw canonical correlation value between two ROIs is compared to a baseline distribution calculated on the same two ROIs, so ROIs with more electrodes do not necessarily have a greater probability of edge detection. On the contrary, due to ceiling effects (canonical correlation values are bounded above by 1) it

may be difficult for ROIs with large numbers of electrodes to achieve significance in comparison of trials to baseline. In our simulations, we considered the case where all ROIs have the same number of electrodes. For the ECoG data, the size of each ROI depended on the number of electrodes that happened to lie over the anatomical regions of interest, so ROIs had varying numbers of component electrodes. While ROIs with large numbers of electrodes do have edges in the networks inferred here, it is possible that there are imbalances in the inference that are affecting the resulting networks. Further investigation into differing electrode counts for different ROIs is left for future study.

4.4 Assessment of uncertainty in network edges and aggregate network measures

An important component of the network inference procedure outlined above is the ability to assess uncertainty in network structure and aggregate network measures. To do so, we resampled the task-related data, and established both the probability of appearance for each individual edge, and confidence intervals for the network density. Although we focused on only a single aggregate measure here (the density) this procedure is easily extended to determine confidence intervals for other aggregate network measures (e.g., clustering coefficient, path length, etc). We note that the resampling procedure described here requires repeated occurrences of the time period of interest, such as occurs in a task-related structure, and may not be appropriate for data lacking this structure. In addition, the resampling procedure is computationally expensive, and for a large number of resamples would require sophisticated computational approaches.

While we focus on brain voltage data (ECoG and EEG) here, there is some precedence of the use of bootstrapping to estimate network variability in the fMRI resting

state functional connectivity literature. For example, bootstrapping has been used to test whether seed-based resting state connectivity is dependent on signal stationarity (Chang and Glover, 2010) and phase coupling (Handwerker et al., 2012). Bootstrapping has also been used to test for group differences in functional networks between schizophrenia patients and controls (Sakoglu et al., 2010) and between genders (Kilpatrick et al., 2006). Allen et al. (2012) used bootstrapping to assess the reliability of estimated network states, defined by clustering dynamic networks occurring over the course of an experimental session.

Bootstrapping has also been used in EEG functional connectivity, albeit in the context of network inference. Murias et al. (2007) used bootstrapping to compute p-values for pairs of electrodes under the null hypothesis that coherences did not differ between children with attention deficit hyperactivity disorder (ADHD) and controls. These p-values were then used to define edges in a functional network, similar to our network inference approach (Section 2.4.3).

4.5 Robustness to persistent correlations

By allowing the null distribution to be different for each potential edge, systematic spatial regularities in the correlations are mitigated. In the case of a speech task, correlations unrelated to speech may exist in the data for many reasons, including the influence of resting state networks, volume conduction, and referencing effects. As long as these regularities are present in the baseline data to the same extent as the task data, these correlations will not appear as edges in the networks since the correlations during speech must be stronger than the correlations during the baseline silent periods in order to

achieve significance. Volume conduction due to deep subcortical sources of electrical activity may also confound inference of functional networks. For the task-related data of interest here, a deep subcortical source not present during silence that emerges during the task could introduce correlations during speech that would be interpreted by the procedure as edges. These edges are spurious, in that the edge does not represent correlated cortical activity appearing locally at each sensor. While this is a valid concern, strategies exist to mitigate volume conduction effects, including re-referencing procedures in scalp EEG. Moreover, invasive recording modalities – such as the ECoG – are proposed to be relatively insensitive to distant electrical sources (Zaveri et al., 2009).

4.6 Scalability and challenges related to experimental data

One of the key difficulties in network construction is the issue of multiple comparisons: a test statistic is computed for each pair of electrodes, so a 100-electrode network (typical in ECoG recordings) represents 4,950 statistical tests. Here, we account for multiple comparisons using the FDR procedure, which controls the proportion of false positive edges relative to the total number of detected edges. This highlights connections between nodes that are extreme relative to the baseline period: the p-values for the test statistics need to be very small in order to be counted as edges. This approach puts more interpretative power on the existence of edges than their absence: if an edge exists in an inferred network it reflects a strong difference between trial and baseline coupling, but if an edge does not exist there might be a difference that was not strong enough to be classified as significant. This could cause problems in situations where the power of the statistical test is different for different pairs of electrodes, for example if the SNR varies

considerably across the set of electrodes. In this case, the variance of the test statistics may be higher in the baseline condition for pairs of electrodes with low SNR, making it more difficult for the trial period coupling to achieve low p-values compared to pairs of electrodes with high SNR. This could ultimately lead to increased false negative edges in the network for edges involving electrodes with relatively low SNR. Hence it is important to take caution in the interpretation of networks in the context of varying SNR, especially with respect to absent edges, since the FDR procedure does not control for false negatives. Note also that the FDR procedure assumes independence between the component tests (Benjamini and Hochberg, 1995). It is possible that a network structure introduces dependencies between the test statistics calculated for each potential edge. If this is the case, a corrected FDR procedure or another control for multiple comparisons may be more appropriate (Dudoit and Laan, 2008).

Additionally, we note that the uncertainty measures we describe are to be interpreted in terms of comparisons between a single epoch and a baseline period, not in terms of comparisons between epochs. Care should be taken in claiming significance in changing network structure through multiple time epochs. The fact that one epoch shows a significant difference in network structure from baseline while another epoch shows none, does not necessarily indicate a significant difference in network structure between epochs. In order to quantify the significance of changes across epochs, similar methods can be developed to those discussed here.

As high density multi-sensor data becomes increasingly common, techniques to characterize these data – and notions of uncertainty in these characterizations – become

essential. In this paper we described a general framework for the inference of functional networks from task-related, multi-sensor data. We proposed a principled approach for assessing uncertainty in network edges and aggregate network measures, as well as a technique to aggregate sensor activity and improve detection of network structure. Within this framework, we made specific choices to construct the functional networks. However, the framework is general and adaptable to choices optimized to specific recording modalities and research questions.

5. Appendix

5.1 Construction of simulated data

The simulation consisted of 600 seconds of synthetic recordings from 9 sensors at a sampling frequency of 1200 Hz. The 9 sensors were split into 3 regions, consisting of 3 sensors each. The first 400 s of the data were a "baseline period" containing no trials, followed by 100 trials lasting 1 s each, separated by one second. The total signal was a sum of four subcomponents: pink noise, white noise, constant correlations, and trial correlations (Figure II.1D). The sensor-space and region-space connectivity networks existing during the trials were the same for all simulations, consisting of two different network topologies present for the first and second halves of the trials, called the Before and After periods, respectively (Figure II.1A). The magnitudes of the different components and the strength of the correlations were manipulated in a series of scenarios designed to test the effects of the signal-to-noise ratio (SNR) and constant background correlations on the ability of our network analysis methodology (described in Section 2.4) to detect task-related changes in the correlation structure.

The simulations described below represent a subset of the possible parameter regimes that may be relevant for a given application. For this reason, the simulation and network analysis code are available online.

5.1.1 Pink Noise. Pink noise, $P(t)$, was generated for each sensor by convolving white noise with a Gaussian kernel of standard deviation 5 ms. The spectrum of the resulting signal (Figure II.1C, top) roughly approximated the decrease in power with frequency commonly observed in neuronal voltage recordings (Miller et al., 2009a; He et al., 2010). After filtering, the signal was normalized by the standard deviation, resulting in a signal with unit sample variance.

5.1.2 White Noise. The white noise component, $W(t)$, for each sensor had sample variance 0.1 μV^2.

5.1.3 Constant Correlations. Constant correlated structure was included by adding a single 2–50 Hz signal lasting the entire duration of the recording, $C(t)$, to every sensor. Because the same signal was added to all sensors, the network topology corresponding to this signal is a complete graph, in which every node is connected to every other node. In order to vary the strength of the constant correlations while holding total variance and spectral characteristics constant, an additional 2–50 Hz signal, $B(t)$, uncorrelated between sensors, was added to each sensor. Signals in the desired passband were generated by filtering white noise using a zero-phase 4th order Butterworth filter (Matlab

filtfilt function). After filtering, the signals were normalized to have unit variance and multiplied by a gain factor that varied by simulation (see Appendix 5.1.5).

5.1.4 Trial Correlations. Task-related correlation structure was introduced into each sensor of the synthetic data in the 8–25 Hz frequency range in order to introduce pre-defined network structures in the first 500 ms ("Before" task onset) and the second 500 ms ("After" task onset) of the trials (Figure II.1A). In the true Before network, edges exist between sensors 1 and 9, 2 and 8, and 3 and 4. This corresponds to a region-space network in which Region 1 (red) is connected to both Region 2 (green) and Region 3 (blue). In the true After network, the sensors in Region 1 (red nodes) are completely connected (cyan edges), and Regions 2 (green nodes) and 3 (blue nodes) have three pairs of connected sensors between them (magenta, yellow, and green edges). The region topology therefore contained only one edge, between Regions 2 and 3.

Figure II.1B shows how the trial correlations between sensors were introduced into the signal, using the After network to illustrate. The 8–25 Hz component was constructed to have constant instantaneous variance during the entire dataset, switching from uncorrelated 8–25 Hz noise outside of trials and transitioning to correlated 8–25 Hz noise during trials. To achieve this, three signals were summed at each sensor: one uncorrelated signal lasting the entire duration of the data that tapers away during the correlated portions of the trials, (Figure II.1B, middle), one correlated signal corresponding to the Before network that appears only during the first half of the trials (not shown), and one correlated signal corresponding to the After network that appears

only during the second half of the trials (Figure II.1B, left). The signal added to each sensor was chosen to introduce a fixed network topology (the colored edges in Figure II.1A, corresponding to colored traces in Figure II.1B) of the correlations during the trials that would not exist in the baseline period.

The window functions governing the transition between uncorrelated activity and correlated activity ($W_U(t)$ and $W_T(t)$, respectively; Figure II.1B, top) were chosen so that the instantaneous variance of the signal (Figure 1B, bottom) would stay constant during trials. In particular, the correlated signal was windowed by the square root of a Gaussian kernel with $\sigma = 50$ ms and $\mu = 250$ ms:

$$W_T(t) = \sqrt{\frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(t-\mu)^2}{2\sigma^2}}}$$

The uncorrelated signal was windowed by:

$$W_U(t) = \sqrt{1 - \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(t-\mu)^2}{2\sigma^2}}}$$

These choices resulted in correlated and uncorrelated signals, $T(t)$ and $U(t)$, respectively, having instantaneous variances that evolve in time according to a Gaussian profile and one minus that Gaussian profile:

$$Var(T(t)) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(t-\mu)^2}{2\sigma^2}}$$

$$Var(U(t)) = 1 - \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(t-\mu)^2}{2\sigma^2}} .$$

Hence the combined signals had unit variance (Figure II.1B, bottom). The combined signal was then multiplied by a gain factor that varied by simulation (see Section 5.1.5). In the resulting signal, the total variance of the correlated signal, $T(t)$, over the entire trial (for electrodes with at least one edge in both the before and after networks) was about one quarter of the total variance, the other three quarters of the total variance being contributed by $U(t)$.

5.1.5 Simulation scenarios. A series of simulation scenarios were constructed in order to investigate the performance of the network analysis techniques under different conditions related to the SNR and to the ratio of trial correlations to constant correlations.

The first set of simulations tested the effects of SNR on inferred trial networks in the absence of constant correlations (Table 1, SNR simulations). SNR was defined as:

$$SNR = \frac{Var(T)}{Var(U)+Var(C)+Var(B)+Var(W)+Var(P)}$$

where the variances are calculated over time during the trials (both the Before epoch and the After epoch), in which the trial correlation component $T(t)$ has nonzero variance. Note that in all simulations, the white noise component $W(t)$ had a variance of 0.1 and the pink noise component $P(t)$ had a variance of 1. In these simulations, $C(t)$ and $B(t)$ were set to zero and the gain on the 8–25 Hz component ($T(t)+U(t)$) was varied to achieve different SNR values (Table 1).

The second set of simulations tested the effects of constant background correlations on the inferred trial networks (Table 1, "Ratio" simulations). By using the

same gains for the 8–25 Hz component, $T(t)+U(t)$, and the 2–50 Hz component, $C(t)+B(t)$, for all of these simulations, the SNR was held the same. What varied in these simulations was the relative variances of $C(t)$ and $B(t)$. This resulted in simulations in which the ratio of the variance of the trial correlations $T(t)$ and constant correlations $C(t)$ varied:

$$CorrelationRatio = \frac{Var(T)}{Var(C)}$$

where, as above, the variances are calculated over time during the trials. Note that $Var(T)$ is about one quarter of the total variance of the 8–25 Hz component.

Finally, simulations were constructed for two example scenarios to illustrate different features of the analysis. These examples are described in more detail in the Results (Section 3.3.2).

Table II.1: Parameters for simulation scenarios

Here we list the settings utilized in the simulation scenarios considered in Results (Section 3). Figures illustrating the inferred functional networks for each scenario are listed in the last column.

Simulation	8–25 Hz variance	2–50 Hz variance		SNR	Correlation ratio	Results Figures
		Uncorrelated	Correlated			
SNR 1	0.00	0	0	0.00	.	Figs II.3, 4, 7
SNR 2	0.26	0	0	0.05	.	Figs II.3, 4, 7
SNR 3	0.63	0	0	0.10	.	Figs II.3, 4, 7
SNR 4	1.20	0	0	0.15	.	Figs II.3, 4, 7
Ratio 1	1.00	0	0.5	0.11	0.5	Figs II.5, 6, 8
Ratio 2	1.00	0.25	0.25	0.11	1.0	Figs II.5, 6, 8
Ratio 3	1.00	0.37	0.13	0.11	2.0	Figs II.5, 6, 8
Example 1	1.00	0	0	0.14	.	Fig II.10
Example 2A	1.00	0	0	0.14	.	Fig II.11A
Example 2B	1.00	0	0	0.14	.	Fig II.11B

5.2 Sparsity of networks and null distribution resolution effects

The smallest possible p-value in the described bootstrap procedure, $p_{(1)}$, is bounded below by the number of bootstrap samples used to calculate the null distribution, here 1/1000 for the simulated data. In a typical network, there may be many potential edges whose test statistics fall above the highest sample in their respective empirical null distributions, leading to a set of p-values $p_{(1)},..., p_{(k_{\min})}$ equal to 1/1000. If the total number of these, $k_{\min}$, satisfies

$$p_{(k_{\min})} = \frac{1}{1000} \leq \frac{qk_{\min}}{N_{MC}}$$

then all of the tests 1,…, $k_{\min}$ will be considered significant. If $p_{(k_{\min})} > qk_{\min} / N_{MC}$, all of the tests will be considered insignificant and the network will be reported as empty. This characteristic, resulting from the interaction of the FDR procedure with empirical estimation of the null distribution, imposes a lower bound on the sparsity of networks that can be detected: anything below this sparsity level will be reported as an empty network. Analytically, the smallest number of edges that can be detected is:

$$\frac{N_{MC}}{qN_{bs}}$$

where N_{bs} is the number of bootstrap samples used for the estimation of the null distributions. Note that this formula also depends on N_{MC}, the number of potential edges in the network, which scales with the square of the number of electrodes. Hence this issue becomes more serious as the size of the networks increases.

There are a variety of possible ways to deal with this issue. The first is to reduce the number of nodes in the networks (hence reducing the number of possible edges) by defining the networks on a larger spatial scale, for example by using the ROI-based analysis using canonical correlation described here. With fewer possible edges, it can become computationally tractable to use enough bootstrap samples to detect even a single edge. When reducing the number of nodes is not possible, a second option is to increase the number of bootstrap samples so that the minimum number of detectable edges is acceptably small. This is a reasonable first approach, since in many applications the networks are sufficiently dense to exceed the threshold. If the resulting inferred networks are not empty, no further analysis is needed. If the networks are empty, however, it could either be because there really is no coupling or because the number of true edges is smaller than the detectable amount. In this case, two other options are available: (1) estimate the null distributions parametrically, or (2) estimate the tails of the null distributions above the highest bootstrap sample using extreme value theory. In the case of correlations, the Fisher transform of the correlation is well approximated by a Gaussian distribution (Fisher, 1915), so the parametric approach is possible. When the sampling distribution of the test statistic is not well behaved, an approach to estimating the p-values only in the tails of the distribution based on extreme value theory may be used (Hill, 1975; Pickands, 1975; Smith, 1987).

For the experimental data used here, we increased the number of bootstrap samples in order to detect networks with very few edges. Specifically, the correlation

networks had 90 nodes and we used 20025 bootstrap samples, so the smallest number of edges that could be detected was

$$\frac{90*89/2}{0.05*20025}=4$$

For the canonical correlation networks, we had 25 regions and used 6000 bootstrap samples, so that we could detect a network with a single edge. Note that by reducing the number of nodes, canonical correlation networks require many fewer bootstrap samples.

5.3 Network inference algorithm structure including uncertainty estimation

In this section we provide a description in pseudo-code of the uncertainty estimation algorithm. Symbols are defined at the end of this section.

Network estimation with uncertainty:
- Estimate the null distribution from baseline data (the same null distribution is used for the observed network and the bootstrapped networks):
- For $j=1:N_{bs}$,
- $I_{baseline} \leftarrow$ a set of L indices chosen uniformly at random (with replacement) from $1:K$
- $Sample_{bs} \leftarrow$ the data for baseline intervals corresponding to $I_{baseline}$: $(T\times N\times L)$
- $\hat{x}_{bs}(j,:,:) \leftarrow (1\times N\times N)$ or $(1\times R\times R)$: the test statistic at each edge calculated from $Sample_{bs}$
- End
- Resample trials to compute bootstrap networks and network measures
- For $i=1:N_{us}$,
- Select a sample of trials over which to compute a surrogate network:
- $I_{us} \leftarrow$ a set of L indices chosen uniformly at random (with replacement) from $1:L$
- $SampleI \leftarrow (E\times T\times N\times L)$: the data for trials corresponding to I_{us}
- Calculate the true $\hat{x}$ for the sample:
- $\hat{x}_i \leftarrow (E\times N\times N)$ or $(E\times R\times R)$: the test statistic at each edge calculated from *SampleI* for each epoch
- **Compare** $\hat{x}_i$, $(E\times N\times N)$ or $(E\times R\times R)$, **at each edge to the empirical null distribution given by** $\hat{x}_{bs}$, $(N_{bs}\times N\times N)$ or $(N_{bs}\times R\times R)$**:**
- For epoch $e=1:E$,
- For node $n_1=1:N$ (or $1:R$)
- For node $n_2=n_1+1:N$ (or $n_1+1:R$)

- $p(e,n_1,n_2) \leftarrow \frac{\#\left(\hat{x}_{bs}(:,n_1,n_2) > \hat{x}_i(e,n_1,n_2)\right)}{N_{bs}}$, where # denotes a count of the number of elements satisfying the expression in parentheses
- If $p(e,n_1,n_2) == 0$ then $p(e,n_1,n_2) \leftarrow 1/N_{bs}$
- End
- End
- End
- Use FDR procedure to identify significant edges:
- For epoch $e = 1:E$,
- $p_{(.)} \leftarrow (1 \times N_{MC})$: Sorted p for the epoch e, for all edges
- $k \leftarrow$ the highest k such that $p_{(k)} \le \frac{qk}{N_{MC}}$
- $threshold \leftarrow p_{(k)}$
- $Net(i,e,:,:) \leftarrow (1 \times 1 \times N \times N)$ or $(1 \times 1 \times R \times R)$: $p(e,:,:) \le threshold$ (binary network consisting of all edges with a p-values less than *threshold*)
- End
- Calculate aggregate network statistics:
- For epoch $e = 1:E$,
- $Density(i,e) \leftarrow \frac{\#\ edges\ in\ Net(i,e,:,:)}{N_{MC}}$
- End
- End
- Return *Net* $(N_{us} \times E \times N \times N)$ and *Density* $(N_{us} \times E)$

To calculate the probability of an edge for each epoch and edge:
- For epoch $e = 1:E$,
- For node $n_1 = 1:N$ (or $1:R$)
- For node $n_2 = n_1 + 1:N$ (or $n_1 + 1:R$)
- $P_{edge}(e,n_1,n_2) \leftarrow \frac{\#\ edges\ in\ Net(:,e,n_1,n_2)}{N_{us}}$
- End
- End
- End

To calculate alpha-level confidence intervals on the density for each epoch:
- For epoch $e = 1:E$,
- $ObservedDensity(e) \leftarrow$ the density of the observed network (calculated using all trials)
- $StandardError(e) \leftarrow$ the standard error of $Density(:,e)$
- $LowerBound(e) \leftarrow ObservedDensity(e) - 1.96 * StandardError(e)$
- $UpperBound(e) \leftarrow ObservedDensity(e) + 1.96 * StandardError(e)$
- End

Constants:
N_{us} : the number of bootstrap samples for uncertainty estimation

E : the number of epochs
T : the number of time points in each epoch
N : the number of sensors
R : the number of regions
L : the number of trials
K : the number of baseline intervals (K>L)
N_{bs} : the number of bootstrap samples for null distribution estimation
q : The FDR threshold for the number of false positives (here, 0.05)
N_{MC} : the number of comparisons performed for each network, equal to the number of possible edges, $N*(N-1)/2$ or $R*(R-1)/2$
α : the desired type-I error rate for confidence intervals on aggregate statistic

CHAPTER III. PAIRWISE COUPLING METRICS FOR ELECTRODE-LEVEL AND REGION-OF-INTEREST LEVEL FUNCTIONAL NETWORK ANALYSIS

1. Introduction

A key problem in functional connectivity analysis is the choice of spatial scale. Theoretical models of functional coupling in the brain exist for multiple spatial scales, e.g. single-neuron, cell assembly, cortical column, and functional area. Functional connectivity analysis, however, often uses a spatial scale dictated by the recording modality. In particular, we focus here on brain voltage recordings such as local field potentials (LFPs), electroencephalography (EEG) and electrocorticography (ECoG), where functional connectivity nodes are typically taken to be the electrode locations. It may be desirable to infer region-of-interest-level functional connectivity by grouping the electrodes based on functional brain areas, but methodologies to infer networks between groups of electrodes are lacking. Simple spatial averaging of the signals in a region, a common technique in the functional magnetic resonance imaging (fMRI) literature, fails for brain voltage signals because it weakens possibly important fine timescale dynamics such as rhythmic activity.

Here we describe two coupling measures, canonical correlation and canonical coherence, that can be used to infer connectivity between groups of signals. While these metrics have a long history in the statistics literature, they are still rare in brain functional connectivity analyses. We hope to make them more accessible for use by the functional connectivity community. To that end, (1) we present canonical correlation and canonical coherence alongside their single-dimensional analogues, correlation and coherence; (2)

we review the literature justifying their use; and (3) we describe and justify two-sample hypothesis tests for differences in mean coupling between two conditions. We also include an appendix with a derivation of the canonical correlation/coherence estimators used here.

We base the discussion of canonical correlation and canonical coherence in a functional connectivity framework that was previously described in Stephen et al. (2014, Chapter II), which infers task-related connectivity with reference to a baseline period. In common with other functional connectivity methodologies (Sporns et al., 2004; Eguiluz et al., 2005; Pereda et al., 2005; Achard et al., 2006; Bassett et al., 2006; Pachou et al., 2008; Astolfi et al., 2009; Bullmore and Sporns, 2009), this framework defines edge weights based on pairwise coupling between nodes and optionally thresholds the weights.

The test statistics presented here represent three improvements relative to those used previously: first, we introduce frequency-domain coupling metrics; second, we make use of theoretical properties of the coupling statistics when appropriate; and third, we structure the tests as two-sample differences of means (Casella and Berger, 1990; Efron and Tibshirani, 1993). While we present the test statistics in the context of a particular network inference framework, they can be used generally to detect coupling differences between two conditions.

The first improvement, inferring networks based on frequency-domain metrics, introduces coherence and canonical coherence as test statistics. Investigating frequency domain coupling in electrophysiological data is motivated by the idea that the mechanisms of functional coupling in the brain may be mediated by rhythmic activity.

For example, the “communication through coherence” hypothesis proposes that the activity in two brain areas may synchronize at a particular frequency during a task in order to open an effective channel for communication between the areas (Fries, 2005). Under the hypothesis, the phase difference of the rhythms between the two areas is constant but the individual phases of each area may vary across trials, so an average across trials could obscure the effect. Furthermore, the rhythms may cohere at a non-zero phase, so the correlation between the two signals may not be positive. Hence coherence and canonical coherence, which look for consistent phase relationships across trials, are better choices to detect his type of coupling than correlation or canonical correlation. Also, given that the true mechanisms of coupling between brain areas are unknown, network inference can be improved by using several different test statistics that target different hypothesized coupling mechanisms.

The second improvement, using theoretical distributions to compute the p-values, takes advantage of the fact that estimates of correlation and coherence are statistically well-behaved (see Methods Section 2.2). When warranted, using a theoretical null distribution rather than a bootstrap estimate improves the power of the statistical test (Efron and Tibshirani, 1993).

The third improvement, using two sample tests, formulates the tests in terms of differences of means between the baseline and trial periods. For example, $\hat{x}_{Corr}$, the correlation-based coupling statistic defined between a pair of electrodes, is based on (1) the mean correlation over a set of trial intervals and (2) the mean correlation over a different set of baseline intervals, incorporating the sampling variance of both the trial

mean and the baseline mean. This is in contrast to the one-sample test used in Chapter II, in which the test statistics and hypothesis tests modeled the sampling distribution of the baseline, but did not incorporate the sampling variability of the trial coupling. Here we describe two different styles of two-sample tests: the first is parametric, using the known distributional properties of correlation and coherence; the second is non-parametric, using bootstrapping to estimate sampling distributions of canonical correlation- and canonical coherence-based test statistics. By incorporating this additional known source of variance, we reduce the number of false positives in the inferred networks.

The bulk of this chapter is the Methods, in which we present and justify the four coupling metrics and their associated hypothesis tests: correlation, coherence, canonical correlation, and canonical coherence. The Methods section also summarizes the network inference methodology, and the construction of the simulated data that was used to test the coupling metrics. In the Results section, we first describe the data transformations that lead to the inferred networks, and then we show the performance of the inference framework on the simulated data. In general, the coupling metrics are well able to detect coupling that is matched to the test statistic. Finally, we discuss the contributions of this work to the literature, and make a case for the use of simple pairwise descriptive statistics in early-stage data analysis.

2. Methods

The network inference methodology used here has a structure common to the functional connectivity literature, namely pairwise estimation of a coupling statistic, possibly followed by thresholding to yield binary networks. In particular, we base this

analysis on the structure laid out in Chapter II (which is Stephen et al., 2014): (1) preprocessing, (2) test statistic estimation, (3) p-value calculation, and (4) multiple comparisons correction[1]. This framework was designed to be modular, so that individual steps could be adapted to the requirements of particular applications. With that in mind, after a brief description of the preprocessing we will focus on the estimation of the test statistic and the p-value calculation. After describing the test statistics in detail, we summarize the network inference methodology with reference to Figure III.1. Finally, we describe the simulated data that we use to test the performance of the test statistics.

2.1 Preprocessing

Although the results presented below are based on simulated data, the preprocessing steps were designed to be similar to common preprocessing steps used in ECoG studies. Each channel was downsampled from 1200 to 400 Hz (Matlab decimate function, which in this case also low-pass filtered the signal below 200 Hz), then high-pass filtered above 2 Hz, and notch filtered at 60, 120, and 180 Hz with a 1 Hz band stop. The high-pass and notch filters used a zero-phase 3rd order Butterworth filter (Matlab filtfilt function)[2].

---

[1] Chapter II additionally develops methodology to characterize uncertainty in the resulting networks, in the form of estimates of the probability of each edge and the variance of aggregate network statistics such as the network density. The test statistics developed here are perfectly compatible with the uncertainty estimation methodology, but for brevity that layer of analysis was not included here. We also suppress the dynamic layer of inference described in that work.

[2] In Chapter II the filter is between 0.1 Hz and 30 Hz. We have increased the lower bound of the filter, since we calculate correlations on epochs of 500 ms. Any activity in the signal below 2 Hz could lead to spurious correlation values (Leonardi and Van De Ville, 2015). Filtering above $1/\mathit{windowlength}$ is also necessary for canonical correlation estimation. While not strictly necessary for the coherence and canonical coherence estimates, we use the same filter for all analyses for the sake of consistency.

The simulated data contained trials lasting 500 ms, and a set of non-overlapping "baseline" intervals lasting 500 ms (see Section 2.4)[3]. Trials were stored in a 3-dimensional matrix with dimensions ($T \times N \times L$), where T is the number of time points, N is the number of electrodes, and L is the number of trials. Baseline intervals were stored in a 3-dimensional matrix with dimensions ($T \times N \times K$), where K is the number of baseline intervals. In the simulations described below, there were 200 time points (500 ms sampled at 400 Hz), 9 electrodes, 100 trials, and 400 baseline intervals.

Both sets of intervals were normalized: first the event-related potential (ERP), or mean over trials, was computed separately for each channel and subtracted. Then the DC component, or mean over time, was computed separately for each interval and subtracted. Removing the ERP reduces (but may not completely remove (Truccolo et al., 2002)) the component of the response that is time-locked to the stimulus, which may induce correlations unrelated to intrinsic network effects (Friston, 1994). The ERP is also a source of nonstationarity, so it is important to remove it before performing spectral inference, such as coherence estimation (Koopmans, 1995). Of course, the ERP often contains relevant task information and can be studied separately. Removing the DC component eliminates 0 Hz artifacts in the correlation and coherence estimates (this was a cautionary step, since the signal was already high-pass filtered above 2 Hz). Both of these normalization steps are consistent with the common practice in stochastic process analysis to treat trend terms separately from residuals, which are then assumed to have zero mean (Thomson, 1982; Koopmans, 1995).

[3] In the language of Chapter II, we had only one "epoch of interest" lasting the entire duration of the trials, since here we do not consider dynamic changes in functional connectivity

2.2 Test statistic estimation and p-value calculation

The test statistics used in Stephen et al. (2014) were the absolute value of the correlation coefficient and the canonical correlation estimated during the task. The p-values were computed by comparing the observed test statistic to a null distribution estimated from the baseline period using bootstrapping. Here we propose three improvements to this approach: (1) defining coupling in the frequency domain, (2) defining the hypothesis test theoretically rather than empirically where possible, and (3) incorporating known sources of variability from the experimental setup.

In this section we describe the four test statistics that incorporate these improvements. The test statistics draw from a long history of work on the distributional properties of pairwise coupling measures, which we summarize for each statistic.

Table III.1: Test Statistics.

Summary of the four coupling statistics used to infer networks.

Measurement	Test Statistic $\hat{x}$	Estimated Variance $\hat{\sigma}$	Distribution
Correlation	$x_{Corr} = \tanh^{-1}\left(Corr_T\right) - \tanh^{-1}\left(Corr_B\right)$	$\sigma_{Corr} = \sqrt{\hat{\sigma}_T^2 + \hat{\sigma}_B^2}$	Gaussian
Canonical Correlation	$\hat{x}_{CCorr} = \tanh^{-1}\left(CCorr_T\right) - \tanh^{-1}\left(CCorr_B\right)$	N/A	Bootstrap over trials
Coherence	$x_{Coh}(f) = \tanh^{-1}\left(\left\|Coh_T(f)\right\|\right) - \frac{1}{2LP-2} - \left(\tanh^{-1}\left(\left\|Coh_B(f)\right\|\right) - \frac{1}{2KP-2}\right)$	$\sigma_{Coh}(f) = \sqrt{\hat{\varsigma}_T^2(f) + \hat{\varsigma}_B^2(f)}$	Gaussian
Canonical Coherence	$\hat{x}_{CCoh}(f) = \tanh^{-1}\left(CCoh_T(f)\right) - \tanh^{-1}\left(CCoh_B(f)\right)$	N/A	Bootstrap over trials

2.2.1 Correlation. The sample correlation for the trial period is calculated as:

$$Corr_T(E_i, E_j) = \frac{\sum_{l=1}^{L}\sum_{t=0}^{T} E_i(\theta_l + t)E_j(\theta_l + t)}{\sqrt{\left(\sum_{l=1}^{L}\sum_{t=0}^{T} E_i(\theta_l + t)^2\right)\left(\sum_{l=1}^{L}\sum_{t=0}^{T} E_j(\theta_l + t)^2\right)}}$$

where $E_i(t)$ and $E_j(t)$ are the signals on electrodes i and j, T is the duration of the intervals, L is the number of trial intervals, and θ_l are the interval onset times. Note that all correlations were computed at zero lag. The sample correlation for the baseline period $Corr_B(E_i, E_j)$ is defined the same way, replacing L with K, the number of baseline intervals. Below, we will fix the choice of electrodes E_i and E_j for simplicity.

The Fisher transform (inverse hyperbolic tangent) of the sample correlation is approximately normally distributed (Fisher, 1915):

$$\tanh^{-1}\left(Corr.\right) \underset{approx}{\sim} N\left(\tanh^{-1}(\rho), \frac{1}{\sqrt{n-3}}\right),$$

where ρ is the true correlation and n is the degrees of freedom in the estimate. Under the null hypothesis that the true correlation is the same during trial and baseline, the statistic

$$x_{Corr} = \tanh^{-1}\left(Corr_T\right) - \tanh^{-1}\left(Corr_B\right)$$

will have mean zero and standard deviation:

$$\sigma_{Corr} = \sqrt{\sigma_T^2 + \sigma_B^2}$$

where σ_T^2 and σ_B^2 are the variances of the trial and baseline periods, respectively. Estimating σ_{Corr} with $\hat{\sigma}_{Corr}$, this leads naturally to a z-statistic:

$$\hat{z}_{Corr} = \frac{x_{Corr}}{\hat{\sigma}_{Corr}}$$

which will be distributed as a standard normal under the null hypothesis.

The estimate of $\hat{\sigma}_{Corr}$ is nontrivial: theoretically, the component variances are $\sigma_T^2 = \frac{1}{d_T - 3}$ and $\sigma_B^2 = \frac{1}{d_B - 3}$, where d_T and d_B are the degrees of freedom for $Corr_T$ and $Corr_B$, respectively. However, the correct choice of degrees of freedom d_T and d_B is unclear. We can assume that the intervals are independent, but each interval consists of multiple timepoints that contribute to the estimate of the correlation. Because the recordings at close timepoints in an interval have some dependence, it is unclear how many degrees of freedom each interval contributes to the correlation estimate. For a white noise Gaussian process the degrees of freedom would be equal to the number of time points, but an autoregressive process with no noise would only contribute one degree of freedom per interval. Our case falls somewhere between these two extremes. Hence the variance of $\hat{z}_{corr}$ is unknown. To solve this, we use a jackknife procedure (Quenouille, 1956; Tukey, 1958) in the spirit of (Bokil et al., 2007) to estimate the variances $\hat{\sigma}_T^2$ and $\hat{\sigma}_B^2$. They use a two-sample extension of the jackknife procedure (Arvesen, 1969) to estimate the variance of the observed test statistic.

The jackknife estimate of variance computes estimates of the test statistic after leaving out one interval of trial or baseline; i.e. x_{i0} is the estimate of x_{corr} with the i^{th} trial interval left out, and x_{0j} is the estimate of x_{corr} with the j^{th} baseline interval left out. Then the variance estimate is given by:

$$\hat{\sigma}^2_{Corr} = \hat{\sigma}^2_T + \hat{\sigma}^2_B = \frac{1}{L(L-1)} \sum_{i=1}^{L} \left[\Delta x_{i0} - \overline{\Delta x_{i0}} \right]^2 + \frac{1}{K(K-1)} \sum_{j=1}^{K} \left[\Delta x_{0j} - \overline{\Delta x_{0j}} \right]^2$$

where

$$\begin{aligned} \Delta x_{i0} &= L x_{corr} - (L-1) x_{i0} \\ \Delta x_{0j} &= K x_{corr} - (K-1) x_{0j} \\ \overline{\Delta x_{i0}} &= \frac{1}{L} \sum_{i=1}^{L} \Delta x_{i0} \\ \overline{\Delta x_{0j}} &= \frac{1}{K} \sum_{j=1}^{K} \Delta x_{0j} \end{aligned} \quad .$$

Hinkley (1978) demonstrates that the jackknife estimate is asymptotically valid as an estimate of the sample error of the (non-jackknife) test statistic, using the sample correlation as a justifying example.

Since $\hat{z}_{Corr}$ has a standard normal distribution under the null hypothesis, the p-values are calculated by comparison to the standard normal cumulative distribution function (CDF) Φ:

$$p = 1 - \Phi(\hat{z})$$

Note that the calculation is set up so that high values of $\hat{z}_{Corr}$ have small p-values, meaning that the correlation during trials is greater than the correlation during silence.

We use one-sided tests throughout for simplicity, however it is a straightforward extension to construct p-values for two-sided tests.

2.2.2 Coherence. In this section, we refer to the sample coherency as:

$$Coh_T(E_i, E_j, f) = \frac{S_{ij}(f)}{\sqrt{S_{ii}(f)S_{jj}(f)}}$$

Where $\hat{S}_{ii}$, $\hat{S}_{jj}$, and $\hat{S}_{ij}$ are the sample spectra and cross-spectrum for the signals on electrodes E_i and E_j, averaged over trial intervals (Brillinger, 2001). Here, the sample spectra and cross-spectra are estimated using the multitaper method. For example, the cross-spectrum is given by:

$$S_{ij}(f) = \frac{1}{LP}\sum_{n=1}^{L}\sum_{k=1}^{P} X_{n,k}(f)Y_{n.k}^{*}(f),$$

where L is the number of trial intervals to be averaged (replace by K for the baseline coherency), P is the number of tapers (here $P = 3$), * denotes the complex conjugate, and $X_{n,k}(f)$ and $Y_{n,k}(f)$ are the Fourier transform at frequency f of the n^{th} time-domain interval of signals E_i and E_j, respectively, tapered by taper k. The spectra are computed similarly. Note that coherency is a complex number with both an amplitude and a phase. The sample coherence is the magnitude squared coherency, $\left|Coh_T(f)\right|^2$.

While the analytical distribution of sample coherence is known, it depends heavily on the value of the true coherence. Also, coherence has a positive bias, since it is constrained to be strictly positive. This bias depends on the sample size and the value of

the true coherence (values close to 0 or 1 require more data to resolve). So, while sample coherence is asymptotically Gaussian (Enochson and Goodman, 1965; Brillinger, 2001), the convergence to Gaussian is slow and the approximation is not good for sample sizes less than about 150 (Enochson and Goodman, 1965).

A natural approach is to try to stabilize the variance of the estimates along the same lines as correlation by applying the Fisher Transform to the magnitude of the coherency (Enochson and Goodman, 1965; Thomson and Chave, 1991; Brillinger, 2001; Bokil et al., 2007). This works well for sample sizes greater than 20 and $0.4<=\left|Coh(f)\right|^2<=0.95$ (Koopmans, 1995). Enochson and Goodman (1965) also point out that for non-extreme values of the true coherence the bias in the estimates can be approximated by $1/(\#\text{ degrees of freedom}-2)$.

Assuming that the signals are weak-sense stationary, the degrees of freedom are equal to the number of trials times the number of tapers. Hence:

$$\tanh^{-1}\left(\left|Coh_T(f)\right|\right) \underset{approx}{\sim} N\left(\tanh^{-1}\left(\left|Coh_T(f)\right|\right)+\frac{1}{2LP-2},\frac{1}{2LP-2}\right)$$

where $Coh_T(f)$ is the true trial coherency. For sample sizes (LP) between 20 and 150, this approximation holds for values of the true coherence between about 0.4 and 0.95 (Enochson and Goodman, 1965). Bokil et al. (2007) use this result to construct a test statistic relating to the difference between two coherences that we will use to compare coherence during speech to coherence during silence:

$$\hat{z}_{Coh}(f)=\frac{x_{Coh}(f)}{\hat{\sigma}_{Coh}(f)}$$

where

$$x_{coh}(f) = \tanh^{-1}\left(\left|Coh_T(f)\right|\right) - \frac{1}{2LP-2} - \left(\tanh^{-1}\left(\left|Coh_B(f)\right|\right) - \frac{1}{2KP-2}\right)$$

$$\sigma_{coh}(f) = \sqrt{\hat{\varsigma}_T^2(f) + \hat{\varsigma}_B^2(f)}$$

Setting $\hat{\varsigma}_T^2 = \frac{1}{2LP-2}$ and $\hat{\varsigma}_B^2 = \frac{1}{2KP-2}$, $\hat{z}_{coh}$ will be approximately standard normally distributed under the null hypothesis that the coherence is the same during speech and silence. While this distribution holds exactly in the limit as $LP, KP \to \infty$, we follow Bokil et al (2007) in accounting for some small-sample variability by estimating the variances $\hat{\varsigma}_T^2$ and $\hat{\varsigma}_B^2$ using a jackknife procedure as described above (see Maris and Oostenveld (2007) for some criticism of this approach). Under the null hypothesis, $\hat{z}_{Coh}(f)$ has a standard normal distribution, so p-values are computed in the same way as for $\hat{z}_{Corr}$.

2.2.3 Canonical Correlation. Canonical correlation is a measure of coupling between two groups of signals $\vec{x} = (x_1, ..., x_p)^T$ and $\vec{y} = (y_1, ..., y_q)^T$, where there are p signals in group $\vec{x}$ and q signals in group $\vec{y}$. It is defined using the linear combinations of $\vec{x}$ and $\vec{y}$ that are maximally correlated:

$$CCorr(\vec{x}, \vec{y}) = \max_{\vec{a},\vec{b}} Corr\left(\vec{a}^T\vec{x}, \vec{b}^T\vec{y}\right)$$

Here, $\vec{x}(t)$ and $\vec{y}(t)$ are the time series on the electrodes in two regions of interest during the time intervals of interest (baseline or task). There are p electrodes in the first region

and q electrodes in the second: in the simulated data, all regions have $p = q = 6$ signals. Canonical correlation determines the linear combination of the signals in the first region and the linear combination of signals in the second region that maximize the correlation at zero lag. The estimate of the canonical correlation $CCorr(\vec{x}, \vec{y})$ is based on the singular value decomposition of the (standardized) covariance matrix between the two groups of signals (see Appendix 5.2). From a theoretical perspective, it is closely related to Principal Component Analysis (PCA), which is based on the covariance matrix of a set of signals with itself. PCA seeks the linear combination of the variables in one group that maximizes variance, and canonical correlation analysis seeks linear combinations of two groups that maximize the correlation. A derivation of the estimate used here and an algorithm for computing it can be found in the Appendix (Section 5.2).

Estimates of canonical correlation have a positive bias that depends on the number of signals in each group (p and q), the number of observations (L task intervals and K baseline intervals), and the population canonical correlation (Lawley, 1959; Thompson, 1990). The bias has been described using functional approximations (Lawley, 1959) and Monte-Carlo simulations (Thompson, 1990), but since it has a complicated dependence on the population value of the canonical correlation it cannot be computed analytically in practice. Several approaches to correcting for the bias have been proposed such as using the Wherry correction (Thompson, 1990) or a modified Jackknife procedure (Dempster, 1966). Here, we instead construct a test statistic in such a way that the bias in the baseline and test intervals will cancel out under the null hypothesis.

In parallel with the test statistics defined above for correlation and coherence, we would like to construct a two-sample test statistic of the form:

$$\hat{x}_{CCorr} = \tanh^{-1}\left(CCorr_T\right) - \tanh^{-1}\left(CCorr_B\right)$$

As with correlation and coherence, the Fisher Transform stabilizes the variance of canonical correlation, so we use it to improve our estimates (Brillinger, 1981). This statistic can be biased:

$$E[\hat{x}_{CCorr}] = CCorr_T + bias(CCorr_T, L, p, q) - \left(CCorr_B + bias(CCorr_B, K, p, q)\right)$$

where $CCorr_T$ and $CCorr_B$ are the population values of the canonical correlation in the two conditions. While the functional form of $bias(\cdot)$ is complicated, we can control its inputs. In our application, p and q are the same in the baseline and task conditions, since the region definitions are assumed to be constant over time. Under the null hypothesis that the population canonical correlation is the same in the two conditions ($CCorr_T = CCorr_B \equiv CCorr$), the expected value of the test statistic will be

$$E[\hat{x}_{CCorr}] = bias(CCorr, L, p, q) - bias(CCorr, K, p, q)$$

Hence, if the number of observations is the same in the baseline and task intervals ($K = L$), the test statistic will have a zero mean. Since we would like to be able to interpret a value of zero as being no difference between $CCorr_T$ and $CCorr_B$, we need to calculate the test statistic using the same number of intervals in the baseline and task conditions.

In order to keep the number of intervals the same in the two conditions, we could throw away excess intervals from whichever condition has a greater number. We will

show below that, using bootstrapping, it is possible without throwing away any data to both (1) estimate the difference in canonical correlation between the two conditions and (2) compute a p-value against the null hypothesis that the difference is zero.

The null hypothesis of interest is that the canonical correlation is the same in the baseline and task intervals. Under the null hypothesis, the test statistic will be zero, and the p-value for a one-sided test against the alternative hypothesis that the canonical correlation is greater during the task than baseline will estimate the probability that the test statistic is less than or equal to zero. To estimate this probability, we construct a two-sample bootstrap procedure (Efron and Tibshirani, 1993) with $N_{BS} = 1000$ iterations. To get the bias to cancel out under the null as described above, we subsample the same number of baseline and task intervals to estimate the test statistic. The algorithm is as follows:

1. Let $M = \min(L, K)$
2. For $i = 1 : N_{BS}$
    a. If $L > M$, randomly subsample M task intervals (without replacement)
    b. If $K > M$, randomly subsample M baseline intervals (without replacement)
    c. $S_{baseline}$ ← a set of M intervals chosen uniformly at random with replacement from the (possibly subsampled) baseline intervals: $(T \times N \times M)$
    d. S_{task} ← a set of M intervals chosen uniformly at random with replacement from the (possibly subsampled) task intervals: $(T \times N \times M)$
    e. For $r_1 = 1 : R$
        i. For $r_2 = 1 : R$
            1. $\underline{\underline{X}}_{baseline}$, $\underline{\underline{X}}_{task}$ ← the data for region r_1 from $S_{baseline}$ and S_{task}, respectively: $(T \times p \times M)$
            2. $\underline{\underline{Y}}_{baseline}$, $\underline{\underline{Y}}_{task}$ ← the data for region r_2 from $S_{baseline}$ and S_{task}, respectively: $(T \times p \times M)$
            3. $\hat{x}_{bs}(i, r_1, r_2)$ ← $\tanh^{-1}\left(CCorr(\underline{\underline{X}}_{task}, \underline{\underline{Y}}_{task})\right) - \tanh^{-1}\left(CCorr(\underline{\underline{X}}_{baseline}, \underline{\underline{Y}}_{baseline})\right)$
        ii. End
    f. End

3. End

Even though the individual samples of the bootstrap distribution use only a subset of the baseline intervals, all of the baseline intervals are represented in the complete distribution.

The resulting bootstrap distribution is an estimate of the sampling distribution of the test statistic (Efron and Tibshirani, 1993). The estimate of the sampling distribution is often used to construct confidence intervals on the test statistic. Here, we use it in two ways: first, to get a robust estimate of the value of the test statistic, and second, to compute a p-value. The mean of the bootstrap distribution is a robust estimate of the value of the test statistic:

$$\bar{x}_{CCorr} = \frac{1}{N_{BS}} \sum \hat{x}_{BS}$$

Note that this estimate will be conservative, because the positive bias in the canonical correlation is stronger for smaller values of the true canonical correlation (Lawley, 1959; Thompson, 1990). We use this estimate for the weighted canonical correlation networks.

To compute a p-value for the hypothesis test with the null hypothesis that the test statistic is less than or equal to zero (the canonical correlation is not greater in the task intervals than the baseline), we use the proportion of the bootstrap samples that fall below zero:

$$p = \frac{\#(\hat{x}_{BS} < 0)}{N_{BS}}$$

Formally, this procedure relies on the duality between confidence intervals and hypothesis tests: the above formula is equivalent to an inversion of the approximate

confidence intervals represented by the percentiles of the bootstrap distribution. Note that these p-values are approximate: they are based on percentile confidence intervals, which are accurate to first order in the number of samples. Basing the p-values on more precise approximate confidence intervals such as BCa intervals (Bias Corrected and Accelerated intervals), which are accurate to second order, would further improve their accuracy (Efron and Tibshirani, 1993, p. 216).

2.2.4 Canonical Coherence. The above argument regarding bias in the estimation of canonical correlation applies equally to the estimation of canonical coherence (James, 1964; Brillinger, 1981), so we use the same method to match the sample sizes in the bootstrap estimation. The test statistic is defined as:

$$\hat{x}_{CCoh}(f) = \tanh^{-1}\left(CCoh_T(f)\right) - \tanh^{-1}\left(CCoh_B(f)\right)$$

The canonical coherences $CCoh_T(f)$ and $CCoh_B(f)$ are estimated as described in the Appendix, and the same bootstrap procedure is used to estimate $\bar{x}_{CCoh}$ and p-values for each region pair.

2.3 Network Inference

The network inference methodology is summarized in Figure III.1. For electrode-level networks, $\hat{x}$ is calculated as described above based on the correlation or coherence between the two electrodes of interest E_i and E_j in the baseline and trial intervals. A jackknife procedure is used to estimate the variance of $\hat{x}$, which is then used to calculate $\hat{z}$. Weighted networks are constructed by defining edge weights to be the value of $\hat{z}$ on

each pair of electrodes. To construct binary networks, a p-value is calculated by comparing $\hat{z}$ to the standard normal CDF Φ, and the False Discovery Rate procedure (FDR, Benjamini and Hochberg, 1995) is used across edges to find a threshold. We use a FDR level of 5% for all tests.

For region-level networks, $\hat{x}$ is calculated based on the canonical correlation or canonical coherence between the two regions of interest R_i and R_j in the baseline and trial intervals. A bootstrap procedure is used to construct an estimate of the sampling distribution of $\hat{x}$. The mean of the distribution, $\bar{x}$, is used for the weighed networks. The proportion of the distribution below zero gives the p-values. The FDR procedure across edges is used to threshold the p-values, resulting in binary networks.

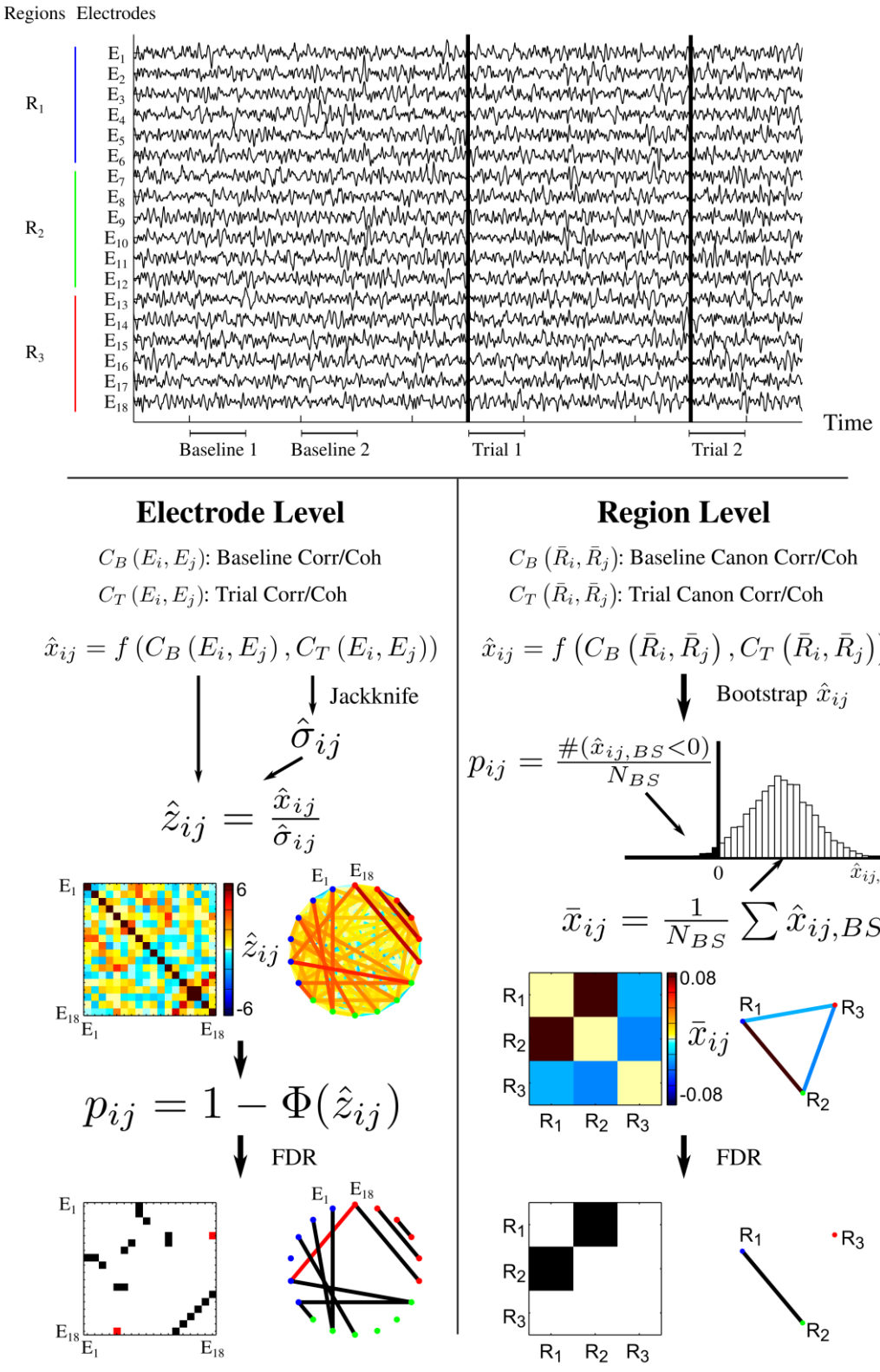

Regions Electrodes
R1
R2
R3
E1 E2 E3 E4 E5 E6 E7 E8 E9 E10 E11 E12 E13 E14 E15 E16 E17 E18
Time
Baseline 1
Baseline 2
Trial 1
Trial 2
Electrode Level
$C_B(E_i, E_j)$: Baseline Corr/Coh
$C_T(E_i, E_j)$: Trial Corr/Coh
$\hat{x}_{ij} = f(C_B(E_i, E_j), C_T(E_i, E_j))$
Jackknife
$\hat{\sigma}_{ij}$
$\hat{z}_{ij} = \frac{\hat{x}_{ij}}{\hat{\sigma}_{ij}}$
6
-6
$\hat{z}_{ij}$
$p_{ij} = 1 - \Phi(\hat{z}_{ij})$
FDR
Region Level
$C_B(\bar{R}_i, \bar{R}_j)$: Baseline Canon Corr/Coh
$C_T(\bar{R}_i, \bar{R}_j)$: Trial Canon Corr/Coh
$\hat{x}_{ij} = f(C_B(\bar{R}_i, \bar{R}_j), C_T(\bar{R}_i, \bar{R}_j))$
Bootstrap $\hat{x}_{ij}$
$p_{ij} = \frac{\#(\hat{x}_{ij,BS} < 0)}{N_{BS}}$
0
$\hat{x}_{ij,BS}$
$\bar{x}_{ij} = \frac{1}{N_{BS}} \sum \hat{x}_{ij,BS}$
0.08
-0.08
$\bar{x}_{ij}$
FDR

Figure III.1: Network inference methodology.

(Top) Simulated voltage traces for the 18 electrodes ($E_1 - E_{18}$) in the time domain coupling simulation. The 18 electrodes are grouped into three regions ($R_1 - R_3$), indicated by the blue, green, and red lines. The simulated data contained trial and baseline intervals lasting 0.5s. (Bottom Left) The construction of electrode-level networks, based on the correlation or coherence (at a frequency of interest) during the baseline and trial intervals. First, the correlations or coherences are used to estimate $\hat{x}$ for each pair of electrodes E_i and E_j. A jackknife procedure is used to estimate the standard deviation of $\hat{x}$, $\hat{\sigma}$. These two components are used to construct the normalized test statistic, $\hat{z}$, which forms the basis of the weighted networks. P-values p are calculated through comparison with the standard normal CDF Φ, and are then thresholded using the FDR procedure to create binary networks. (Bottom Right) The construction of region-level networks, based on the canonical correlation or canonical coherence (at a frequency of interest) during the baseline and trial intervals. The canonical correlations or canonical coherences are used to estimate a bootstrap distribution of $\hat{x}$ for each pair of regions R_i and R_2. The mean of the bootstrap distribution, $\bar{x}$, is used as the basis of the weighted networks. P-values p are calculated as the proportion of the bootstrap samples that fall below zero. Finally, the p-values are thresholded using the FDR procedure to create binary networks. See Methods for more detail.

The FDR procedure sorts the p-values in ascending order and chooses the largest p-value that satisfies:

$$p_e \leq \frac{e}{N(N-1)/2}\alpha$$

where e indexes the sorted edges (the smallest p-value having an e of 1), α is the level of the FDR, and N is the number of electrodes (replace with the number of regions for canonical correlation and canonical coherence networks). The denominator, $N(N-1)/2$, is the number of edges before thresholding, which is the number of hypothesis tests being performed. The procedure then rejects the null hypothesis for all tests having a p-value less than or equal to the chosen p-value.

Note that while all of the above hypothesis tests are formulated as one-sided tests for a positive difference between task and baseline (an increase in the coupling measure

during the task relative to baseline), it is a straightforward extension to use two-sided tests to look for either positive or negative effects.

For coherence and canonical coherence networks, all of the steps are repeated for each frequency of interest.

2.4 Simulations

Simulated data were generated to test the performance of the network inference methodology in the presence of time domain and frequency domain coupling. The basic structure of the simulated data was the same as described in Chapter II, with the signal on each electrode consisting of the sum of four components: pink noise, white noise, a background coupled signal, and trial coupled signal. The pink noise, white noise, and background coupling were the same for the time domain and frequency domain scenarios, but the trial coupling differed.

2.4.1 Pink Noise. The pink noise was generated for each electrode by convolving white noise with a Gaussian kernel of standard deviation 5 ms. After filtering, the signals were normalized to have unit variance.

2.4.2 White Noise. The white noise component had sample variance 0.1.

2.4.3 Background coupling. Constant correlated structure was included by adding a single 2–50 Hz signal lasting the entire duration of the recording, $C(t)$, to every electrode. Because the same signal was added to all electrodes, the network topology

corresponding to this signal is a complete graph, in which every node is connected to every other node. Signals in the desired passband were generated by filtering white noise using a zero-phase 4th order Butterworth filter (Matlab filtfilt function). After filtering, the signals were normalized to have a variance of 0.5. Note that unlike in Chapter II, there was no uncorrelated component at 2–50 Hz.

2.4.4 Time domain trial coupling. The simulations with time-domain coupling consisted of 18 electrodes, split in to 3 regions. In addition to the pink noise, white noise, and background coupling, correlated structure in the 8–25 Hz frequency range was introduced based on a pre-defined network structure (Figures III.4 and III.5, Panel A, True Networks). The 8–25 Hz component was constructed to have constant instantaneous variance during the entire dataset, switching from uncorrelated 8–25 Hz noise outside of trials and transitioning to correlated 8–25 Hz noise during trials. That is, the total variance of the stochastic process on each electrode was constant as a function of time, while the covariance structure between electrodes changed during trials to reflect the desired network structure. To achieve this, two signals were summed at each electrode: one uncorrelated signal lasting the entire duration of the data that tapers away during the correlated portions of the trials, and one correlated signal that appears only during the trials.

The window functions governing the transition between uncorrelated activity and correlated activity ($W_U(t)$ and $W_T(t)$, respectively) were chosen so that the instantaneous

variance of the signal would stay constant during trials. In particular, the correlated signal was windowed by the square root of a Gaussian kernel with $\sigma = 50$ ms and $\mu = 250$ ms:

$$W_T(t) = \sqrt{\frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(t-\mu)^2}{2\sigma^2}}}$$

The uncorrelated signal was windowed by:

$$W_U(t) = \sqrt{1 - \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(t-\mu)^2}{2\sigma^2}}}$$

These choices resulted in correlated and uncorrelated signals, $T(t)$ and $U(t)$, respectively, having instantaneous variances that evolve in time according to a Gaussian profile and one minus that Gaussian profile:

$$Var(T(t)) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(t-\mu)^2}{2\sigma^2}}$$

$$Var(U(t)) = 1 - \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(t-\mu)^2}{2\sigma^2}} .$$

Hence the combined signals had unit variance. In the resulting signal, the total variance of the correlated signal, $T(t)$, over the entire trial (for electrodes with at least one edge in both the before and after networks) was about one quarter of the total variance, the other three quarters being contributed by $U(t)$.

In the true time-domain coupling networks, each electrode of Region 1 is connected to two electrodes in Region 2 and vice versa. Each electrode in Region 3 is connected to one other electrode in the same region. The correlated signals were introduced in such a way that the inter-region coupling (between Regions 1 and 2) was

half as strong as the intra-region coupling (within Region 3). Since each electrode in Regions 1 and 2 had two inter-region edges and each electrode in Region 3 had one intra-region edge, the combined signal-to-noise ratio was the same for all electrodes (Table 1).

2.4.5 Frequency domain trial coupling. The simulations with frequency domain coupling were constructed the same way as the time domain coupling, switching from incoherent noise to a coherent signal during trials, except that two frequency bands were used: 20–30 Hz (beta) and 35–45 Hz (gamma). The signal in each frequency band consisted of incoherent filtered white noise that transitioned into coherent signals during the trials, according to the same windowing functions described above. The coherent signals, however, were not filtered white noise but sinusoids at 25 and 40 Hz, respectively, with a uniform random starting phase ϕ on each trial:

$$T(t) = \sqrt{2}\sin(2\pi ft + \phi)$$

where f is the frequency of the signal and t is the time since the start of the trial, in seconds. The factor of $\sqrt{2}$ ensures that $T(t)$ has unit variance, so that after windowing it and adding it to $U(t)$ the signal will have a total variance of 1. Adding the same T(t) to multiple electrodes made the electrodes coherent at frequency f.

The structure of the networks for the frequency domain coupling is shown in Figure III.5, Panel A. Here there were 12 electrodes split into 2 regions. In the beta band, each electrode in Region 1 is connected to two electrodes in Region 2 and vice versa. Superimposed on this coupling is within-region coupling in the gamma band, where all

electrodes in Region 1 share the same coherent signal (resulting in all-to-all coupling) and each electrode in Region 2 is coupled with one other electrode in the same region. As with the time-domain coupling, the inter-region coupling (in the beta band) was half as strong as the intra-region coupling (in the gamma band). Since each electrode had two inter-region signals (in the beta band) and one intra-region signal (in the gamma band), the signal-to-noise ratio between the total coupling in each frequency band and the total noise was the same for all electrodes (Table 1).

Table III.2: Parameters for simulation scenarios

Here we list the settings for the simulated data. The figures showing the results for each simulation scenario are listed in the last column. Abbreviations: uncorr. = uncorrelated component; corr. = correlated component; SNR = signal to noise ratio; correlation ratio = (variance of trial correlated component)/(variance of baseline correlated component).

Simulation	Trial Frequency		Trial Variance		2–50 Hz variance	SNR	Correlation ratio	Results Figures
	Uncorr.	Corr.	Uncorr.	Corr.				
Time Domain	8–25 Hz	8–25 Hz	0.75	0.25	0.5	0.11	0.5	Figs III.4, 6
Frequency Domain	20–30 Hz	25 Hz	0.75	0.25	0.5	0.11	0.5	Figs III.5, 7
	35–45 Hz	40 Hz	0.75	0.25	0.5	0.11	0.5	

3. Results

We tested the performance of the test statistics using two simulation scenarios: the correlation and canonical correlation networks were inferred for a simulation with time-domain coupling (filtered white noise), and the coherence and canonical coherence networks were inferred for a simulation with frequency-domain coupling. Figures III.4 and III.5 show the results of these tests.

3.1 Visualization of data transformations

Before describing these results, we will first describe the data transformations that take place in the network inference, so that the connection between the inferred networks and the source data is clear. We do so here in detail for the time domain networks; similar transformations occur for frequency domain networks.

3.1.1 Correlation Networks. Figure III.2 shows how the data changes with each step of the correlation network inference procedure. The figures show the edges (pairs of nodes) on the x axis, sorted by their final p-values, with index e. The sorting of the edges is the same for all of the panels, therefore the correlation for an edge in Panel 1 is located vertically above the $\hat{x}$ for the same edge in Panel 2, for example. The edges that will ultimately pass the thresholding for the binary network are colored red, and the rest are colored black.

The first panel shows the average correlation for each edge over baseline intervals ($+$) and trial intervals ($\times$). The baseline correlations are nonzero because the simulated data included some all-to-all coupling between electrodes, and the intent of the networks is to detect increases in coupling related to the trial period. The trial correlations are higher on some edges due to the introduced coupling, and there is some random jitter in both the trial and baseline correlations due to sampling variability.

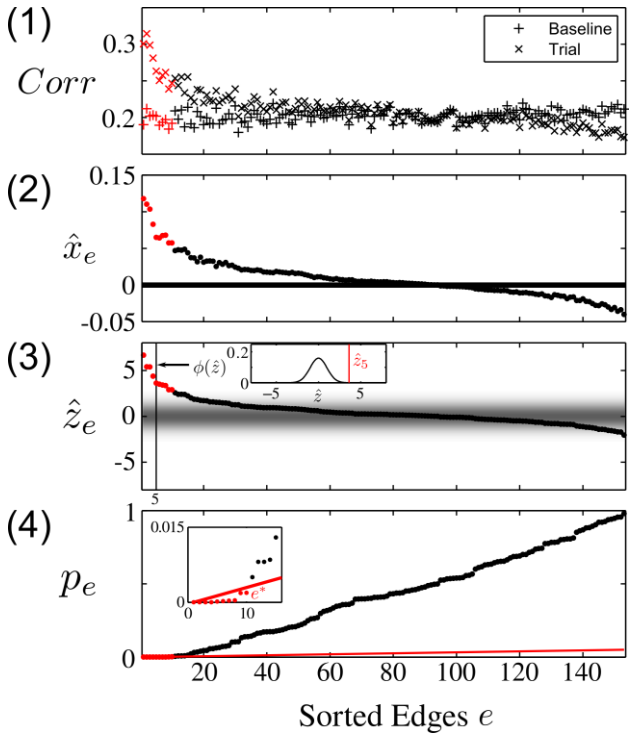


Figure III.2: Data transformations for the correlation networks.

Panels show the steps that the data undergo in the estimation of the correlation networks. In all panels, the x-axis represents the edges, sorted by their ultimate p-value. Edges that pass the FDR threshold for inclusion in the final binary networks are indicated in red in all panels. (Panel 1) The baseline and trial correlations for all edges. (Panel 2) The test statistic $\hat{x}$ calculated from the baseline and trial correlations. (Panel 3) The normalized test statistic $\hat{z}$, shown superimposed on the standard normal probability density function $\phi(\hat{z})$ in greyscale. The inset shows $\phi(\hat{z})$ versus $\hat{z}$, for comparison with $\hat{z}$ for edge number 5. (Panel 4) The p-values for each edge, calculated by comparing $\hat{z}$ to the standard normal distribution. The red line is $FDR(e) = \frac{e}{N(N-1)/2}\alpha$, where e is the edge index, $\alpha = 0.05$ is the FDR level, and $N = 18$ is the number of electrodes. Inset: close-up of the few edges with the smallest p-values to show the last edge having $p_e \le \frac{e}{N(N-1)/2}\alpha$, which is labeled e^*. For more detail, see Results.

The difference in the Fisher transform of these two correlations yields $\hat{x}$ for each edge, shown in the second panel. Notice that the edges with high $\hat{x}$ are not strictly the same as the edges with high trial correlations: there exist (black) edges with high trial correlations but also high baseline correlations, so their (transformed) difference is smaller than (red) edges with lower trial correlations.

The third panel shows $\hat{z}$, which is $\hat{x}$ divided by its estimated variance. Again, some edges with increased $\hat{x}$ have reduced $\hat{z}$. This is because the estimated variance was higher for those edges, making their z-score smaller. The $\hat{z}$ for each edge is compared to a standard normal distribution: this is represented in the third panel by superimposing the $\hat{z}$ values on a heatmap representation of the standard normal probability density function $\phi(\hat{z})$, where the grayscale intensity represents the probability density for the value of $\hat{z}$ on the y-axis. The inset shows this comparison for edge number 5. Integrating the probability density from positive infinity to $\hat{z}$ yields the p-value (this is equivalent to one minus the cumulative distribution function evaluated at $\hat{z}$).

The resulting p-values are shown in the fourth panel. The FDR procedure compares the sorted p-values to the line $FDR(e) = \frac{e}{N(N-1)/2}\alpha$, which is shown in red. As can be seen in the inset, the edge with the largest p-value falling below this line, e^*, passes the thresholding (rejecting the null hypothesis of no coupling), and all edges with smaller p-values also pass[1].

[1] It is not necessary that all of these edges also have p-values below the FDR line, they just need to be below the p-value of e^*. It is possible (though it does not happen here), with increasing e, for the p-values to go above the FDR line and then back below it. It is the *last, biggest* p-value that is below the FDR line that matters.

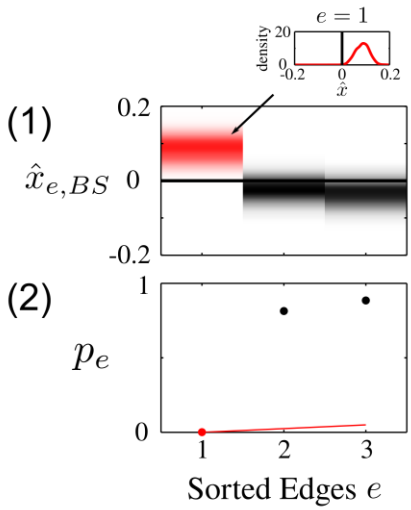


Figure III.3: Data transformations for the canonical correlation networks.

Panels show the steps that the data undergo in the estimation of the canonical correlation networks. In all panels, the x-axis represents the edges, sorted by their ultimate p-value. Edges that pass the FDR threshold for inclusion in the final binary networks are indicated in red in all panels. (Panel 1) The estimated sampling distribution of the test statistic $\hat{x}$ (estimated using a two-sample bootstrap procedure, see Methods 2.2), which is compared to zero (black line). The inset shows the estimated probability density of $\hat{x}$ for the edge with the smallest p-value. (Panel 2) The p-values for each edge, calculated by comparing the estimated distribution of $\hat{x}$ to zero. The red line is $FDR(e) = \frac{e}{R(R-1)/2}\alpha$, where e is the edge index, $\alpha = 0.05$ is the FDR level, and $R = 3$ is the number of regions. For more detail, see Results.

3.1.2 Canonical Correlation networks. The canonical correlation based networks do not involve estimating a single value for the test statistic $\hat{x}$. Rather, a distribution of $\hat{x}$ values is estimated for each edge using a two-sample bootstrap as described in Methods 2.2. Figure III.3 shows these distributions in Panel 1. Note that region-level networks have many fewer nodes (here, only 3), so there are fewer possible edges in the networks ($R(R-1)/2 = 3$). The p-values (shown in Panel 2) are the proportion of bootstrap observations that are less than zero; or, equivalently, the integral of the estimated sampling distribution from negative infinity to zero. The FDR threshold line is shown in

red, $FDR(e) = \frac{e}{R(R-1)/2}\alpha$. Only one edge passes the threshold: the p-value for the first edge is 0.001, which is less than the threshold of $\frac{0.05}{3} = 0.0167$.

3.2 Time Domain Coupling

Figure III.4 shows the results of inferring correlation (left column) and canonical correlation (right column) networks on the time-domain coupling simulation. The coupling in this simulation consisted of white noise filtered into the 8–25 Hz frequency band, windowed to occur during trials, and added to the signals on the electrodes in such a way as to create the network architectures shown in Panel A. The inter-regional connections between region 1 (blue) and region 2 (green) were half as strong as the intra-regional connections within region 3 (red), which led to very weak trial correlations (Panel B), almost indistinguishable from the background baseline correlations. This weak signal of inter-regional connections was poorly detected at the electrode level in both the weighted and binary networks (Panel C). The intra-regional connections, being stronger, were all detected. Canonical correlation, on the other hand, was able to detect the connection between region 1 and region 2, because of the large number of (weak) connections. Canonical correlation is able to combine evidence across the electrodes in the regions, making it well-suited to detect larger-scale coupling, even in the presence of non-task-related coupling.

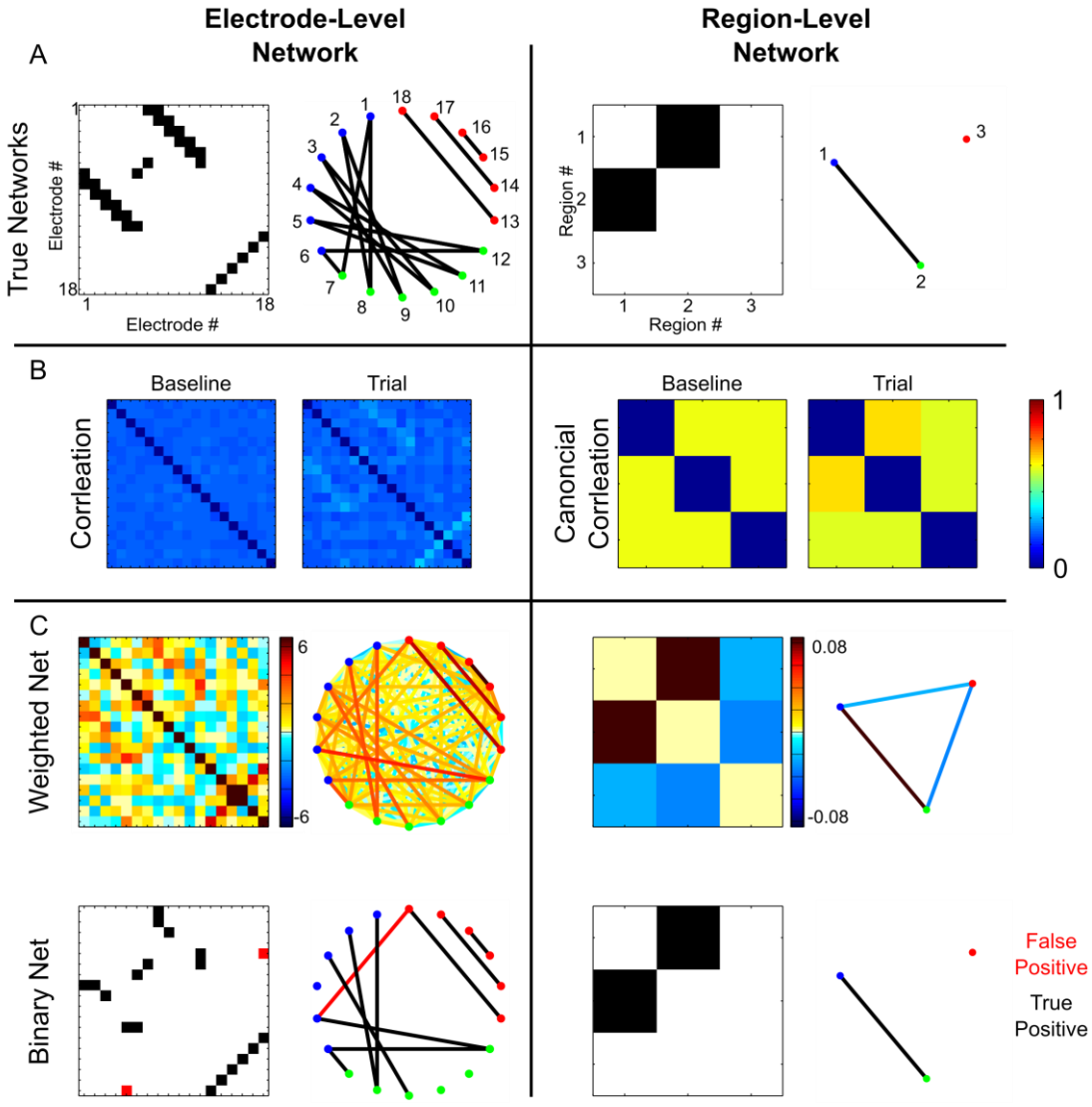


Figure III.4: Time-domain networks.

Correlation and canonical correlation networks inferred from the simulation with time-domain coupling. In all panels, the electrode-level correlation networks are in the left column and the region-level canonical correlation networks are in the right column. In the network diagrams, the nodes are colored by region identity. (Panel A) The binary true networks, shown as both a connectivity matrix (left) and a network diagram (right). (Panel B) The observed coupling matrices for the baseline and trial. (Panel C, top) The weighted networks, shown in matrix format and network diagram format. For correlation networks the weights are given by $\hat{z}$, and for canonical correlation networks the weights are given by $\bar{x}$ (see Figure III.1 and Methods). (Panel C, bottom) The inferred binary networks.

3.3. Frequency Domain Coupling

Figure III.5 shows the results of inferring coherence (left column) and canonical coherence (right column) networks on the frequency-domain coupling simulation. The frequency domain coupling simulation had twelve electrodes, split into two regions. Frequency domain coupling was introduced as windowed sinusoids at two different frequencies: 25 Hz (beta) and 40 Hz (gamma), each in a background of band-limited incoherent noise at 20–30 Hz and 35–45 Hz, respectively. The beta network consisted of inter-regional connections between region 1 and region 2, while the gamma network consisted of intra-regional connections within the two regions. As before, the inter-regional connections were half as strong as the intra-regional connections.

We show inferred networks at three frequencies of interest: 24, 32, and 40 Hz. Since the bandwidth of the coherence analysis was 8 Hz, both coherence and canonical coherence do not detect any coupling at 32 Hz. Interestingly, the coherence networks were better able to detect inter-regional connections than the correlation networks, even though the SNR was the same in the two simulations. The coherence networks detected all but one connection at 24 Hz and all of the connections at 40 Hz, with one false positive at 32 Hz. The canonical coherence networks detect the connection between region 1 and 2, with no false positives.

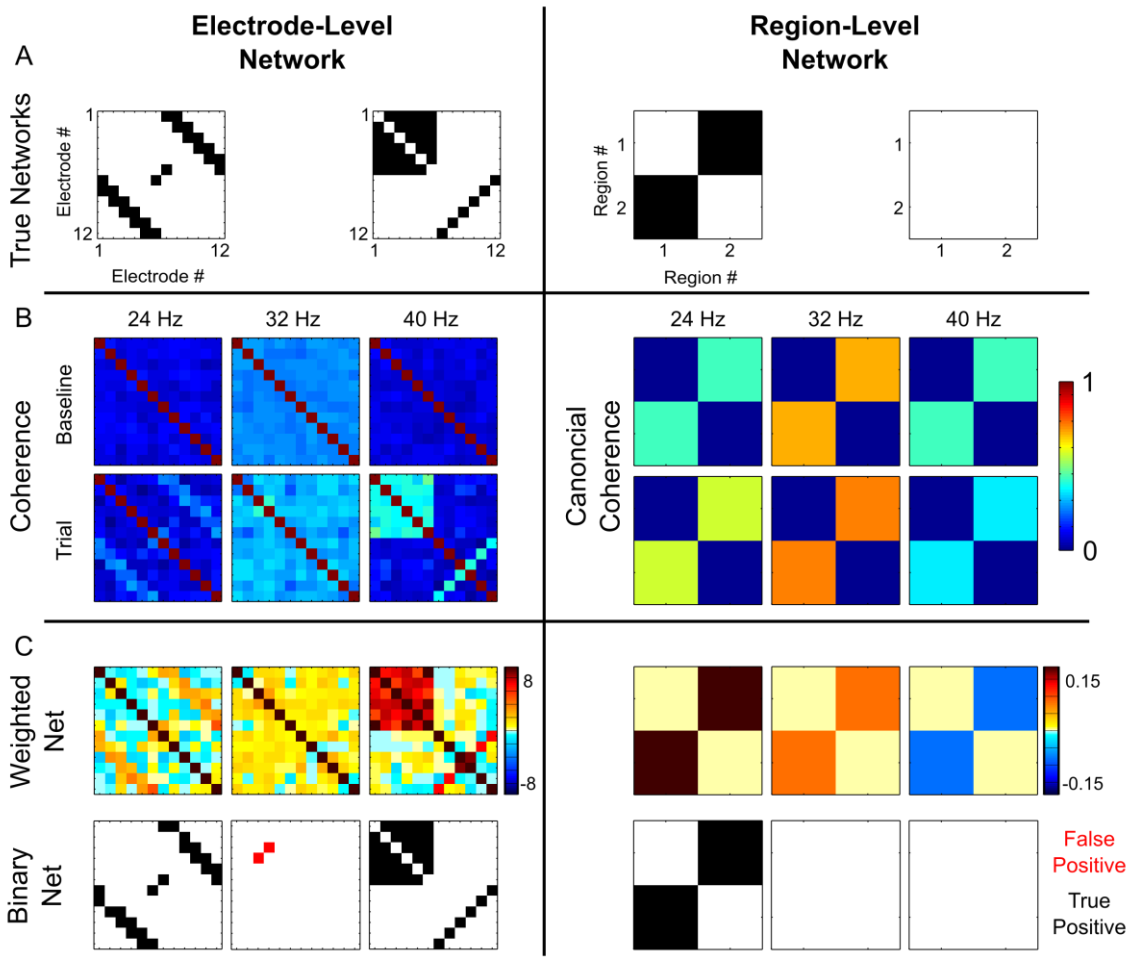


Figure III.5: Frequency-domain networks.

Coherence and canonical coherence networks inferred from the simulation with frequency-domain coupling. In all panels, the electrode-level coherence networks are in the left column and the region-level canonical coherence networks are in the right column. (Panel A) The binary true networks for the two frequencies of coupling, 25 Hz (left) and 40 Hz (right). (Panel B) The observed coupling matrices for the baseline and trial, at three frequencies: 24 Hz, 32 Hz, and 40 Hz. The frequency resolution of the coherence analysis was 8 Hz (±4 Hz). (Panel C, top) The weighted networks. For coherence networks the weights are given $\hat{z}$, and for canonical coherence networks the weights are given by $\overline{x}$ (see Figure III.1 and Methods). (Panel C, bottom) The inferred binary networks.

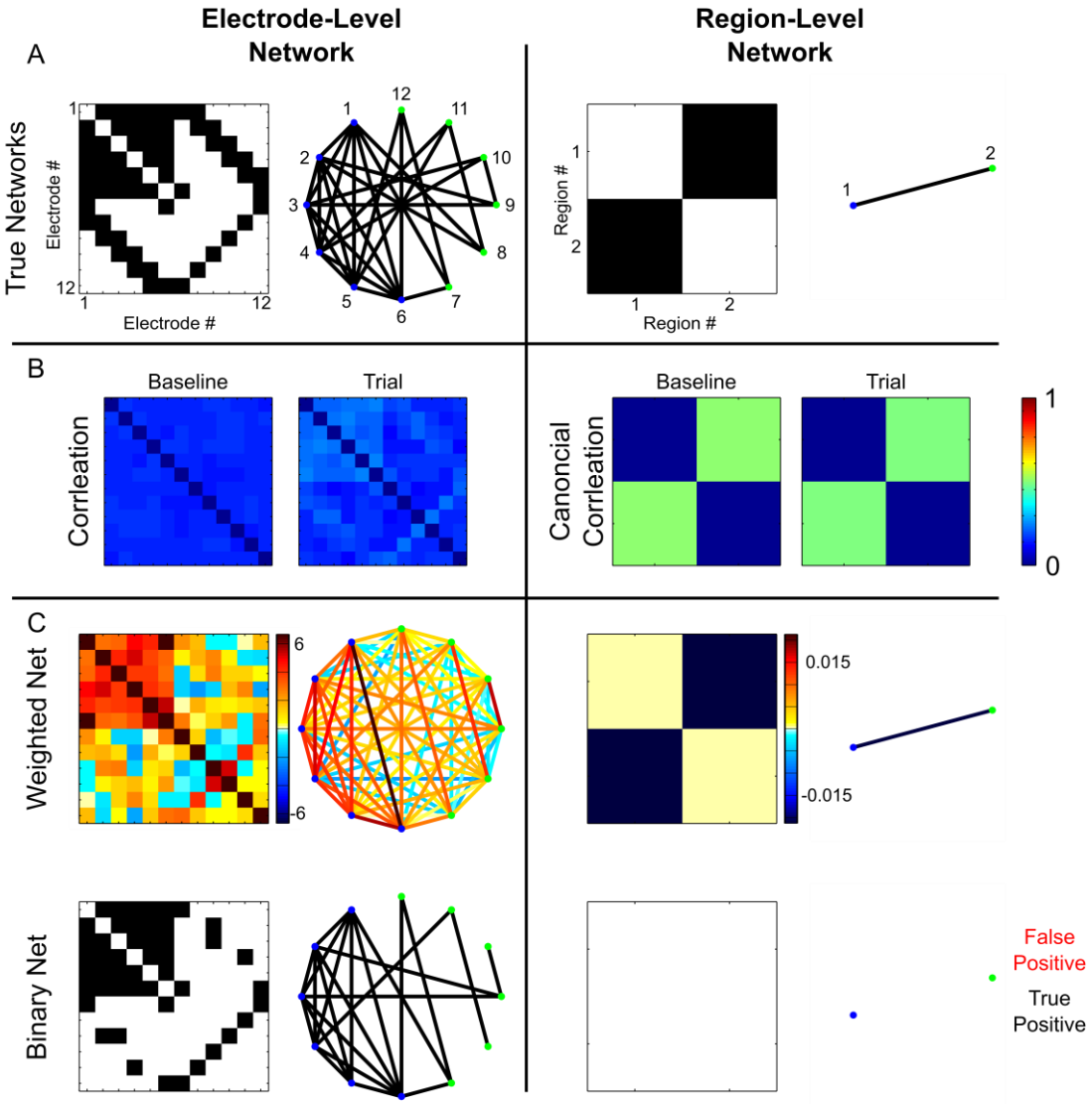


Figure III.6: Ability of time-domain networks to detect frequency-domain coupling.

Correlation and canonical correlation networks inferred from the simulation with *frequency*-domain coupling, where the connections between edges are introduced through the addition of a sinusoid windowed by a Gaussian function (see Methods 2.4). The coupling is introduced at two frequencies, 25 Hz and 40 Hz, with the patterns given in Figure III.5. Here, the two frequencies of coupling are superimposed since they both introduce time-domain correlations. The inter-regional connections are half as strong as the intra-regional connections. Layout as in Figure III.4.

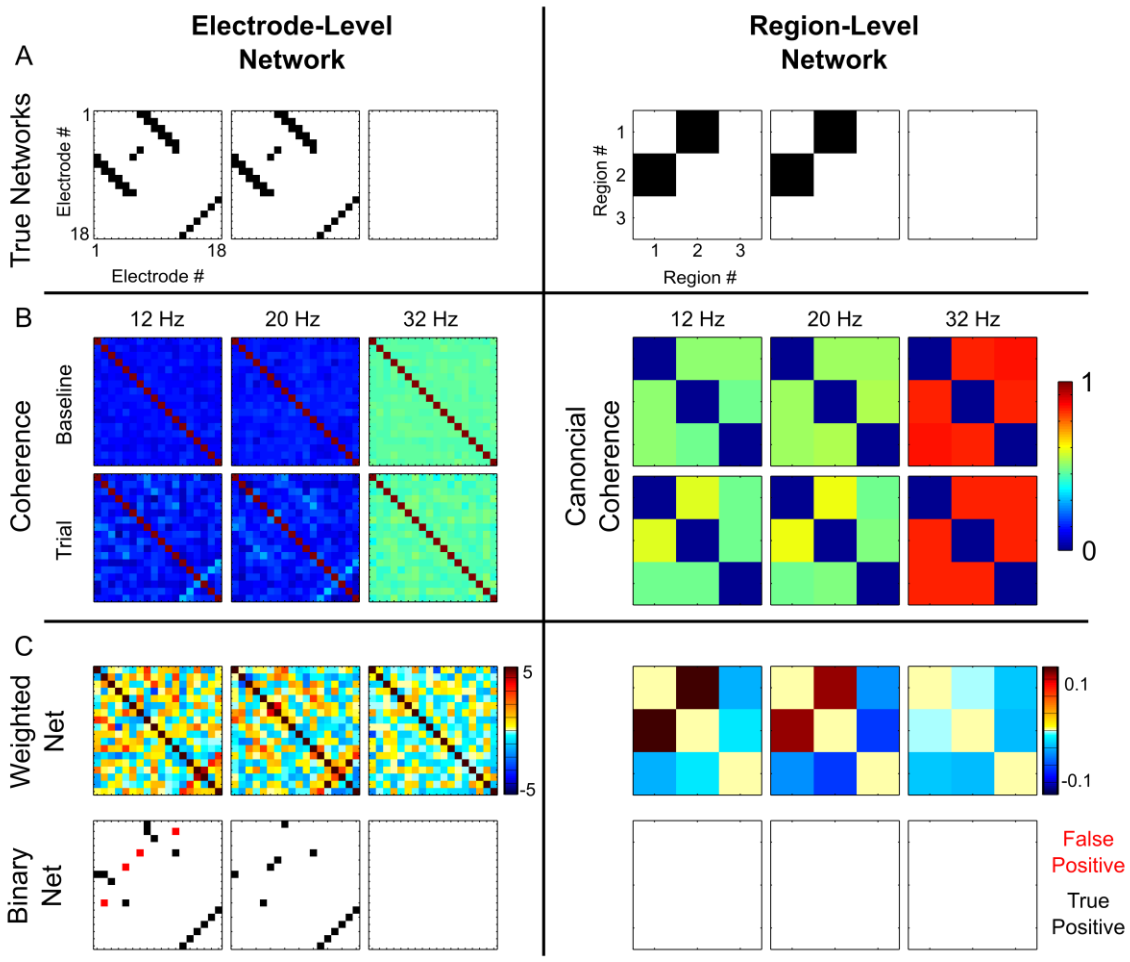


Figure III.7: Ability of frequency-domain networks to detect time-domain coupling.

Coherence and canonical coherence networks inferred from the simulation with *time*-domain coupling, where the connections between edges are introduced through the addition of white noise filtered in the 8-25 Hz band (see Methods 2.4). Layout as in Figure III.5. The frequency resolution of the coherence analysis was 8 Hz (±4 Hz).

3.4. Mismatches between coupling and test statistic

Figures III.6 and III.7 show the ability of the procedures to detect coupling that is not matched to the test statistic: Figure III.6 shows the results of correlation-based networks on the simulation with frequency-domain (sinusoidal) coupling, and Figure III.7 shows the results of coherence-based networks on the simulation with time-domain coupling.

Figure III.6 shows inferred correlation networks for the simulated data in Figure III.5, in which the coupling is introduced using Gaussian-windowed sinusoids at two frequencies: 25 and 40 Hz. The first thing to notice about Figure III.6 is that the frequency-domain coupling is superimposed in the time-domain. In other words, whereas there are two separable networks in the frequency domain, it is impossible to uncouple them in the time domain. In the simulated data used here, the networks are divided so that the 25 Hz edges occur between regions and the 40 Hz networks occur within regions, so that no edge occurs in both networks, but this is not a requirement of the frequency-domain coupling. This highlights a feature of frequency domain coupling: multiple networks can be represented on the same set of edges, potentially allowing for multiplexing of different network functions.

The correlation networks correctly detect all of the within-region edges but fail to detect some of the between-region edges, which are weaker. They perform more poorly than the coherence networks for these inter-region edges, since correlation is unable to separate the coupling signals at different frequencies: the intra-region edges obscure the inter-region edges. Unlike in Figure III.5, the region-level (canonical correlation)

networks are not capable of detecting the inter-region connections. Even the weighted network for canonical correlation has a value very close to zero for the inter-region edge. Like with the electrode-level network, this is because the signals on the nodes contain activity related to both inter-region and intra-region connections, and canonical correlation is not able to decouple these effects with limited data.

Figure III.7 shows the inference of coherence networks from the simulated data in Figure III.4. Again, when the coupling statistic does not match the underlying coupling, the network inference performs more poorly. In this case, coherence networks are inferred from simulated data in which the coupling is introduced as white noise, filtered between 8 and 25 Hz. The electrode-level networks detect all of the (stronger) intra-region connections, but they do not detect all of the (weaker) inter-region connections. For the inter-region connections, the electrode-level networks have both more false positives and more false negatives than those in Figure III.4 using correlation-based networks. Again, the region-level networks are not able to detect the inter-regional coupling, although the weighted networks do show a positive effect of about 0.1 between Region 1 and Region 2. The p-value on the edge is 0.030 for 12 Hz networks and 0.048 for the 20 Hz networks: these are not significant given the FDR correction for multiple comparisons across edges, which is $0.05/3 = 0.0167$, so in this situation there would need to be more data in order for the network inference to have sufficient power to detect the edge.

In general, mismatches between the test statistic and coupling lead to reduced performance of the network inference. While correlation-based networks are capable of

detecting sinusoidal coupling and coherence-based networks are capable of detecting filtered noise coupling, performance is reduced when the test statistic mismatches the coupling. As a result, some true edges are missed and some false edges are detected. In this case, the inference procedure is capable of detecting the true network given sufficient statistical power, which in practice would mean that in the case of a mismatch collecting more data could improve network inference.

4. Discussion

We presented here four coupling statistics for use in task-related brain network analyses. The coupling statistics were motivated by improvements to the coupling statistics presented in Chapter II, intending to improve the statistical power of the network inference methodology and to match the coupling statistics to potential forms of functional coupling. Specifically, the improvements took three forms: introducing frequency domain coupling statistics, using theoretical distributional properties where possible, and formulating tests as two-sample differences of means.

Introducing frequency domain test statistics, coherence and canonical coherence, allows the functional network inference to be sensitive to oscillatory coupling. Rhythmic dynamics have been implicated in wide range of neural functions and are thought to coordinate activity within and between brain areas, using potentially overlapping frequency bands for different purposes (Fries, 2005; Kopell et al., 2010; Wang, 2010; Miller et al., 2012). We try to mimic this kind of coupling in the frequency-domain coupling simulation here: the signals creating the networks at the beta and gamma frequencies were superimposed on the same electrodes. This highlights an advantage of

frequency-domain coupling in general: introducing coupling at different frequencies enables multiplexing of multiple networks on the same set of nodes. The additional flexibility in implementing functional coupling can cause problems for inference, however, since it introduces another dimension into the statistical analysis. Inferring networks along the frequency dimension can require more data to achieve sufficient statistical power. One approach to mitigating this problem is to focus on specific frequency bands of interest, as we do here. In many tasks, we have prior knowledge about the frequency bands that are likely to be involved, and this prior knowledge can be used to constrain the dimensionality of the statistical inference.

Using theoretical distributions is another way to incorporate prior knowledge about the statistical properties of the estimators. We advocate using theoretical distributions when possible, i.e. when the parameters are known and they are adequate models of the underlying process. For example, even though the exact distributional properties of the sample correlation and coherence are known, their dependence on the population correlation and coherence made them impractical to use for estimation. Instead, it was more practical to use the approximate distribution that removes the dependence on the population values by making use of the Fisher transform (Fisher, 1915). In addition, even though we could assume that the test statistic $\hat{x}$ was normally distributed with zero mean under the null hypothesis, we did not know the variance of the null distribution since the sample consisted of timeseries data. The variance would depend on the degrees of freedom in the timeseries, which was unknown. Hence even though we knew some aspects of the distribution of the test statistic under the null

hypothesis, we needed to estimate the variance using the jackknife procedure. This approach was proposed for comparisons of coherences by Bokil et al. (2007), combined with the bias correction for coherence that we use here. We follow the same methodology for comparisons of correlations, with the exception that the correlation estimates do not need to be bias corrected.

We formulate the test statistics as two-sample differences between trial and baseline coupling in order to capture the sampling variability of both coupling statistics. This approach also lends itself well to contrast-style experimental designs which compare neural activity between two different tasks. For example, a network could be inferred using the difference in coupling between an auditory and a visual task. In this case, we would expect edges with a strong positive weight to be more important for the auditory task and edges with a strong negative weight to be more important for the visual task.

Another advantage to structuring the test statistics as differences between the task and baseline is that it provides a convenient and interpretable edge weight. Using the absolute correlation (or coherence, etc) during the trials as the edge weight is often insufficient because it does not capture whether the coupling changes during the task. When trying to infer task-related networks, we are often less interested in the absolute coupling levels than in the increase or decrease in coupling during a task. This is because there may be sources of coupling that are not related to the task. These sources of coupling may be neural, such as background activity, or artifactual, such as muscle activity that is picked up on multiple electrodes. Since these types of coupling occur during both the baseline and the task period, they may tend to cancel out in the

difference. Hence the difference between the coupling statistic during tasks and during baseline may reflect task-related changes in coupling.

Of course, there are many possible sources for task-related changes in coupling. For example, electrodes may become coupled because they have a common driver (e.g. thalamus) that causes similar activity on the electrodes in the absence of a direct connection. Electrodes can also have increased correlations because they respond to the task in similar ways, even if they have no need to "connect" in the sense of functional connectivity. This is a common problem in functional connectivity inference (Friston, 2011; Greenblatt et al., 2012), and the only step we take here to mitigate it is to remove the event-related potentials in preprocessing (which may not remove all stimulus-evoked activity (Truccolo et al., 2002)).

We believe that despite these drawbacks, simple linear pairwise metrics have value precisely because of their simplicity. For example, researchers commonly use partial correlations rather than pairwise correlations to try to account for the common driver problem described above. In a situation in which the common driver is observed, e.g. on another electrode, partial correlation can sometimes more accurately detect the underlying connectivity[2]. If the common driver is not observed however, e.g. if it is subcortical, the partial correlation will fail. When it fails, it can spread out the "blame"

[2] Partial correlation may seem like a simple extension of correlation, but as a network inference technique it suffers much more from problems of high dimensionality. Gaussian graphical models (GGMs) are a popular simplification of the full partial correlation model (Kolaczyk, 2009). Even for GGMs, though, there is a lot of work being done on regularization to control the dimensionality of the inference. (Bolstad et al., 2011), for example, use group lasso regularization on a timeseries-based network model. They show that, even assuming all possible sources are observed, the ability of the inference to uncover the true network depends strongly on the architecture of the true network itself. In particular, cycles and redundancy in the true network can cause state-of-the-art network inference techniques to fail, even if the generative model is perfect.

for the coupling over the parameters of the model in a way that is difficult to trace back to the original data. More complicated models can be constructed with hidden states to account for these kinds of dependencies, but they require strong intuitions about the structure of the system in order to avoid dimensionality explosions and under-constrained models (and they can fail if the models are wrong).

This is not to say that such models should not be used – on the contrary, we believe they are critical for understanding the mechanisms of neural dynamics. We argue, however, that simple descriptive statistics like correlations are helpful because they can detect the signatures of strong interesting features of the data, which can be pursued more fully in later analyses such as detailed statistical modeling. Even many nonlinear effects can be detected first using lower-dimensional linear test statistics. This is especially valuable for researchers approaching a new dataset, when the nature of the data is still unclear and there may be artifacts or noise to be understood. In that situation, test statistics are more valuable if (1) their behavior can be easily linked to the source data, and (2) the effects of model misspecification are local, in the sense that inferences about a part of the system (e.g. one edge) have limited functional dependence on inferences about other parts of the system. Pairwise correlations and coherences have these features. Hence simple robust pairwise statistics as we describe here can be helpful at the early stages of data screening and processing, to discover problems with the data, failures of our assumptions, and features that deserve more in-depth analysis.

With that in mind, in the presentation of the simulation results we made an effort to convey intuition about the behavior of the test statistics, the nature of the source data,

and the effect sizes of the coupling. While the binary networks may be a simple data summary, they have little value without a sense for the strength of the coupling and the features in the data that create the significant effects. In the binary networks presented here the edges represent statistically significant differences in coupling between the task period and the baseline period. The actual difference in coupling between the two periods is thrown away. For this reason, it can be valuable to display weighted networks in addition to binary networks.

Furthermore, since the networks inferred using this framework are based on the differences in correlation between the task and baseline periods, the value of the test statistic (the effect size) on an edge does not necessarily reflect the correlation during the baseline period or the correlation during the tasks. For example, an edge could have a high correlation in both the baseline and test intervals, resulting in a low value of the test statistic. Conversely, an edge with a small negative correlation during the baseline period and a small positive correlation during the trial period could have a relatively large test statistic. It is therefore important to visualize the steps taken over the course of the network inference in order to understand the nature of the data at the different levels of analysis. In the Results and Figure III.2 we demonstrated how the steps of analysis transformed the data in the correlation-based network inference. In the context of brain network analysis, this provides context regarding the effect size of the observed coupling, which can have implications for its biological relevance.

The goal of this chapter was to present a set of test statistics that are tailored to a few of the second-order (between-electrode) effects that may be present in neural

voltages. Two of the test statistics, canonical correlation and canonical coherence, are rare in functional connectivity analysis. We presented these measures alongside the more commonly used single-dimensional analogues to further accessibility to the neuroscience community.

5. Appendix: Derivation of Canonical Coherence

This appendix presents the derivation of canonical coherence. In the first section, we present the optimization problem that is solved by canonical coherence. In the second section we derive the solution and present an algorithm for estimating the solution from data. In the third section, we modify the estimation algorithm for computational efficiency and stability. Finally, in the fourth section, we explain how canonical correlation can be estimated using the same method. We include at the end summaries of a few concepts that are used in the derivation: the method of Lagrange multipliers, complex gradients, properties of singular value decompositions, and properties of positive definite matrices.

5.1 Statement of problem

Suppose we have two multidimensional data vectors (random variables) represented in the frequency domain, $\underline{X}(f)$ and $\underline{Y}(f)$, where $\underline{X}$ is $p \times 1$ and $\underline{Y}$ is $q \times 1$. For each frequency, we would like to find the "optimum linear combination of the measurements within each set such that the two resulting series are maximally coherent" (Thomson et al., 2007). Restricting ourselves to a single frequency f, the goal is to find two vectors $\underline{\alpha}$ ($p \times 1$) and $\underline{\beta}$ ($q \times 1$) that will be used to define linear combinations of $\underline{X}$ and $\underline{Y}$:

$$\begin{aligned} a &= \underline{\alpha}^{\dagger} \underline{X} \\ b &= \underline{\beta}^{\dagger} \underline{Y} \end{aligned} \tag{5.1.1}$$

so that the (scalar) correlation between a and b is maximized:

$$
\begin{aligned}
C(\underline{\alpha},\underline{\beta}) &= \frac{E\{ab^*\}}{\sigma_a\sigma_b} \\
&= \frac{E\{\underline{\alpha}^\dagger \underline{X}\underline{Y}^\dagger \underline{\beta}\}}{E\{\underline{\alpha}^\dagger \underline{X}\underline{X}^\dagger \underline{\alpha}\}^{1/2} E\{\underline{\beta}^\dagger \underline{Y}\underline{Y}^\dagger \underline{\beta}\}^{1/2}} \\
&= \frac{\underline{\alpha}^\dagger \underline{\underline{S}}_{XY} \underline{\beta}}{(\underline{\alpha}^\dagger \underline{\underline{S}}_{XX} \underline{\alpha})^{1/2} (\underline{\beta}^\dagger \underline{\underline{S}}_{YY} \underline{\beta})^{1/2}}
\end{aligned}
\tag{5.1.2}
$$

where $\underline{\underline{S}}_{XY}$, $\underline{\underline{S}}_{XX}$, and $\underline{\underline{S}}_{YY}$ are the cross-spectral matrices of the elements of $\underline{X}$ and $\underline{Y}$ ($\underline{\underline{S}}_{XY}$, $\underline{\underline{S}}_{XX}$, and $\underline{\underline{S}}_{YY}$ are all functions of frequency f). The symbol † represents the Hermitian transpose. We have assumed for simplicity in equation (5.1.2) that $E\{\underline{X}\} = \underline{0}$ and $E\{\underline{Y}\} = \underline{0}$, so that linear combinations of their elements will also be centered at zero: $E\{a\} = 0$ and $E\{b\} = 0$. Since $C(\underline{\alpha},\underline{\beta})$ can be complex, we will try to maximize its square. Hence the problem can be written:

$$
\underline{\alpha},\underline{\beta} = argmax_{\underline{\alpha}',\underline{\beta}'} |C|^2 = argmax_{\underline{\alpha}',\underline{\beta}'} CC^* \tag{5.1.3}
$$

Note that the solution is not unique. We constrain it further by requiring that[3]:

$$
\begin{aligned}
\underline{\alpha}^\dagger \underline{\underline{S}}_{XX} \underline{\alpha} &= 1 \\
\underline{\beta}^\dagger \underline{\underline{S}}_{YY} \underline{\beta} &= 1
\end{aligned}
\tag{5.1.4}
$$

Under this constraint,

$$
C(\underline{\alpha},\underline{\beta}) = \underline{\alpha}^\dagger \underline{\underline{S}}_{XY} \underline{\beta} \tag{5.1.5}
$$

5.2 Solution

Let:

$$
\begin{aligned}
\underline{u} &= \underline{\underline{S}}_{XX}^{1/2} \underline{\alpha} \\
\underline{v} &= \underline{\underline{S}}_{YY}^{1/2} \underline{\beta}
\end{aligned}
\tag{5.2.1}
$$

[3] (Thomson et al., 2007) uses the same constraint, but (Brillinger, 1981) uses $\underline{\alpha}^\dagger \underline{\alpha} = 1$ and $\underline{\beta}^\dagger \underline{\beta} = 1$.

The square roots of $\underline{\underline{S}}_{XX}$ and $\underline{\underline{S}}_{YY}$ exist and are positive definite because the spectral matrices are positive definite. This assumes that none of the elements of $\underline{X}$ is a linear combination of the other elements of $\underline{X}$ (and the same is true for the elements of $\underline{Y}$). The square roots of a positive definite matrix are also positive definite, so they are invertible and Hermitian (See Section 5.8 for properties of positive definite matrices). Note that $\underline{u}^{\dagger}\underline{u}=1$ and $\underline{v}^{\dagger}\underline{v}=1$ ($\underline{u}$ is $p\times 1$ and $\underline{v}$ is $q\times 1$). Using substitution (5.2.1) in equation(5.1.5),

$$\tilde{C}(\underline{u},\underline{v})=\underline{u}^{\dagger}\underline{\underline{S}}_{XX}^{-1/2}\underline{\underline{S}}_{XY}\underline{\underline{S}}_{YY}^{-1/2}\underline{v}=\underline{u}^{\dagger}\underline{\underline{Q}}_{XY}\underline{v} \tag{5.2.2}$$

where we have let $\underline{\underline{Q}}_{XY}=\underline{\underline{S}}_{XX}^{-1/2}\underline{\underline{S}}_{XY}\underline{\underline{S}}_{YY}^{-1/2}$. In order to find the maximum of $\tilde{C}$ under the constraints in equation (5.1.4), we will use the method of Lagrange Multipliers (see Section 5.5).

The Lagrange function is:

$$\Lambda(\underline{u},\underline{v},\lambda)=\tilde{C}\tilde{C}^{*}-\lambda_1\cdot(\underline{u}^{\dagger}\underline{u}-1)-\lambda_2\cdot(\underline{v}^{\dagger}\underline{v}-1) \tag{5.2.3}$$

where we have used the constraints from (5.1.4), using substitution (5.2.1). To find the stationary points, we need to take the complex gradient (see Section 5.6)

$$\begin{aligned}\nabla_u\tilde{C}\tilde{C}^{*}-\lambda_1(\nabla_u\underline{u}^{\dagger}\underline{u})&=\underline{0}\\ \nabla_v\tilde{C}\tilde{C}^{*}-\lambda_2(\nabla_v\underline{v}^{\dagger}\underline{v})&=\underline{0}\end{aligned} \tag{5.2.4}$$

We are only looking for extrema, so equation (5.6.4) tells us we can simplify this to:

$$\begin{aligned}\frac{\partial}{\partial\underline{u}^{*}}\tilde{C}\tilde{C}^{*}-\lambda_1\left(\frac{\partial}{\partial\underline{u}^{*}}\underline{u}^{\dagger}\underline{u}\right)&=\underline{0}\\ \frac{\partial}{\partial\underline{v}^{*}}\tilde{C}\tilde{C}^{*}-\lambda_2\left(\frac{\partial}{\partial\underline{v}^{*}}\underline{v}^{\dagger}\underline{v}\right)&=\underline{0}\end{aligned} \tag{5.2.5}$$

We use the product rule (5.6.8) combined with equations (5.6.9) to get[4]:

$$\begin{aligned}\frac{\partial}{\partial \underline{u}^*}\tilde{C}\tilde{C}^* &= \frac{\partial \tilde{C}}{\partial \underline{u}^*}\tilde{C}^* + \tilde{C}\frac{\partial \tilde{C}^*}{\partial \underline{u}^*} = \tilde{C}^*\underline{\underline{Q}}_{XY}\underline{v} \\ \frac{\partial}{\partial \underline{v}^*}\tilde{C}\tilde{C}^* &= \frac{\partial \tilde{C}}{\partial \underline{v}^*}\tilde{C}^* + \tilde{C}\frac{\partial \tilde{C}^*}{\partial \underline{v}^*} = \tilde{C}\underline{\underline{Q}}_{XY}^{\dagger}\underline{u}\end{aligned} \tag{5.2.6}$$

Using these and equation (5.6.5), we can simplify equations (5.2.5) to:

$$\begin{aligned}\tilde{C}^*\underline{\underline{Q}}_{XY}\underline{v} &= \lambda_1\underline{u} \\ \tilde{C}\underline{\underline{Q}}_{XY}^{\dagger}\underline{u} &= \lambda_2\underline{v}\end{aligned} \tag{5.2.7}$$

We need to find a $\underline{u}$ and $\underline{v}$ that satisfy these equations.

What we will do is turn equations (5.2.7) into an eigenvalue problem that will justify using the left and right first singular vectors of $\underline{\underline{Q}}_{XY}$ as $\underline{u}$ and $\underline{v}$. Start by multiplying both sides of the first part of (5.2.7) by $\underline{\underline{Q}}_{XY}^{\dagger}$:

$$\underline{\underline{Q}}_{XY}^{\dagger}\underline{\underline{Q}}_{XY}\underline{v} = \frac{\lambda_1}{\tilde{C}^*}\underline{\underline{Q}}_{XY}^{\dagger}\underline{u} \tag{5.2.8}$$

Now substituting the second part of (5.2.7) in the right hand side of (5.2.8):

$$\underline{\underline{Q}}_{XY}^{\dagger}\underline{\underline{Q}}_{XY}\underline{v} = \frac{\lambda_1\lambda_2}{\tilde{C}^*\tilde{C}}\underline{v} \tag{5.2.9}$$

Similar reasoning from leads to:

$$\underline{\underline{Q}}_{XY}\underline{\underline{Q}}_{XY}^{\dagger}\underline{u} = \frac{\lambda_1\lambda_2}{\tilde{C}^*\tilde{C}}\underline{u} \tag{5.2.10}$$

This setup (equations (5.2.8) and (5.2.9)) has the form of a singular value decomposition (Section 5.7) of $\underline{\underline{Q}}_{XY}$, where $\underline{u}$ and $\underline{v}$ are a pair of left and right singular vectors with the corresponding singular value $\frac{\lambda_1\lambda_2}{\tilde{C}^*\tilde{C}}$:

[4] Recall that $\tilde{C}^*$ is a scalar, so $\underline{\underline{Q}}_{XY}\underline{v}\tilde{C}^* = \tilde{C}^*\underline{\underline{Q}}_{XY}\underline{v}$.

$$\underline{\underline{Q}}_{XY} = \underline{\underline{U}}\,\underline{\underline{\Sigma}}\,\underline{\underline{V}}^{\dagger} \tag{5.2.11}$$

Picking any $\underline{u}_k$, $\underline{v}_k$, and σ_k set from the singular vectors and values of $\underline{\underline{Q}}_{XY}$, we can calculate a value of $\tilde{C}(\underline{u},\underline{v})$ that corresponds to a stationary point of the Lagrange function (5.2.3). Since there are k such values, one for each set of $\underline{u}_k$, $\underline{v}_k$, and σ_k, we will call them $\tilde{C}_k$:

$$\begin{aligned}\tilde{C}_k &= \underline{u}_k^{\dagger}\underline{\underline{Q}}_{XY}\underline{v}_k \\ &= \underline{u}_k^{\dagger}\underline{\underline{U}}\,\underline{\underline{\Sigma}}\,\underline{\underline{V}}^{\dagger}\underline{v}_k \\ &= \underline{e}_k^{\dagger}\underline{\underline{\Sigma}}\underline{e}_k \\ &= \sigma_k\end{aligned} \tag{5.2.12}$$

where $\underline{e}_k$ is the k^{th} standard basis vector (all zeros with a 1 at index k). Note that σ_k (and by extension $\tilde{C}_k$) is real. Since we want the value that maximizes the magnitude squared coherence $\tilde{C}_k\tilde{C}_k^{*} = \sigma_k^2$, we pick $\underline{u}_1$, $\underline{v}_1$, and σ_1, the first left and right singular vectors and singular value[5], where σ_1^2 is the Magnitude Squared Coherence. Finally, $\underline{\alpha}$ and $\underline{\beta}$ can be calculated:

$$\begin{aligned}\underline{\alpha} &= \underline{\underline{S}}_{XX}^{-1/2}\underline{u} \\ \underline{\beta} &= \underline{\underline{S}}_{YY}^{-1/2}\underline{v}\end{aligned} \tag{5.2.13}$$

where the inverse square roots exist because the spectral matrices are positive definite (see Section 5.8). $\underline{\alpha}$ and $\underline{\beta}$ can be used to calculate the Canonical Series $a(t)$, $b(t)$ by plugging them into equations (5.1.1) and taking the inverse Fourier transform.

---

[5] (Brillinger, 1981) refers to the σ_k as the "k^{th} canonical coherence" and the corresponding $a_k(t)$, $b_k(t)$ as the "k^{th} pair of canonical series".

So the steps for estimating the canonical coherence between two signals $\underline{X}$ and $\underline{Y}$ are:

1. Estimate the cross-spectral matrices $\underline{\underline{S}}_{XX}(f)$, $\underline{\underline{S}}_{XY}(f)$, $\underline{\underline{S}}_{YX}(f)$, and $\underline{\underline{S}}_{YY}(f)$ using the multitaper method for all frequencies. For each frequency, do the following steps.
2. Calculate $\underline{\underline{Q}}_{XY} = \underline{\underline{S}}_{XX}^{-1/2}\underline{\underline{S}}_{XY}\underline{\underline{S}}_{YY}^{-1/2}$
3. Calculate the SVD $\underline{\underline{Q}}_{XY} = \underline{\underline{U}}\,\underline{\underline{\Sigma}}\,\underline{\underline{V}}^{\dagger}$
4. Pick the first singular value σ_1 and the first left and right singular vectors $\underline{u}_1$ and $\underline{v}_1$
5. The estimate of the magnitude squared coherence is σ_1^2. The canonical series $a(t)$ and $b(t)$ can be calculated from $\underline{u}_1$ and $\underline{v}_1$ using equations (5.2.13) and (5.1.1).

5.3 Simplification

Assuming we are using the multitaper method to estimate the spectral matrices in the previous section, there is a shortcut we can take to calculate $\underline{\underline{Q}}_{XY}$ without explicitly calculating the spectral matrices. This section will derive the shortcut.

In the previous section we started with the true frequency-domain data vectors $\underline{X}(f)$ $(p\times 1)$ and $\underline{Y}(f)$ $(q\times 1)$. In a practical context we will estimate these using the multitaper method, tapering the time-domain data and taking the Fourier transform, resulting in matrices $\underline{\underline{X}}(f)$ $(p\times k)$ and $\underline{\underline{Y}}(f)$ $(q\times k)$ of tapered Fourier coefficients

where k is the number of trials times the number of tapers. The spectra would then be estimated as follows:

$$\begin{aligned}
\hat{\underline{\underline{S}}}_{XX}(f) &= \frac{1}{k}\underline{\underline{X}}\,\underline{\underline{X}}^{\dagger} \\
\hat{\underline{\underline{S}}}_{XY}(f) &= \frac{1}{k}\underline{\underline{X}}\,\underline{\underline{Y}}^{\dagger} \\
\hat{\underline{\underline{S}}}_{YX}(f) &= \frac{1}{k}\underline{\underline{Y}}\,\underline{\underline{X}}^{\dagger} \\
\hat{\underline{\underline{S}}}_{YY}(f) &= \frac{1}{k}\underline{\underline{Y}}\,\underline{\underline{Y}}^{\dagger}
\end{aligned} \tag{5.3.1}$$

It is important for the analysis that the spectra be invertible, which can only be true if $k \geq p, q$. That is, the number of tapers times trials must be larger than the dimension of both $\underline{X}$ and $\underline{Y}$. Ideally k should be much larger than p and q, since the number of parameters estimated in the spectral matrices is $p(p-1)/2$ and $q(q-1)/2$, and k represents the number of observations. Variance in the estimates of the spectral matrices will propagate to the canonical coherence estimate.

To simplify, we start by plugging these estimates into the definition of $\underline{\underline{Q}}_{XY}$:

$$\begin{aligned}
\underline{\underline{Q}}_{XY} &= \hat{\underline{\underline{S}}}_{XX}^{-1/2}\hat{\underline{\underline{S}}}_{XY}\hat{\underline{\underline{S}}}_{YY}^{-1/2} \\
&= \left(\frac{1}{k}\underline{\underline{X}}\,\underline{\underline{X}}^{\dagger}\right)^{-1/2}\left(\frac{1}{k}\underline{\underline{X}}\,\underline{\underline{Y}}^{\dagger}\right)\left(\frac{1}{k}\underline{\underline{Y}}\,\underline{\underline{Y}}^{\dagger}\right)^{-1/2}
\end{aligned} \tag{5.3.2}$$

If we perform a SVD on $\underline{\underline{X}} = \underline{\underline{U}}_X\underline{\underline{\Sigma}}_X\underline{\underline{V}}_X^{\dagger}$ and $\underline{\underline{Y}} = \underline{\underline{U}}_Y\underline{\underline{\Sigma}}_Y\underline{\underline{V}}_Y^{\dagger}$, we get:

$$\begin{aligned}
\underline{\underline{Q}}_{XY} &= \left(\underline{\underline{U}}_X\underline{\underline{\Sigma}}_X\underline{\underline{V}}_X^{\dagger}\underline{\underline{V}}_X\underline{\underline{\Sigma}}_X\underline{\underline{U}}_X^{\dagger}\right)^{-1/2}\left(\underline{\underline{U}}_X\underline{\underline{\Sigma}}_X\underline{\underline{V}}_X^{\dagger}\underline{\underline{V}}_Y\underline{\underline{\Sigma}}_Y\underline{\underline{U}}_Y^{\dagger}\right)\left(\underline{\underline{U}}_Y\underline{\underline{\Sigma}}_Y\underline{\underline{V}}_Y^{\dagger}\underline{\underline{V}}_Y\underline{\underline{\Sigma}}_Y\underline{\underline{U}}_Y^{\dagger}\right)^{-1/2} \\
&= \left(\underline{\underline{U}}_X\underline{\underline{\Sigma}}_X^2\underline{\underline{U}}_X^{\dagger}\right)^{-1/2}\left(\underline{\underline{U}}_X\underline{\underline{\Sigma}}_X\underline{\underline{V}}_X^{\dagger}\underline{\underline{V}}_Y\underline{\underline{\Sigma}}_Y\underline{\underline{U}}_Y^{\dagger}\right)\left(\underline{\underline{U}}_Y\underline{\underline{\Sigma}}_Y^2\underline{\underline{U}}_Y^{\dagger}\right)^{-1/2} \\
&= \left(\underline{\underline{U}}_X\underline{\underline{\Sigma}}_X^{-1}\underline{\underline{U}}_X^{\dagger}\right)\left(\underline{\underline{U}}_X\underline{\underline{\Sigma}}_X\underline{\underline{V}}_X^{\dagger}\underline{\underline{V}}_Y\underline{\underline{\Sigma}}_Y\underline{\underline{U}}_Y^{\dagger}\right)\left(\underline{\underline{U}}_Y\underline{\underline{\Sigma}}_Y^{-1}\underline{\underline{U}}_Y^{\dagger}\right) \\
&= \underline{\underline{U}}_X\underline{\underline{V}}_X^{\dagger}\underline{\underline{V}}_Y\underline{\underline{U}}_Y^{\dagger}
\end{aligned} \tag{5.3.3}$$

Note that in order to get the dimensions of the matrices to line up in the simplification, it is necessary to use the "thin" singular value decomposition ("economy size" in Matlab), in which for example $\underline{\underline{U}}_X$ is $(p \times p)$, $\underline{\underline{\Sigma}}_X$ is $(p \times k)$, and $\underline{\underline{V}}_X$ is $(k \times p)$ (see Section 5.7).

This suggests another method for estimating the canonical coherence:

1. Calculate the matrices of tapered Fourier coefficients $\underline{\underline{X}}(f)$ and $\underline{\underline{Y}}(f)$ from the data.
2. Compute the SVDs $\underline{\underline{X}} = \underline{\underline{U}}_X \underline{\underline{\Sigma}}_X \underline{\underline{V}}_X^\dagger$ and $\underline{\underline{Y}} = \underline{\underline{U}}_Y \underline{\underline{\Sigma}}_Y \underline{\underline{V}}_Y^\dagger$
3. Calculate $\underline{\underline{Q}}_{XY} = \underline{\underline{U}}_X \underline{\underline{V}}_X^\dagger \underline{\underline{V}}_Y \underline{\underline{U}}_Y^\dagger$
4. Calculate the SVD $\underline{\underline{Q}}_{XY} = \underline{\underline{U}}\, \underline{\underline{\Sigma}}\, \underline{\underline{V}}^\dagger$
5. Pick the first singular value σ_1 and the first left and right singular values $\underline{u}_1$ and $\underline{v}_1$
6. The estimate of the magnitude squared coherence is σ_1^2. The canonical series $a(t)$ and $b(t)$ can be calculated from $\underline{u}_1$ and $\underline{v}_1$ using equations (5.2.13) and (5.1.1).

The advantage to using this method over estimating the spectra explicitly is that it does not involve matrix inversion, which can be inaccurate for matrices with both very large and very small eigenvalues (matrices that are close to singular).

5.4 Canonical Correlation

Canonical correlation is the same as canonical coherence, applied to real-valued time domain data. This corresponds to the replacements listed in Table 3.

Table III.3: Replacements for canonical correlation.

Variables: p is the number of signals in region X, q is the number of signals in Y, k is the number of trials times the number of tapers used in the multitaper analysis, and l is the number of trials times the number of time points in each trial.

Frequency Domain	Interpretation	Time Domain	Interpretation
$\underline{X}(f):(p\times 1)$ $\underline{Y}(f):(q\times 1)$	Data vectors as a function of frequency	$\underline{X}(t):(p\times 1)$ $\underline{Y}(t):(q\times 1)$	Data vectors as a function of time
$\underline{\underline{S}}_{XX}(f)$ $\underline{\underline{S}}_{XY}(f)$ $\underline{\underline{S}}_{YY}(f)$	Cross-spectral matrices as a function of frequency	$\mathrm{cov}(\underline{X},\underline{X})$ $\mathrm{cov}(\underline{X},\underline{Y})$ $\mathrm{cov}(\underline{Y},\underline{Y})$	Covariance matrices
$\underline{\underline{X}}(f):(p\times k)$ $\underline{\underline{Y}}(f):(q\times k)$	Data matrices as a function of frequency: rows are electrodes, columns are Fourier coefficients for each trial and taper	$\underline{\underline{X}}:(p\times l)$ $\underline{\underline{Y}}:(q\times l)$	Data matrices: rows are electrodes, columns are stacked time-domain recordings for each trial

5.5 Lagrange Multipliers

The method of Lagrange Multipliers can be used to find a local maximum or minimum of a function $f(\underline{x},\underline{y})$ subject to constraints of the form $g(\underline{x},\underline{y})=c$, where $\underline{x}$ and $\underline{y}$ are vector-valued. It introduces a variable λ (the Lagrange Multiplier) and evaluates the function:

$$\Lambda(\underline{x},\underline{y},\lambda)=f(\underline{x},\underline{y})+\lambda\cdot(g(\underline{x},\underline{y})-c) \tag{5.5.1}$$

If there exists a solution $f(\underline{x}_0,\underline{y}_0)$ to the original problem, then a stationary point of Λ will exist at $(\underline{x}_0,\underline{y}_0,\lambda_0)$ for some λ_0. However, not all stationary points of Λ correspond to solutions of the original problem. We find the stationary points of Λ by taking the gradient with respect to the free parameters $\underline{x}$, $\underline{y}$, and λ, and setting it to zero. The reason this method works has to do with finding a point where the contours of f are tangent to the contours of the constraints. This occurs when the gradients of the functions are parallel.

5.6 Complex Gradients

Complex gradients (see Schreier and Scharf (2010) Appendix 2 on Wirtinger calculus) extend the concept of gradients to complex variables. The idea is to treat the complex variable of interest $x = u + jv$ as two independent variables x and x^*, and take the partial derivatives with respect to each independently. The partial derivatives (Wirtinger derivatives) are defined as:

$$\begin{aligned} \frac{\partial}{\partial x} &= \frac{1}{2}\left(\frac{\partial}{\partial u} - j\frac{\partial}{\partial v} \right) \\ \frac{\partial}{\partial x^*} &= \frac{1}{2}\left(\frac{\partial}{\partial u} + j\frac{\partial}{\partial v} \right) \end{aligned} \tag{5.6.1}$$

These derivatives exist whenever $Re\{f(x)\}$ and $Im\{f(x)\}$ are both real-differentiable. x can be vector-valued[6] ($n \times 1$), leading to a complex gradient ($2n \times 1$):

$$\nabla_x f = \left[\frac{\partial f}{\partial \underline{x}}; \frac{\partial f}{\partial \underline{x}^*} \right] \tag{5.6.2}$$

where

$$\begin{aligned} \frac{\partial f}{\partial \underline{x}} &= \left[\frac{\partial f}{\partial x_1} \frac{\partial f}{\partial x_2} \cdots \frac{\partial f}{\partial x_n} \right]^T \\ \frac{\partial f}{\partial \underline{x}^*} &= \left[\frac{\partial f}{\partial x_1^*} \frac{\partial f}{\partial x_2^*} \cdots \frac{\partial f}{\partial x_n^*} \right]^T \end{aligned} \tag{5.6.3}$$

When looking for extrema of real-valued functions (note that our Λ from equation (5.2.3) is real-valued) it is sufficient to look for solutions just to one of $\frac{\partial f}{\partial x}$ or $\frac{\partial f}{\partial x^*}$, since:

$$\nabla_x f = 0 \Leftrightarrow \frac{\partial f}{\partial \underline{x}} = 0 \Leftrightarrow \frac{\partial f}{\partial \underline{x}^*} = 0 \tag{5.6.4}$$

[6] In Schreier and Scharf (2010) they use rows for vector-valued x ($1 \times n$) because they make use of a 2-element column vector description of x, $[x\ x^*]^T$, in the math leading up to the definitions in (5.6.1). Here I have transposed everything for simplicity in applying the theory to our case.

For our purposes it is also useful to know (Example A2.1 in (Schreier and Scharf, 2010)):

$$\frac{\partial}{\partial x} xx^* = x^*, \frac{\partial}{\partial x^*} xx^* = x \tag{5.6.5}$$

which leads trivially to[7]:

$$\begin{aligned} \frac{\partial}{\partial \underline{x}} \underline{x}^\dagger \underline{x} &= \frac{\partial}{\partial \underline{x}} \left[x_1^* x_1 + \ldots + x_k^* x_k \right] = \underline{x}^* \\ \frac{\partial}{\partial \underline{x}^*} \underline{x}^\dagger \underline{x} &= \frac{\partial}{\partial \underline{x}^*} \left[x_1^* x_1 + \ldots + x_k^* x_k \right] = \underline{x} \end{aligned} \tag{5.6.6}$$

Similar reasoning leads us to[8]:

$$\begin{aligned} &\frac{\partial}{\partial \underline{x}} \underline{x}^\dagger \underline{a} = \underline{0}, \frac{\partial}{\partial \underline{x}^*} \underline{x}^\dagger \underline{a} = \underline{a} \\ &\frac{\partial}{\partial \underline{x}} \underline{a}^\dagger \underline{x} = \underline{a}^\dagger, \frac{\partial}{\partial \underline{x}^*} \underline{a}^\dagger \underline{x} = \underline{0} \end{aligned} \tag{5.6.7}$$

Also, the product rule applies to complex gradients, so:

$$\begin{aligned} \frac{\partial}{\partial \underline{x}} (f \cdot g) &= \frac{\partial f}{\partial \underline{x}} \cdot g + f \cdot \frac{\partial g}{\partial \underline{x}} \\ \frac{\partial}{\partial \underline{x}^*} (f \cdot g) &= \frac{\partial f}{\partial \underline{x}^*} \cdot g + f \cdot \frac{\partial g}{\partial \underline{x}^*} \end{aligned} \tag{5.6.8}$$

Equations (5.6.7) lead us to some useful facts that are used in the derivations of Section 5.2:

[7] This shows up in (Schreier and Scharf, 2010) Equation A2.31, correcting for the fact that we used column vectors instead of row vectors.
[8] In (Schreier and Scharf, 2010) Equations A2.29 and A2.30.

$$\frac{\partial \tilde{C}}{\partial \underline{u}^*} = \frac{\partial}{\partial \underline{u}^*}\left(\underline{u}^\dagger \underline{\underline{Q}}_{XY} \underline{v}\right) = \frac{\partial}{\partial \underline{u}^*} \underline{u}^\dagger \left(\underline{\underline{Q}}_{XY} \underline{v}\right) = \underline{\underline{Q}}_{XY} \underline{v}$$
$$\frac{\partial \tilde{C}^*}{\partial \underline{u}^*} = \frac{\partial}{\partial \underline{u}^*}\left(\underline{u}^\dagger \underline{\underline{Q}}_{XY} \underline{v}\right)^* = \frac{\partial}{\partial \underline{u}^*} \underline{u}^T \left(\underline{\underline{Q}}_{XY} \underline{v}\right)^* = \frac{\partial}{\partial \underline{u}^*}\left(\underline{\underline{Q}}_{XY} \underline{v}\right)^\dagger \underline{u} = \underline{0}$$
$$\frac{\partial \tilde{C}}{\partial \underline{v}^*} = \frac{\partial}{\partial \underline{v}^*}\left(\underline{u}^\dagger \underline{\underline{Q}}_{XY} \underline{v}\right) = \frac{\partial}{\partial \underline{v}^*}\left(\underline{u}^\dagger \underline{\underline{Q}}_{XY}\right) \underline{v} = \underline{0}$$
$$\frac{\partial \tilde{C}^*}{\partial \underline{v}^*} = \frac{\partial}{\partial \underline{v}^*}\left(\underline{u}^\dagger \underline{\underline{Q}}_{XY} \underline{v}\right)^* = \frac{\partial}{\partial \underline{v}^*}\left(\underline{u}^\dagger \underline{\underline{Q}}_{XY}\right)^* \underline{v}^* = \frac{\partial}{\partial \underline{v}^*} \underline{v}^\dagger \left(\underline{u}^\dagger \underline{\underline{Q}}_{XY}\right)^\dagger = \frac{\partial}{\partial \underline{v}^*} \underline{v}^\dagger \left(\underline{\underline{Q}}_{XY}^\dagger \underline{u}\right) = \underline{\underline{Q}}_{XY}^\dagger \underline{u}$$

(5.6.9)

5.7 Singular Value Decompositions

The singular value decomposition of a matrix $\underline{\underline{M}}$ $(m \times n)$ is:

$$\underline{\underline{M}} = \underline{\underline{U}}\, \underline{\underline{\Sigma}}\, \underline{\underline{V}}^\dagger \tag{5.7.1}$$

where $\underline{\underline{U}}$ is a $(m \times m)$ real or complex unitary matrix ($\underline{\underline{U}}^{-1} = \underline{\underline{U}}^\dagger$), $\underline{\underline{\Sigma}}$ is a $(m \times n)$ real rectangular diagonal matrix, and $\underline{\underline{V}}$ is a $(n \times n)$ real or complex unitary matrix. The columns of $\underline{\underline{U}}$ and $\underline{\underline{V}}$ are called the left and right singular vectors of $\underline{\underline{M}}$, respectively, and the values on the diagonal of $\underline{\underline{\Sigma}}$ are called the singular values of $\underline{\underline{M}}$. The left singular vectors are the eigenvectors of $\underline{\underline{M}}\,\underline{\underline{M}}^\dagger$, and the right singular vectors are the eigenvectors of $\underline{\underline{M}}^\dagger \underline{\underline{M}}$. The square of the singular values are the corresponding eigenvalues for both $\underline{\underline{M}}\,\underline{\underline{M}}^\dagger$ and $\underline{\underline{M}}^\dagger \underline{\underline{M}}$.

Note that $\underline{\underline{\Sigma}}$ is diagonal, so if $m < n$, the columns beyond index m will contain zeros (this is the case for our $\underline{\underline{X}}$ and $\underline{\underline{Y}}$ matrices in section 5.3). In this case, the product $\underline{\underline{\Sigma}}\,\underline{\underline{V}}^\dagger$ will ignore the bottom $n-m$ rows of $\underline{\underline{V}}^\dagger$. Thus it is sometimes computationally efficient not to calculate those rows, which correspond to the rightmost $n-m$ columns

of $\underline{\underline{V}}$. In this "thin" singular value decomposition, $\underline{\underline{U}}$ is $(m \times m)$, $\underline{\underline{\Sigma}}$ is $(m \times m)$, and $\underline{\underline{V}}$ is $(n \times m)$.

5.8 Positive Definite Matrices

A symmetric square matrix $\underline{\underline{M}}$ is called positive definite if $\underline{z}^{\dagger}\underline{\underline{M}}\underline{z}$ is real and positive for all non-zero vectors $\underline{z}$. If the result is real and non-negative the matrix is said to be positive semi-definite.

Useful facts:

1. Covariance matrices are always positive semi-definite, and they are positive definite if the variables are linearly independent.
2. Positive definite matrices are invertible and the inverse is positive definite.
3. If a matrix has all positive eigenvalues, then it is positive definite (and vice versa). The eigenvalue decomposition is $\underline{\underline{P}}^{-1}\underline{\underline{D}}\underline{\underline{P}}$ ($\underline{\underline{P}}$ unitary, $\underline{\underline{D}}$ diagonal), where $\underline{\underline{D}}$ has all real positive values on the diagonal.
4. Positive definite matrices have exactly one positive definite square root, called the principal square root. If $\underline{\underline{P}}^{-1}\underline{\underline{D}}P$ is the eigenvalue decomposition of $\underline{\underline{M}}$, then $\underline{\underline{P}}^{-1}\underline{\underline{D}}^{1/2}\underline{\underline{P}}$ is the principal square root of $\underline{\underline{M}}$, where only the positive square roots have been chosen for the diagonal of $\underline{\underline{D}}^{1/2}$.
5. All positive definite matrices (and their inverses) are also Hermitian, so $\underline{\underline{M}}^{\dagger} = \underline{\underline{M}}$. Hence the principal square root of a positive definite matrix (and its inverse) is also Hermitian.

CHAPTER IV. CHARACTERIZING THE DYNAMICALLY EVOLVING FUNCTIONAL NETWORKS OF SPEECH

1. Introduction

The production of speech involves a large network of brain regions that spans several lobes of the cerebral cortex along with numerous subcortical structures (e.g., Guenther et al., 2006). Fluent speech requires very rapid movements of the tongue and other articulators. For example, a typical speaker can easily produce 10 phonemes in 1 second; this involves the precise sequencing of individual articulatory gestures that each last approximately 100 ms. The fine temporal scale of the behavior of speech makes it important to analyze the corresponding neural activity with at least as fine a timescale. Investigating the neural mechanisms of these fast movements requires a recording modality with high temporal resolution and coverage of speech related areas, hence the use of electrocorticography (ECoG), with its fine temporal resolution and wide spatial coverage, has become popular in the study of speech.

ECoG electrodes record local field potentials (LFPs) across several millimeters of cortex. The range of receptivity in ECoG signals has been estimated to encompass on the order of half a million neurons underlying each ECoG electrode (Ojemann et al., 2013). As with other LFP recordings, ECoG signals represent the superposition of two interrelated types of dynamics: oscillatory activity and population activations.

Narrow-band oscillatory dynamics are thought to coordinate activity within and between brain areas (Fries, 2005; Kopell et al., 2010; Wang, 2010; Miller et al., 2012). With regard to event-related changes in dynamics, activity has been historically

conceptualized as either “evoked” or “induced” (Pfurtscheller and Lopes da Silva, 1999): evoked activity is phase-locked to the event, whereas induced activity consists of oscillations that change in amplitude around the time of the event but with varying phase. Classic event-related potentials (ERPs) are often associated with evoked oscillations that exhibit a phase reset related to the event (Pfurtscheller and Lopes da Silva, 1999; Polich, 2007), although they typically also consist of activity that is not explicitly oscillatory (Shah et al., 2004).

Speech production exhibits many of the rhythmic phenomena that have been classically associated with motor activity in general. For example, beta frequency oscillations (15–30 Hz) are common over motor cortex, and exhibit a characteristic depression in power during movement, followed by a rebound after movement offset (Pfurtscheller and Lopes da Silva, 1999). The beta rhythm is coherent between electromyography (EMG) activity in muscles and primary somatosensory cortex, leading to a theory that it acts as a calibration signal (Baker, 2007). Other groups have proposed that the beta rhythm is an idling rhythm or that it coordinates preparatory activity, inhibitory control, or maintenance of the status quo (Engel and Fries, 2010). It has also been implicated in sensorimotor integration, top-down attentional control (Wang, 2010), or long-range communication between cortical areas (Jones et al., 2009). In reality there may be many different roles for beta in different contexts and brain areas, and there is evidence that it may be subdivided in to several smaller frequency bands (Kopell et al., 2010; Weiss and Mueller, 2012).

In the context of speech production, beta rhythms show classic depression and rebound over motor cortex during the task. They are also involved in a wide range of language-related processing across different areas of the frontal, parietal, and temporal lobes, even in the absence of a motor task (review: Weiss and Mueller, 2012). Weiss and Mueller (2012) interpret the diversity of power and coherence results as pointing to a role for different beta rhythms in (1) semantic processing of action verbs, (2) attention and violation of expectations, (3) the binding of processing in cortical language areas and (4) the maintenance of working memory.

Since the advent of ECoG there has been a focus on the observation of modulation in spectral power in the "high gamma" (>100 Hz) range during tasks (Crone et al., 2006). Recently there has been a growing consensus that broadband modulation in the high gamma range arises from changes in population spiking activity (Manning et al., 2009; Ray and Maunsell, 2011; Buzsaki et al., 2012; Miller et al., 2014)[1]. The asynchronous population spiking theoretically contributes power at all frequencies, but at low frequencies the broadband modulation is obscured by so-called "true rhythms", or narrowband phase-coherent rhythmic activity (Manning et al., 2009; Miller et al., 2014).

Hence there has been a trend toward interpreting ECoG signals as a superposition of narrowband effects, representing large-scale coordination within and between brain areas, and broadband effects, representing local population spiking activity[2]. This

[1] The connection between population firing and high gamma power is not perfect: some narrow band modulation in the high gamma range has been observed in the hippocampus, and it is interpreted as a true rhythm (Scheffer-Teixeira et al., 2013). However, no true rhythms in the high-gamma range have been observed in areas covered in this study.

[2] Note that in situations in which the population activity is rhythmic it will contribute to narrowband power, so it is important not to assume that broadband and narrowband signals are

interpretation can be very powerful, because it allows us to describe population firing at a moderate temporal resolution (10s of milliseconds) using ECoG data, which has a much broader spatial extent than typical electrophysiology. ECoG studies differ in whether they choose to quantify this effect using the full spectrum ("broadband") or power across a broad, high frequency range ("high gamma"): while we chose the former, the latter is also common, and the results and interpretations are very similar.

ECoG studies have started to use high gamma activity to map out the dynamics of cortical processing during speech. A number of studies have investigated the dynamics of high-gamma activity during spoken speech of auditorily or visually presented words (Towle et al., 2008; Edwards et al., 2010; Flinker et al., 2010; Pei et al., 2011; Leuthardt et al., 2012). Some particularly informative work uses high-density ECoG recordings. Bouchard et al. (2013) for example used a high-density grid over ventral rolandic cortex and found a somatotopic map of the speech articulators where the subareas activate in concert to precisely control the muscles necessary to form the diverse range of phonemes. Similarly, the superior temporal gyrus (STG) exhibits functional organization that exhibits a high degree of variability across subjects: Flinker et al. (2011) found spatial selectivity below 4mm for distinctions of words versus non-words and self-generated speech versus heard speech. These studies confirm the local nature of the high gamma signal, and their close relationship to intracortical electrophysiological results provides

---

mechanistically unrelated. Both are thought to result from synaptic potentials and, to a lesser extent, action potentials. Synaptic potentials can arise both from local connections and from incoming projections from other brain areas, and so contain both local and incoming activity. Mechanistic modeling of broadband/narrowband relationships will be necessary to characterize the process more fully (e.g. Miller et al., 2012), and such modeling will have to take into account the different sources of cortical rhythms.

more evidence of the connection between high gamma power and population activity, specifically in cortical areas involved in speech processing. The ECoG studies with lower electrode density but greater spatial coverage are being used to bridge these local effects with systems-level speech processing theory (Edwards et al., 2010; Flinker et al., 2010).

In addition to the studies of first-order effects, a recent focus in the study of speech processing has been functional and effective connectivity, or how the different regions involved in speech processing dynamically change their coupling patterns. Two recent fMRI investigations, for example, used effective connectivity analysis to study speech networks: the first used structural equation modeling to identify feedback influences of auditory cortex on frontal areas during perturbed speech (Tourville et al., 2008), and the second used dynamic causal modeling to identify a parsimonious model of feed-forward speech production under normal conditions (Eickhoff et al., 2009). ECoG studies have the potential to extend these results to a much finer timescale, at least over regions that are typically covered by ECoG grids.

In one early approach, Matsumoto et al. (2004) used electrical stimulation in ECoG patients to identify functional language areas (a common clinical procedure used for mapping eloquent cortex, cortical areas important for basic sensation, motor control, and communication), and then assessed connectivity by recording the effects of the stimulation on the other ECoG electrodes. They found that stimulation-identified language, face motor, and auditory areas were interconnected, and further that neighboring areas were also recruited. These connections were inferred from early and late negative potentials with unclear mechanisms: they may reflect combinations of

cortico-cortical association pathways and subcortical pathways. While this procedure does not identify task-related connectivity, it does quantify connectivity based on the strength of pathways between areas.

Korzeniewska et al. (2011) targeted speech-dependent changes in effective connectivity by assessing event-related causality time-locked to stimulus onset in the voltage signals band-passed in the high gamma range (70–115 Hz). Interestingly, the patient they studied was fluent in sign language, so they were able to compare spoken and signed responses in picture naming and auditory word repetition tasks. They found that the event-related causal networks for spoken and signed responses were very different across tasks, response modality, and time in the trials. Electrodes with task-related increases in high gamma activity tended to be source hubs of event-related causal increases, with a linear relationship between high gamma power increases and event-related causal outflows.

Another approach to identifying task-related changes in connectivity is based on the "communication through coherence" hypothesis, that cortical rhythms facilitate communication between brain areas (Fries, 2005), with different rhythms potentially subserving different types of computations (Siegel et al., 2012). If this is the case, coherence between ECoG electrodes can be used to detect communication between brain areas (Miller et al., 2012). Rather than look for direct cause and effect relationships between regions, this approach targets oscillations as a putative mechanism for cortical communication, so that by detecting coherent oscillations, communication may be

inferred. To our knowledge, this study is the first to assess large-scale coherence-based functional connectivity during speech.

The purpose of this study was to expand on the speech dynamics literature, comprehensively characterizing the temporal dynamics of ECoG signals during an overt reading task, specifically around the time of speech onset. In particular, we aim to describe both the narrow-band rhythmic changes as well as the broadband putative population signal. First, we characterize the signals on the level of individual electrodes (first-order) using event-related potentials, analyses of band-limited spectral power, and time courses of broadband power, and characterize the relative latencies of effects that time-locked to speech onset. Then we focus on pairs of electrodes (second-order), presenting dynamic time domain and frequency domain functional connectivity networks related to speech.

2. Results

The following analysis represents results from 9 patients undergoing treatment for intractable epilepsy. Details about the subjects, experimental protocol, and data analysis can be found in the Methods. The placement of the ECoG electrodes for each subject is shown in Figure IV.1. In the figure, electrodes are colored based on three anatomical regions of interest: perirolandic cortex (“rolandic”), STG (“auditory”), and anterior supramarginal gyrus (“aSMG”). We highlight these areas because they are associated with speech motor control, auditory feedback processing, and somatosensory feedback processing (Golfinopoulos et al., 2010), and therefore would be likely to have activity that is time-locked to the onset of speech.

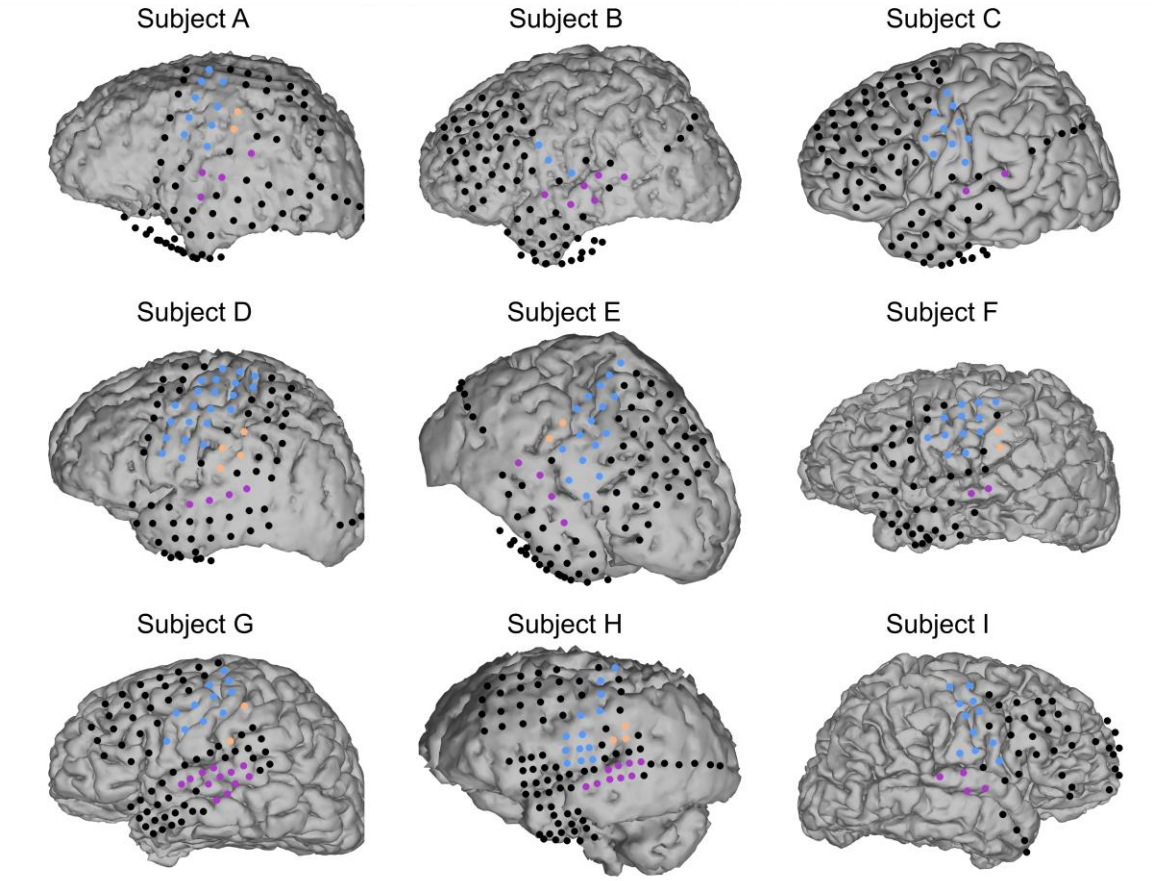


Figure IV.1: Electrode Grids for each subject.

For each subject, the electrodes used for analysis are shown coregistered with the individual subject MRI scan with the exception of Subject C which uses a standard MRI template brain (see Methods). Electrodes that were excluded due to noise or epileptic activity are not shown. Colors indicate regions of interest: Rolandic (blue), aSMG (peach), Auditory (purple), unlabeled (black). Ambiguous electrodes and electrodes lying outside of the three regions of interest were left unlabeled (see Methods).

During the experiment, subjects read aloud from famous speeches or nursery rhymes as the text scrolled across a computer screen. The slow passage of the text meant that subjects frequently paused while waiting for new text to show up, a feature which we used to define trials: trials were defined as one second intervals in which 500 ms of silence was followed by 500 ms of speech, with time 0 indicating speech onset (see

Methods for details). In addition, we chose baseline silent periods using segments of silence that were at least 500 ms away from any speech. Note that even though the first 500 ms of the trials technically contains silence, it is likely to contain preparatory activity related to the upcoming speech. Hence when we refer to "silence", we specifically mean the periods of silence that are separated from speech by at least 500 ms. We will refer to the first half of the trials as "before speech onset".

The analyses described below fall into two categories: first order (within-electrode) analyses and second order (between-electrode) analyses. Both types of analyses have frequency domain and time domain components. The first-order analyses echo both historical studies of speech processing based on EEG and more recent metrics developed for ECoG in particular. The second-order analyses extend the classical descriptions to the domain of dynamic relationships between cortical subregions. By presenting both types of analyses for the same subjects, we connect our work to previous results and provide the full context for our new observations.

2.1 First Order, Time Domain: Event-related potentials are most consistent over rolandic cortex

One classic result in motor EEG studies is the readiness potential (Kornhuber and Deecke, 1964), a negative event-related potential (ERP) over central electrodes preceding motor tasks including speech (Deecke et al., 1986; Wohlert, 1993). During voluntary movement, the negative deflection in the ERP can begin as much as two seconds before the motor movement, and can be decomposed into and early and late phase. The early phase starts about 2 s before movement onset over the pre-supplementary motor area,

extends to the supplementary motor area, and then engages premotor cortex. The late phase starts about 400 ms before movement onset in primary motor cortex (Shibasaki and Hallett, 2006). Here we confirm the existence of the late-phase readiness potential on the cortical surface over rolandic cortex, accompanied by a positive deflection in the ERP after speech onset that has been observed previously in EEG recordings (Wohlert, 1993).

Figure IV.2 summarizes the results of the ERP analysis. The detailed ERPs on each electrode for each subject are in Appendix 1. For each subject and region, the average ERP was calculated across all electrodes in the region and displayed as a z-score with respect to silence (see Methods). Hence the y-axes represent the number of standard errors away from the mean the ERPs are, under the silence distribution.

The rolandic ERPs have the same basic shape for all subjects, namely a negative deflection in the ERP starting about 100–250 ms before speech onset, followed by a zero-crossing at about 0-50 ms, and a positive deflection lasting until about 250–350 ms after speech onset. The relatively late onset of the negative deflection relative to the late-phase readiness potential can be explained by the continuous nature of the task: the readiness potential literature typically uses self-initiated movements (Kornhuber and Deecke, 1964; Deecke et al., 1986; Wohlert, 1993). Interestingly, the two subjects with right hemisphere electrodes (E and I) also had significant negative ERPs preceding speech, consistent with Wohlert (1993), who found that the readiness potential for speech is not lateralized.

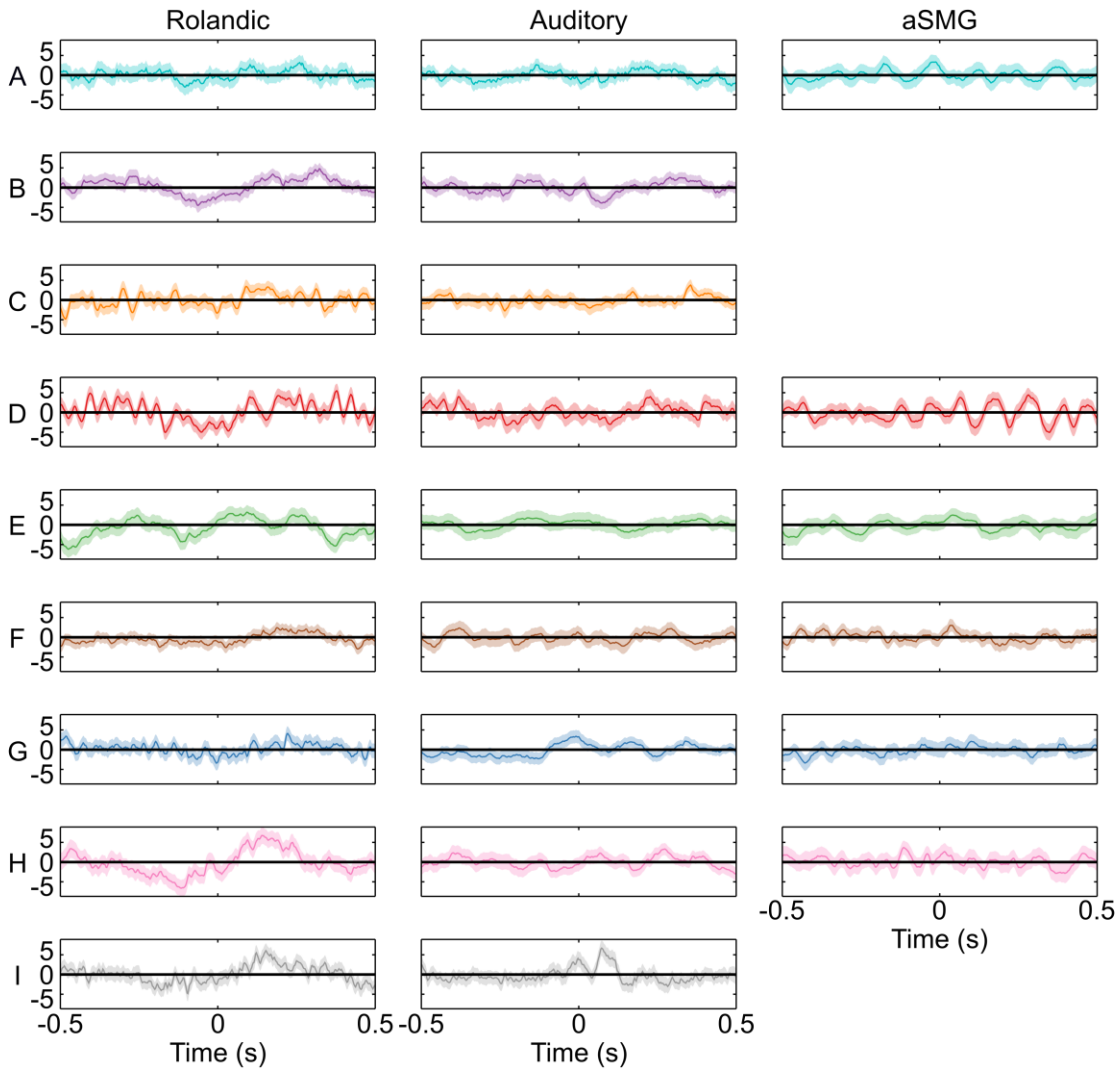


Figure IV.2: ERPs by region.

ERPs averaged over all rolandic, auditory, and aSMG electrodes for each subject, in z-score units. In order to bring the ERPs for all subjects into the same scale, the ERP values are displayed as a z-score with respect to the voltages during silence. Hence the y-axis should be interpreted as the number of standard deviations away from the mean, under the distribution during silence (see Methods). The confidence intervals in this figure are 95% confidence intervals based on the standard error during trials, rescaled to match the z-score.

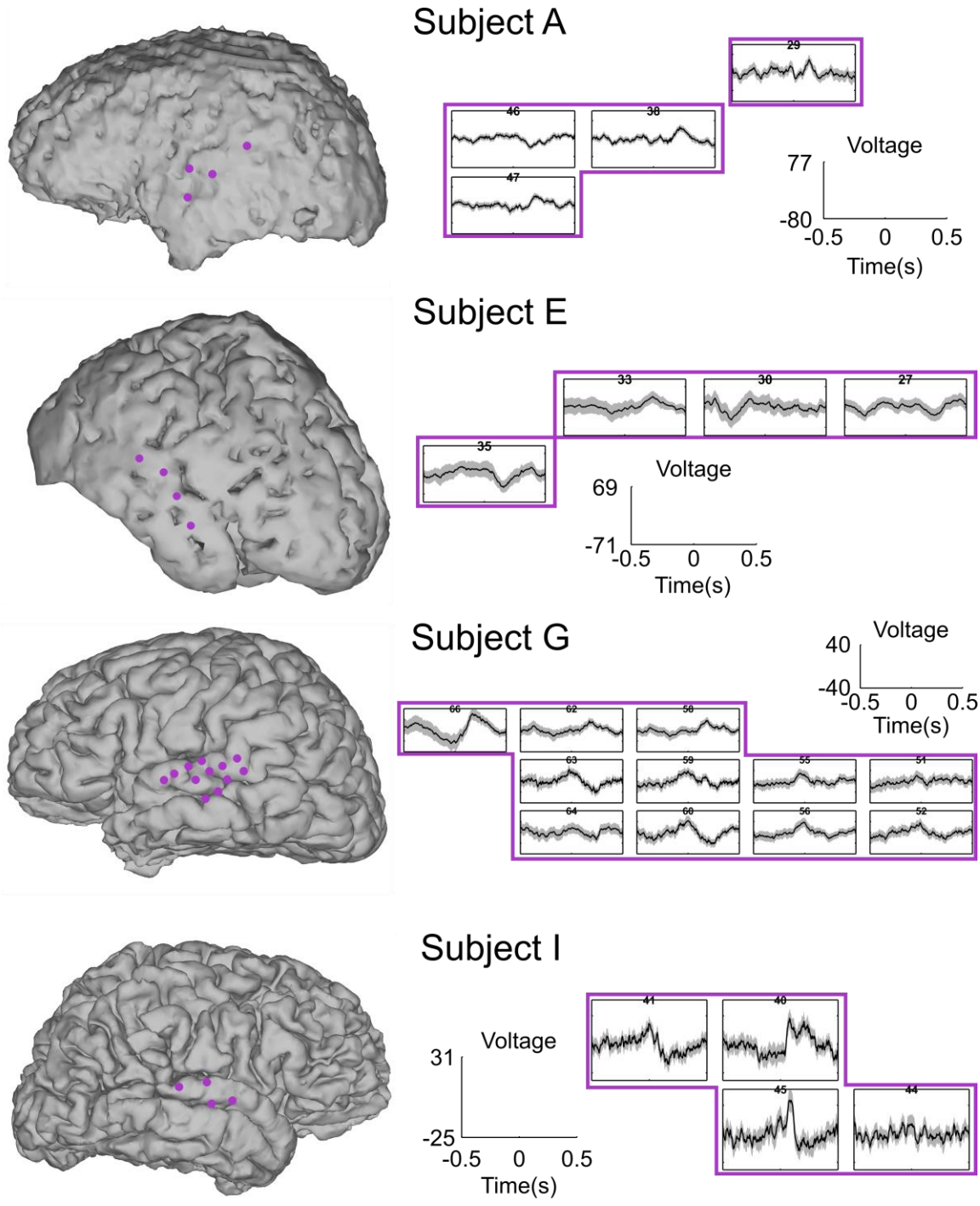


Figure IV.3 Auditory ERPs for Subjects A, E, G, and I.

ERPs for electrodes over auditory cortex. Confidence intervals are 95% confidence intervals based on the standard error during trials.

The ERPs averaged over auditory and aSMG electrodes are not as consistent across subjects. Note that averaging across electrodes in a large region can be problematic when the ERPs have different shapes. In particular, electrodes in auditory cortex had varying ERPs, which often canceled out in the average. Figure IV.3 shows the auditory ERPs for electrodes in four subjects: A, E, G, and I (ERPs for the rest of the subjects and areas are in Appendix 1). For example, Subject E showed an early negative deflection on electrode 30 over the middle STG, and a much later negative deflection on electrode 35 over the posterior STG. Conversely, Subject G has a large negative deflection and later positive deflection on electrode 66 over middle STG, with earlier positive deflections over more posterior electrodes. This variability in ERPs over auditory cortex has been observed before: Flinker et al. (2010) observed substantial spatial variability in ERPs across auditory cortex with some electrodes exhibiting consistent ERPs across trials. This variability is likely due to modular processing in subareas of auditory cortex, where nearby areas perform different functions and have correspondingly different response profiles.

Interestingly, while most subjects do not have apparent rhythmic activity in the ERP, Subject D shows fluctuations in the rolandic and aSMG ERPs that may reflect phase alignment during trials. Specifically, the rolandic region-level ERP (Figure IV.2, red trace, left) has beta band fluctuations (about 22 peaks over the course of 1 s) that are not apparent in the individual electrode ERPs (Appendix 1, Subject D, electrodes labeled "Rolandic"). This suggests that there may be beta coherence across rolandic electrodes that is not consistent in phase across trials. Additionally, Subject D has fluctuations in the

ERP for aSMG in the alpha range (Figure IV.2, red trace, right), especially after speech onset (about 5 peaks from time 0 to 500 ms). This rhythm, in contrast to the rolandic rhythm, does show up in three of the four individual electrode ERPs (Appendix 1, Subject D, electrodes 51, 80, and 83), suggesting that these electrodes have consistent alpha phase across trials. It is unclear why only one subject would have these strong rhythmic effects in the ERP.

2.2 First Order, Frequency Domain: Spectral effects are large and consistent over rolandic and auditory cortex

The first order frequency analyses consist of three parts. First, we show the spectrograms for all electrodes, which show clear beta depression and broadband enhancement over rolandic and auditory electrodes. Second, we quantify narrow-band power changes between the pre-onset and post-onset portions of the trials in traditional frequency bands (delta, theta, alpha, beta, gamma, and high gamma) in bubble plots. Third, we describe the time course of beta and broadband power changes across the trials and quantify the latency of the changes for electrodes with significant effects.

2.2.1 Spectrograms. Figure IV.4 shows the set of spectrograms for Subject D, displaying the log ratio of the power at a give time during the trial relative to the power during silence: red colors indicate greater power than silence, and blue colors indicate lower power. This subject shows two types of effects that were observed for many of the subjects, which we analyze more thoroughly in subsequent analyses. Specifically, a subset of electrodes over areas that are expected to be involved in the task show a

depression in beta power and/or a very broad increase in higher frequency (>50 Hz) power relative to silence and time-locked to speech onset. For example, electrodes 89, 63, and 90, which are over primary motor cortex (rolandic), have a beta depression and an apparently simultaneous high gamma enhancement starting about 300 ms before speech onset. Electrodes 58 and 27 do not show as strong beta modulation, but they do have a very broad increase in high gamma that starts about 150 ms before speech onset and about 100 ms after speech onset, respectively. The precise location of electrode 58 is ambiguous: it is on the border of ventral primary somatosensory cortex (rolandic) and anterior supramarginal gyrus (aSMG), two areas known to be active during speech (Golfinopoulos et al., 2010). Electrode 27 is solidly in the posterior STG, which is auditory. Note that the time resolution on these plots is 200 ms and the frequency resolution is 10 Hz (±5 Hz).

Because the increases in high frequency power are very broad (for Subject D the bandwidth is at least 120 Hz), we do not interpret them as coherent rhythms. Rather, as discussed above, there is a growing consensus among ECoG researchers that broad modulation at high frequencies reflects population synaptic and spiking activity that contributes power at all frequencies, but that is obscured at low frequencies due to narrow-band true rhythms. For example, on electrode 89 of Subject D, the strong narrow-band modulation in the beta frequency range would obscure any superimposed broadband effect.

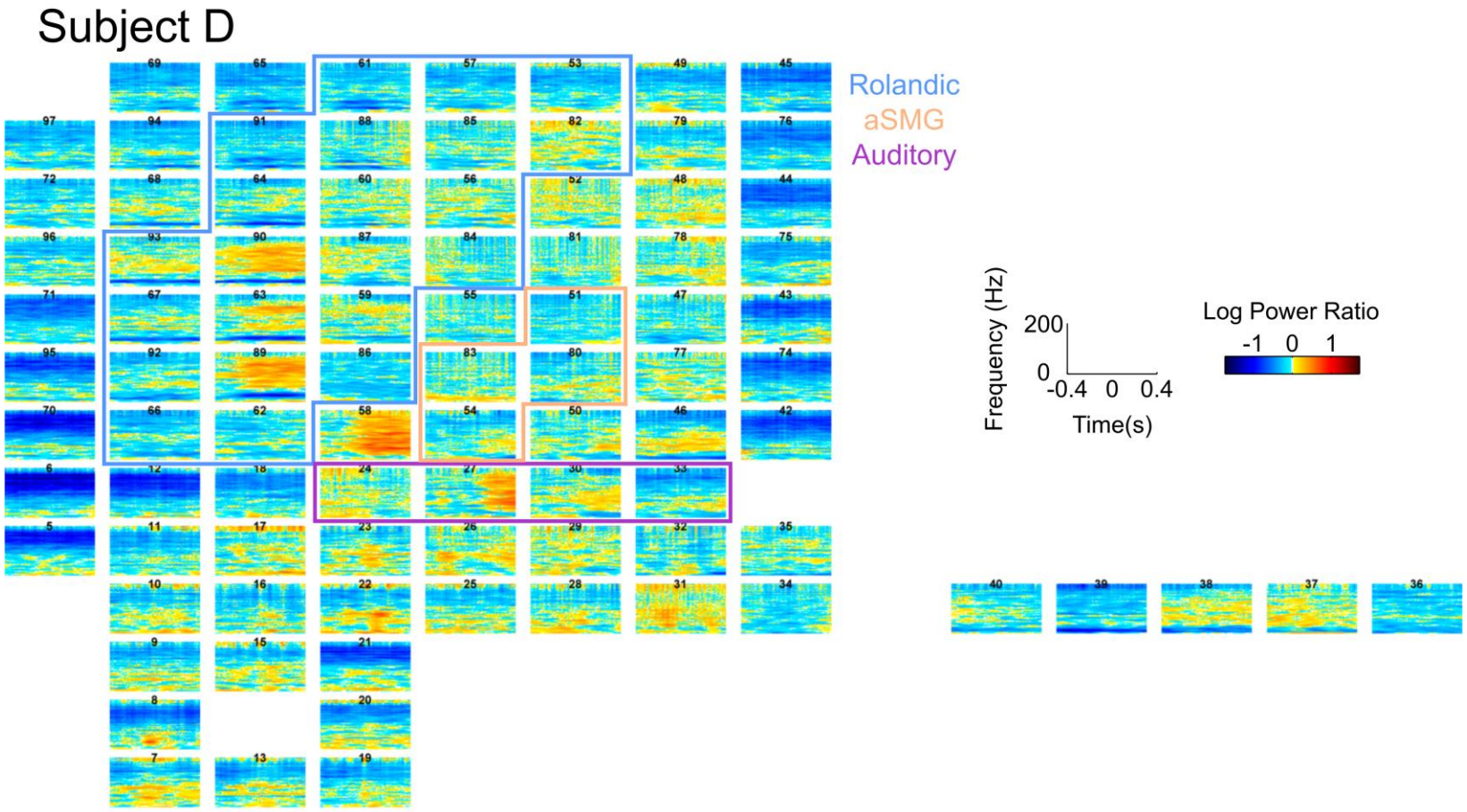


Figure IV.4: Spectrogram grid for Subject D.

For each electrode, the spectrogram is shown relative to speech onset (time 0). Spectrograms are shown in log units relative to the spectrum during silence (see Methods). Electrodes in regions of interest are bounded by colored boxes: Rolandic (blue), aSMG (peach), or Auditory (purple). Time resolution: 200 ms; Frequency resolution: 10 Hz (±5 Hz).

There are other interesting effects in Figure IV.4. For example, the broad high gamma range is negative throughout the trial for many electrodes around the edges of the grid. The fact that the blue region extends throughout the trial means that these areas have lower high gamma power during the entire trial relative to silence. This effect is curious, but it does not show up in the recordings for the other subjects, so we are tempted to view it as an artifact related to the edges of the electrode grid.

Appendix 2 contains the spectrograms for the other subjects. While the detailed timing and magnitude of the effects vary by subject, all subjects have electrodes showing spectral modulation that is time-locked to the onset of speech. Many of these electrodes show modulation that can be categorized using the two general effects described above: beta depression and broadband enhancement. We investigate the timing of this modulation in more detail below.

2.2.2 Bubble Plots. Task-related spectrograms often show modulation in a set of common frequency bands (Pfurtscheller and Lopes da Silva, 1999; Buzsaki and Draguhn, 2004). For example, here many electrodes have a depression in beta power (15–30 Hz) relative to silence that begins before speech onset and extends to the end of the trial (Figure IV.4 and Appendix 3). To identify electrodes with significant changes in common frequency bands specifically around the time of speech onset, we compared the power in the 500 ms after speech onset to the power in the 500 ms before speech onset in the delta (1–4 Hz), theta (4–8 Hz), alpha (8–12 Hz), beta (15–30 Hz), gamma (30–80 Hz), and high gamma (100–200 Hz) bands. While the frequency bands of relevant modulation may

vary by subject, brain region, and task (Pfurtscheller and Lopes da Silva, 1999; Magri et al., 2012), we find these plots informative as a summary of the more detailed spectrogram analysis.

Figure IV.5 shows bubble plots of the log ratio of the band power in predefined frequency bands between the second and first half of the trials for Subject D. The color and size of the bubbles represents the log ratio, with blue representing decreases in power for the second half compared to the firts and red representing increases. This visualization of the spectral modulation is intended to provide an anatomical perspective on the electrodes that have frequency modulation during the task. For example, large increases in high gamma power occur over the precentral gyrus, the supramarginal gyrus, and the STG. These are all speech related areas that would be expected to modulate their activity in a consistent way relative to the timing of speech onset (Golfinopoulos et al., 2010). The precentral gyrus and parts of the temporal lobe also show significant beta depression.

Because this metric compares the 500 ms before speech onset to the 500 ms after speech onset, it does not reflect effects that occur in both periods. For example, on some electrodes there is beta depression throughout the trial (see Electrode 90 in Figure IV.4), meaning the total power in the beta band is very similar before speech relative to after speech. These electrodes may not have significant beta depression according to the before/after analysis.

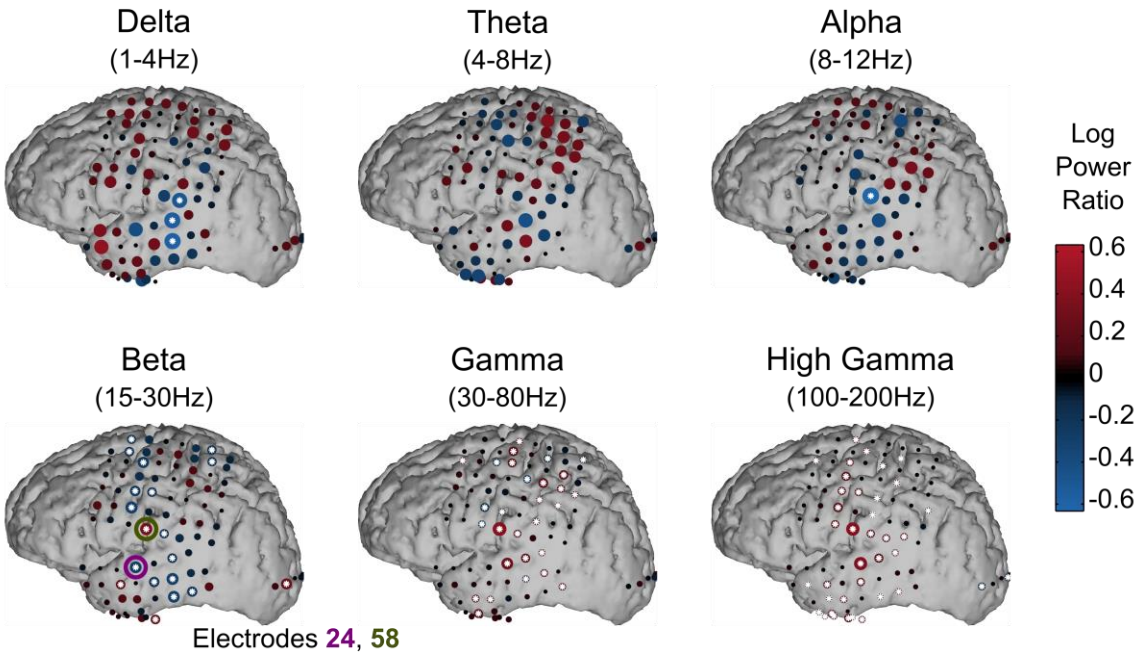


Figure IV.5: Bubble plots for Subject D.

For each frequency band of interest (delta, theta, alpha, beta, gamma, high gamma), the bubble plot represents the comparison between power in the 500 ms after speech onset and the 500 ms before speech onset. The size and color of the bubbles represents the log power ratio (after/before), and significant differences are indicated with a white asterisk (see Methods). For the beta band, electrodes 24 and 58 have been circled in magenta and green, respectively. Bubble plots for the other subjects are included in Appendix 3.

There are also electrodes that do not have striking differences relative to silence, but that do have strong modulation when comparing before speech onset to after. For example, Electrode 24 in Figure IV.4 does not have high power relative to silence in the beta band, but the difference in beta between the first half of the trial and the second half is quite strong and significant according to Figure IV.5.

It is important to take caution in interpreting the frequency bands on electrodes that have large high gamma modulation: since the high gamma modulation is unlikely to reflect true rhythms in this context, we expect it to be a marker of broadband modulation

that affects all frequencies, as discussed above[1]. This can lead to confusion at lower frequencies: for example, an increase in gamma power could reflect a true rhythm or it could simply be an extension of a broadband power increase. This effect can extend even lower: electrode 58 has a significant increase in beta power in Figure IV.5, but it is unclear if this increase is a result of a true beta rhythm or if it is an extension of the increases in the gamma and high gamma band. In fact, looking back at the spectrogram for this electrode in Figure IV.5 there is little reason to believe that the beta band is behaving differently than higher frequencies. On the other hand, electrodes over the precentral gyrus have significant decreases in beta power that cannot be attributed to broadband effects since the broadband power increases. Hence the robust beta depression that occurs over many electrodes cannot be a result of broadband contamination.

Appendix 3 contains the bubble plots for the other subjects. There are many different significant effects in different subjects, but it is notable that most subjects have beta decreases around rolandic cortex and high gamma increases around rolandic and auditory cortices.

2.2.3 Broadband and beta time courses and latency. Because there were consistent broadband and beta effects across subjects, we characterized these effects in more depth, specifically by (1) looking at finer timescale changes in the average broadband and beta and (2) estimating the latency of the changes.

[1] We chose not to try to correct the spectrograms and bubble plots for the broadband effect, as described in (Miller et al., 2009b). See Methods for a discussion of this choice.

In order to quantify the changes in broadband as a function of time, we used the method of principal spectral components proposed by Miller et al (2009b), which uses principal component analysis (PCA) to decompose a normalized spectrogram into linear combinations of frequencies that covary. A main result of that study was that the component of the spectrum that explains the most variance of the signal during a motor task is almost completely flat across frequencies. This effect was robust across electrodes and subjects in that study and generalized to other tasks (Miller et al., 2014), and we had the same result when applying the technique to our data. The authors interpreted this "broadband" component as resulting from arrhythmic neuronal population activity changes during the task, which are commonly observed in the high gamma range. For a summary of the technique, see the Methods.

The broadband traces, averaged across trials, are shown for each subject in Appendix 4. Noticing that many electrodes showed an increase in the broadband signal during trials, we wanted to quantify the latency of this increase. First, we selected electrodes with significantly greater broadband power in the second half of the trials compared to the first half (0 to 500 ms and -500 to 0 ms, respectively). In this way we hoped to capture electrodes with a wide range of latencies.

In order to quantify the latency of the broadband change, we fit a sigmoid function to the average trace using nonlinear regression (see the Methods for details). These fits are shown superimposed on the broadband traces for electrodes with significant increases in broadband (Appendix 4). The latency was defined as the inflection point of the sigmoid function (the red vertical line in the figures).

Figure IV.6 summarizes the broadband latency analysis results. Within subjects, the latency of the broadband increase was always shorter for rolandic electrodes than for auditory electrodes. Across subjects there is quite a bit of variability in the timing of these changes relative to speech onset, but rolandic electrodes had latencies either before speech onset or in the first 100 ms after speech onset, and auditory electrodes had latencies exclusively after speech onset. There were no aSMG electrodes with a significant broadband change, so it is impossible to draw conclusions about the latency of aSMG using this metric. The mean latency for rolandic electrodes was 52 ms before speech onset and the mean latency for auditory electrodes was 92 ms after speech onset. The difference between the rolandic and auditory latencies was significant (two-sided, two-sample t-test with unequal variances; $p<0.0001$).

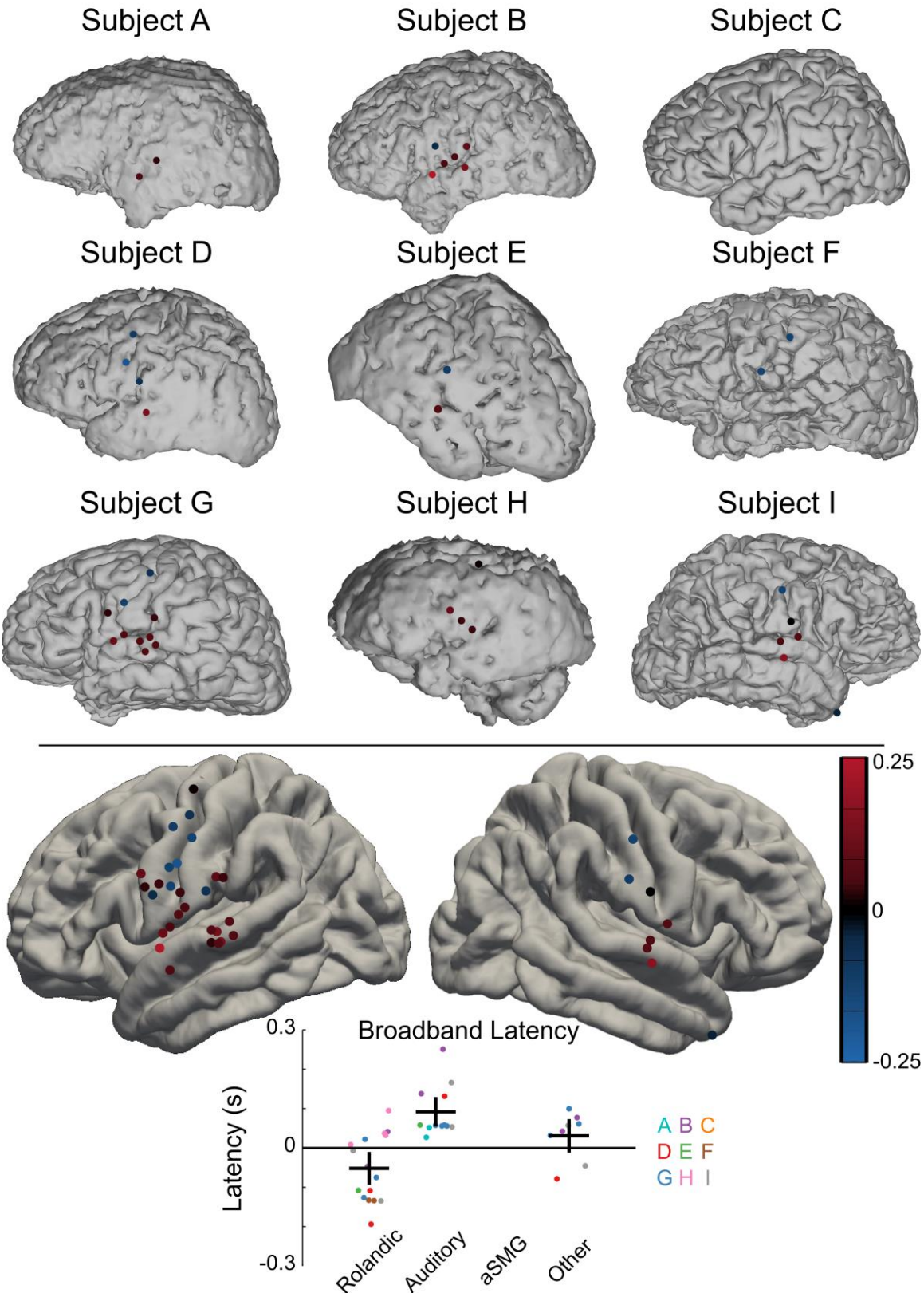

Subject A
Subject B
Subject C
Subject D
Subject E
Subject F
Subject G
Subject H
Subject I
0.25
0
-0.25
Broadband Latency
Latency (s)
0.3
0
-0.3
A B C
D E F
G H I
Rolandic
Auditory
aSMG
Other

Figure IV.6: The latency of the broadband increase.

A latency analysis was performed for electrodes with significant increases in broadband power between the 500 ms before speech onset and the 500 ms after speech onset (see Methods and Appendix 4). The upper panel shows these electrodes, colored by the broadband onset latency. The lower panel aggregates results across subjects: showing electrodes from all subjects on the Freesurfer average surfaces (above), and a scatterplot of the latencies by region (below), with mean latency and 95% confidence intervals shown in black.

Aside from the broadband effect, many electrodes also showed beta depression during the trials. Hence we performed a similar analysis to compute the average beta power as a function of time in the trials and to quantify the latency of the beta depression. Appendix 5 contains the average log beta power as a function of time in the trials for all subjects (see Methods for details). Many electrodes showed a beta depression before speech onset, which in some cases recovered by the end of the trials. Note that this recovery in the beta power is different from the well-studied post-movement beta rebound (Pfurtscheller and Lopes da Silva, 1999), because speech offset did not occur until after the end of the trials. In order to capture the beta depression, we defined electrodes with significant beta depression by comparing the first 200 ms of the trials (-500 to -300 ms) with the second-to-last 200 ms (100 to 300 ms). Then, we performed the same sigmoid fitting as the broadband latency analysis but leaving out the last 200 ms of the trials. As can be seen in the fits (Appendix 4), this setup plausibly identifies electrodes with beta depression and quantifies the latency of the depression. Another approach could have been to fit a more complicated model with both a depression and a recovery, but we found the current method to be adequate for our purpose (i.e. quantifying the latency of the beta depression).

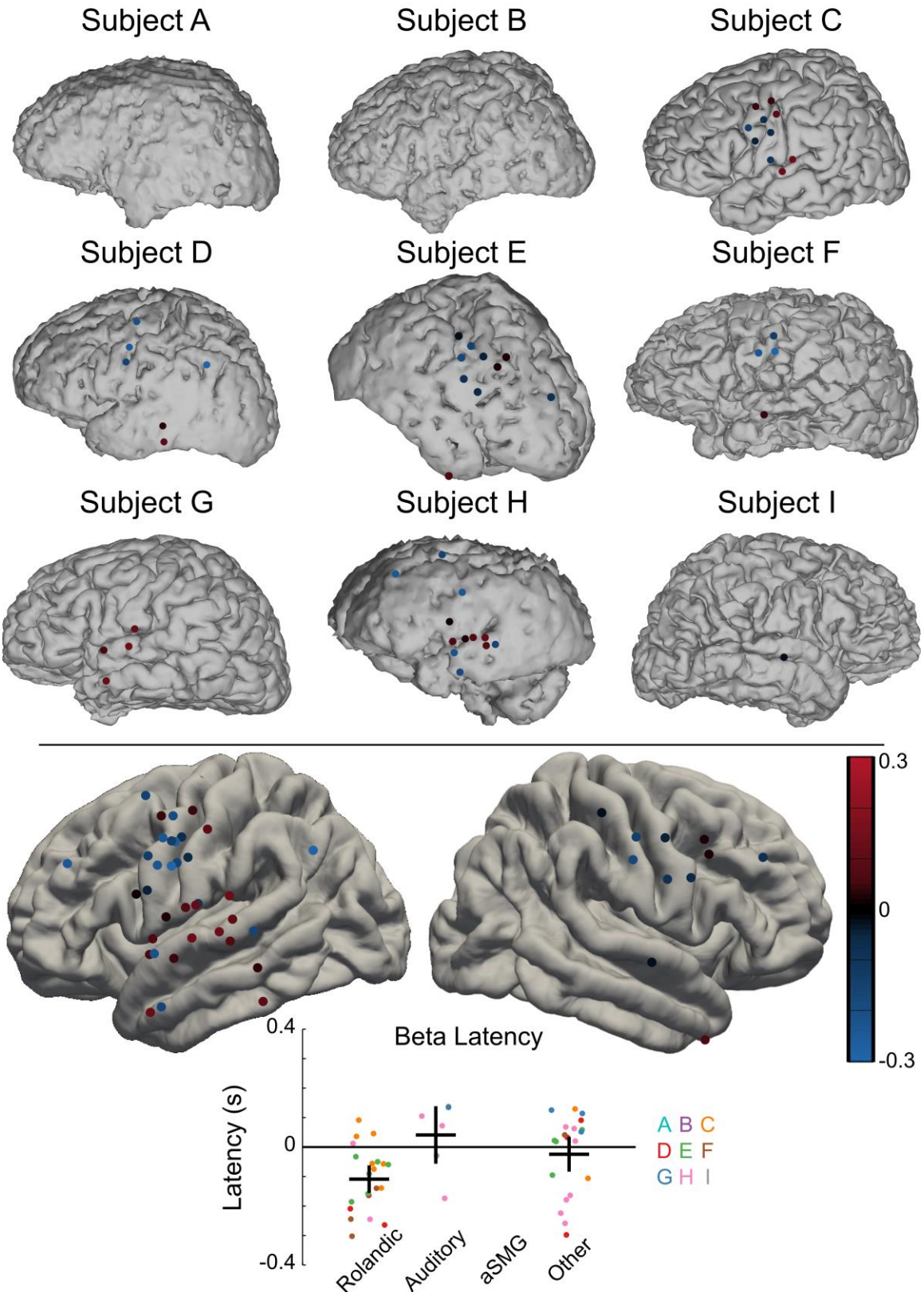

Subject A
Subject B
Subject C
Subject D
Subject E
Subject F
Subject G
Subject H
Subject I
0.3
0
-0.3
Beta Latency
Latency (s)
0.4
0
-0.4
A B C
D E F
G H I
Rolandic
Auditory
aSMG
Other

Figure IV.7: The latency of the beta depression.

Electrodes with significant beta depression between the first 200 ms of trials (-500 to 300 ms) and the second-to-last 200 ms (200 to 400 ms) were included in the beta latency analysis. Layout as in Figure IV.6.

Figure IV.7 summarizes the beta latency analysis. The latencies for the beta depression had greater variability than the latencies of the broadband increase. In particular, while rolandic electrodes tended to have earlier beta depression than auditory electrodes, even within subjects there were some exceptions. For example, Subject H (pink) had one auditory electrode with a very early beta depression (almost 200 ms before speech onset) and one rolandic electrode with beta depression starting after speech onset. These extreme values may be a result of poor fit, as the beta depression is very small for both of these electrodes (Appendix 5, Subject H, electrodes 70 and 67). The average beta latency on rolandic electrodes was 113 ms before speech onset, and the average latency for auditory electrodes was 41 ms after speech onset. There is more overlap between the two groups than there was for the broadband latency analysis, but the difference between the groups still achieved significance (two-sided, two-sample t-test with unequal variances; $p<0.05$). No aSMG electrodes showed significant beta depression.

2.3 Second order: Two subjects show significant changes in correlations between electrodes during speech that may be associated with changes in narrowband coherence

Having characterized the dynamic behavior of individual electrodes and found consistent time- and frequency-domain effects, we turn to relationships between electrodes. If there are robust changes in the coupling of LFPs at the scale of population activity, it may be possible to detect them through network analysis. With that in mind,

we inferred functional connectivity networks based on the correlations and coherences between electrodes, specifically trying to detect pairs of electrodes whose correlations or coherences during speech were significantly greater or less than silence. Unfortunately, the resulting networks were not consistent across subjects. However, we present here the detailed results for two subjects that have interesting coupling dynamics during speech, in the hope that these results may inspire further attempts to pin down large-scale patterns of speech-related coupling (Figures IV.8, 9, and 10). This is by nature a very high-dimensional exploratory analysis, so while we have made an effort to distinguish significant effects from spurious ones, any results will need to be replicated before they should be taken too seriously. For completeness, we also present summaries of the region-level correlation and coherence dynamics (Appendix 7) and electrode-level summaries of the correlation networks (Appendix 6). In addition, to make the high-dimensional second-order effects more digestible, we developed an interactive visualization of all of the subjects' networks online, SpectraVis. We encourage the reader to use the interactive visualization while reading the text to aid comprehension (see Section 2.4 for a description, http://ericdeno.com/research/SpectraVis).

Networks were constructed using a 200 ms rectangular sliding window over the course of the trials. Test statistics were variance-stabilized measures of the difference in the correlation (or coherence) between the trial window and silence. Under the null hypothesis that there is no difference between speech and silence, the test statistic should be normally distributed with mean 0 and standard deviation 1. Positive and negative edges were inferred based on a two-sided test with a FDR of 5% over electrode pairs

(note that we did not correct for multiple comparisons over time or frequency). The complete network inference methodology is described in the Methods.

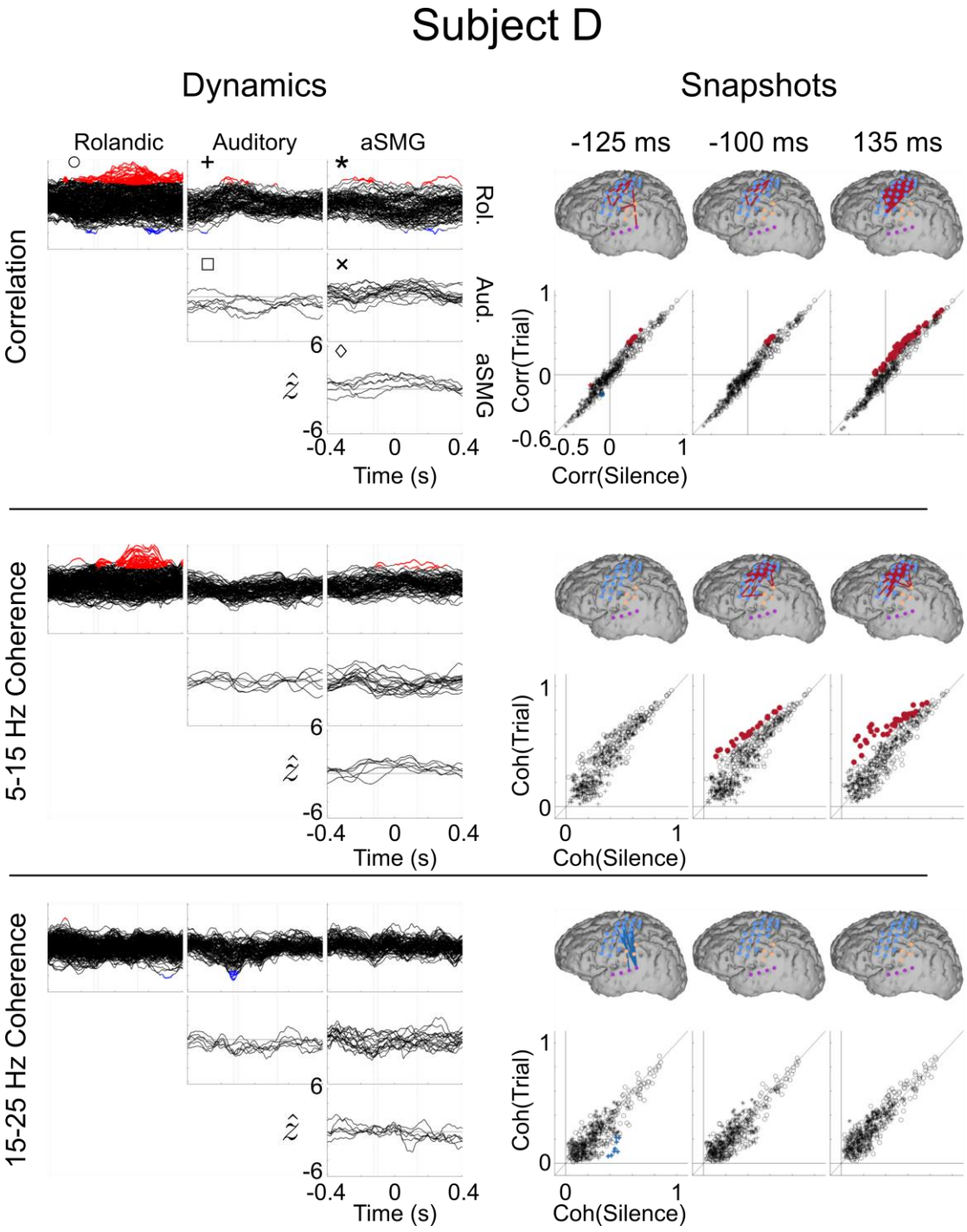


Figure IV.8: Summary of second-order effects for Subject D.

The correlation networks, shown on the top row. In the right column (labeled “Snapshots”), the networks are shown at three time points in an anatomical view with significant edges indicated as red (greater correlation during speech) or blue (greater correlation during silence). In order to convey the effect sizes involved, below each network is a scatter plot of the correlation during silence versus during speech, for all electrode pairs, with color indicating significance. The left columns (labeled “Dynamics”) show the time courses of the test statistic for every pair of electrodes, grouped by their incident regions. For example, the top left figure shows all pairs of electrodes where both electrodes are over rolandic cortex. The y-axis is a z-score for a test of the difference in correlation between speech and silence: positive values mean the correlation was greater during a 200 ms window centered at that time point in speech compared to silence. Time points of significance are shown in red and blue. The symbols in the upper left of the plots are a legend for the correlation scatter plots: for example, a circle represents a rolandic-rolandic electrode pair. The other rows show the same set of figures for coherence networks at selected frequency bands (note that the frequency resolution of the analysis was 10 Hz).

The two subjects that we discuss in detail are D and E (Figures IV.8 and IV.9). Subject E has both positive and negative rolandic-rolandic correlation changes. These changes are dramatically split, with positive correlation changes over dorsal rolandic cortex and negative changes over ventral rolandic cortex. This dorsal-ventral split may be related to the positioning of the electrodes with respect to the central sulcus: all of the positive edges are between electrodes on or very near the precentral gyrus, primary motor cortex (M1); the negative edges, in contrast, either cross the central sulcus or connect electrodes in the postcentral gyrus, primary somatosensory cortex (S1). Evoked potentials show a voltage reversal across the central sulcus (Wood et al., 1988), which could lead to negative correlations between electrodes on either side of the sulcus. A first hint that there may be two different mechanisms at play, however, is that the positive deflections precede the negative deflections: this can be seen both in the time courses and in the snapshots, where at 70 ms before speech onset there are already positive edges over dorsal rolandic cortex, but the negative edges do not become significant until the 35 ms snapshot. Further evidence that these are distinct effects comes from the coherence

networks: the 5–15 Hz coherence networks have positive edges over dorsal rolandic cortex that largely overlap with the positive correlation edges, while the 20–30 Hz coherence networks have negative edges over ventral rolandic cortex that largely overlap with the negative correlation edges. Note that these effects are graded: while we show the positive/negative edge classifications for convenience, the trajectories of the correlation and coherence values occur at intermediate values as well.

The results for Subject E suggest (1) an increase in theta or alpha coherence starting about 150 ms before speech onset primarily over M1 and (2) a loss of coherence at beta frequencies in S1 and between S1 and M1, occurring right at speech onset, roughly coincident with the beta power decrease. It is worth noting that the electrodes for this subject are over the right hemisphere and the subject's left hemisphere was dominant for language, so the effects may be related to the non-dominant hemisphere.

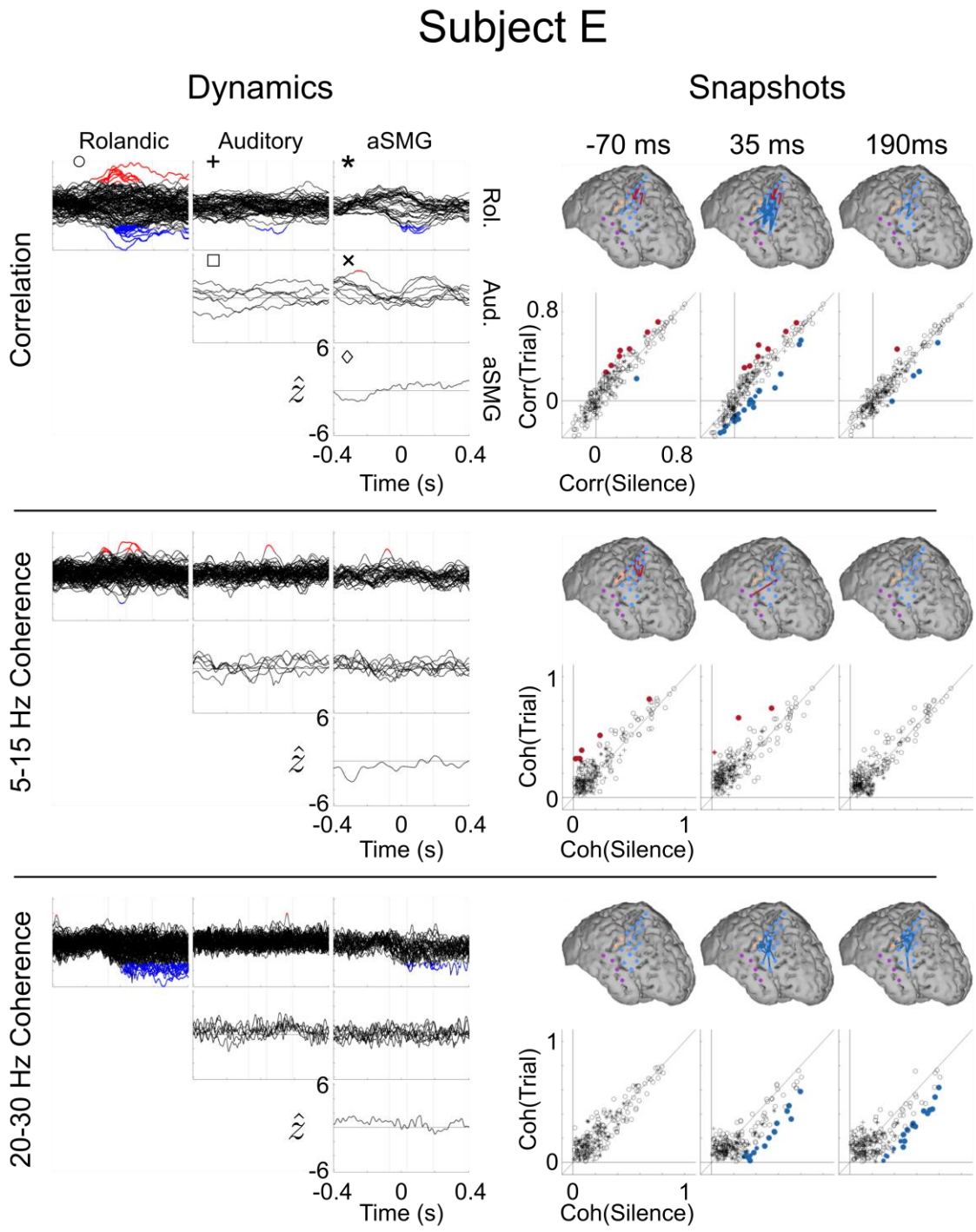


Figure IV.9: Summary of second-order effects for Subject E.

Layout as in Figure IV.8.

Moving on to Subject D (Figure IV.8), there are similar positive correlation changes to Subject E, but there are fewer negative changes. The positive edges form a connected component that at first (-300 to -45 ms) includes primarily M1-M1 connections, then starting at -45 ms some M1-S1 connections are included. At 135 ms after speech onset the positive correlations form a dense cluster between 15 electrodes in M1 and S1. Interestingly, the S1 electrodes in the cluster correlate with many M1 electrodes but they do not correlate significantly with each other. Consistent with Subject E, the positive correlations co-occur and overlap with positive coherence changes in the 5–15 Hz band. The cluster of positive 5–15 Hz coherence edges includes M1-S1 connections earlier than for correlation, with S1 nodes being pulled in as early as 100 ms before speech onset, around the same time that the cluster initially forms. These patterns (viewable interactively in SpectraVis) lend further support to the interpretation that theta or alpha coherence increases in M1 prior to speech onset, with additional coherence with parts of S1.

There are no significant rolandic-to-rolandic decreases in beta coherence for Subject D. However, there is a short interval from 150 ms until 100 ms before speech onset in which some rolandic electrodes (both M1 and S1) decohere with electrodes 30 and 33 in auditory cortex (specifically pSTG) in the beta band (15–25 Hz). These connections reflect a decrease in coherence from about 0.4 during silence to less than about 0.2 during speech. While the loss of coherence during speech is expected given the motor-related depression in beta power, the high coherences during the baseline silence

are interesting and may be related to attention or working memory processes (Weiss and Mueller, 2012) between sensorimotor and auditory cortices.

The region-level networks for Subjects D and E (Figure IV.10) generally reflect the inter-region effects from the electrode-level networks (Figures IV.8 and IV.9). There are few inter-region connections, with the exception of a time period about 150 ms until 50 ms before speech onset in Subject D when the test statistic for canonical correlation between rolandic and auditory cortex is significantly negative, meaning that the canonical correlation between trials was significantly lower than silence. This effect occurs at the same time as a negative trend in the test statistic for the beta band (15-25 Hz) canonical coherence, so it may be related to a loss in beta coherence between rolandic and auditory cortex prior to speech onset. This is consistent with the electrode-level observations, where there were a few edges between rolandic and auditory cortex that had significantly lower beta coherence in the same time period relative to silence. The fact that the result extends to the canonical test statistics suggests that it may reflect a change in the structure of coupling between rolandic and auditory cortex on a large scale, rather than spatially specific coupling between sub-regions in the two areas.

For completeness, we include in the appendix summaries of the correlation networks of all subjects (Appendix 6). While they tend to follow the same general trends as subjects D and E, the results are not as dramatic. Furthermore, the few electrodes that have significant correlation changes for these subjects do not also have significant coherence changes (results not shown), making it difficult to trace the origin of the time domain network results.

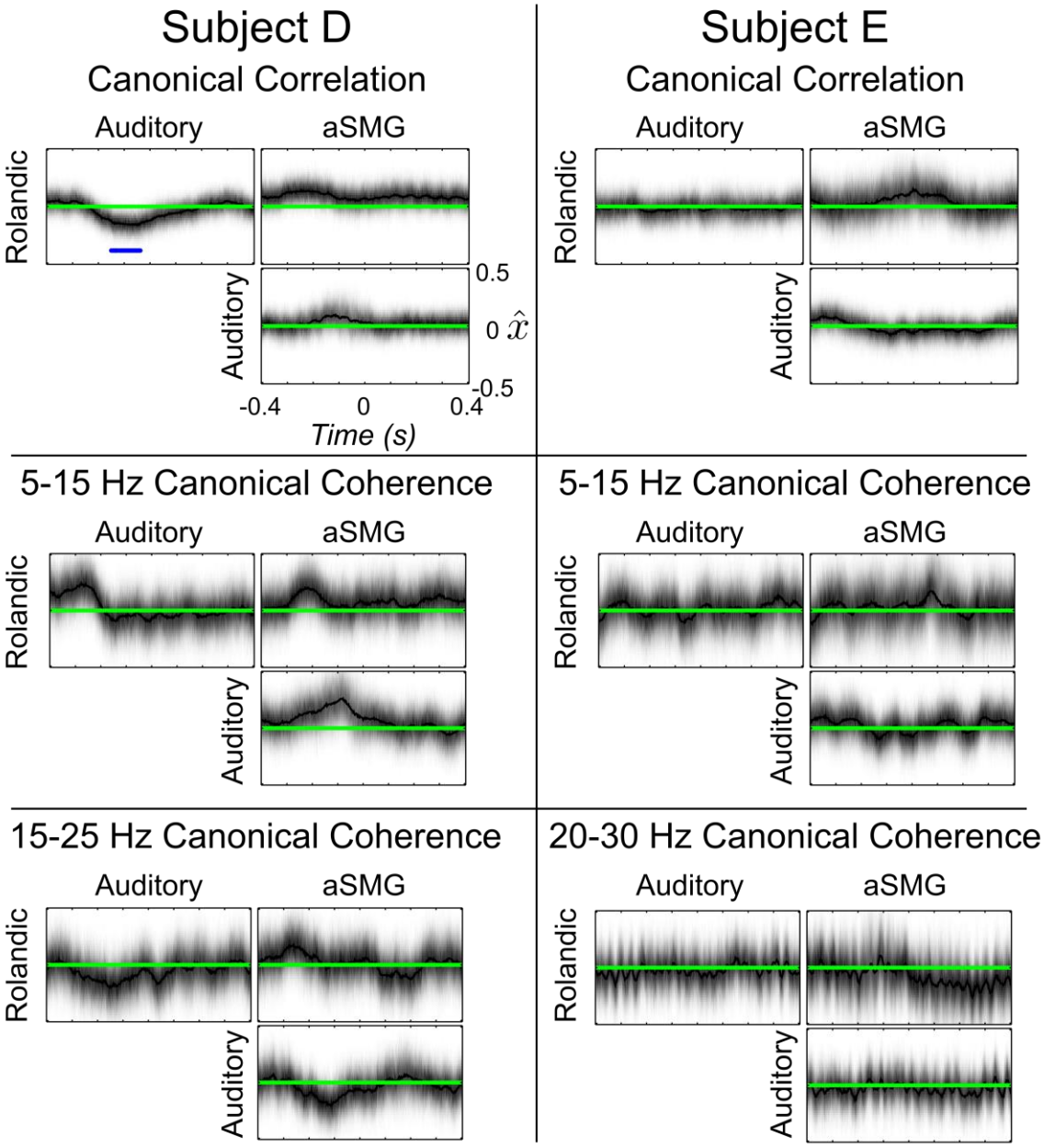


Figure IV.10: Region-level second-order effects for Subjects D and E.

Region-level coupling time courses for Subject D (left column) and Subject E (right column), computed using canonical correlation (first row) and canonical coherence (second and third rows). Time courses are shown between all pairs of regions (Rolandic-Auditory, Rolandic-aSMG, and Auditory-aSMG) over the course of trials (time 0 is speech onset). The estimated distribution of the test statistic ($\hat{x}_{CCorr}$ for canonical correlation and $\hat{x}_{CCoh}$ for canonical coherence) is shown in greyscale as a function of time, and significant deviations from baseline (silence) occur where this distribution deviates significantly from zero, marked with a green line (see Methods 4.5). In

Subject D, the canonical correlation between rolandic and auditory cortex during trials is significantly less than silence for a part of the interval: the timing of this effect is indicated in blue.

2.4 SpectraVis

In addition to the static display of the networks, we developed an interactive tool to represent the high-dimensional network data in a more digestible way. The tool, SpectraVis, shows the time-domain and frequency domain network results for all subjects, dynamically linked to the spectrograms and coherograms across all edges. A user can display networks in a binary format or weighted by test statistic or raw correlations, using either an anatomical display or a force-directed graph diagram. In addition, the tool can be used to play movies of the networks over time. This tool is open-source and available online at http://ericdeno.com/research/SpectraVis. Figure IV.11 shows a screenshot of the tool.

3. Discussion

This study presents a comprehensive analysis of the dynamics of neural activity surrounding speech onset. The first order effects largely confirm and expand findings from previous studies, firstly by linking time-domain findings with classic EEG phenomena and secondly by replicating spectral results from more recent ECoG work in speech. The second order effects represent new results regarding the dynamic coupling between populations during speech production.

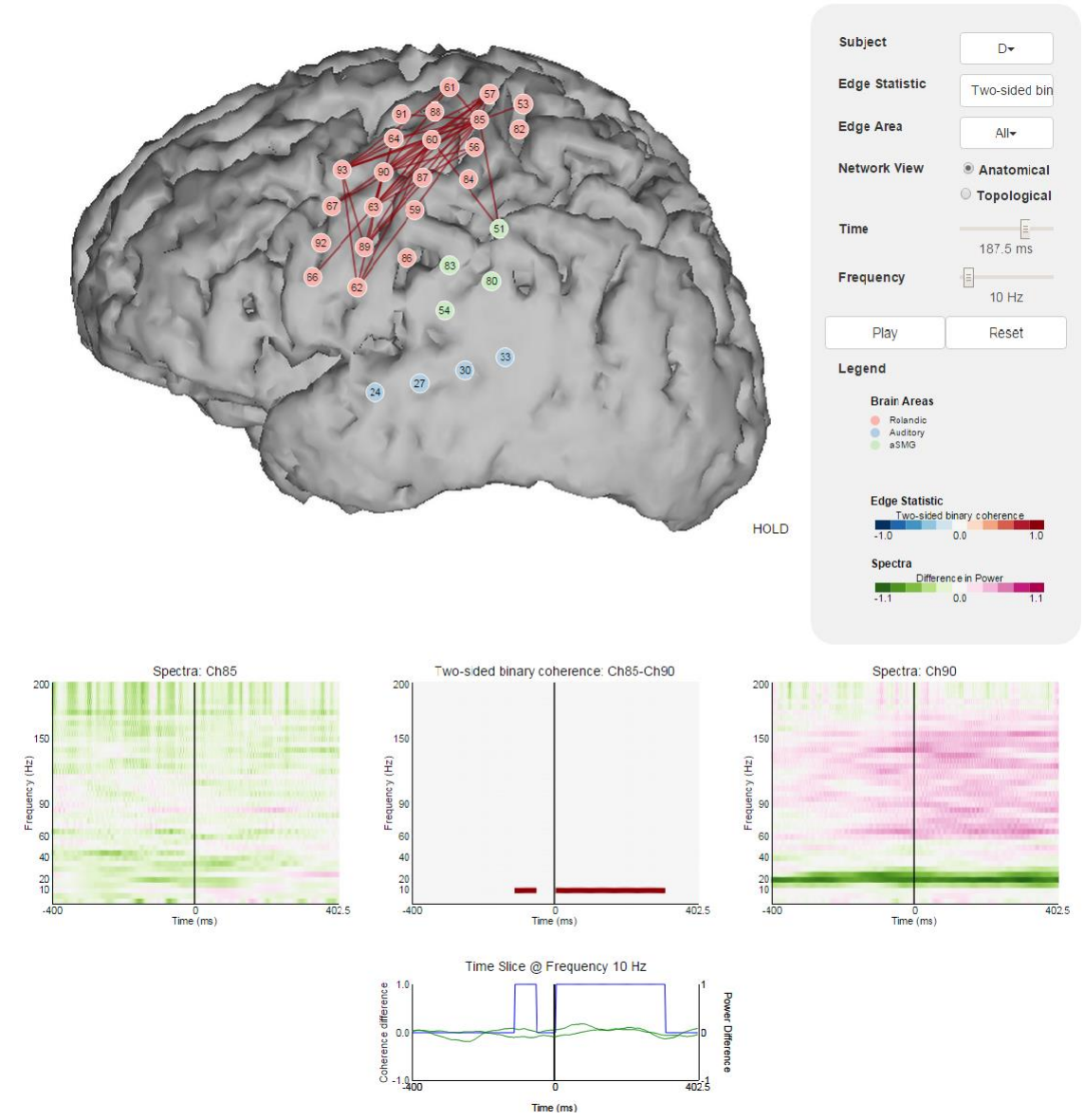


Figure IV.11: SpectraVis Screenshot.

3.1 First Order Effects

Across subjects, ERPs in rolandic areas have a negative deflection before speech onset followed by a positive deflection after speech onset. Auditory cortex did not show strong ERPs in the average, but individual electrodes had ERPs that varied significantly

over auditory cortex. The difference in consistency between rolandic and auditory ERPs may be related to functional differences: sub-areas of primary motor cortex are likely to activate with similar latencies relative to speech onset because of the need to coordinate effector muscle engagement. Electrodes labeled as auditory in this study, on the other hand, cover sub-areas that are heterogeneous in function and therefore are likely to have different activation profiles.

One of the main results of this study is the estimation of the latency of spectral effects with respect to speech onset time. In particular, broadband activity over rolandic electrodes increased on average 52 ms before speech onset, and over auditory electrodes on average 92 ms after speech onset. With the interpretation that broadband activity reflects population firing activity, the motor latency is consistent with theory: Guenther et al. (2006) predict that a motor command from primary motor cortex will take about 40 ms to induce a response in the speech articulator muscles. On the other hand, Edwards et al. (2010) observed high gamma ECoG activity over rolandic cortex in a verb generation task as early as 300 ms before speech onset, and as early as 550 ms before speech onset for a picture naming task. These earlier response latencies may reflect greater task difficulty due the semantic nature of the tasks.

The auditory broadband latency also matches previous results: the verb generation study of Edwards et al. (2010) observed latencies to the peak high gamma response over the STG of about 150–500 ms after speech onset. The lack of auditory population activity immediately after speech onset is likely due to the auditory areas covered by this study: the ECoG grids did not enter the Sylvian Fissure, so the electrodes are unlikely to pick up

much primary auditory activity. Rather, all subjects had electrodes over STG, which would be expected to have longer response latencies due to its role in higher-level auditory processing. According to the DIVA model, the pSTG is the location of the Auditory Error Map, which compares an auditory target signal to the incoming auditory signal (Guenther et al., 2006).

A beta depression during speech is expected, given that our baseline was active: beta rhythms are most prominent in motor cortex during motor preparation, and desynchronize during movement (Pfurtscheller and Lopes da Silva, 1999; McFarland et al., 2000). The fact that the beta depression preceded the broadband is notable: on electrodes with a significant beta depression, the average rolandic latency was 113 ms before speech onset and the average auditory latency was 41 ms after speech onset. This is consistent with a conception of the beta dynamics as signaling spatially broad control mechanisms which set conditions for local processing (Baker, 2007; Jones et al., 2009; Engel and Fries, 2010; Wang, 2010; Weiss and Mueller, 2012).

3.2 Second Order Effects

By necessity we have been cautious in drawing conclusions from the second-order effects: looking for dynamic effects across electrode pairs and frequencies is a very high-dimensional task. One way to manage the high dimensionality is to facilitate exploration: to this end, we developed an interactive network visualization tool, SpectraVis, which links the binary and weighted time- and frequency-domain networks to the corresponding spectrograms and coherograms. SpectraVis is available online at

http://ericdeno.com/research/SpectraVis, and the source code is hosted at https://github.com/edeno/SpectraVis for general use.

Beyond visualizing the high dimensional results in a more intuitive way, we have made an effort to manage the dimensionality (1) by inferring networks only over electrodes in regions of interest, (2) by controlling the false discovery rate of the networks, and (3) by characterizing the uncertainty of the region-level averages. It is reassuring that the effects we report in the main text occur (1) in subjects who have a large number of trials and silent periods, which should lead to higher statistical power, and (2) over regions where the subjects have good spatial coverage. Possibly the subjects who do not display such striking second-order effects are simply lacking in statistical power or spatial coverage.

The network analyses are also limited by resolution: in order to have a time resolution as small as 200 ms, the frequency resolution of the analysis was limited to 10 Hz (±5 Hz). As a result, we cannot distinguish between the theta and alpha bands, and present networks for the combined 5–15 Hz frequency band. For effects that last longer than 200 ms, such as the 5–15 Hz rolandic-rolandic coherence in Subject D, it should be possible to improve the frequency resolution by decreasing the time resolution to 500 ms, which would give a 4 Hz frequency resolution (±2 Hz), enough to distinguish the two bands. This would not be possible for Subject E, however, where the 5–15 Hz coherence changes are more transient. We leave this multi-scale analysis to future work, in which we hope to have more trials for each subject and hence greater statistical power.

For now, we can combine our observation with previous classic results, which point to robust alpha rhythms over rolandic cortex, commonly referred to as “mu” rhythms (Pfurtscheller and Lopes da Silva, 1999). Mu rhythms were originally observed in EEG recordings, but have since been reported in ECoG (Arroyo et al., 1993; Toro et al., 1994; Szurhaj et al., 2003). There is evidence that the EEG mu rhythm has a lower peak frequency in epileptic subjects than healthy subjects, possibly due to anticonvulsant medications (Kuhlman, 1978; Derambure et al., 1997), and this low mu (5–12 Hz) is confirmed in some of the ECoG studies with epileptic patients (Arroyo et al., 1993; Szurhaj et al., 2003).

Jones et al. (2009) propose a feed-forward thalamic source for rolandic alpha rhythms based on previous findings that cortical alpha is coherent with alpha oscillations in the thalamus (review: Hughes, 2005). This could explain the dense connectivity observed in the 5–15 Hz coherence networks over rolandic cortex, since a common thalamic driver of rolandic alpha frequency oscillations would show up as dense pairwise connectivity. On the other hand, it is notable that the increase in alpha coherence in both Subject D and Subject E is not accompanied by an increase in power: the power in the 5–15 Hz band is remarkably constant throughout the trials (if anything it decreases, as would be expected based on classic studies of the mu rhythm), and is very close to the alpha power during silence. Hence the coherence at 5–15 Hz must correspond to the rhythms on the electrodes coming into phase. This is not true of the beta coherence effect: beta loses coherence at the same time that the beta power drops.

Interestingly, the strongest second-order effects occurred within regions rather than between them. It is possible that this is due to proximity and volume conduction, although ECoG electrodes have relatively small regions of receptivity that are thought to be dominated by local sources (Zaveri et al., 2009). Also, there were many long-distance within-ROI connections and few short between-ROI connections, despite the existence of nearby electrodes across region boundaries. For example, the correlation network for Subject D has dense connectivity that extends all the way from ventral to dorsal rolandic cortex, with only sparse connections to the nearby aSMG electrodes. Another possibility is that the low frequency oscillations involved in the task span the region such that they are coherent across space. This would result in all-to-all pairwise coherence, but there may be more direct ways to detect such a process. For example, global coherence analysis quantifies spatially coherent oscillations across many sensors as the proportion of variance explained by the first principal component of the cross-spectral matrix (Mitra and Bokil, 2008; Cimenser et al., 2011). For example, using global coherence we could quantify the tendency of all rolandic electrodes to cohere spatially at beta frequencies.

The beta rhythm could be spatially coherent across rolandic cortex without having precisely the same phase at each location: for example, (Rubino et al., 2006) found evidence of beta-frequency traveling waves in motor cortex during movement planning. The observed traveling waves had consistent but non-zero phase relationships across motor cortex, with a consistent directionality of propagation. Because the latency and amplitudes of the traveling waves reflected task parameters, the authors proposed that the wave structure of the beta could reflect information transfer within rolandic cortex.

It is notable that Subject D had a decrease in beta frequency coherence on several electrode pairs between rolandic and auditory cortex prior to speech onset. This may point to coherence between the two regions during silence, which is lost during the motor-related beta desynchronization over rolandic cortex at speech onset. Supporting this interpretation, the canonical correlation was significantly lower during the same time period relative to silence, and this effect corresponded to a trend in the same direction for canonical coherence. Hence the effect may represent a spatially broad restructuring of the beta coherence between rolandic and auditory cortex prior to the onset of speech.

In addition to replicating the current results in a more controlled paradigm, future work will aim to elucidate the fine-timescale dynamics of effects that have so far been limited to fMRI analyses. For example, previous work in our group has developed a computational model of speech production called DIVA that exhibits realistic compensatory responses to auditory or somatosensory perturbations (Guenther et al., 2006; Tourville et al., 2008). Analyzing ECoG recordings during speech in the presence of such perturbations could test the dynamics of our theoretical model and direct us toward improvements. The broadband signal is a good place to start, since it is the component of the ECoG signal that is most closely coupled with fMRI BOLD signals (Ojemann et al., 2013). In addition, the model could be made more dynamically realistic by incorporating the rhythmic changes that occur during speech.

4. Methods

4.1 Subject information and experimental protocol

Twelve patients undergoing treatment for intractable epilepsy participated in this study. As a part of their treatment, electrocorticographic grids were implanted subdurally to monitor neural activity over the course of several days. During this period, subjects gave informed consent to participate in this study, which was approved by the local institutional review boards. Three subjects were eliminated from analysis due to experimental noise or excessive interictal activity over most electrodes, and the data from the remaining nine subjects, five female and four male, was subsequently analyzed. Table 1 summarizes the clinical profiles of these subjects.

Figure IV.1 shows the electrode grid placement for all subjects, obtained via preoperative MRI and postoperative CT scan (Kubanek et al., 2013). For one subject, C, the individual MRI scan was unusable so the electrodes were coregistered with the "Colin Brain". An anatomy expert (J Tourville) used the coregistered 3D images for each subject to define the anatomical regions underlying each electrode, with special emphasis on rolandic cortex (primary motor and primary somatosensory), auditory cortex (STG), and the anterior supramarginal gyrus (SMG). Note that the coregistration of the electrodes with the individual subject MRI scans has some uncertainty, related to differences in the brain morphology between the preoperative MRI scan and the postoperative CT scan (Hermes et al., 2010), which could lead to ambiguous or incorrect region labeling in some cases. In cases where an electrode appeared close to a border between regions, it was left unlabeled. In addition, J Tourville created a manual representation of the

electrodes for all subjects on the Freesurfer average surface template (e.g. the bottom images in Figure IV.6), placing the electrodes for all subjects in their approximate anatomical locations on the template brain. Note that due to poor surface reconstruction from the MRI scans, it was impossible to do this step automatically. These approximate locations on the template brains were used only for display purposes.

Table IV.1: Summary of Subject Clinical Profiles

Subject ID	Age	Sex	Hand.	Lang. Dom.	Grid Hem.	Seizure focus	# trials	# silences (200 ms)	# silences (500 ms)
A	29	F	R	L	L	Left Temporal	63	43	15
B	30	M	R	Bi	L	Left Temporal	38	88	27
C	29	F	R	L	L	Left Temporal	85	154	46
D	56	M	R	L	L	Left Temporal	98	202	59
E	26	M	R	L	R	Right Temporal	75	194	62
F	45	M	R	L	L	Left Temporal	86	45	13
G	29	F	R	Bi	L	Left Temporal	77	273	87
H	25	F	R	Bi	L	Left Temporal	83	72	23
I	25	F	R	Bi	R	Right Frontal	50	114	41

Neural signals were recorded using g.tec g.USBamp amplifiers (sampling rate 9600 Hz). Data acquisition and stimulus presentation were handled using BCI2000 software (Schalk et al., 2004).

Further details about the experimental paradigm and data collection can be found in (Lotte et al., 2015): four of the subjects described here are also described in that study. The experimental paradigm and data collection were the same for all data described here

and in (Lotte et al., 2015), but the trial definitions and data analysis differ. The data from another subject described here, D, was also described in (Stephen et al., 2014).

4.2 Task and Trial definitions

Subjects performed an overt reading task in which they read aloud from a historical speech or nursery rhyme as it scrolled across a computer screen. Subjects performed 2–3 of these sessions (consisting of one entire speech or nursery rhyme, about 5 minutes), and one session was selected for analysis for each subject based on absence of artifacts or epileptic activity. The session was then manually checked for artifacts, and electrodes or time periods with substantial artifacts or interictal activity were eliminated from subsequent analysis.

Because the words scrolled slowly across the computer screen, subjects paused during the performance of the task while they waited for the next word or phrase to appear. We used these pauses to impose a post-hoc trial structure on the data: trials were defined as any time during the session when at least 500 ms of silence was followed by at least 500 ms of speech. Note that the particular word spoken varied by trial. Time during the trial is described relative to speech onset, so trials extend from -500 ms to +500 ms. To identify such trials, the audio recording was first used to automatically identify time periods of silence as well as voicing onset times. Specifically, the smoothed root mean square audio intensity was obtained using a 20 ms Hamming window, and this smoothed intensity was thresholded based on the intensity of the background noise.

These automatically generated voicing onset times were then manually verified and refined to reflect the times of speech onset. For example, in a word starting with an

unvoiced fricative or stop consonant, voicing onset can occur over 100 ms after speech onset (Lisker and Abramson, 1967). Speech onset times on such trials were adjusted manually using the Praat software package (www.praat.org) by observing the audio signal in the time and frequency domains while listening to the signal around the speech onset time. Trials were also manually screened to ensure that they included 500 ms of silence followed by at least 500 ms of continuous speech.

In addition to trials, time periods of silence were also identified in the signal based on the audio recording. In order for silences to be included in downstream analyses, they needed to be separated from speech by at least 500 ms before and after. Because the baseline silent periods were during the performance of the task, they should not be interpreted as an inactive baseline: rather, these silences represent an active baseline in which the subject had finished reading the words on the screen was waiting for more words to come up. Hence, for example, the baseline likely contains activity related to vision and attention.

4.3 Preprocessing of the ECoG data

The neural recordings were preprocessed by downsampling from 9600 to 400 Hz (Matlab decimate function, which in this case also low-pass filtered the signal below 200 Hz), high-pass filtering above 1 Hz, and notch filtering at 60, 120, and 180 Hz with a 1 Hz band stop. The high-pass and notch filters used a zero-phase 3rd order Butterworth filter (Matlab filtfilt function). The signals were then referenced to the common average by subtracting the mean over all electrodes. For five subjects (A, B, F, G, and H) the

common average reference was applied to the two ECoG grids separately to account for differences in the shared noise between the grids.

4.4 First-order Analyses

4.4.1. Time Domain. Event-related potentials (ERPs) were calculated by averaging the preprocessed voltages for each electrode across trials (Appendix 1). Also displayed are 95% confidence intervals based on the standard error of the mean.

In addition to the ERPs on the individual electrodes, Figure IV.2 shows the ERPs averaged over all rolandic, auditory, and SMG electrodes for each subject. In order to bring the ERPs for all subjects into the same scale, the ERP values are displayed as a z-score with respect to the voltages during silence. That is, the mean voltage during silence was subtracted from the ERP and then this differenced signal was divided by the estimated standard error during silence:

$$ERP_R(t) = \frac{\frac{1}{MN_R}\sum_{e=1}^{N_R}\sum_{i=1}^{M} x_{e,i}(t) - \hat{\mu}_s}{\hat{\sigma}_s / \sqrt{MN_R}}$$

where M is the number of trials, N_R is the number of electrodes in region R, and $\hat{\mu}_s$ and $\hat{\sigma}_s$ are the mean and standard deviation of the voltage during silence across all electrodes in region R. In this way, the y-axes of Figure IV.2 can be interpreted as the number of standard errors away from the mean, under the silence distribution. The confidence intervals in this figure are 95% confidence intervals based on the standard error during trials, rescaled to match the z-score.

For all of the analyses described below, the (unnormalized) ERPs were first subtracted from each trial on each electrode (See Chapter III for a justification of this step).

4.4.2. Frequency Domain

4.4.2.1 Spectrograms. The first-order (single electrode) frequency-domain analyses were based on spectrograms of the signal, averaged over trials (Figure IV.4 and Appendix 2). Before computing the spectrograms, the ERP was subtracted from the trials and the DC component was removed from each time window. The multitaper method (Thomson, 1982) was used to estimate spectrograms with a sliding window of 200 ms, a time-half-bandwidth product of 1 Hz*sec, and 1 taper. The resulting spectrograms have a time resolution of 200 ms and a frequency resolution of 10 Hz (±5 Hz). The spectrograms are displayed relative to the average spectrum on the electrodes during silence:

$$P(t, f) = \log\left(S_{trials}(t, f)\right) - \log\left(S_{silence}(f)\right)$$

where $S_{trials}(t, f)$ is the estimated power spectrum for the 200 ms window centered at time t relative to speech onset (t ranges from -400 to 400 ms), and $S_{silence}(f)$ is the average power spectrum during silence, where the silences are broken up into non-overlapping segments of length 200 ms.

In an effort to distinguish true rhythms from broadband power changes due to population activity, Miller et al (2009b) proposed a method of estimating spectrograms with broadband effects removed. While we used their method for quantifying the broadband timecourses (see Section 4.4.2.3), in our dataset we found it difficult to

completely remove the broadband effect from the spectrograms without disrupting the narrowband modulation[2]. This is likely related to the fact that electrodes with narrowband modulation (e.g. beta depression) often have broadband enhancement at about the same time. As a result, it was difficult to remove the broadband effects from the spectrograms, so we chose to show here spectrograms without broadband effects removed.

4.4.2.2 Bubble plots. As with the spectrograms, we used the multitaper method to perform the bubble plot analysis. Using 500 ms windows (the first and second half of the trials), we manipulated the choice of time-half-bandwidth product to produce estimates of the power in the frequency bands of interest during the two time windows. For example, to estimate the beta power we used a time-half-bandwidth product of 3.75 Hz*sec (with 6 tapers), resulting in a frequency resolution of 15 Hz (±7.5 Hz). The spectrum estimate at 22.5 Hz therefore reflected power between 15 and 30 Hz[3]. We then computed the ratio of the power after speech onset to the power before speech onset, for each electrode and each frequency band.

Figure IV.5 and Appendix 3 show the results of this analysis. The size and color of the bubbles in the figure represent the log of the power ratio. To test the significance of the effect size at each electrode, an f-test was used based on the fact that multitaper spectral estimates are constructed from the sum of Chi-squared distributed spectral

[2] Specifically, while the first spectral component typically did not contain substantial narrowband activity and was therefore suitable as an estimate of the broadband modulation, subsequent spectral components often contained a mix of narrowband and residual broadband effects. See Section 4.4.2.3 for a definition of principal spectral components.
[3] The timeseries were zero-padded before spectral estimation in order to obtain estimates in the centers of the frequency bands of interest.

estimates, each with 2 degrees of freedom (Thomson, 1982). Using this fact, it can be shown that the ratio of two multitaper estimates is f-distributed with parameters $2KM_1$ and $2KM_2$, where K is the number of tapers, M_1 is the number of trials used for the numerator, and M_2 is the number of trials used for the denominator. White stars indicate electrodes with significant differences under a two-sided test, corrected for a False Discovery Rate (FDR) of 5% across electrodes within each band (split evenly between positive and negative effects). One characteristic of this test is that it has greater statistical power for broad frequency bands since they were estimated using more tapers. Hence the high gamma band test required a very small effect size for significance. Gamma and beta required larger effect sizes, and delta, theta, and alpha required the largest effect sizes for significance. The precise threshold varied by subject, since each subject had a different number of trials.

4.4.2.3 Broadband timecourses and latency analysis. We quantify the broadband signal using the method of principal spectral components developed by Miller et al (2009b). Here we briefly summarize the method of principal spectral components, as it was used in this study to quantify the broadband signal as a function of time. Using the preprocessed session data, trials were split into five 200 ms non-overlapping windows (-500 to -300 ms, -300 to -100 ms, -100 to 100 ms, 100 to 300 ms, and 300 to 500 ms). Additionally, the silences were split into 200 ms-long windows. All of the windows were pooled while keeping track of their identity (i.e. silence or time in trial). The DC component was subtracted from each window, and the spectrum $P(f,\tau_q)$ was calculated

using the multitaper method with 1 taper and a time-half-bandwidth product of 1 Hz*sec. Here τ_q indexes the time of the 200 ms windows being analyzed, of which there are N_q (which is five times the number of trials plus the number of silence windows). The mean spectrum over all windows was then divided from each individual spectrum and the log was taken (notation is from (Miller et al., 2009b)):

$$\tilde{P}(f,\tau_q) = \ln\left(P(f,\tau_q)\right) - \ln\left(\frac{1}{N_q}\sum_{p=1}^{N_q} P(f,\tau_p)\right)$$

PCA was then used on these normalized spectra. In particular, the covariance matrix between frequencies was estimated:

$$C(f,f') = \sum_{\tau_q} \tilde{P}(f,\tau_q)\tilde{P}(f',\tau_q)$$

and the eigenvectors and eigenvalues of this matrix were computed. An interesting feature of our data, which is consistent with (Miller et al., 2009b) and true for all electrodes and subjects, is that the eigenvector with the largest eigenvalue (the eigenvector which explains the most variance), e_1, is positive and generally flat across frequencies. We interpret this eigenvector as representing spectral changes due to broadband effects.

To estimate the amount of broadband activity in the signal as a function of time, we estimate a spectrogram of the entire session $P(f,t)$ using wavelet analysis (using a complex Morlet wavelet), divide out the mean spectrum over the entire session and take the log yielding $\tilde{P}(f,t)$, then project the normalized spectrum onto the first eigenvector:

$$B(t) = \sum_f e_1(f)\tilde{P}(f,t)$$

The resulting scalar timeseries, the first spectral component, represents the broadband component of the signal as a function of time in the session (projections onto the other eigenvectors are also spectral components, numbered by their corresponding eigenvectors). We used this signal to calculate the average broadband as a function of time in the trials. These average traces are shown for all subjects in Appendix 4, with 95% confidence intervals based on the standard error of the mean.

The broadband latency analysis was performed on electrodes with significantly greater broadband power in the second half of the trials compared to the first half (0 to 500 ms and -500 to 0 ms, respectively). We assessed significance using a two-sided permutation test using 5000 permutations, corrected for a FDR of 5%.

For the electrodes with significantly greater broadband in the second half of the trials, we fit a sigmoid function to the broadband trace using nonlinear regression (Matlab function nlinfit). In particular, we modeled the broadband timecourses on each electrode using the following function:

$$B(t) = C_1 + \frac{C_2 - C_1}{1 + e^{-k(t-t_0)}}$$

This model has four parameters: C_1, the broadband at the start of the trial, C_2, the broadband at the end of the trial, k, the rate of increase of the transition, and t_0, the mid-point of the transition, which we interpret as the latency of the transition. The resulting fits and their corresponding latency estimates are shown superimposed on the broadband timecourses in Appendix 4.

4.4.2.4 Beta timecourses and latency analysis. Beta power time courses were computed using a sliding window of length 200 ms. For this analysis we used the multitaper method with a time-half-bandwidth product of 1.5 hz*sec and 2 tapers to get beta estimates within the 15–30 Hz frequency band. To be consistent with the broadband power analysis we then took the log of the signal. We used this signal to calculate the average beta power as a function of time in the trials.

The beta latency analysis had the same structure as the broadband latency analysis, except that many electrodes had a recovery in the beta power towards the end of the trials, which obscured the timing of the initial beta depression under our simple sigmoid model. Since we were primarily interested in the latency of the beta depression we modified the analysis to use a time window that mostly avoided the beta recovery. In particular, we defined electrodes with significant beta depression by comparing the first 200 ms of the trials (-500 to -300 ms) with the second-to-last 200 ms (100 to 300 ms), using an FDR-corrected f-test as described above for the bubble plots. Then, when fitting a sigmoid function to the mean beta timecourses, we limited the fits to the window -500 to 300 ms.

4.5 Second-order Analyses

The second-order (between-electrode) analyses consisted of correlation and coherence networks inferred between electrodes in the three regions of interest: rolandic, auditory, and aSMG. As described above, electrodes that were clearly over one of these three regions were identified manually based on their anatomical location in the co-

registered MRI images. The network inference methodology was described previously in (Stephen et al., 2014 and Chapter 3).

4.5.1. Time Domain: Correlation networks. First, we inferred time-domain correlation networks on a 200 ms sliding window over the course of the trials (Chapters II and III). Because correlations can be contaminated by power in frequencies below the inverse of the window length (Leonardi and Van De Ville, 2015), we modified the preprocessing of the neural signals for this analysis, high-pass filtering above 5 Hz (1/200 ms) rather than 1 Hz. All other preprocessing was the same as for the first-order analyses, including the notch filters at 60, 120, and 180 Hz.

The correlations between all pairs of electrodes for each 200 ms window during trials, averaged over trials, were compared to the average correlation on that pair of electrodes during 200 ms epochs of silence. The coupling statistic was the variance-stabilized difference between trials and silence used in Chapter III. In particular, $\hat{z}_{corr}$ between two electrodes x and y when the sliding window is centered at time τ relative to speech onset is:

$$\hat{z}_{corr}(x,y,\tau) = \frac{\tanh^{-1}\left(\rho_{speech}(x,y,\tau)\right) - \tanh^{-1}\left(\rho_{silence}(x,y)\right)}{\sqrt{\hat{\sigma}^2_{speech}(x,y,\tau) + \hat{\sigma}^2_{silence}(x,y)}}$$

The correlation $\hat{\rho}_{speech}(x,y,\tau)$ is calculated as:

$$\hat{\rho}_{speech}(x,y,\tau) = \frac{\sum_{l=1}^{L}\sum_{t=-T/2}^{T/2} x(\theta_l+\tau+t)\,y(\theta_l+\tau+t)}{\sqrt{\left(\sum_{l=1}^{L}\sum_{t=-T/2}^{T/2} x(\theta_l+\tau+t)^2\right)\left(\sum_{l=1}^{L}\sum_{t=-T/2}^{T/2} y(\theta_l+\tau+t)^2\right)}}$$

where T is the duration of the sliding window (200 ms), τ takes on values ranging from -400 ms to 400 ms, and θ_l are the speech onset times (the trial midpoints). The calculation of $\hat{\rho}_{silence}(x,y)$ is analogous, except that there is no need for the sliding window:

$$\hat{\rho}_{silence}(x,y) = \frac{\sum_{m=1}^{M}\sum_{t=-T/2}^{T/2} x(\theta_m+t)\,y(\theta_m+t)}{\sqrt{\left(\sum_{m=1}^{M}\sum_{t=-T/2}^{T/2} x(\theta_m+t)^2\right)\left(\sum_{m=1}^{M}\sum_{t=-T/2}^{T/2} y(\theta_m+t)^2\right)}}$$

Here θ_m are the midpoints of the non-overlapping 200-ms long epochs of silence. The variances $\hat{\sigma}^2_{speech}$ and $\hat{\sigma}^2_{silence}$ were estimated using a two-sample jackknife procedure as described in Chapter III (Arvesen, 1969).

Under the null hypothesis of no difference in correlation between the trial epoch and silence, this test statistic is distributed approximately as a standard normal (see Chapter III), with mean 0 and standard deviation 1. P-values for the two-sided alternative hypothesis that the test statistic was non-zero were computed by comparison to the standard normal, and binary edges were defined using a FDR of 5% over edges (note that we did not correct for false discoveries across time).

4.5.2. Frequency Domain: Coherence Networks. The structure of the frequency domain networks was the same as the time domain networks, except that the test statistic was based on the coherence between electrodes. In particular, the test statistic was a bias-corrected, variance-stabilized difference between the trial and silence coherence (Bokil et al., 2007):

$$\hat{z}_{coh}(x,y,\tau,f)=\frac{\tanh^{-1}\left(\left|C_{speech}(x,y,\tau,f)\right|\right)-\frac{1}{2LK-2}-\left(\tanh^{-1}\left(\left|C_{silence}(x,y,f)\right|\right)-\frac{1}{2MK-2}\right)}{\sqrt{\hat{\varsigma}^2_{speech}(x,y,\tau,f)+\hat{\varsigma}^2_{silence}(x,y,f)}}$$

where f indexes frequency, K is the number of tapers used to estimate the coherence, and L and M are the number of trials and silence intervals, as above. Here we use the term coherence to refer to the magnitude of the coherency (in some contexts coherence refers to the magnitude *squared* coherency), which is statistically better behaved (see Chapter III):

$$\left|C(f)\right|=\left|\frac{S_{XY}(f)}{\sqrt{S_{XX}(f)S_{YY}(f)}}\right|$$

where S_{XX}, S_{YY}, and S_{XY} are the sample spectra and cross-spectrum for the signals of interest (X and Y refer to the signals recorded from two electrodes), averaged over epochs on a sliding window as described above (Brillinger, 2001). The spectra and cross-spectra were estimated using the multitaper method with time-half-bandwidth product of 1 hz*sec and 1 taper. This resulted in a time resolution of 200 ms and a frequency resolution of 10 Hz (±5 Hz). The variances $\hat{\varsigma}^2_{speech}$ and $\hat{\varsigma}^2_{silence}$ were estimated using the jackknife procedure as described in Chapter III and (Bokil et al., 2007). The resulting

test-statistic is approximately distributed as a standard normal, so p-values were computed and corrected for false discoveries in the same way as for the correlation networks.

4.5.3. Region-level: Canonical correlation and canonical coherence. Region-level networks were inferred using canonical correlation and canonical coherence (described in detail in Chapter III), under the rolandic, auditory, and aSMG region definitions. As with the networks described above, canonical networks were inferred on a 200 ms sliding window during trials. The test statistic for canonical correlation was:

$$\hat{x}_{CCorr} = \tanh^{-1}\left(CCorr_T\right) - \tanh^{-1}\left(CCorr_B\right)$$

where $CCorr_T$ and $CCorr_B$ are the canonical correlations estimated for the trial window and the baseline (silence) window, respectively (see Chapter III Appendix for a description of how canonical correlation and canonical coherence were calculated). The test statistic for canonical coherence was:

$$\hat{x}_{CCoh}(f) = \tanh^{-1}\left(CCoh_T(f)\right) - \tanh^{-1}\left(CCoh_B(f)\right)$$

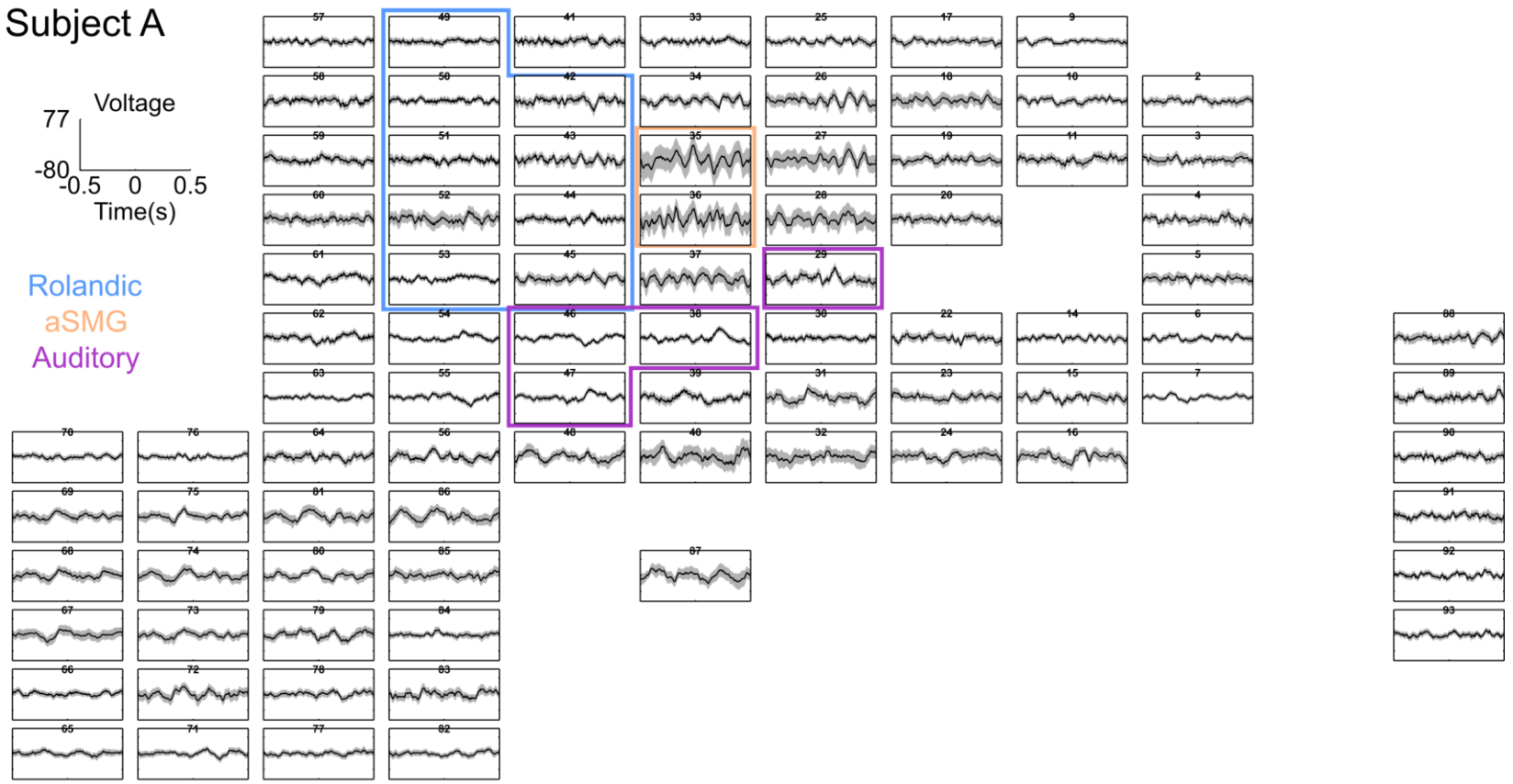
where $CCoh_T$ and $CCoh_B$ are the canonical coherences estimated for the trial window and the baseline (silence) window, respectively.

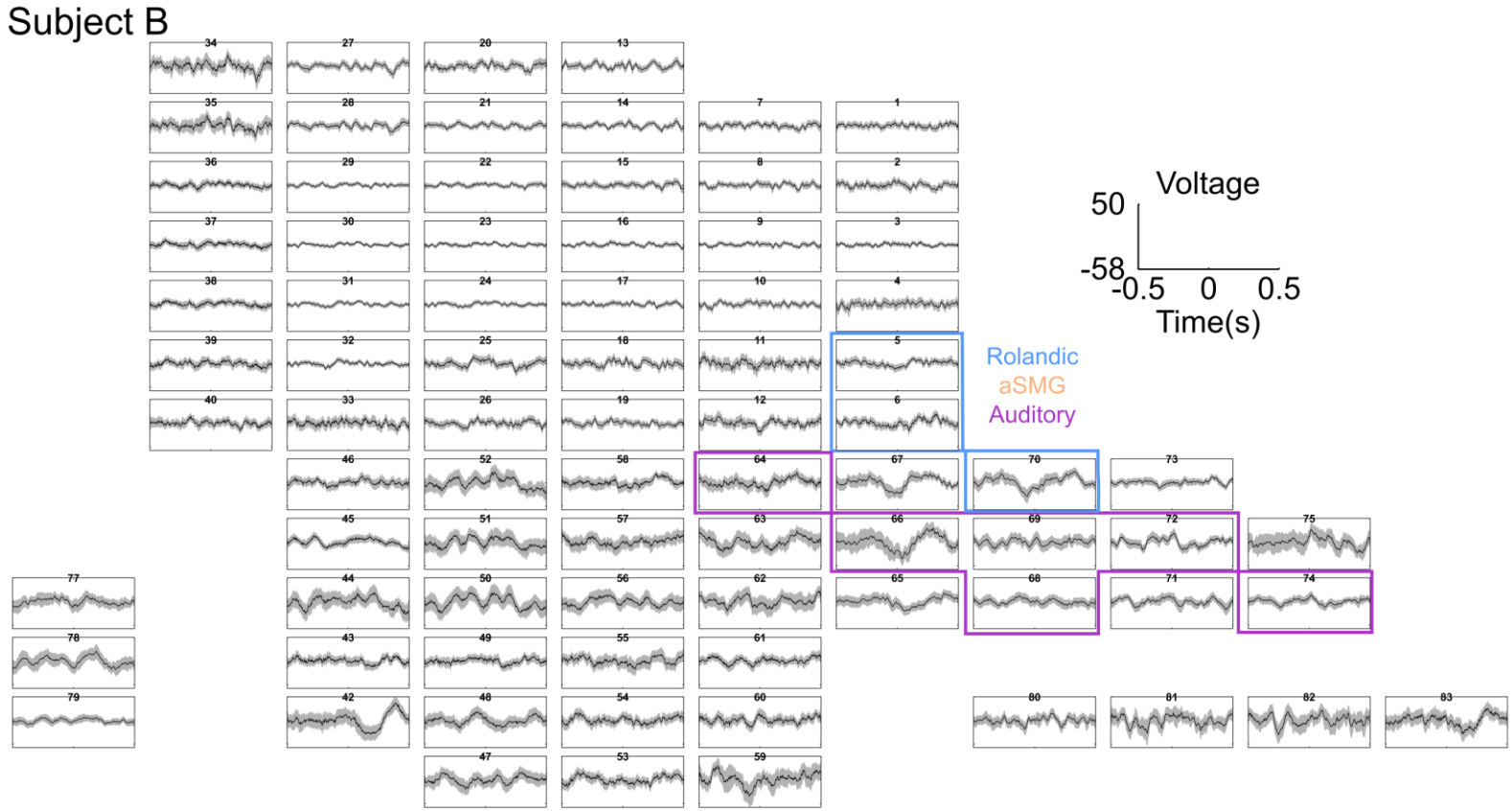
As described in Chapter III, bootstrapping was used to estimate a p-value for a hypothesis test with the null hypothesis that the canonical correlation (or coherence) was the same during the trial window and silence. The p-values were then corrected for multiple comparisons using the FDR procedure across region pairs.

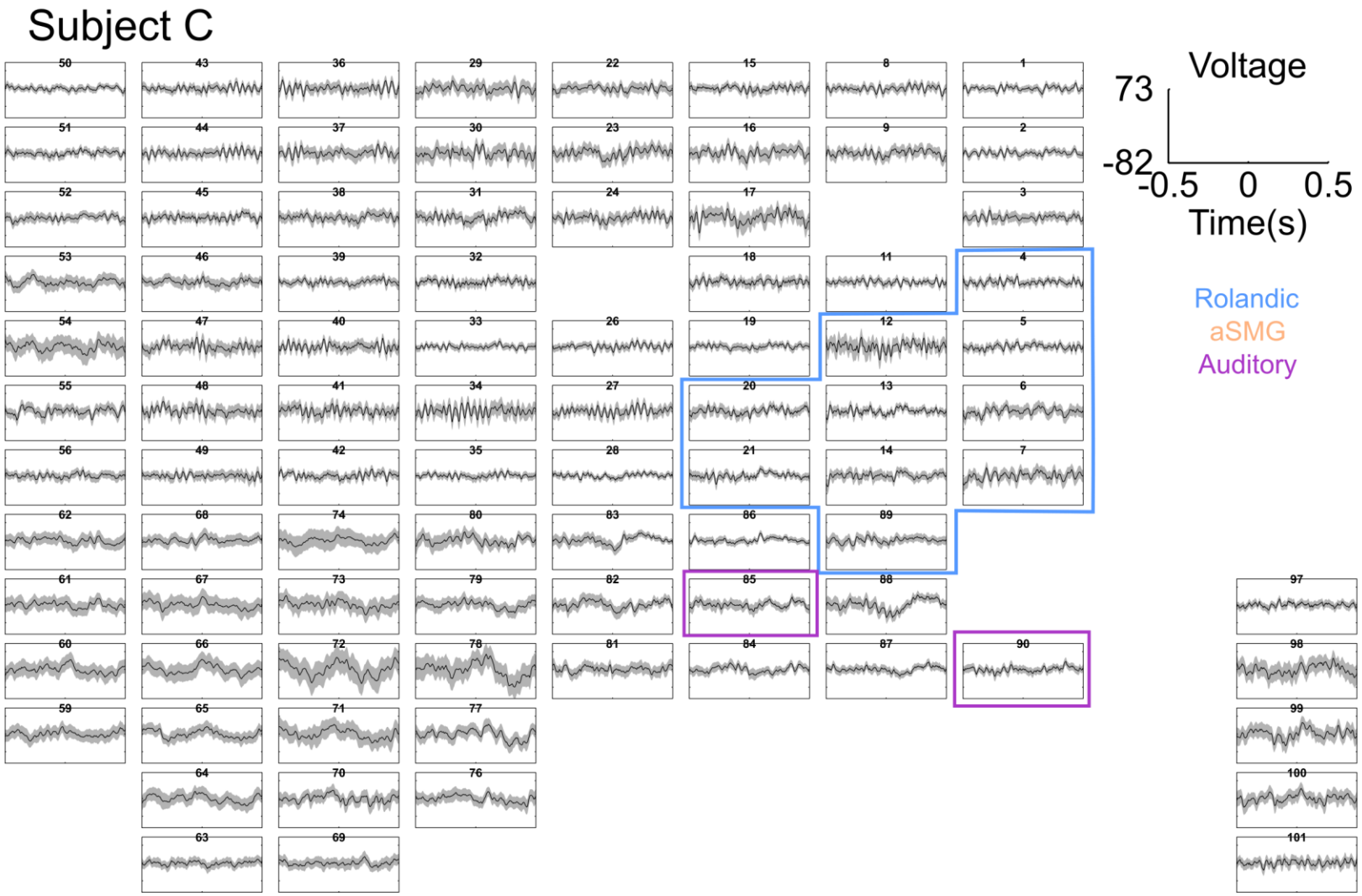
4.5.4. Interactive visualization of network results. SpectraVis was implemented by E Denovellis using the D3 JavaScript library (Bostock et al., 2011). The code is open source and available at https://github.com/edeno/SpectraVis, with complete specifications.

5. Appendices

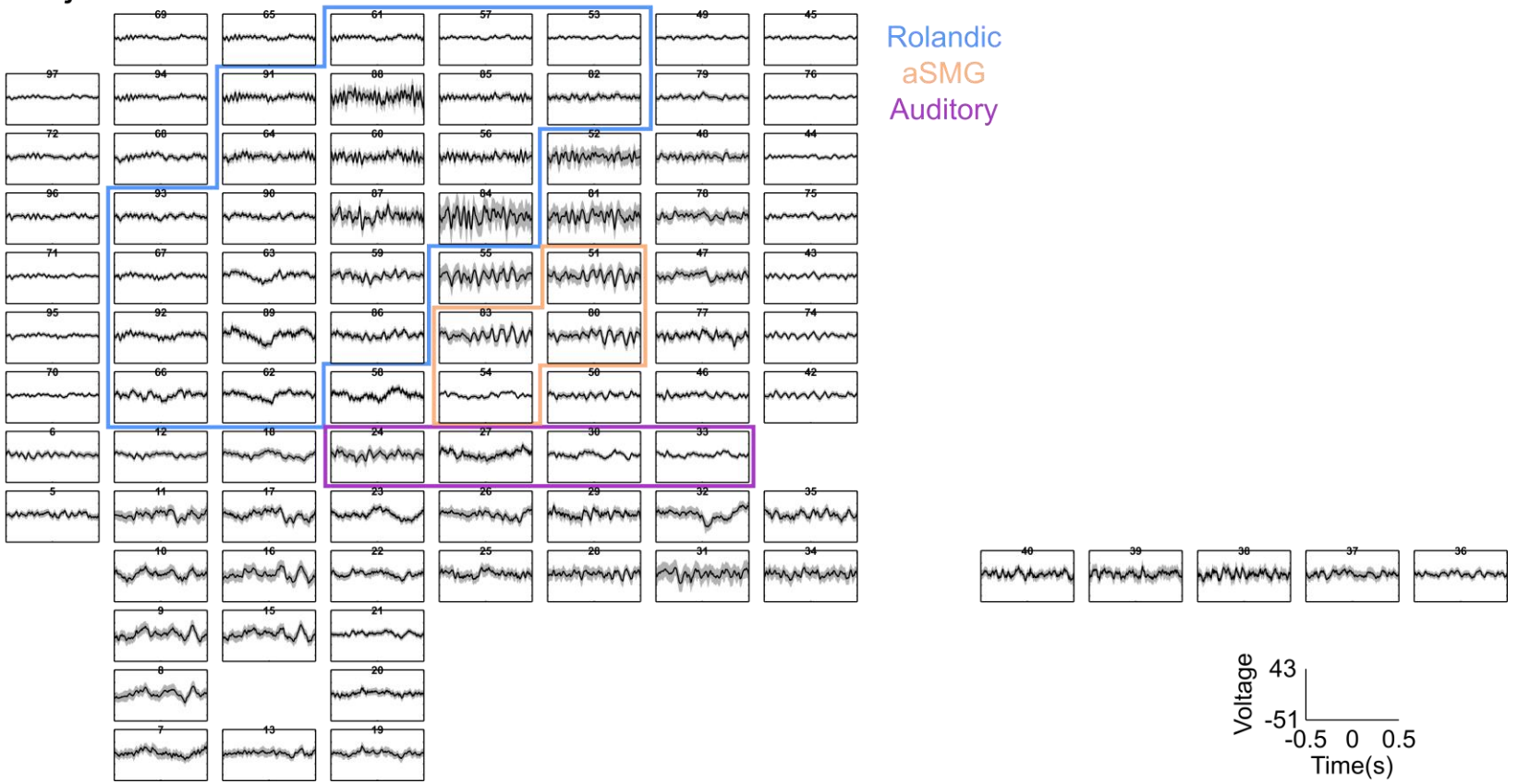
5.1: ERP grids for all subjects



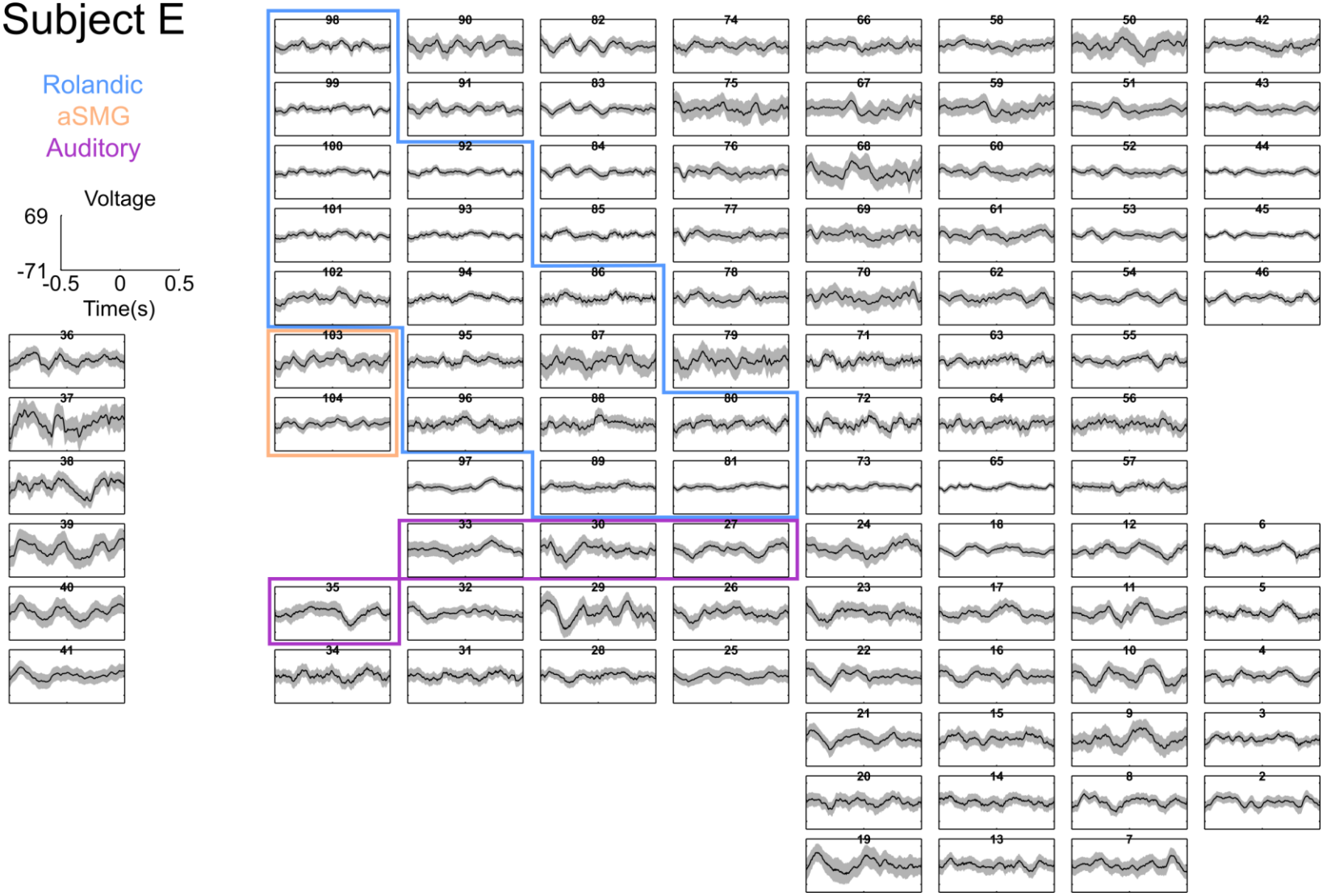

Subject B
Voltage
50
-58
-0.5
0
0.5
Time(s)
Rolandic
aSMG
Auditory

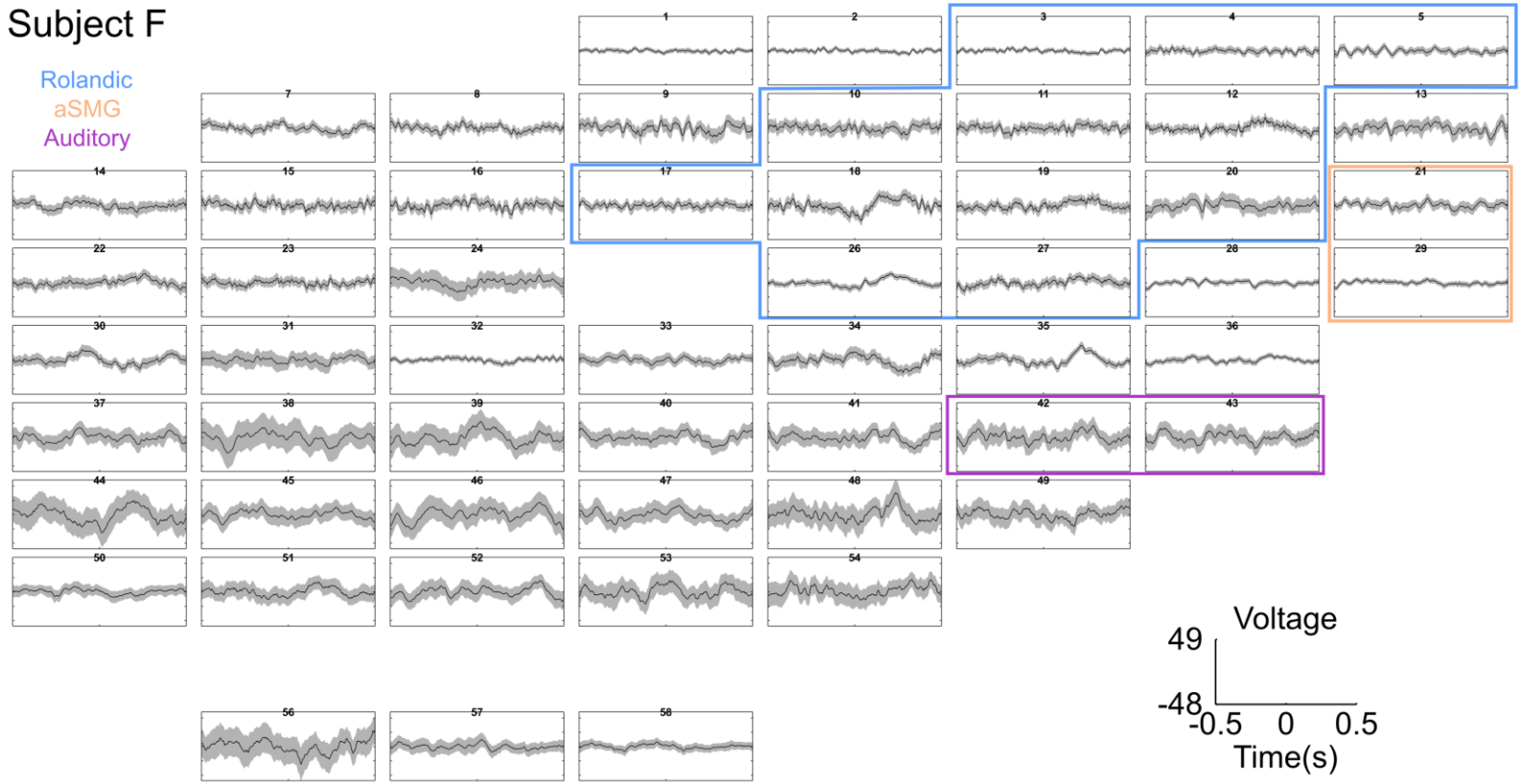

Subject C
Voltage
73
-82
-0.5
0
0.5
Time(s)
Rolandic
aSMG
Auditory

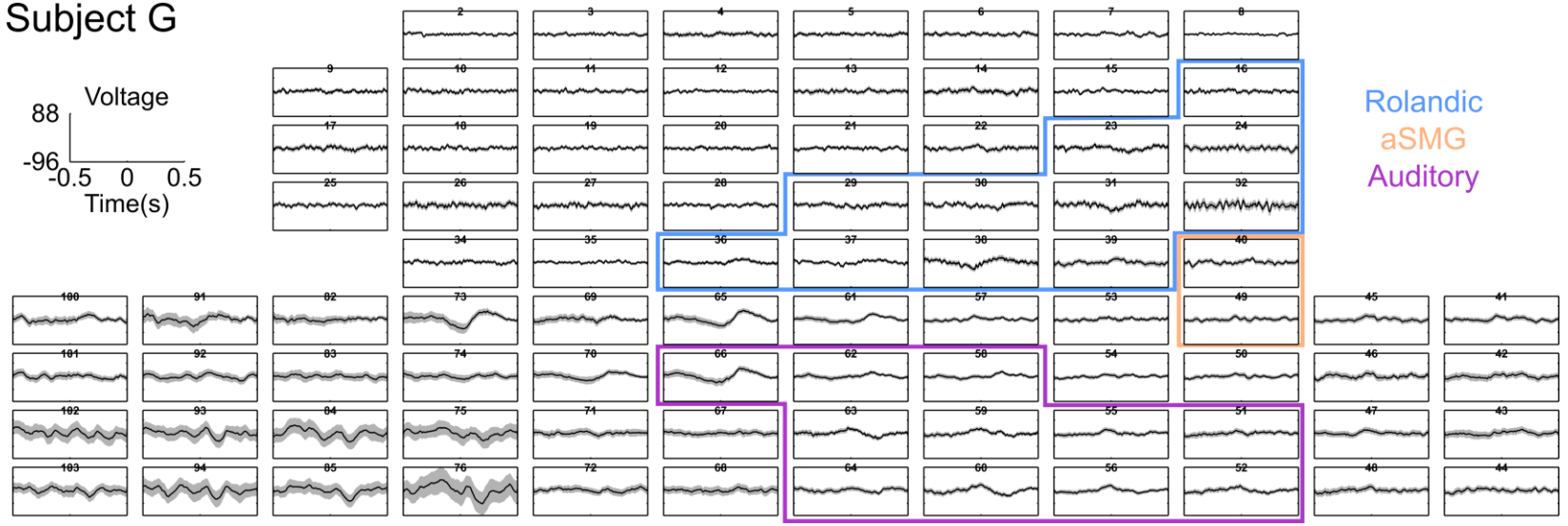
Subject D



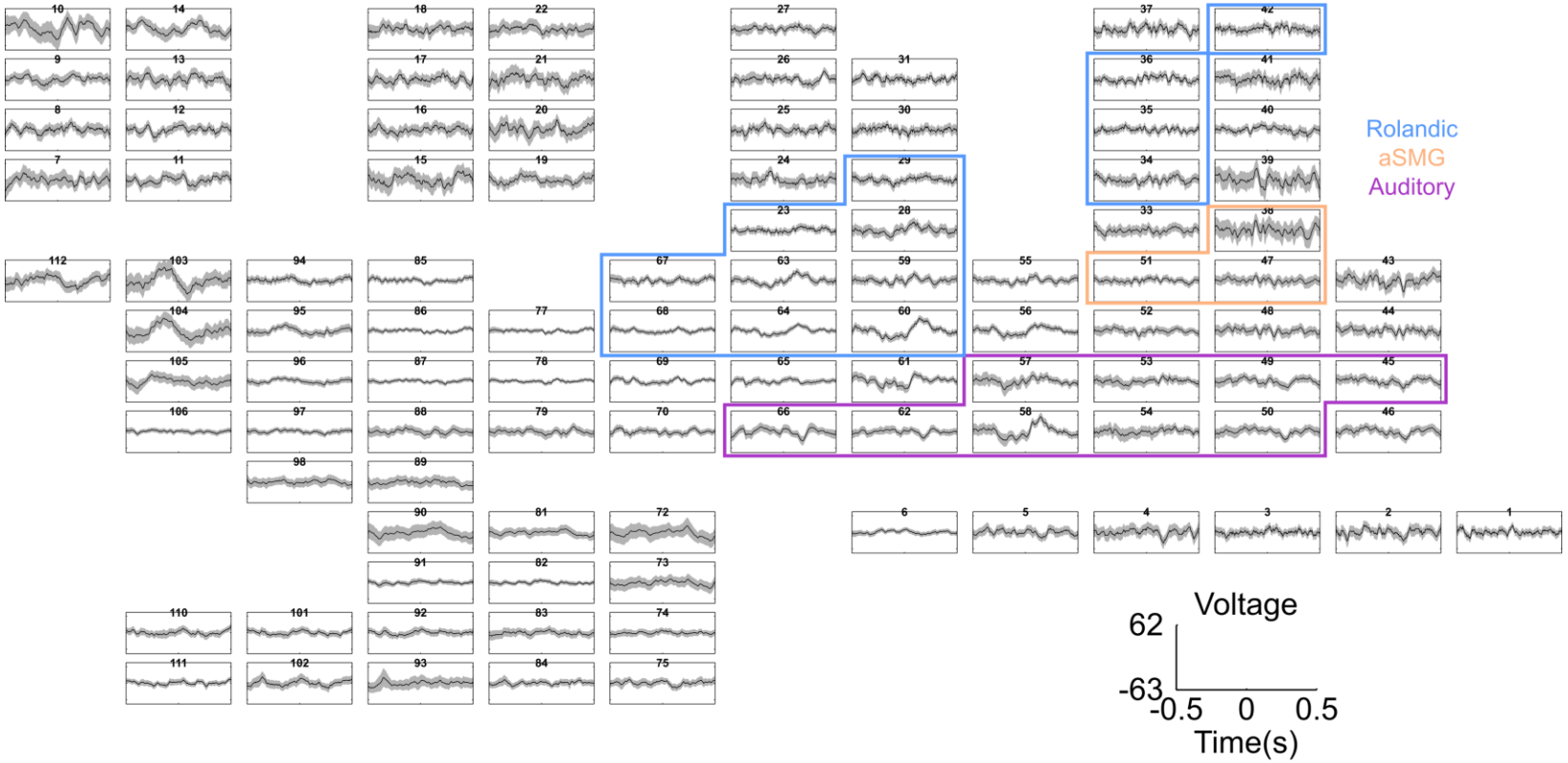
Subject E

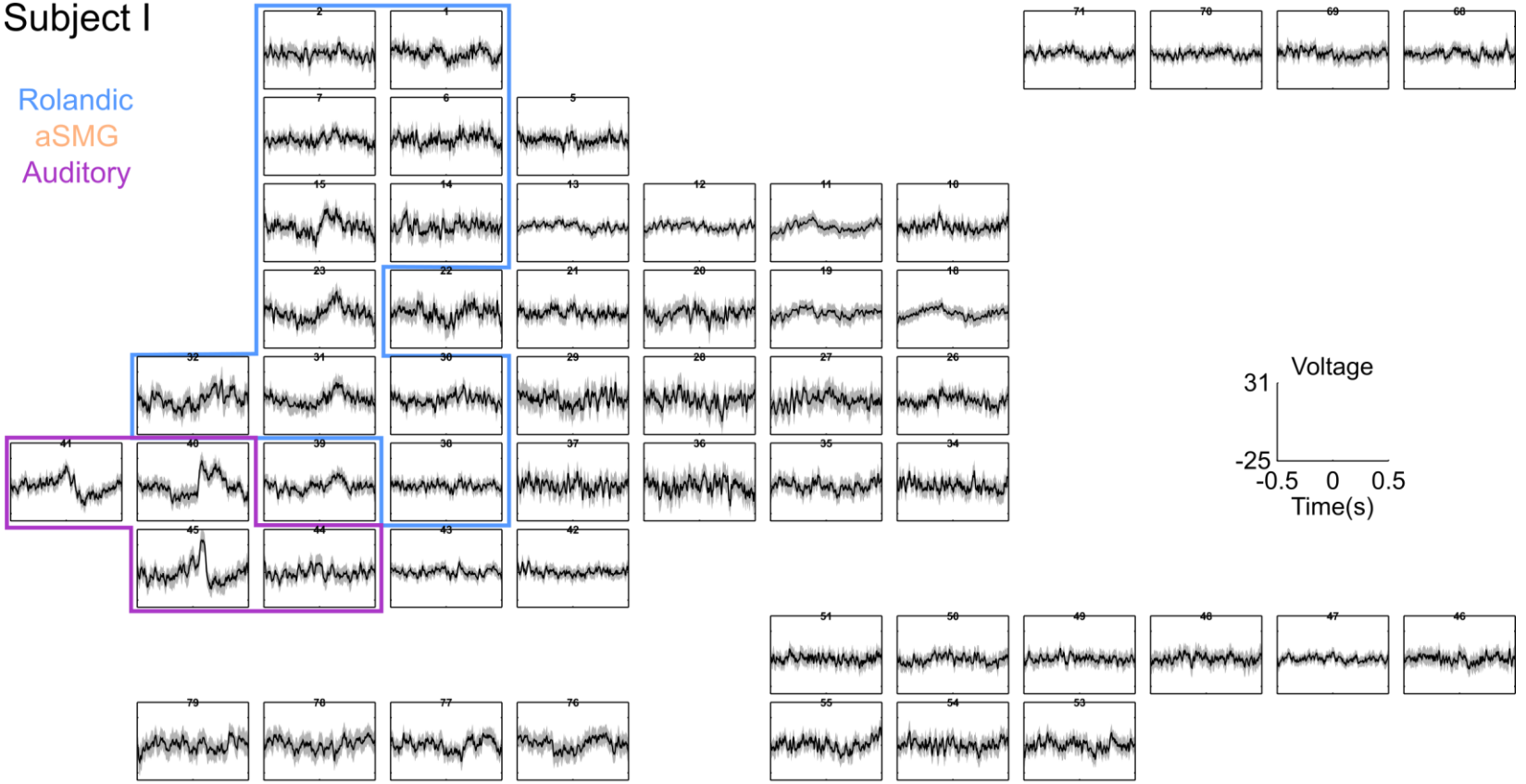

Rolandic
aSMG
Auditory
Voltage
69
-71
-0.5
0
0.5
Time(s)
98
90
82
74
66
58
50
42
99
91
83
75
67
59
51
43
100
92
84
76
68
60
52
44
101
93
85
77
69
61
53
45
102
94
86
78
70
62
54
46
103
95
87
79
71
63
55
104
96
88
80
72
64
56
97
89
81
73
65
57
33
30
27
24
18
12
6
35
32
29
26
23
17
11
5
34
31
28
25
22
16
10
4
21
15
9
3
20
14
8
2
19
13
7
36
37
38
39
40
41


Subject F
Rolandic
aSMG
Auditory
1
2
3
4
5
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
56
57
58
Voltage
49
-48
-0.5
0
0.5
Time(s)


Subject G
Voltage
88
-96
-0.5
0
0.5
Time(s)
Rolandic
aSMG
Auditory

Subject H



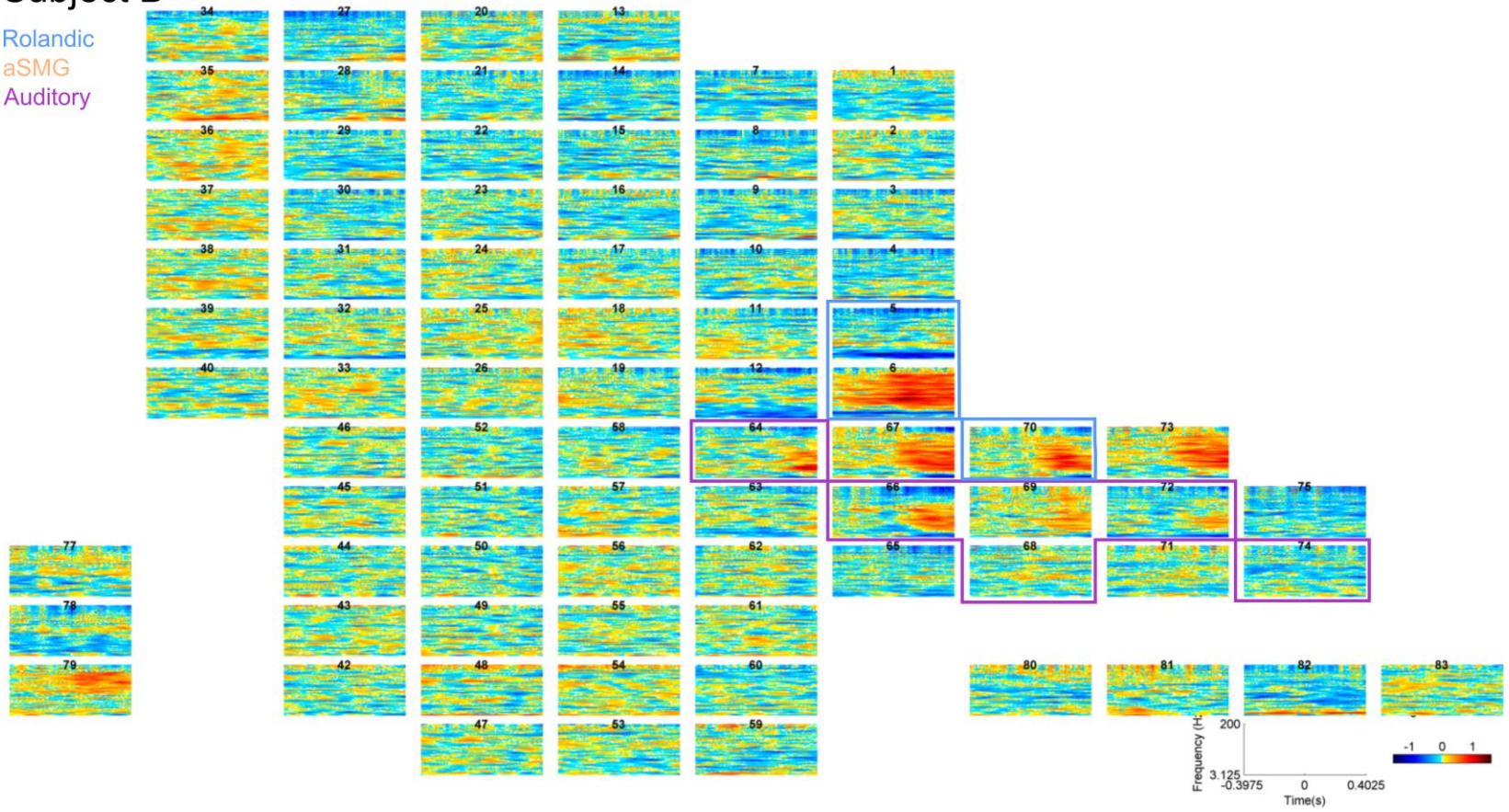

Subject I
Rolandic
aSMG
Auditory
Voltage
31
-25
-0.5
0
0.5
Time(s)

5.2: Spectrogram grids for all subjects

For Subject D, see Figure IV.4

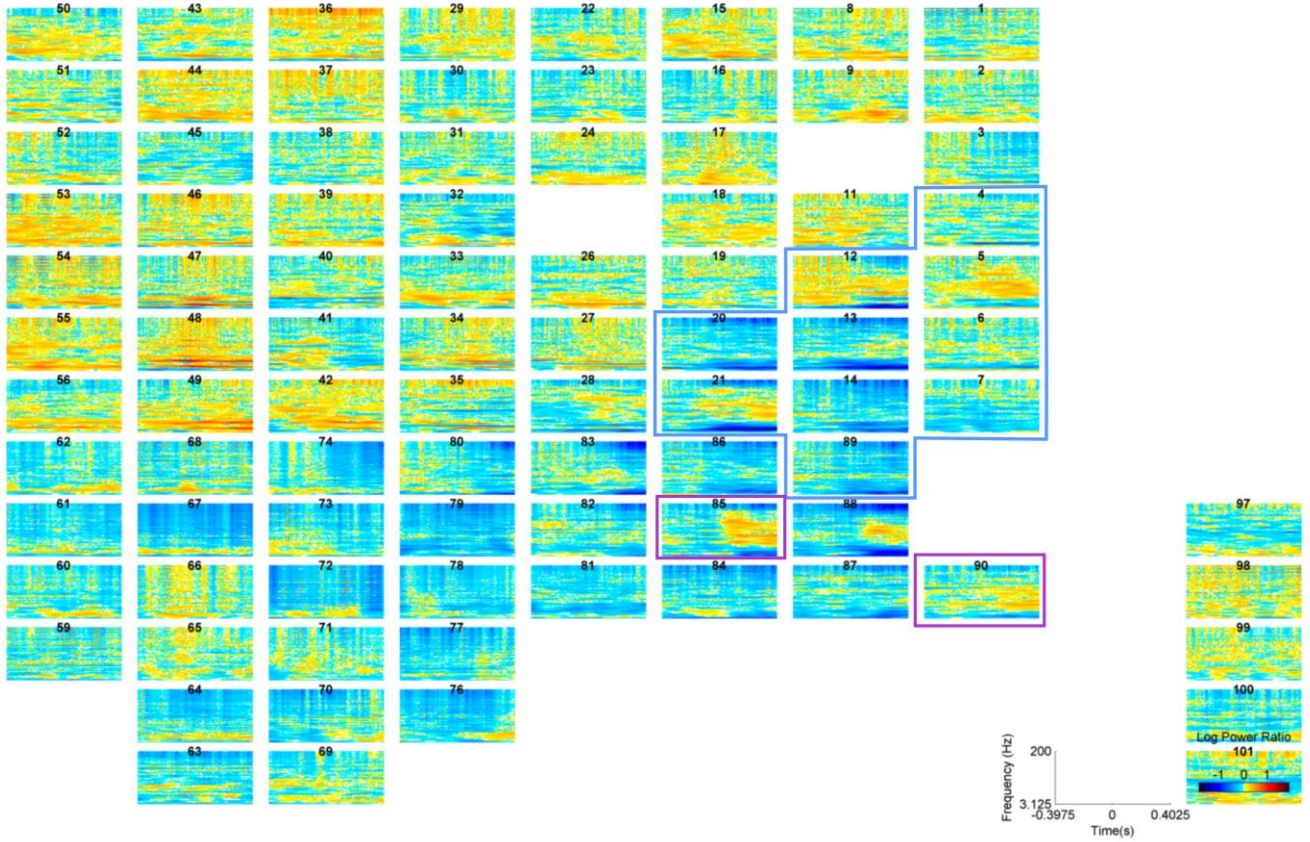
Subject B

Rolandic
aSMG
Auditory



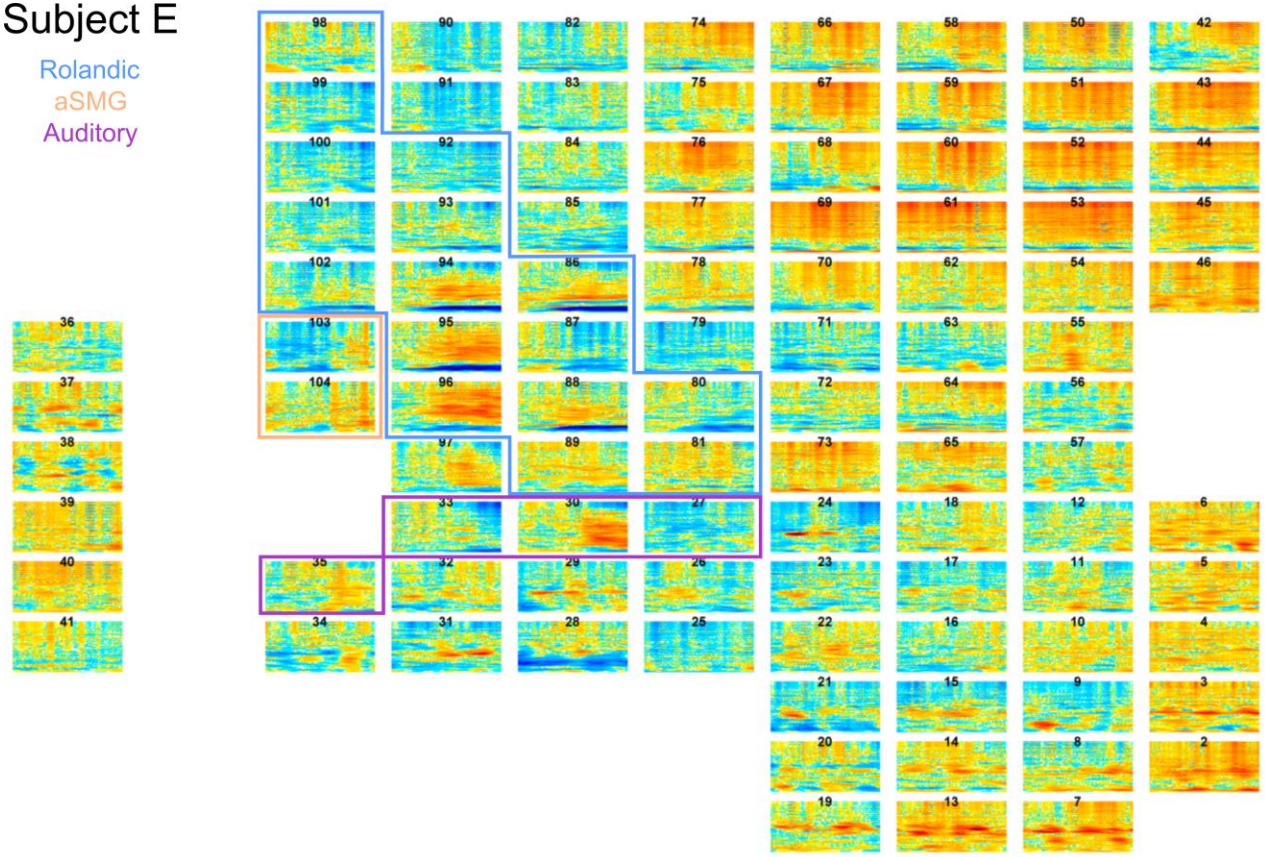
Subject C

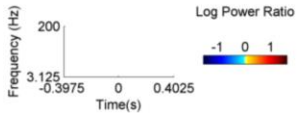
Rolandic
aSMG
Auditory

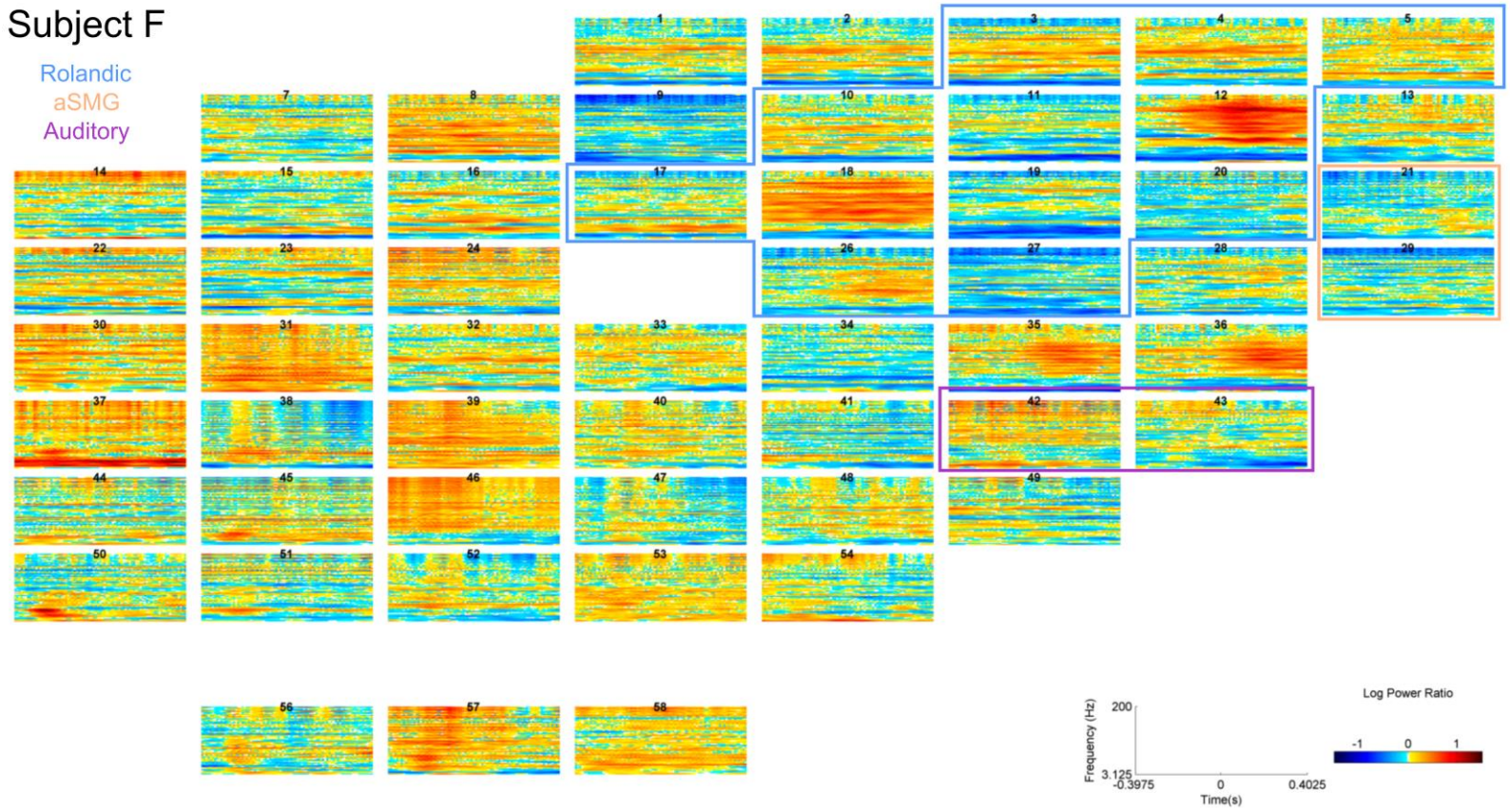


Subject E

Rolandic
aSMG
Auditory

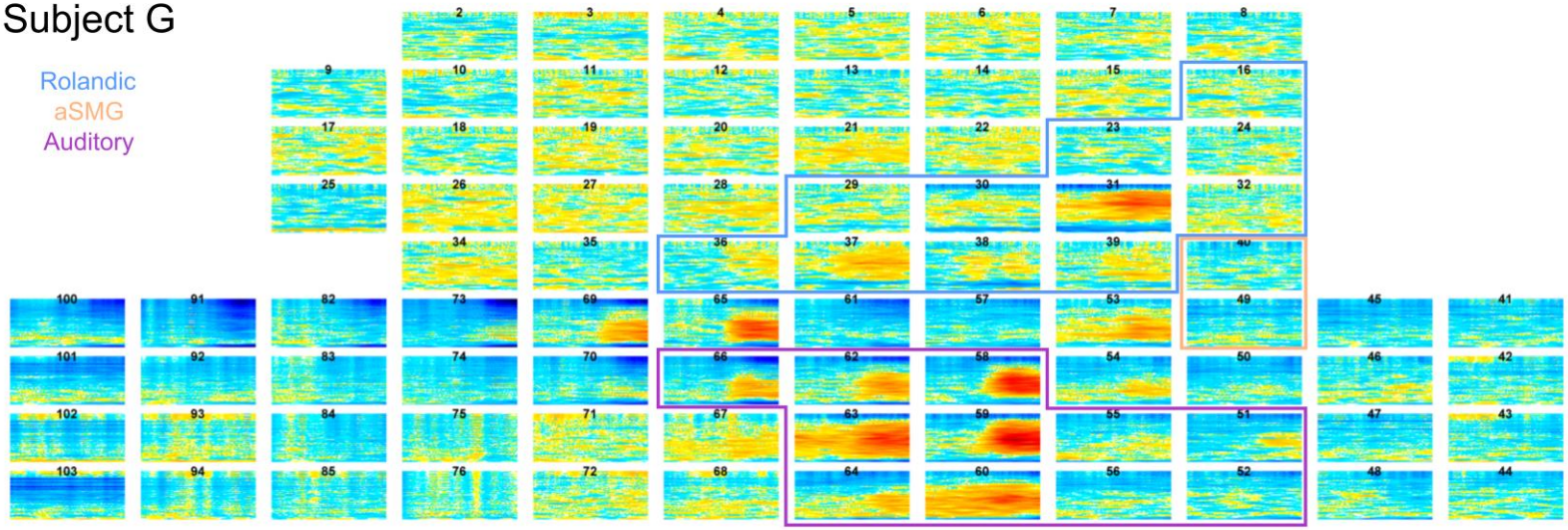


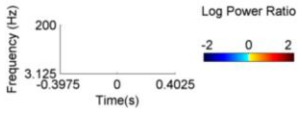



Subject F
Rolandic
aSMG
Auditory
1
2
3
4
5
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
56
57
58
Log Power Ratio
-1
0
1
Frequency (Hz)
200
3.125
-0.3975
0
0.4025
Time(s)

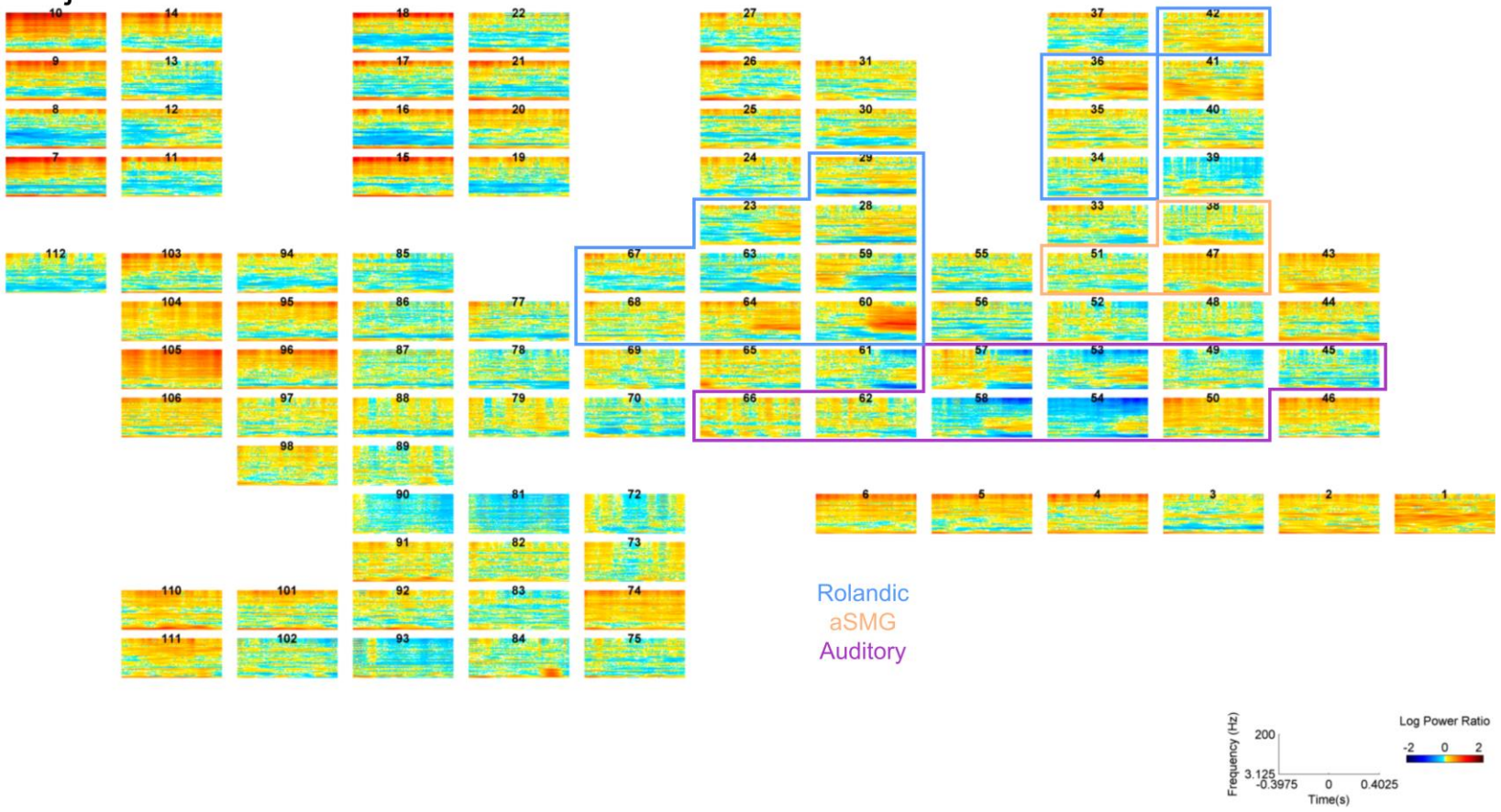
Subject G

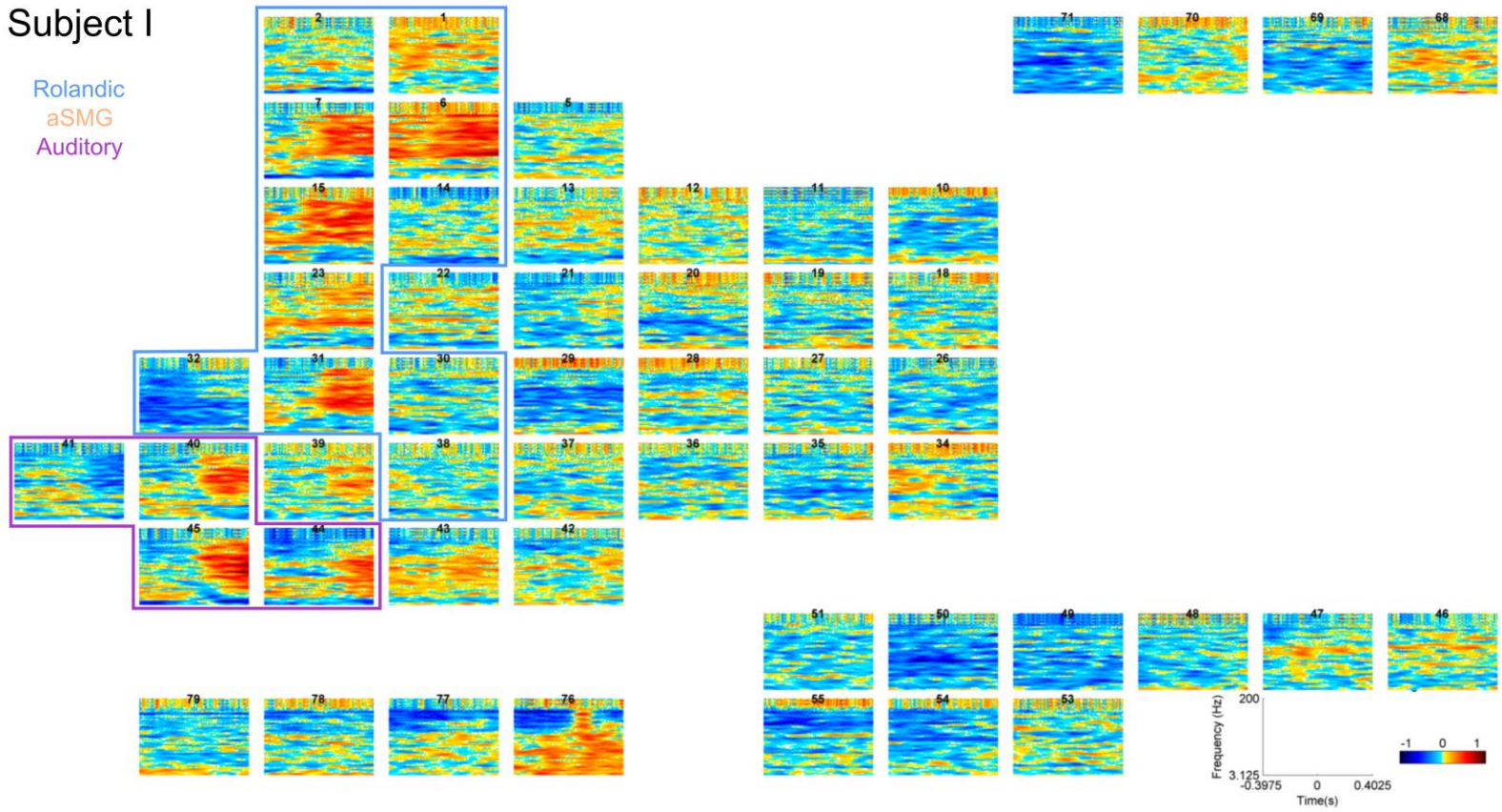
Rolandic
aSMG
Auditory





Subject H

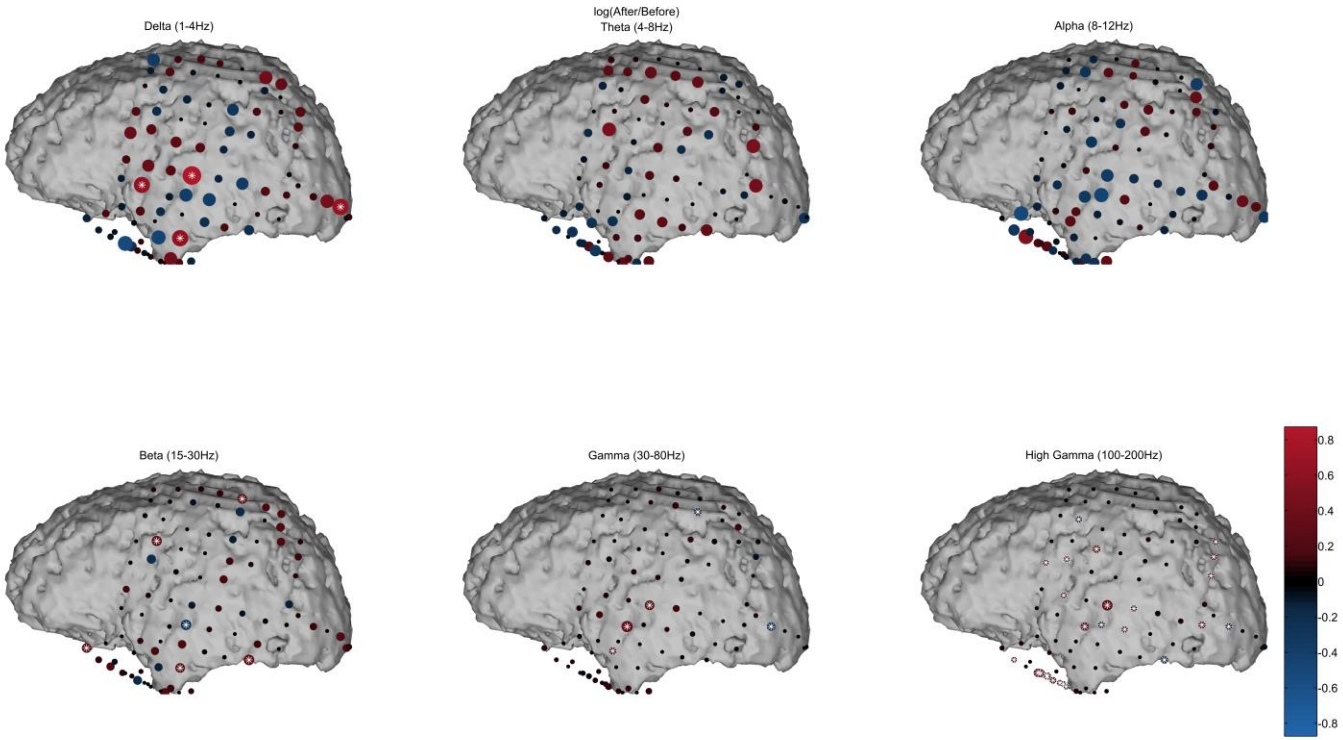

Rolandic
aSMG
Auditory
Log Power Ratio
-2
0
2
Frequency (Hz)
200
3.125
-0.3975
0
0.4025
Time(s)


Subject I
Rolandic
aSMG
Auditory
Frequency (Hz)
200
3.125
-0.3975
0
0.4025
Time(s)
-1
0
1

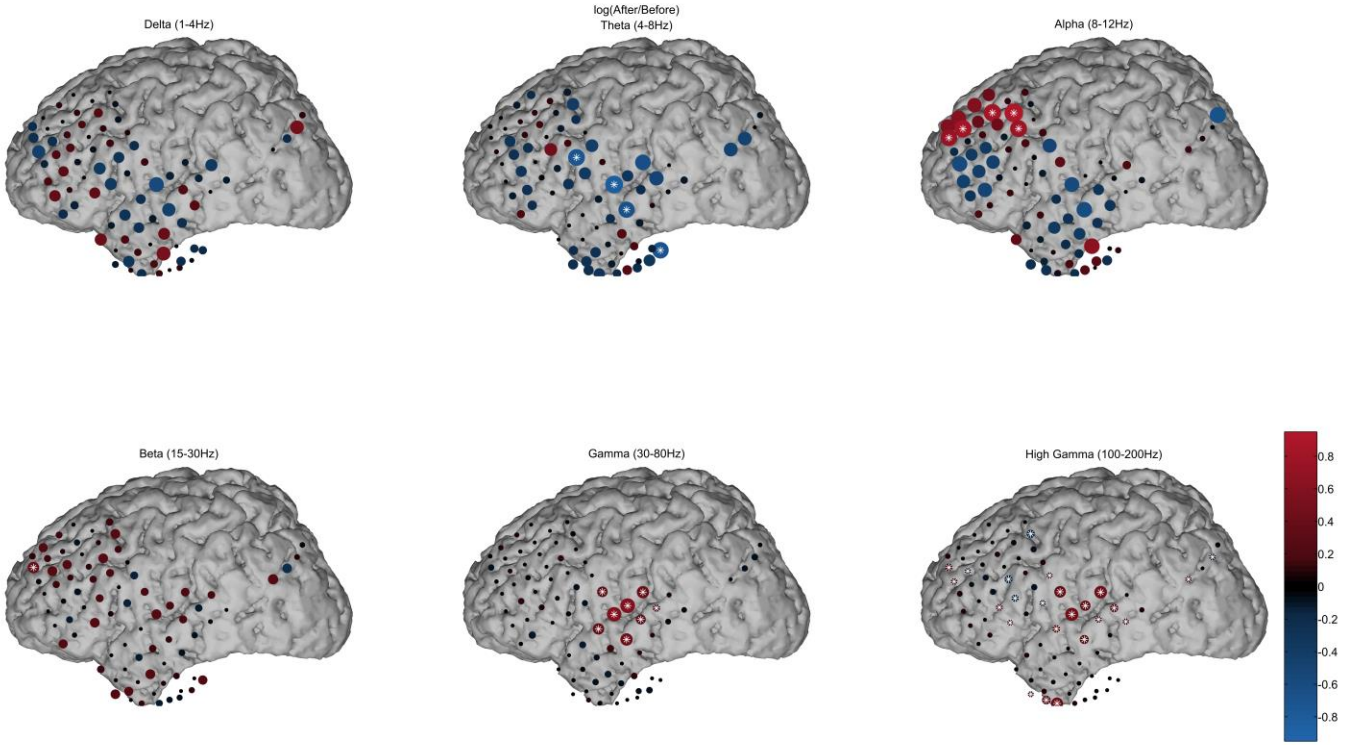
5.3: Bubble plots by frequency band for all subjects

For Subject D, see Figure IV.5

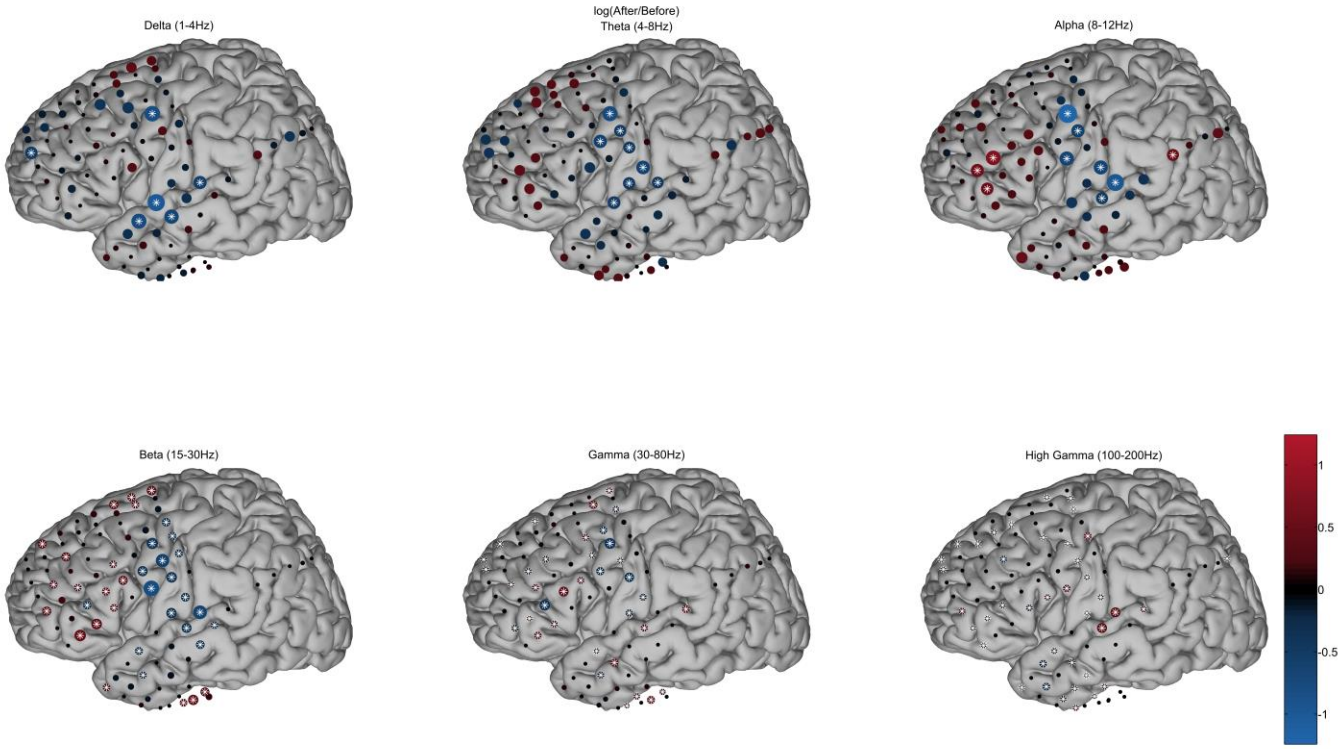
Subject A


log(After/Before)
Delta (1-4Hz)
Theta (4-8Hz)
Alpha (8-12Hz)
Beta (15-30Hz)
Gamma (30-80Hz)
High Gamma (100-200Hz)
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8

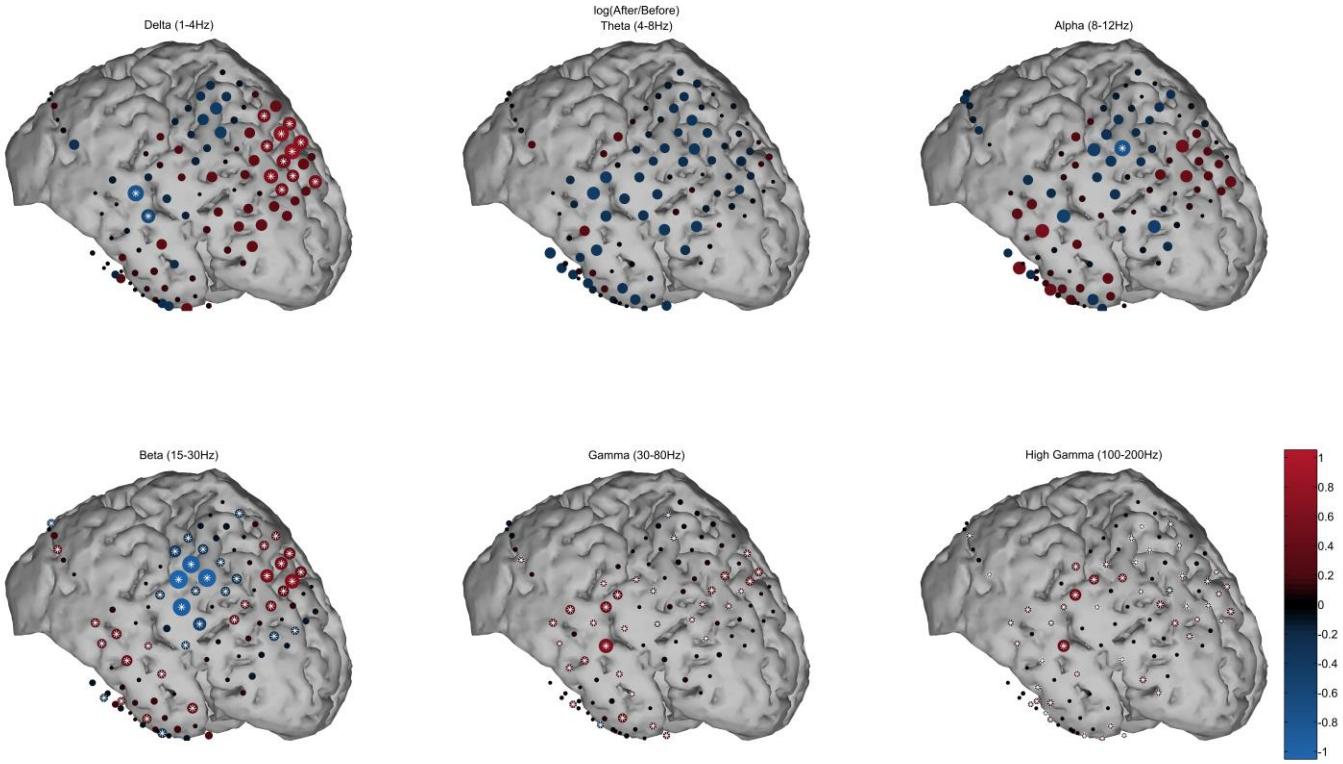
Subject B


log(After/Before)
Delta (1-4Hz)
Theta (4-8Hz)
Alpha (8-12Hz)
Beta (15-30Hz)
Gamma (30-80Hz)
High Gamma (100-200Hz)
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8

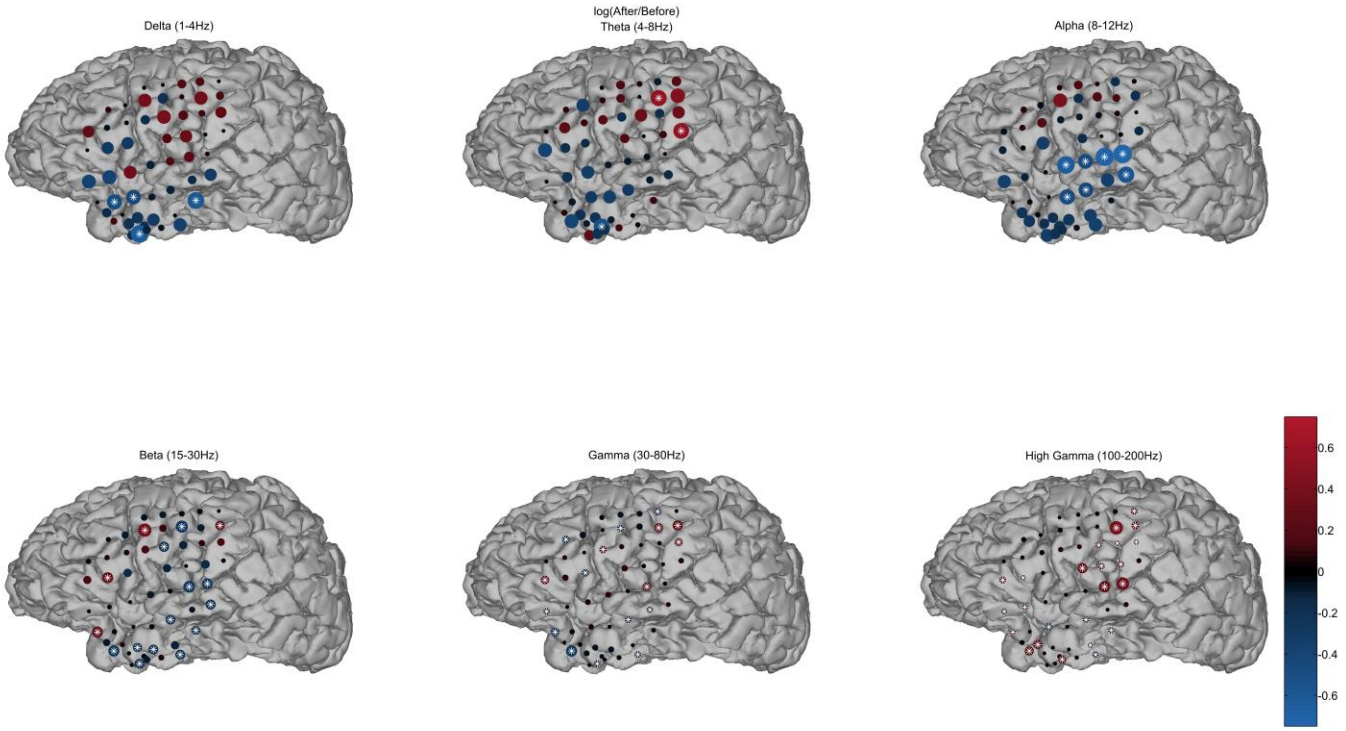
Subject C


Delta (1-4Hz)
log(After/Before)
Theta (4-8Hz)
Alpha (8-12Hz)
Beta (15-30Hz)
Gamma (30-80Hz)
High Gamma (100-200Hz)
1
0.5
0
-0.5
-1

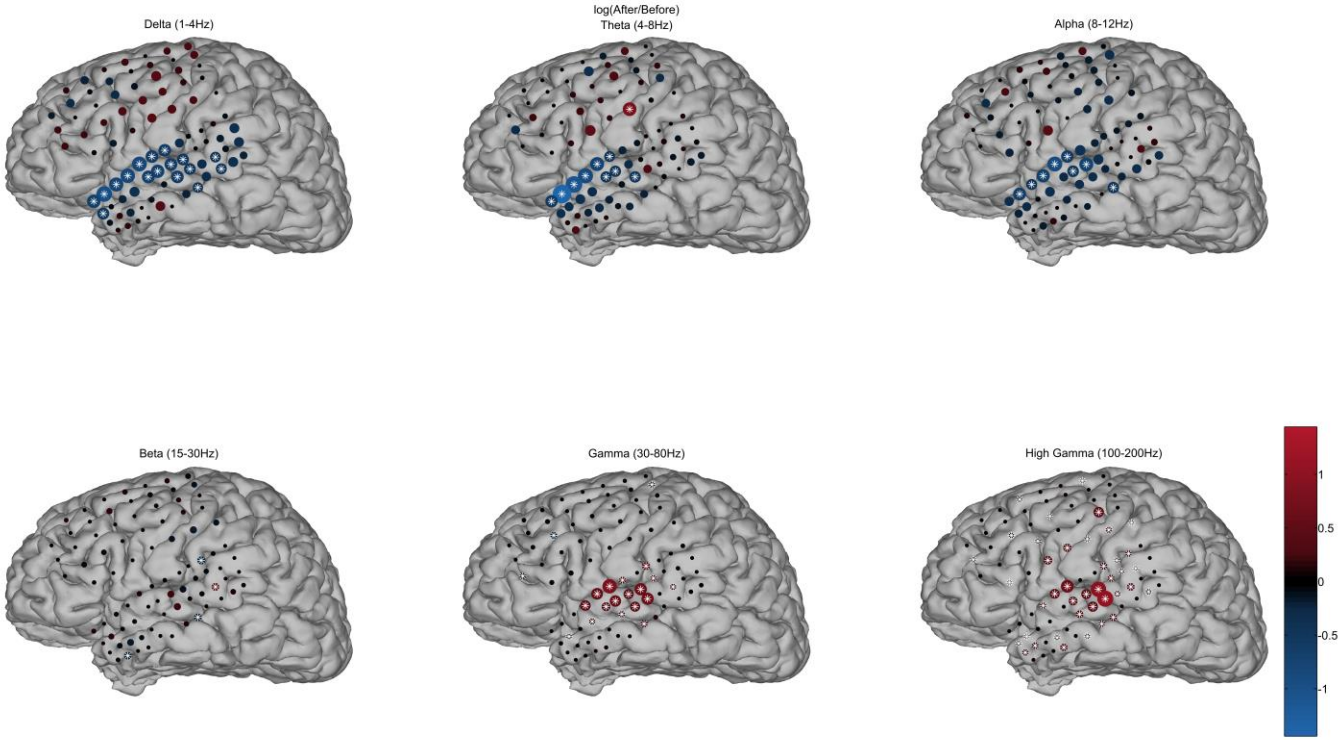
Subject E


log(After/Before)
Delta (1-4Hz)
Theta (4-8Hz)
Alpha (8-12Hz)
Beta (15-30Hz)
Gamma (30-80Hz)
High Gamma (100-200Hz)
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1

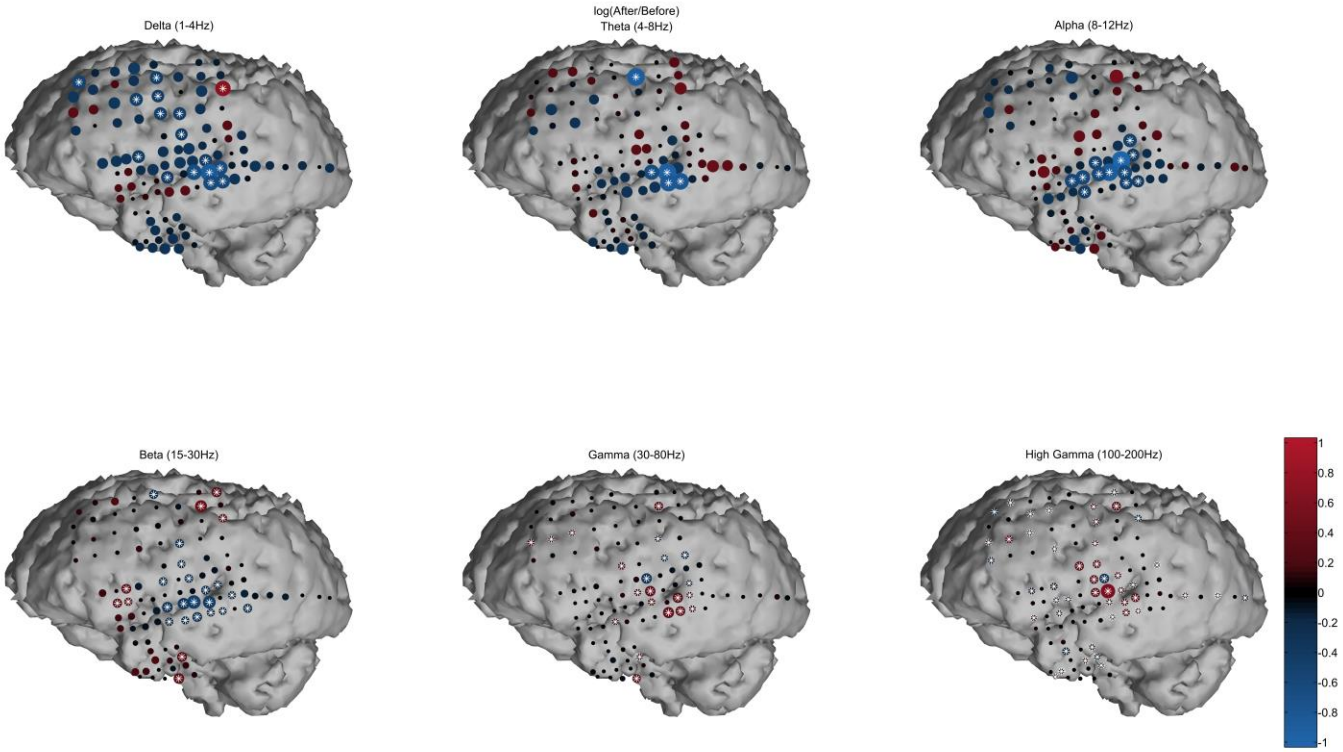
Subject F


log(After/Before)
Delta (1-4Hz)
Theta (4-8Hz)
Alpha (8-12Hz)
Beta (15-30Hz)
Gamma (30-80Hz)
High Gamma (100-200Hz)
0.6
0.4
0.2
0
-0.2
-0.4
-0.6

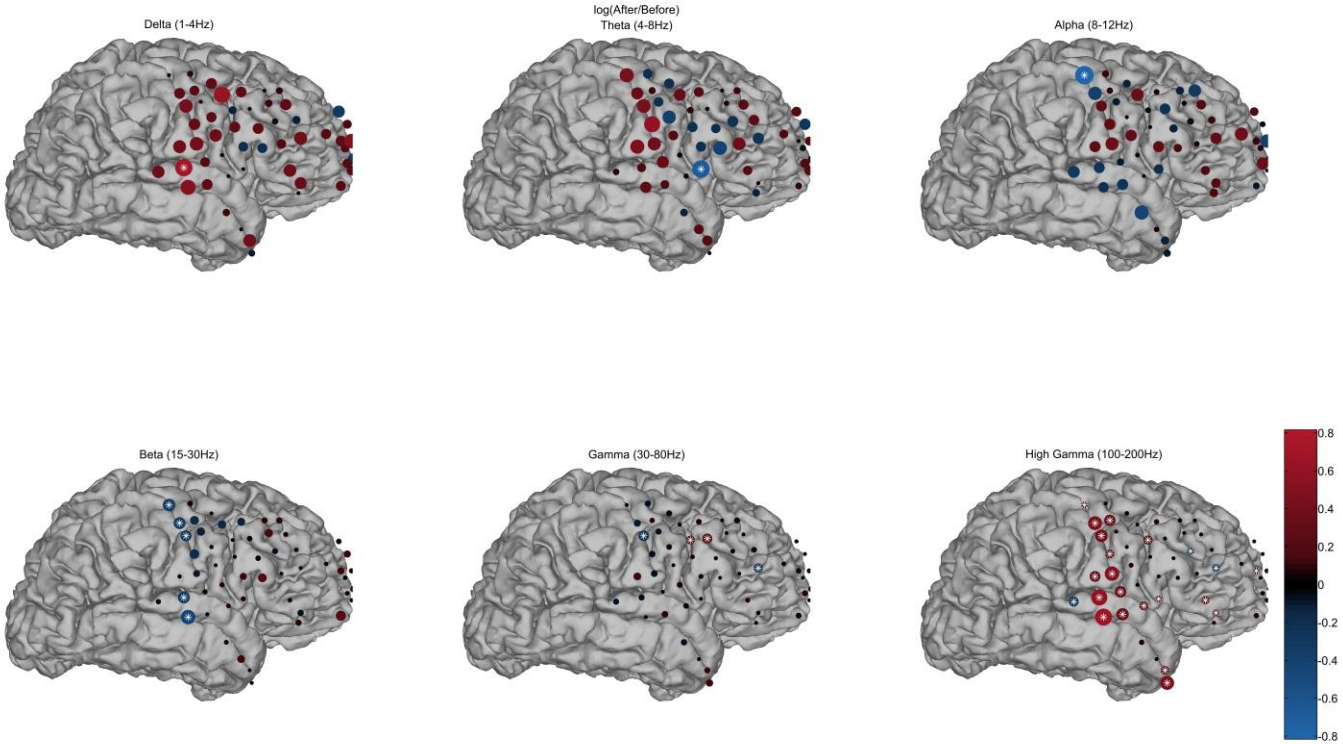
Subject G


Delta (1-4Hz)
log(After/Before)
Theta (4-8Hz)
Alpha (8-12Hz)
Beta (15-30Hz)
Gamma (30-80Hz)
High Gamma (100-200Hz)
1
0.5
0
-0.5
-1

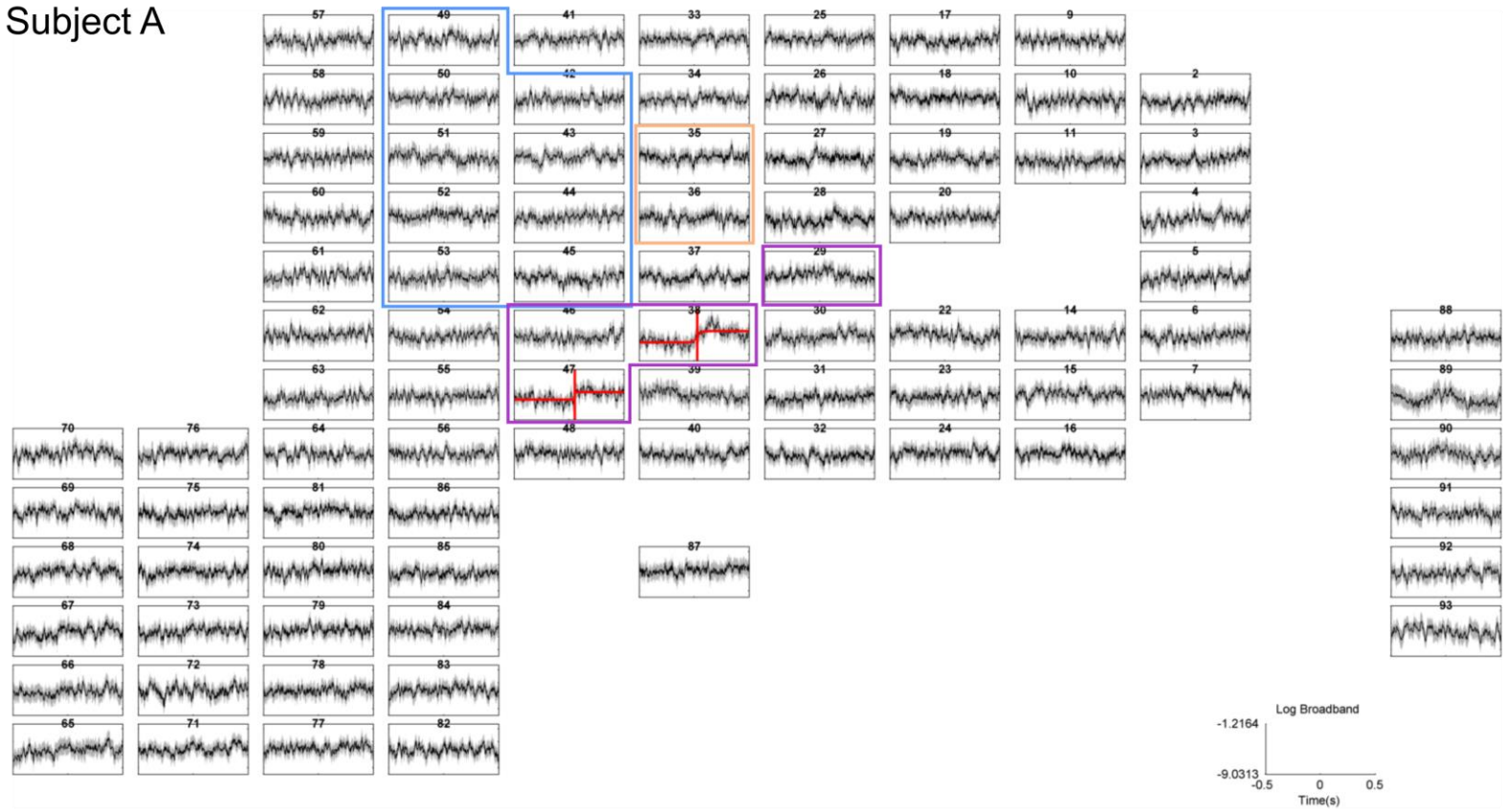
Subject H


log(After/Before)
Delta (1-4Hz)
Theta (4-8Hz)
Alpha (8-12Hz)
Beta (15-30Hz)
Gamma (30-80Hz)
High Gamma (100-200Hz)
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1

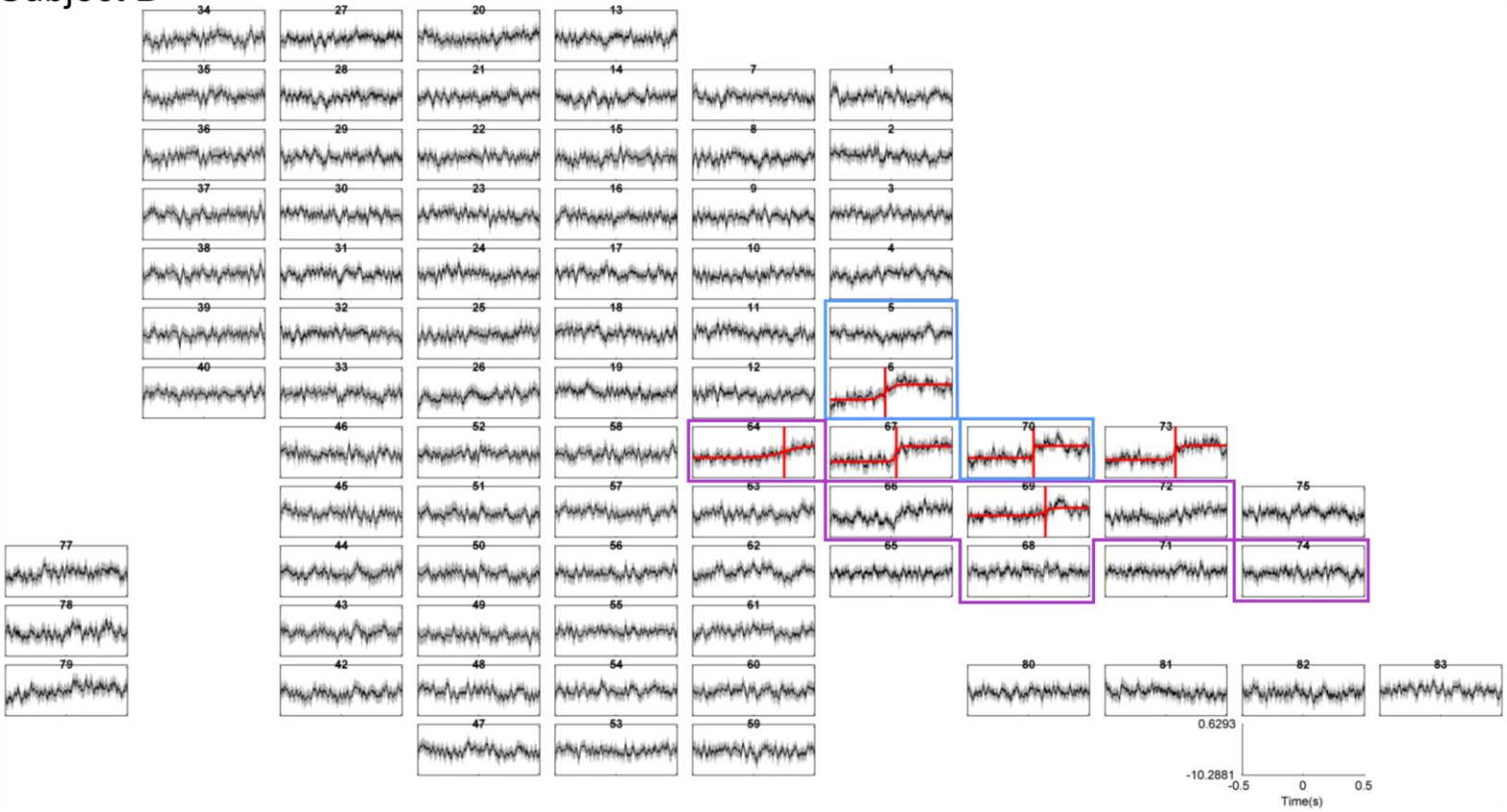
Subject I


Delta (1-4Hz)
log(After/Before)
Theta (4-8Hz)
Alpha (8-12Hz)
Beta (15-30Hz)
Gamma (30-80Hz)
High Gamma (100-200Hz)
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8

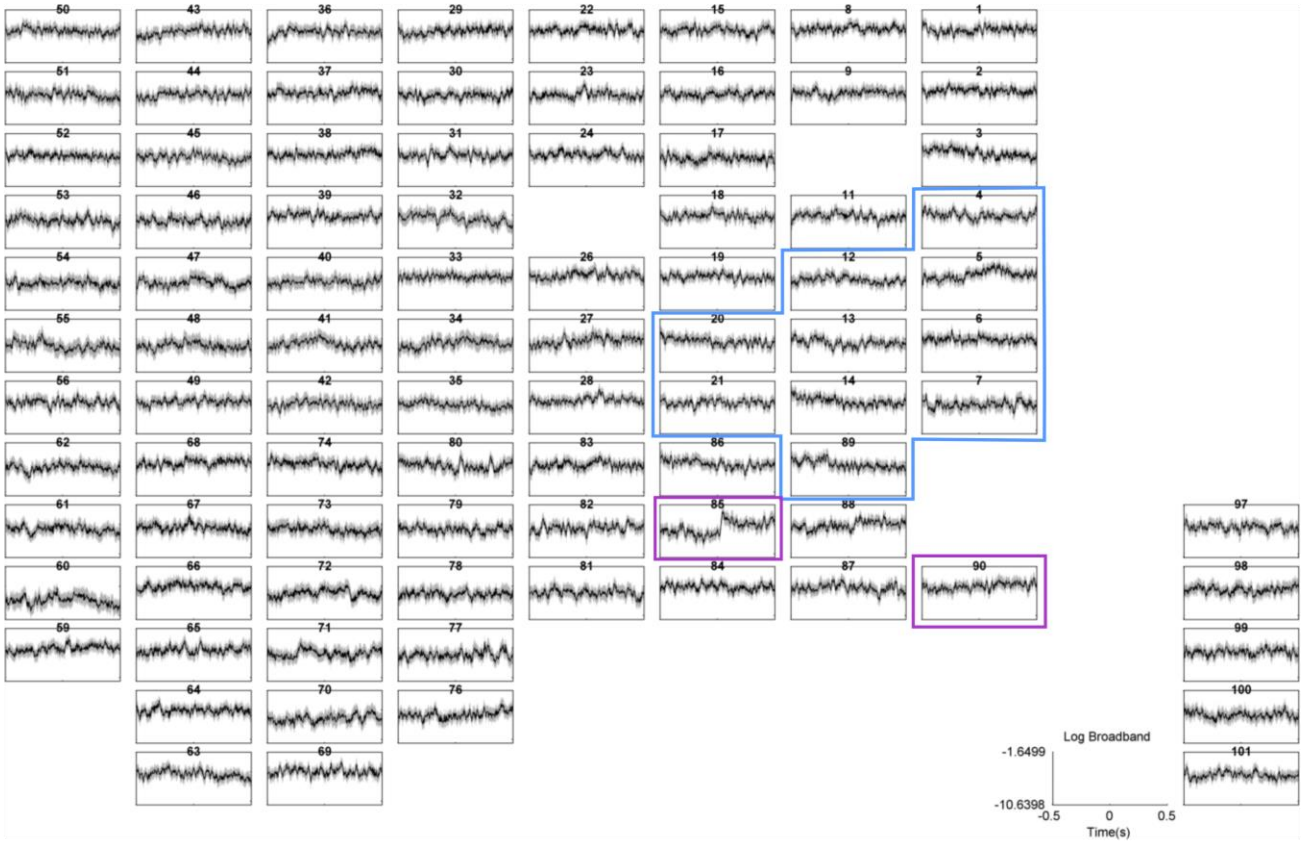
5.4: Broadband time course grids with latency for all subjects



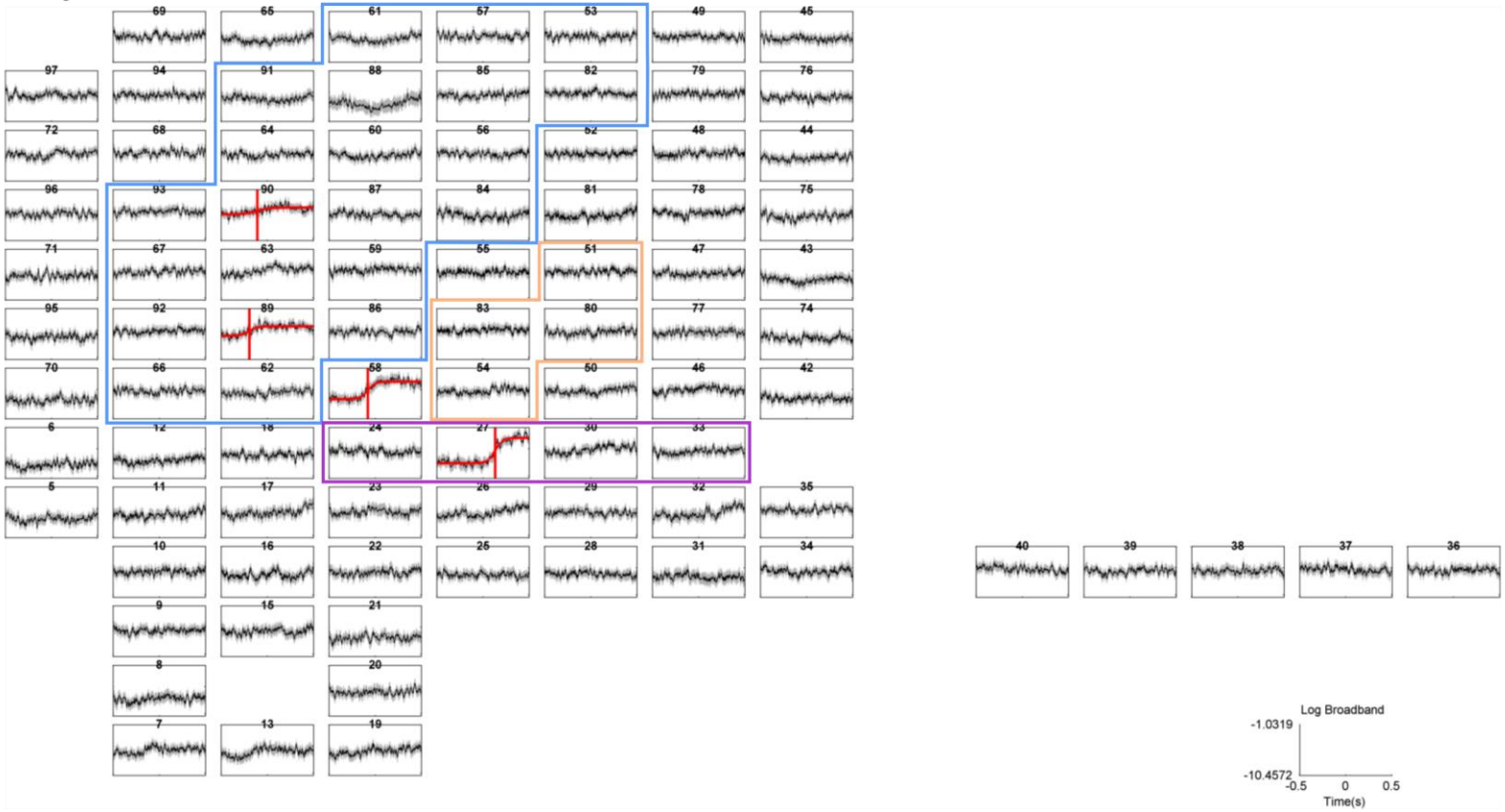
Subject B


34
27
20
13
35
28
21
14
7
1
36
29
22
15
8
2
37
30
23
16
9
3
38
31
24
17
10
4
39
32
25
18
11
5
40
33
26
19
12
6
46
52
58
64
67
70
73
45
51
57
63
66
69
72
75
77
44
50
56
62
65
68
71
74
78
43
49
55
61
79
42
48
54
60
80
81
82
83
47
53
59
0.6293
-10.2881
-0.5
0
0.5
Time(s)

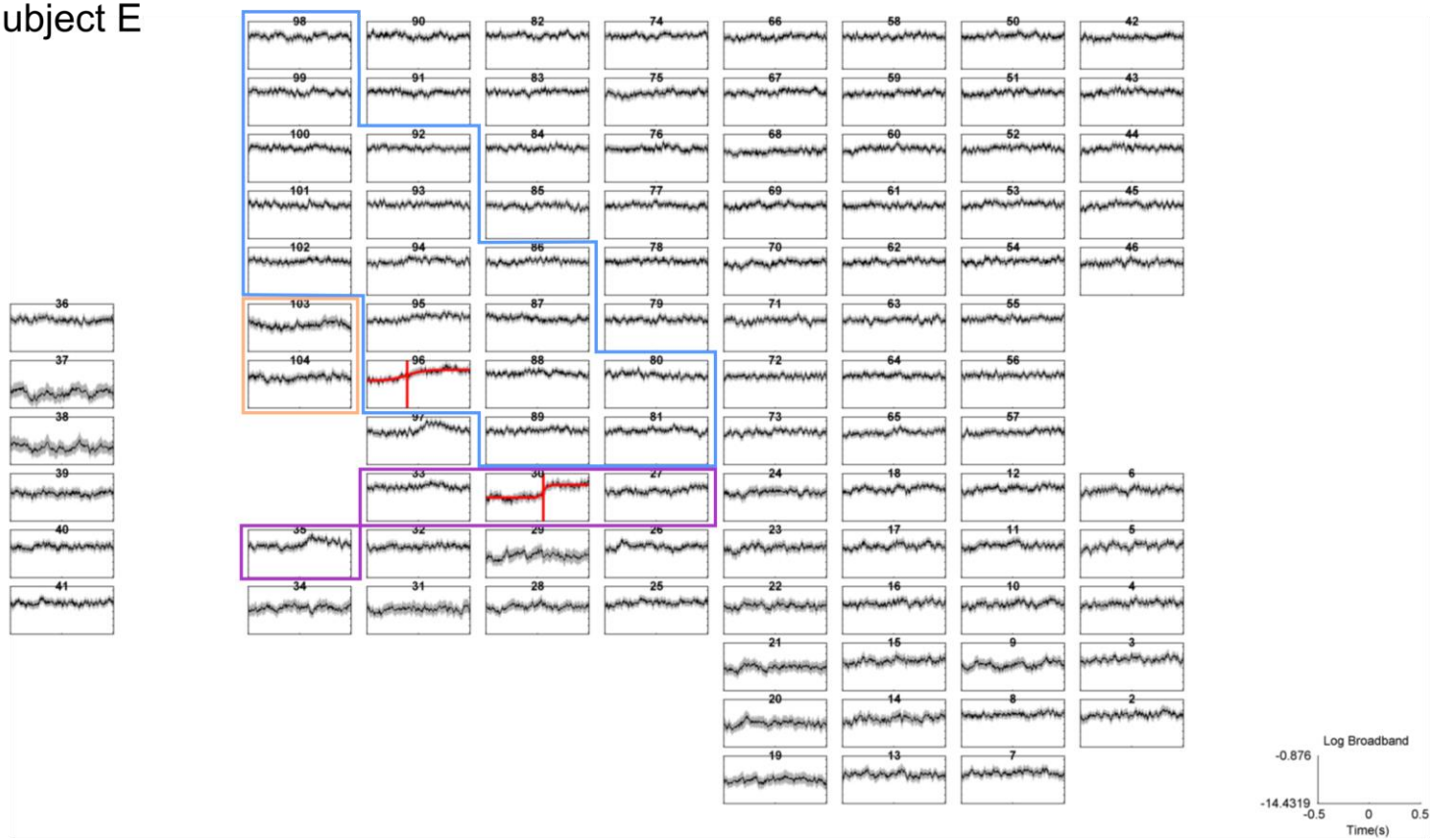
Subject C



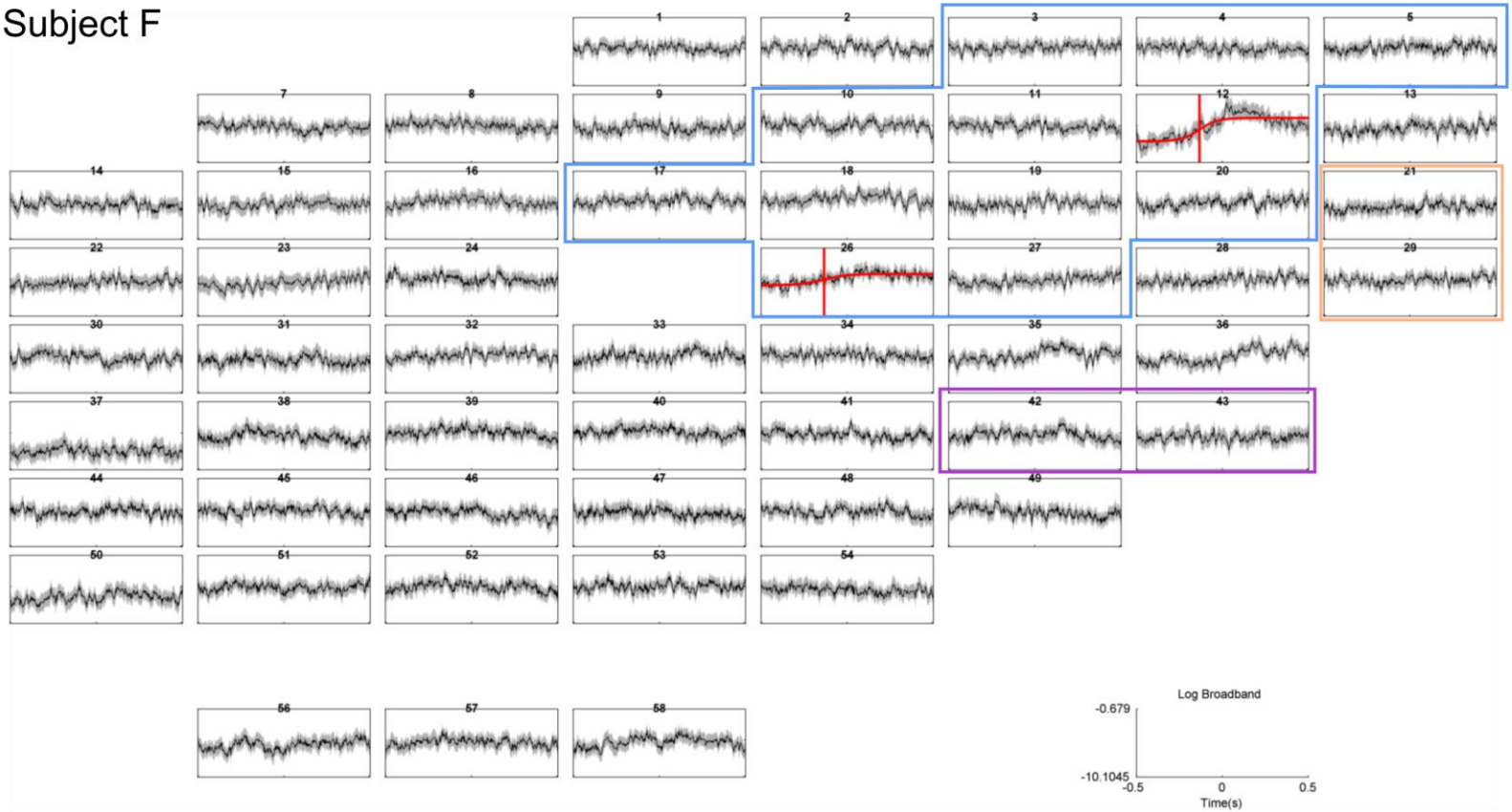
Subject D



Subject E

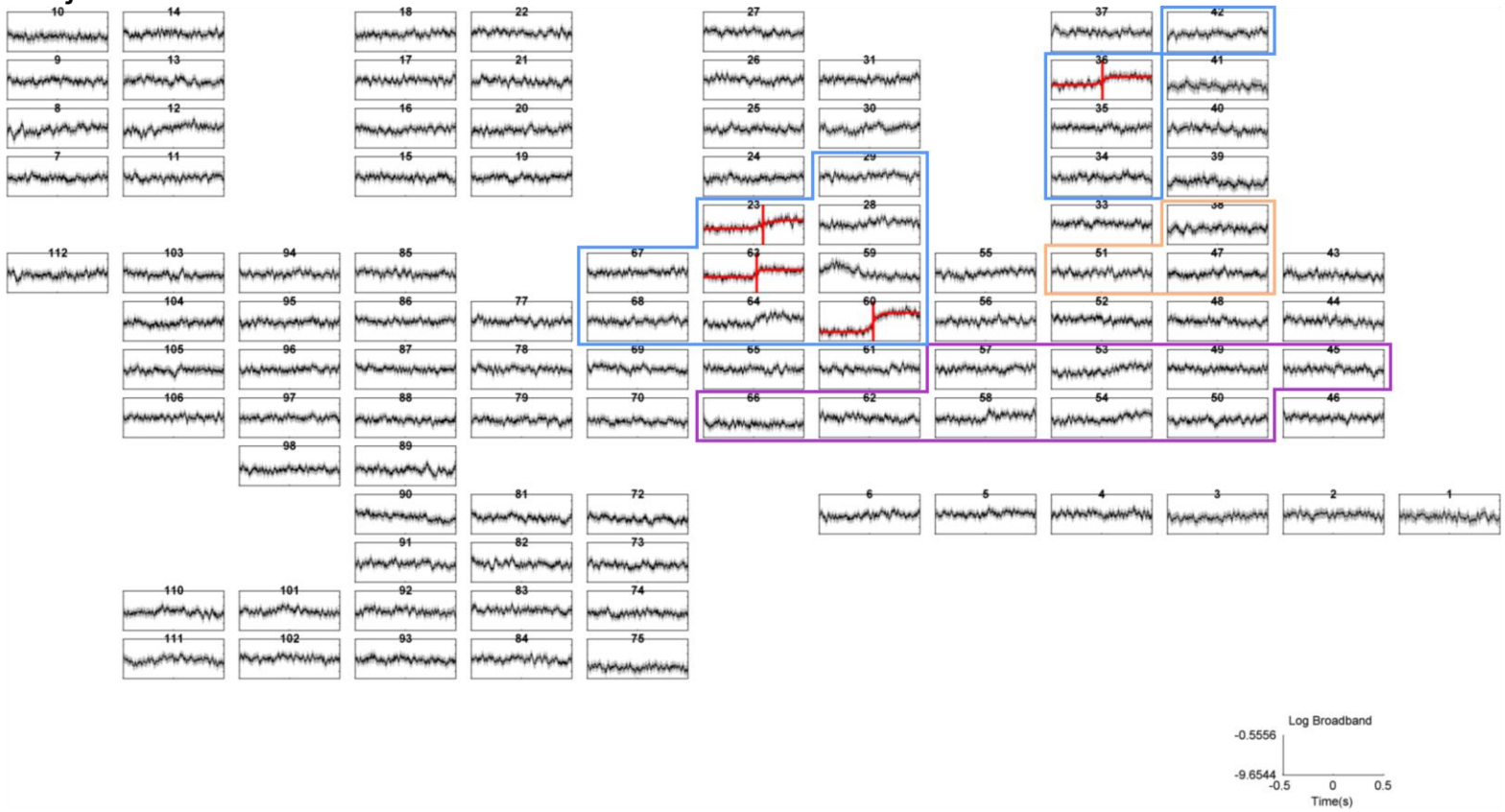


Subject F

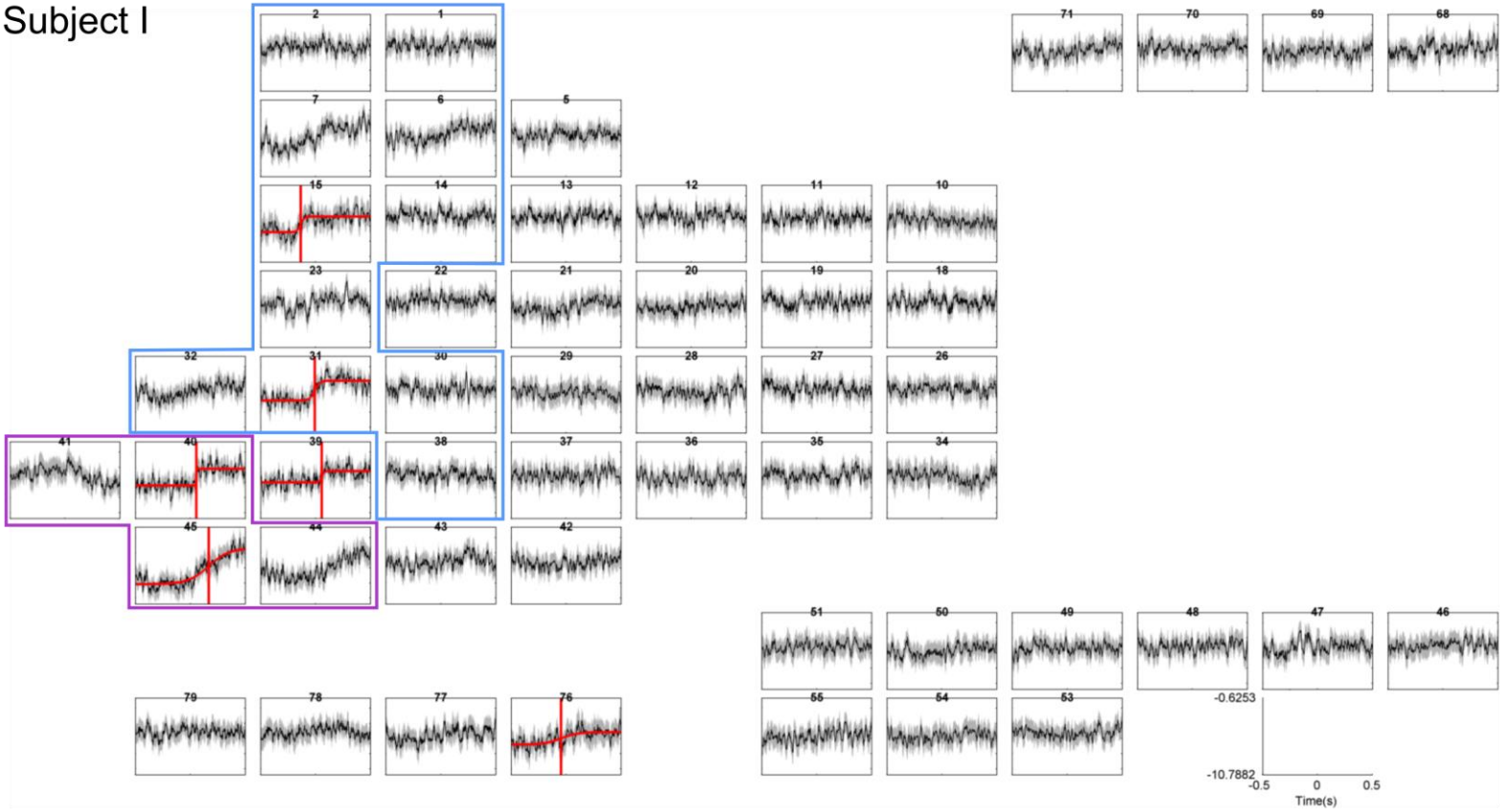

1
2
3
4
5
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
56
57
58
Log Broadband
-0.679
-10.1045
-0.5
0
0.5
Time(s)

Subject G

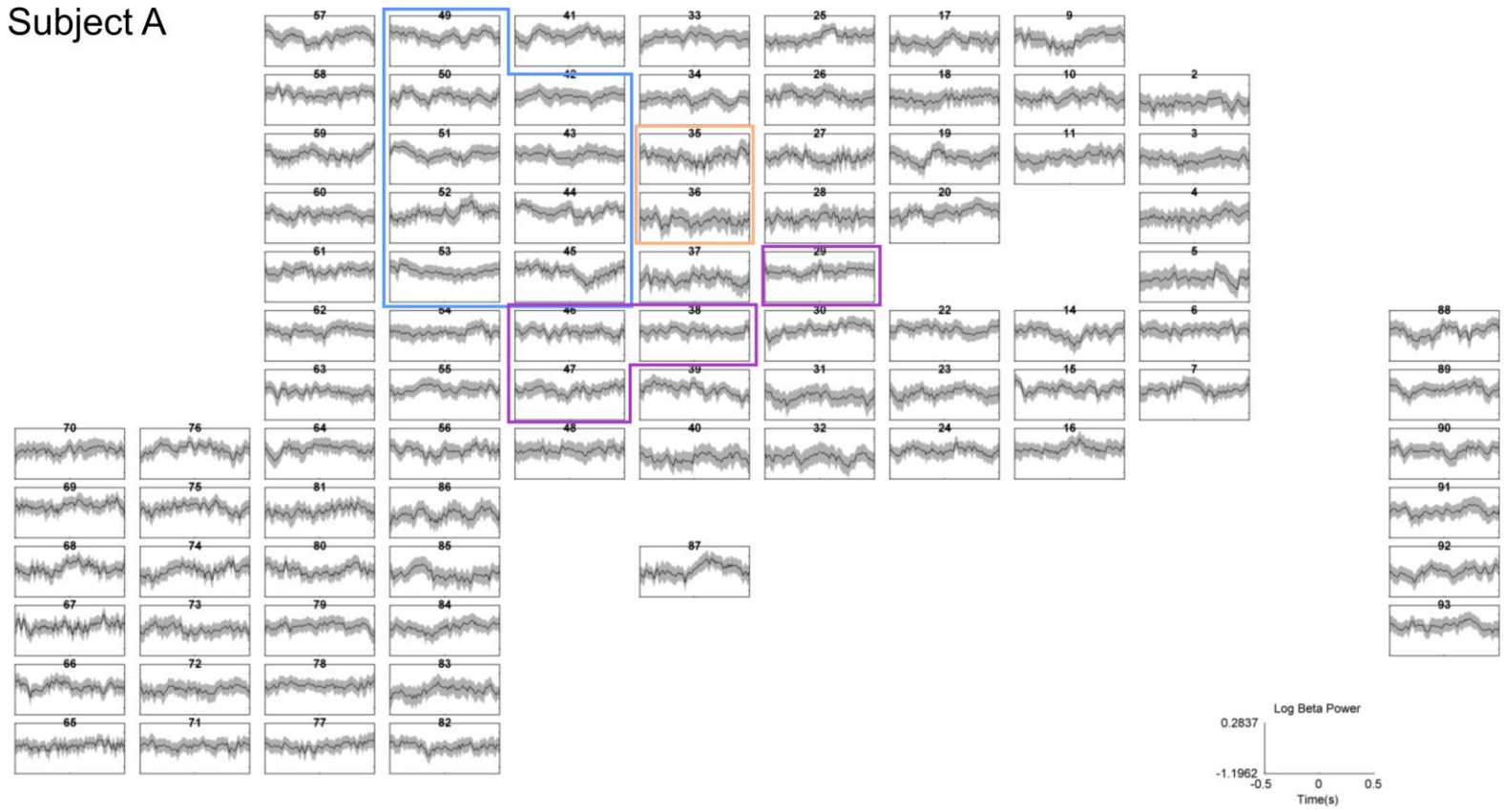
Subject H



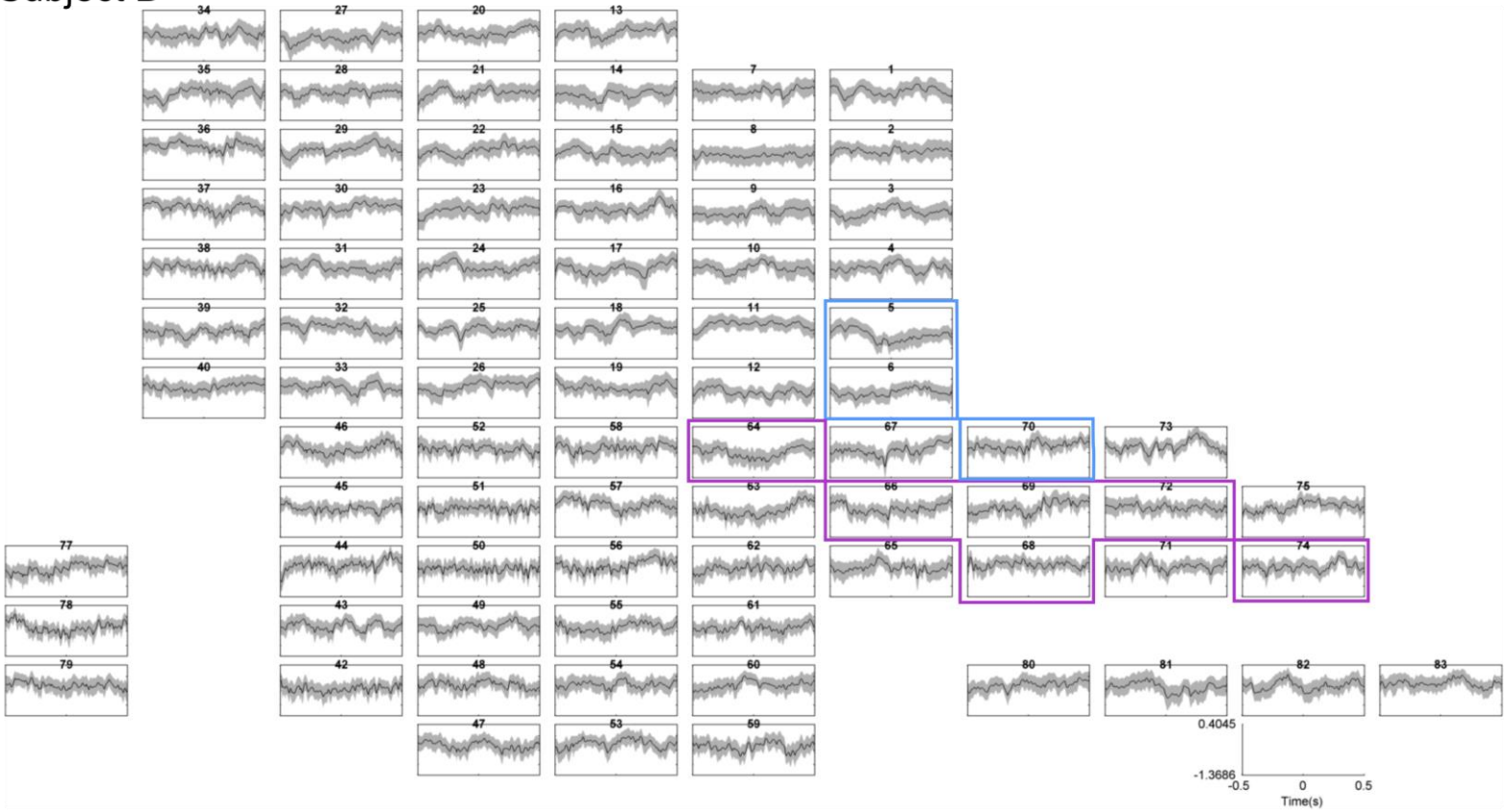
Subject I


2
1
71
70
69
68
7
6
5
15
14
13
12
11
10
23
22
21
20
19
18
32
31
30
29
28
27
26
41
40
39
38
37
36
35
34
45
44
43
42
51
50
49
48
47
46
79
78
77
76
55
54
53
-0.6253
-10.7882
-0.5
0
0.5
Time(s)

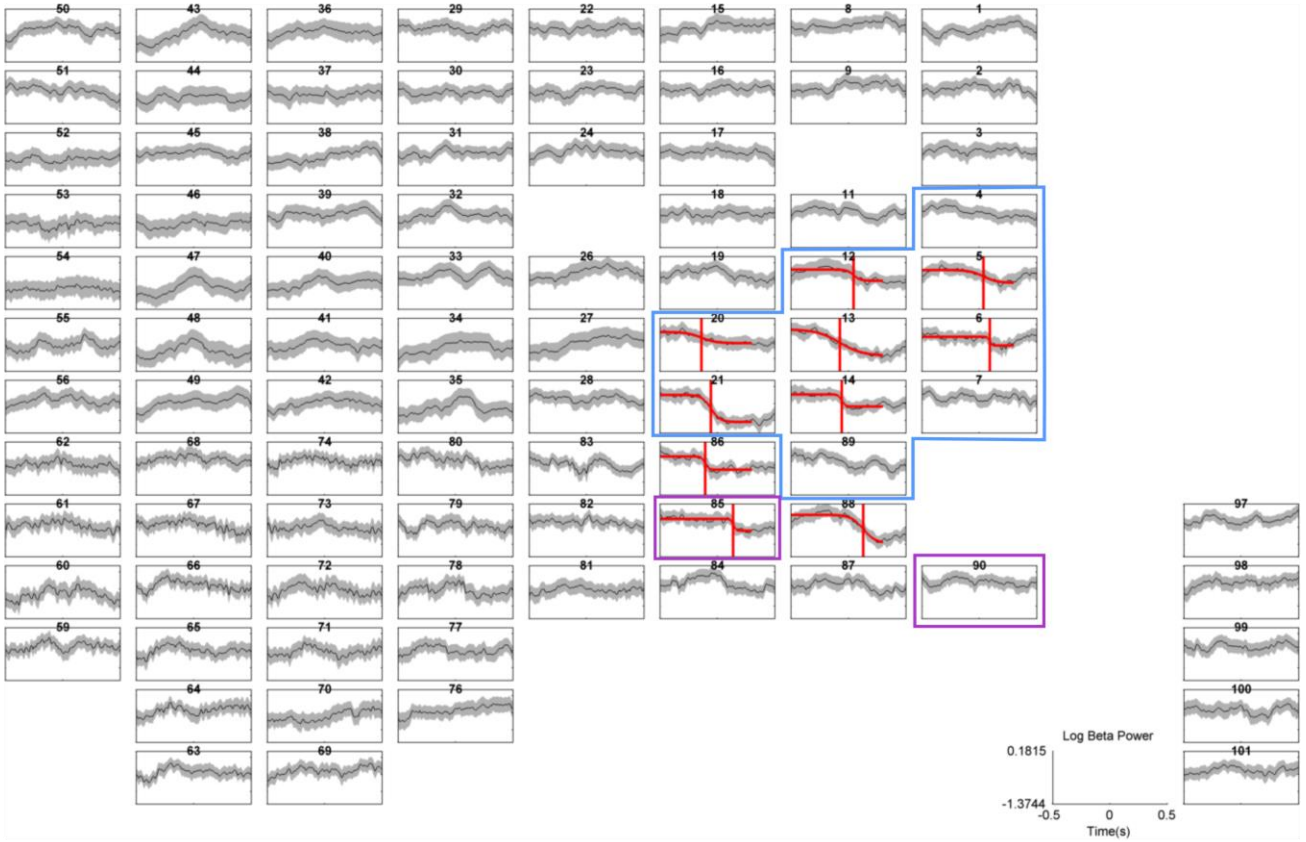
5.5: Beta time course grids with latency for all subjects



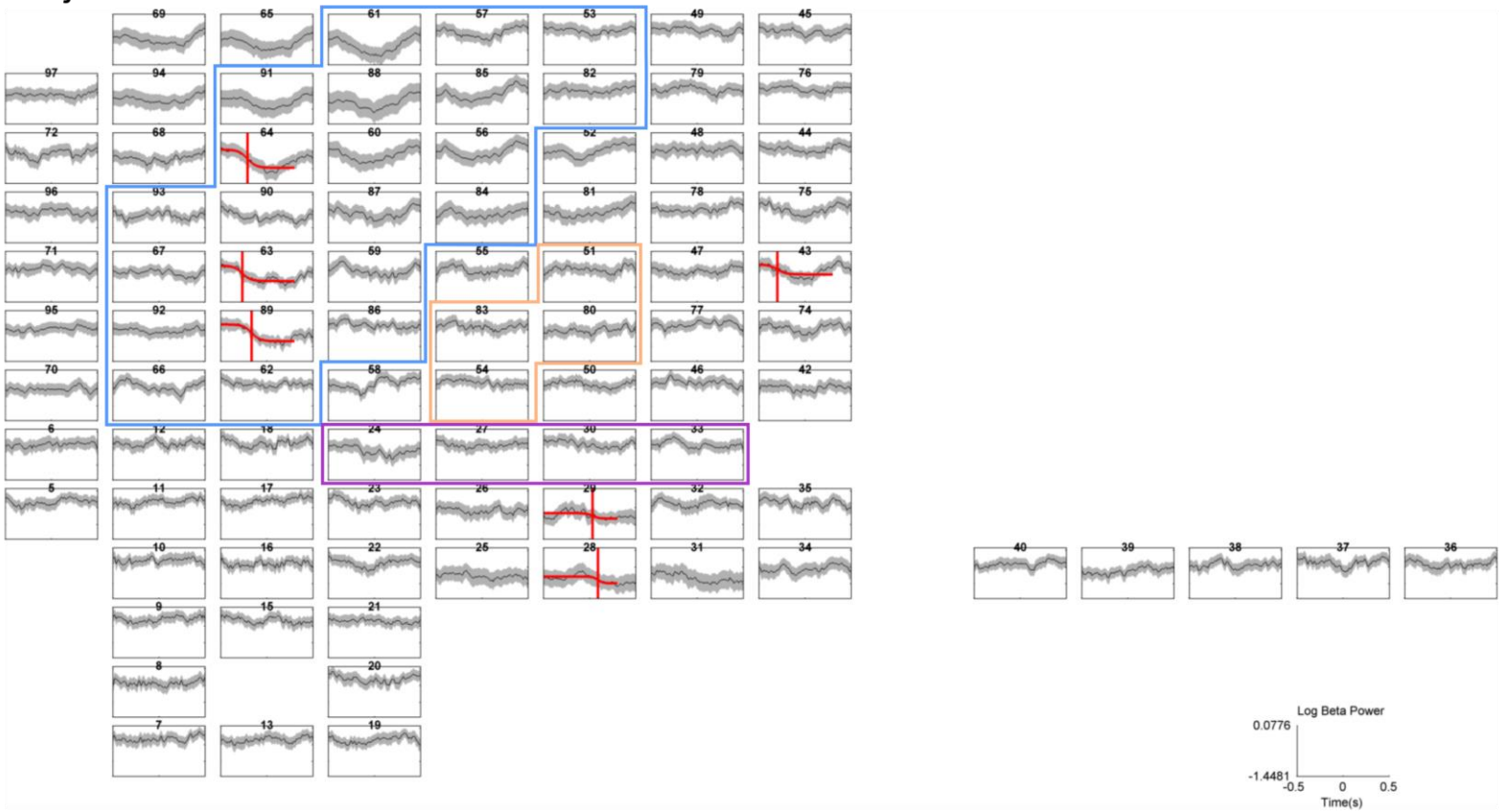
Subject B


34
27
20
13
35
28
21
14
7
1
36
29
22
15
8
2
37
30
23
16
9
3
38
31
24
17
10
4
39
32
25
18
11
5
40
33
26
19
12
6
46
52
58
64
67
70
73
45
51
57
63
66
69
72
75
77
44
50
56
62
65
68
71
74
78
43
49
55
61
79
42
48
54
60
80
81
82
83
47
53
59
0.4045
-1.3686
-0.5
0
0.5
Time(s)

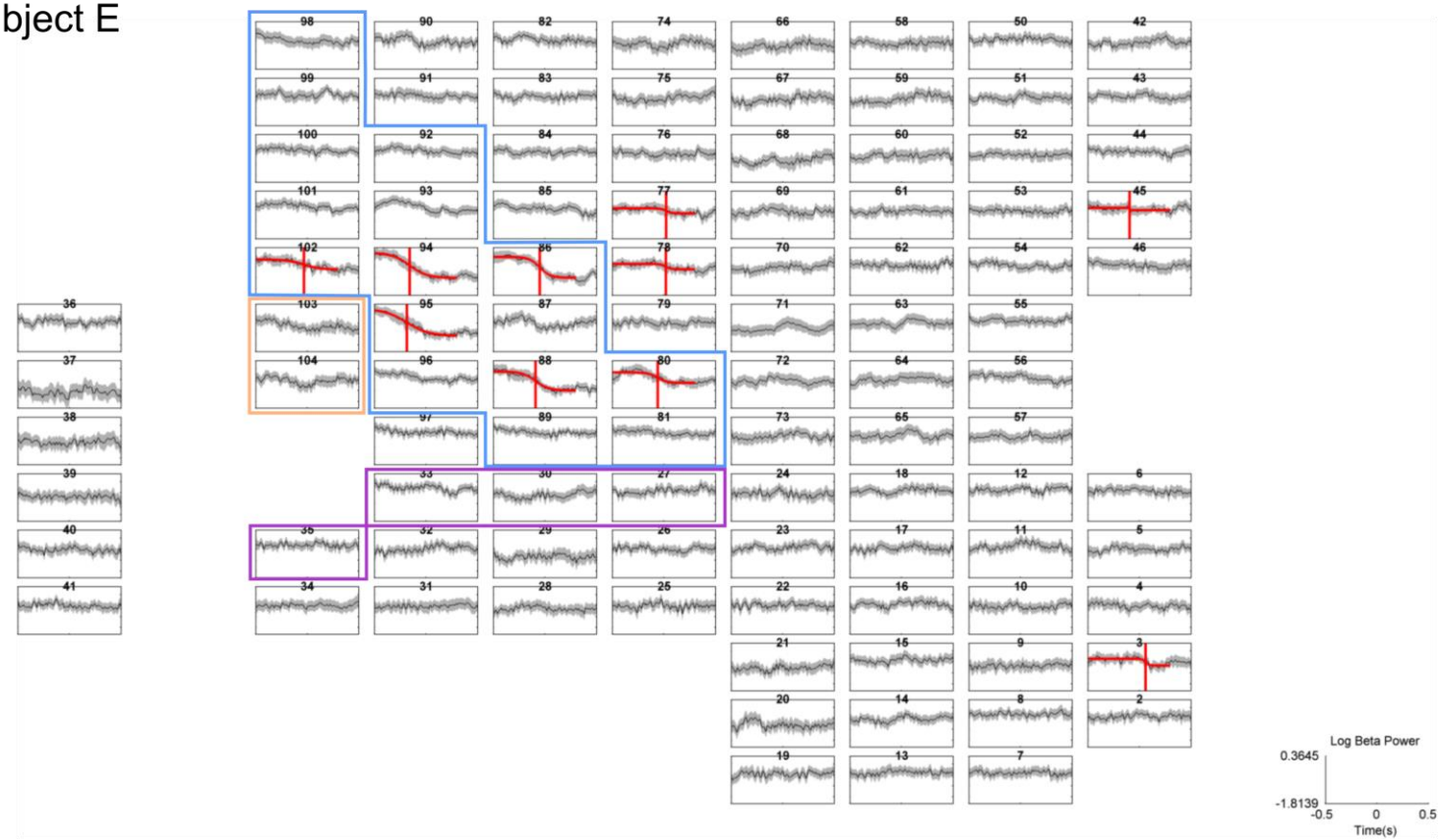
Subject C



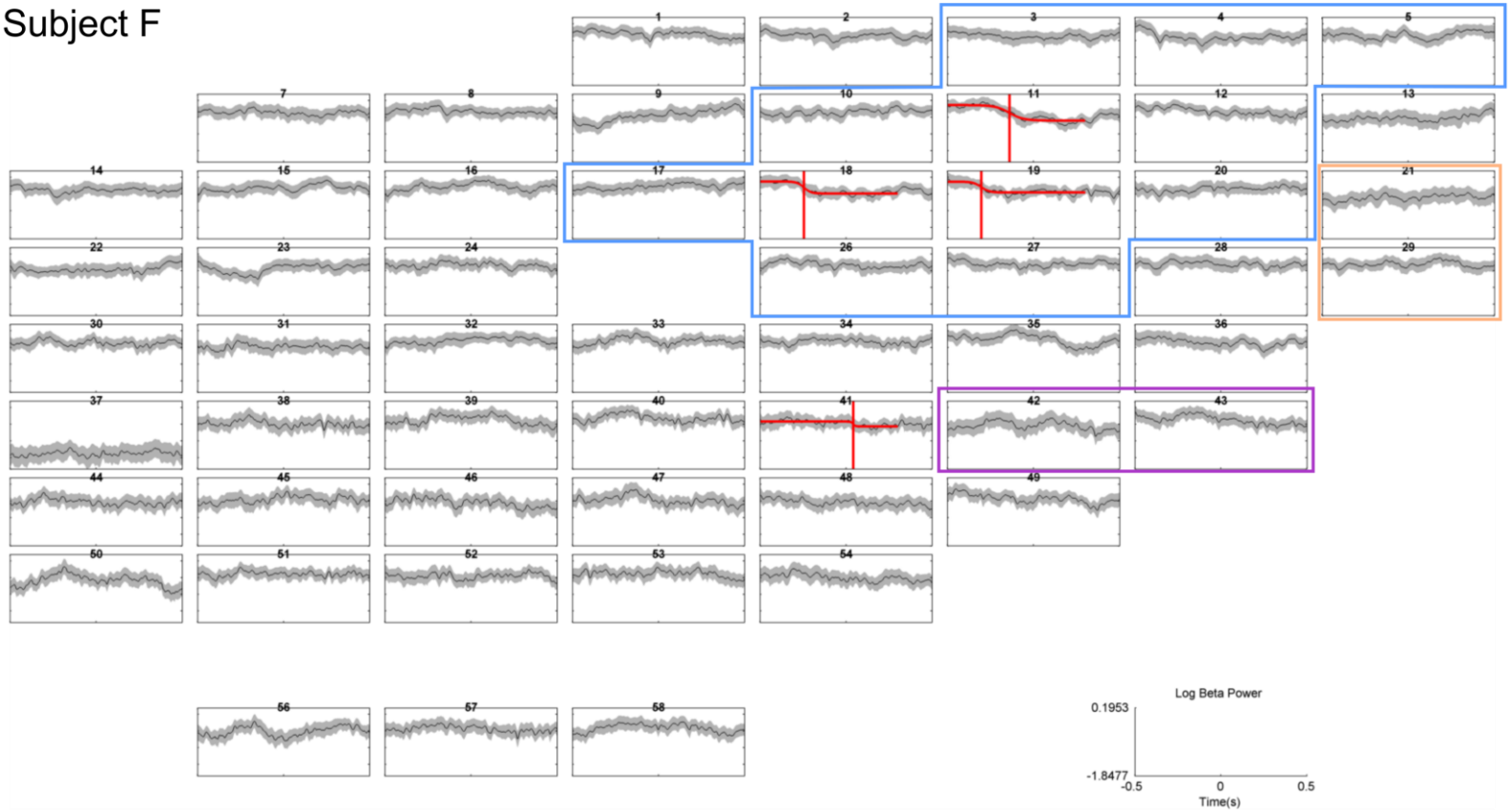
Subject D



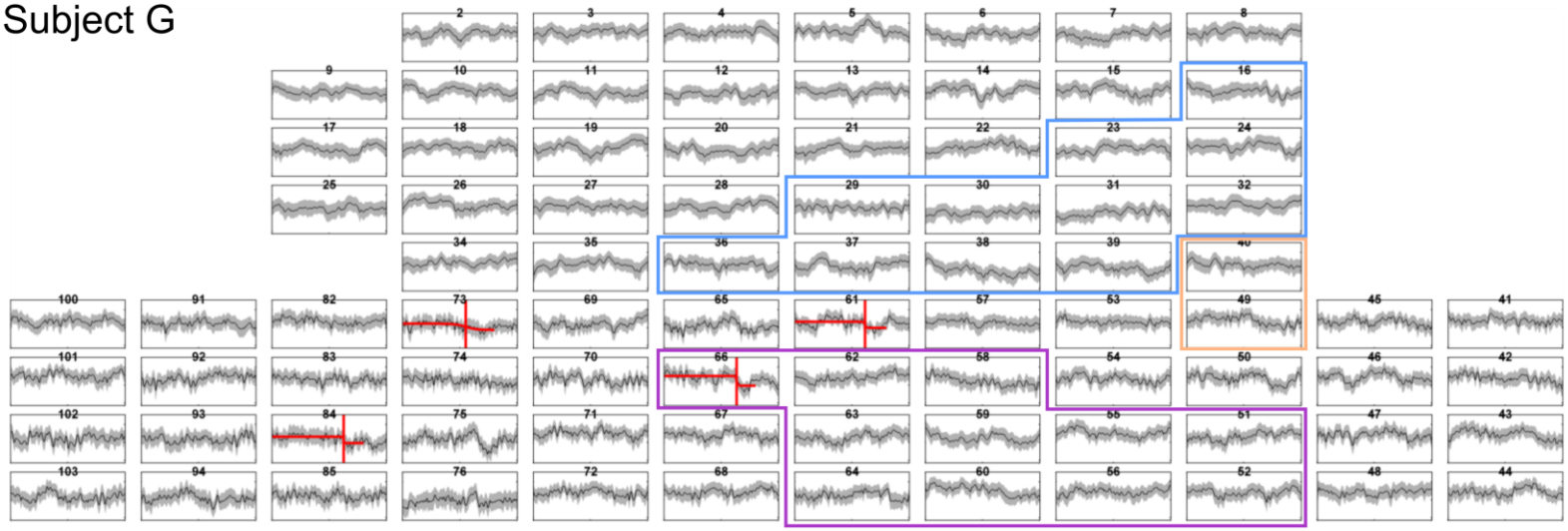
Subject E

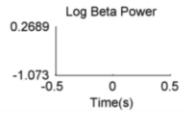


Subject F

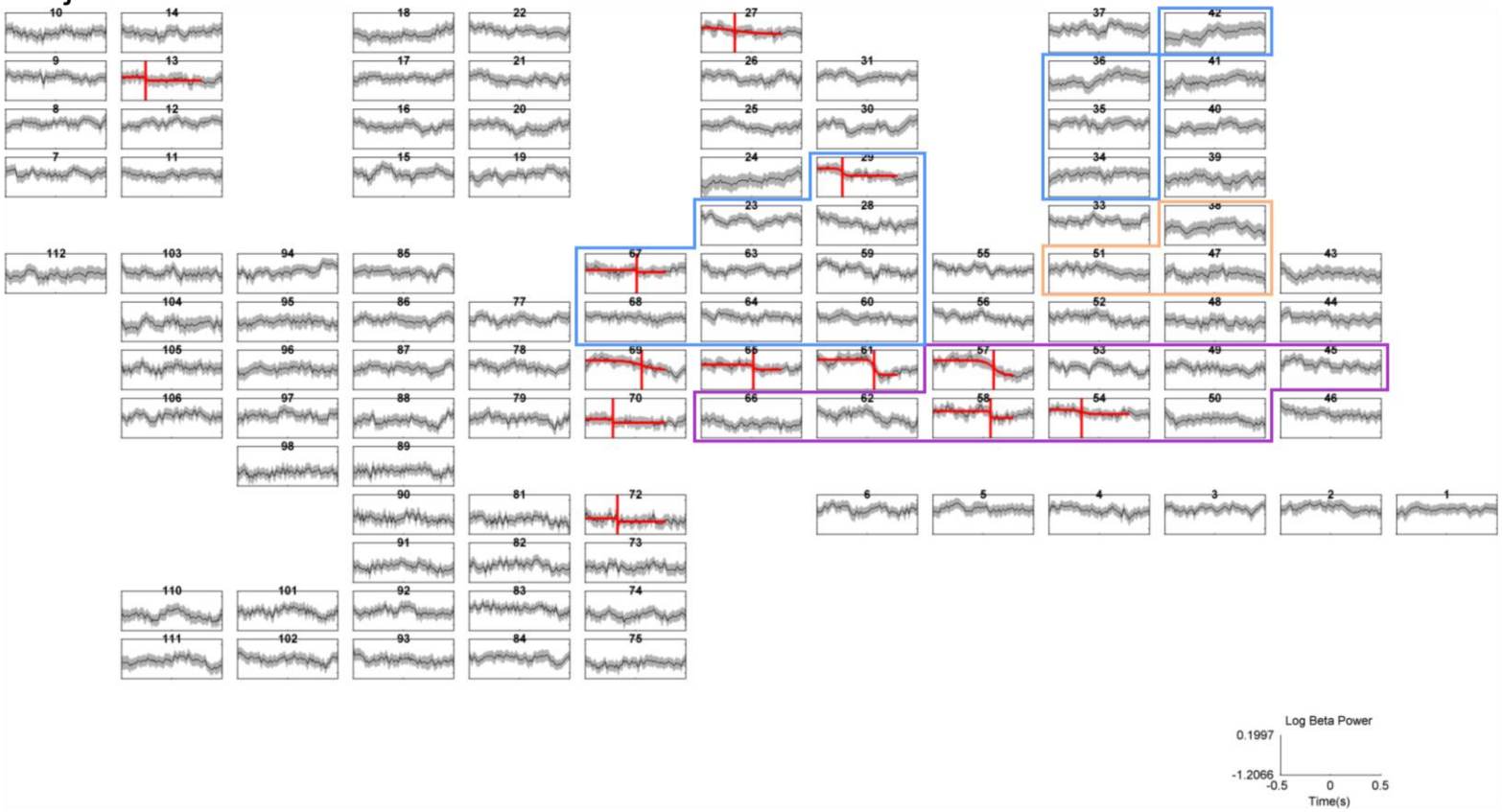

1
2
3
4
5
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
56
57
58
Log Beta Power
0.1953
-1.8477
-0.5
0
0.5
Time(s)

Subject G

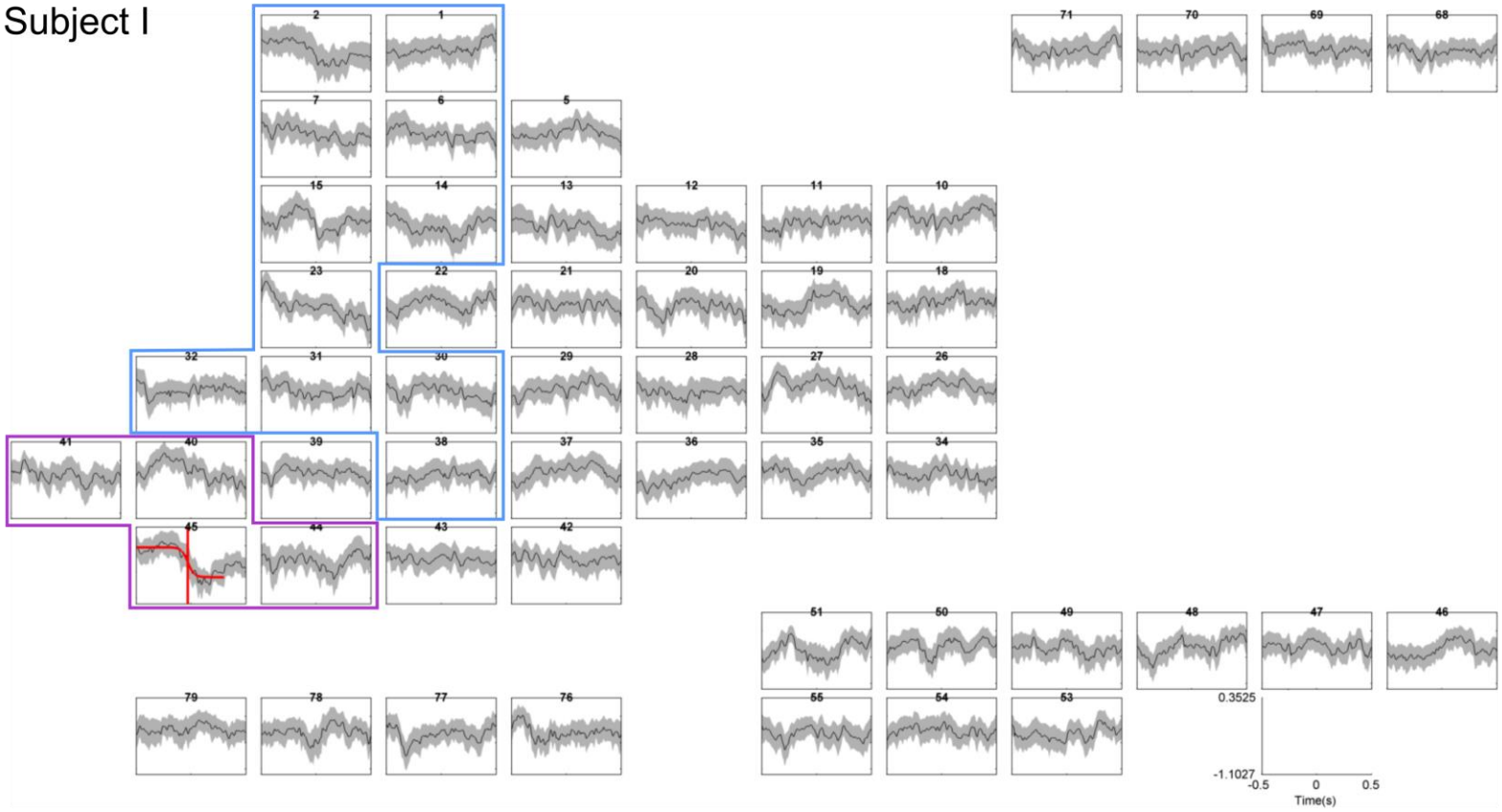

2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
34
35
36
37
38
39
40
100
91
82
73
69
65
61
57
53
49
45
41
101
92
83
74
70
66
62
58
54
50
46
42
102
93
84
75
71
67
63
59
55
51
47
43
103
94
85
76
72
68
64
60
56
52
48
44


Log Beta Power
0.2689
-1.073
-0.5
0
0.5
Time(s)

Subject H



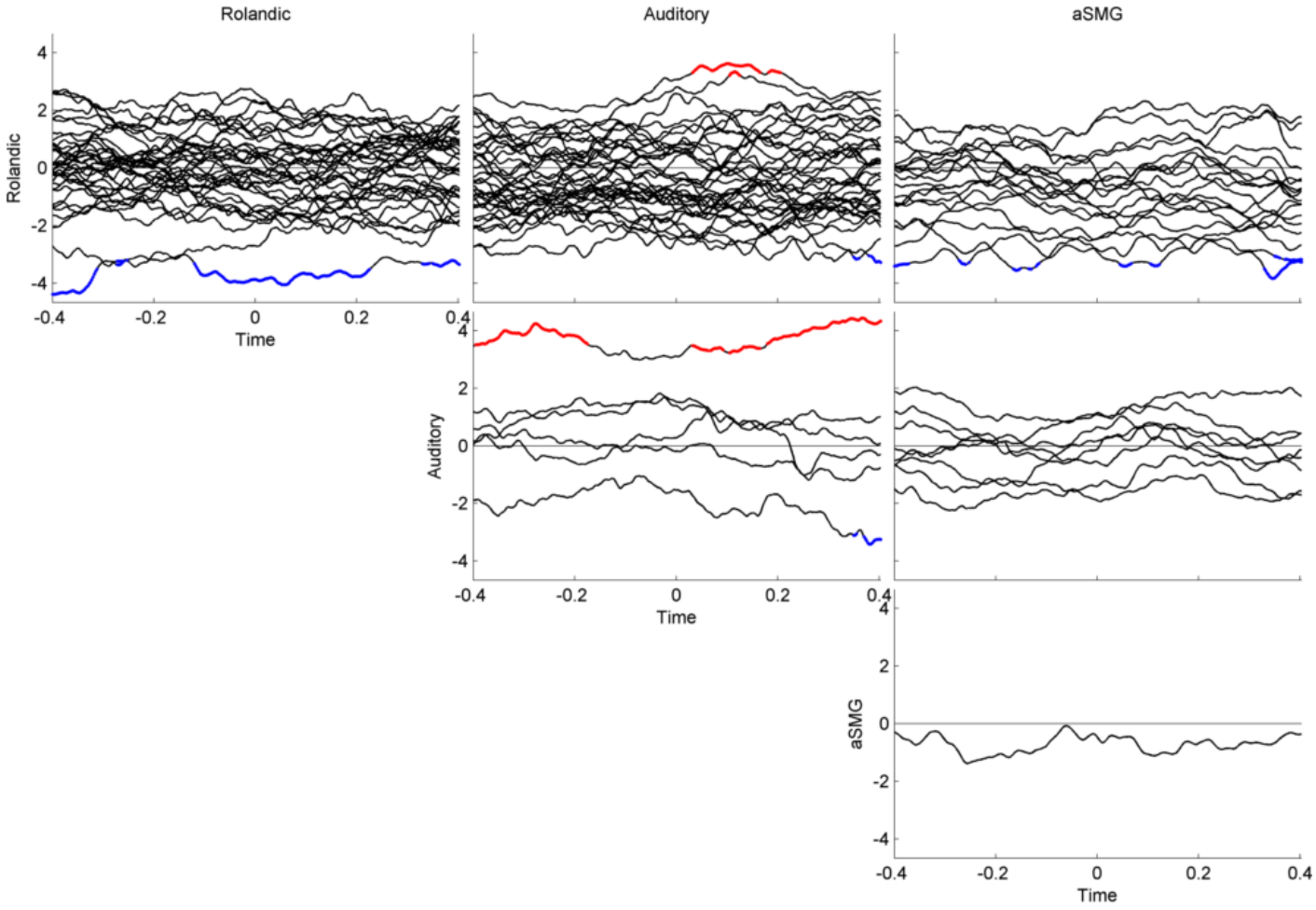
Subject I


2
1
71
70
69
68
7
6
5
15
14
13
12
11
10
23
22
21
20
19
18
32
31
30
29
28
27
26
41
40
39
38
37
36
35
34
45
44
43
42
51
50
49
48
47
46
79
78
77
76
55
54
53
0.3525
-1.1027
-0.5
0
0.5
Time(s)

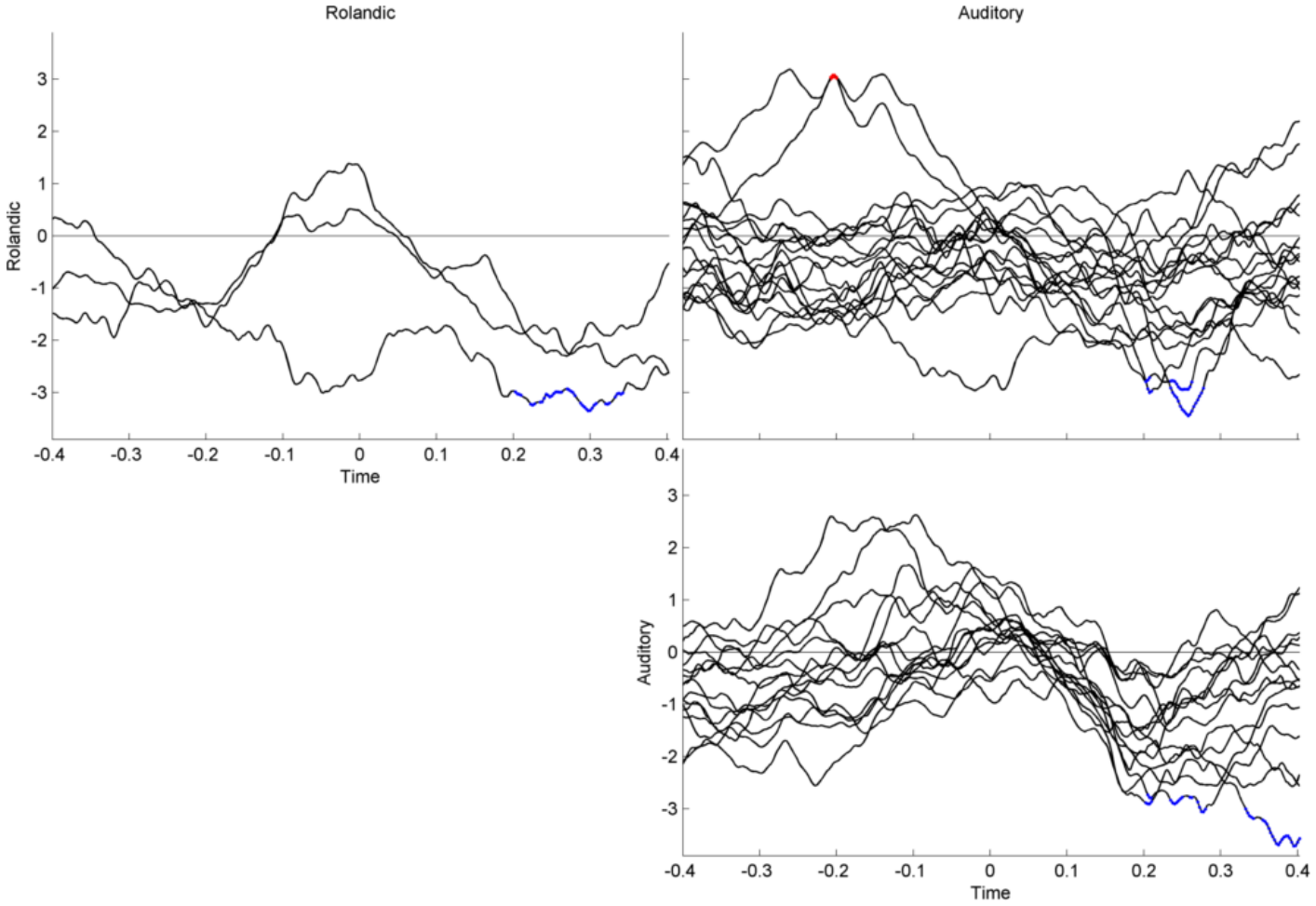
5.6: Correlation time courses for all subjects

For Subjects D and E, see Figures IV.8 and IV.9

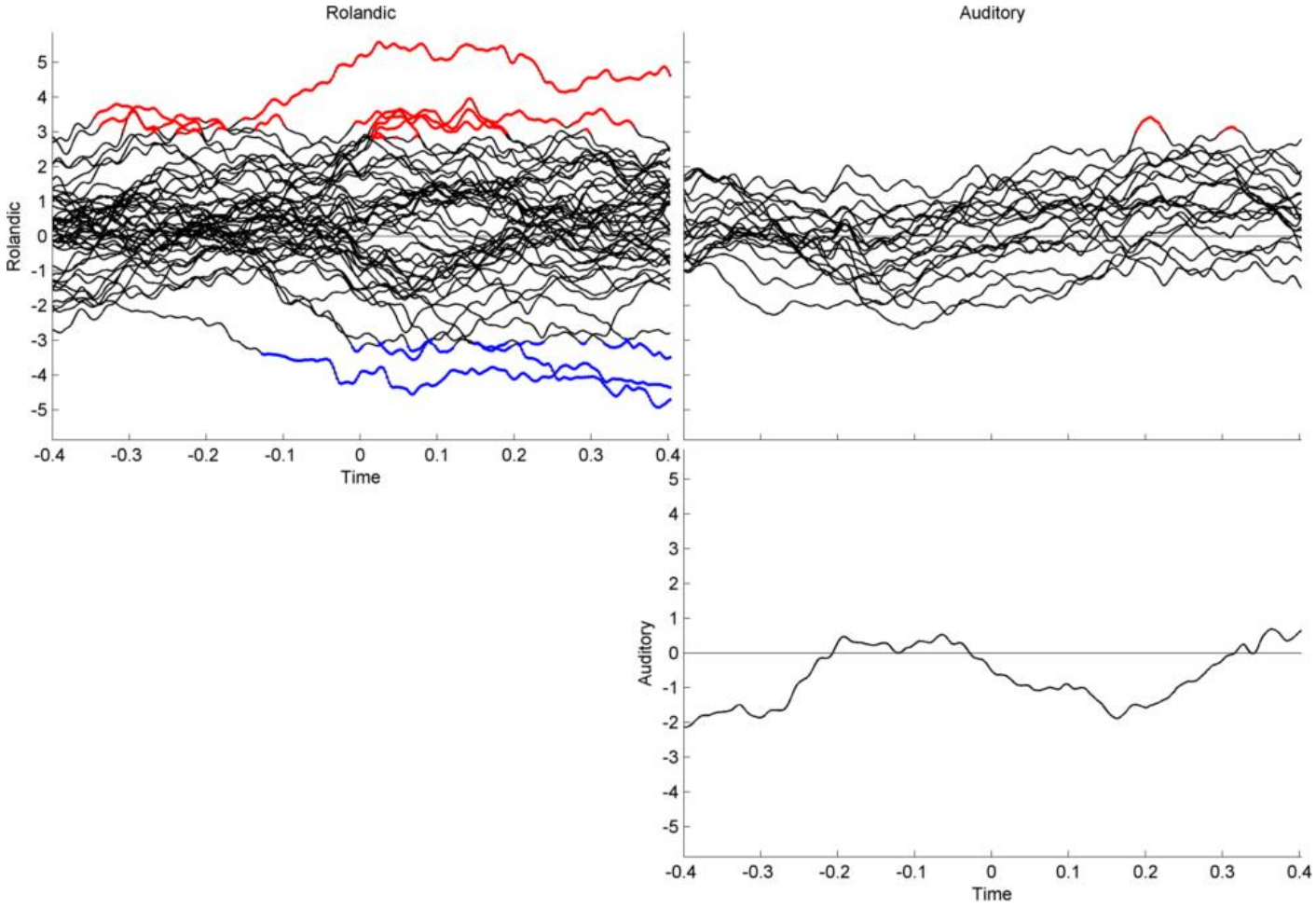
Subject A


Rolandic
Auditory
aSMG
Rolandic
Auditory
aSMG
Time
Time
Time
4
2
0
-2
-4
-0.4
-0.2
0
0.2
0.4

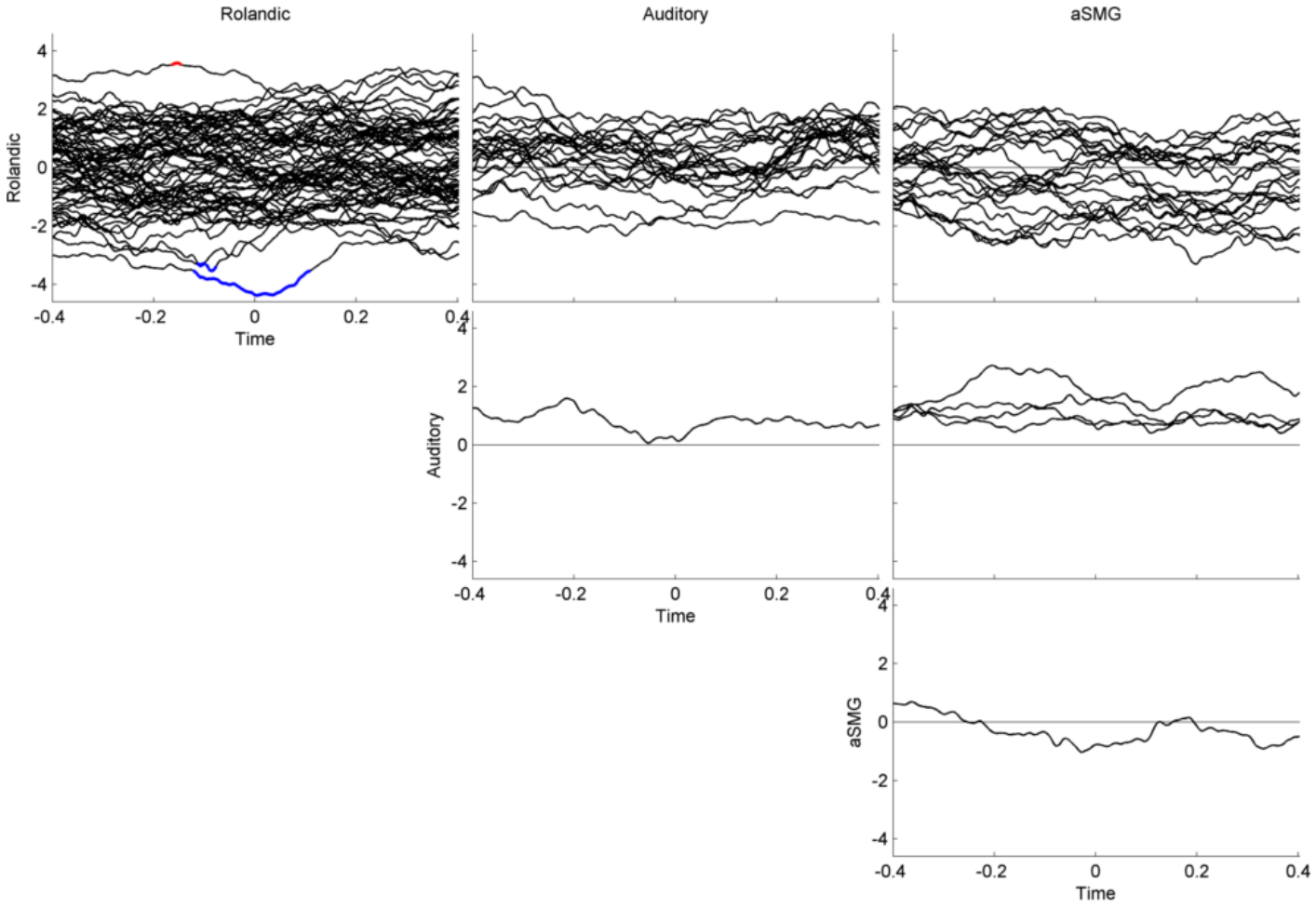
Subject B


Rolandic
Auditory
Rolandic
Auditory
Time
Time
3
2
1
0
-1
-2
-3
-0.4
-0.3
-0.2
-0.1
0
0.1
0.2
0.3
0.4

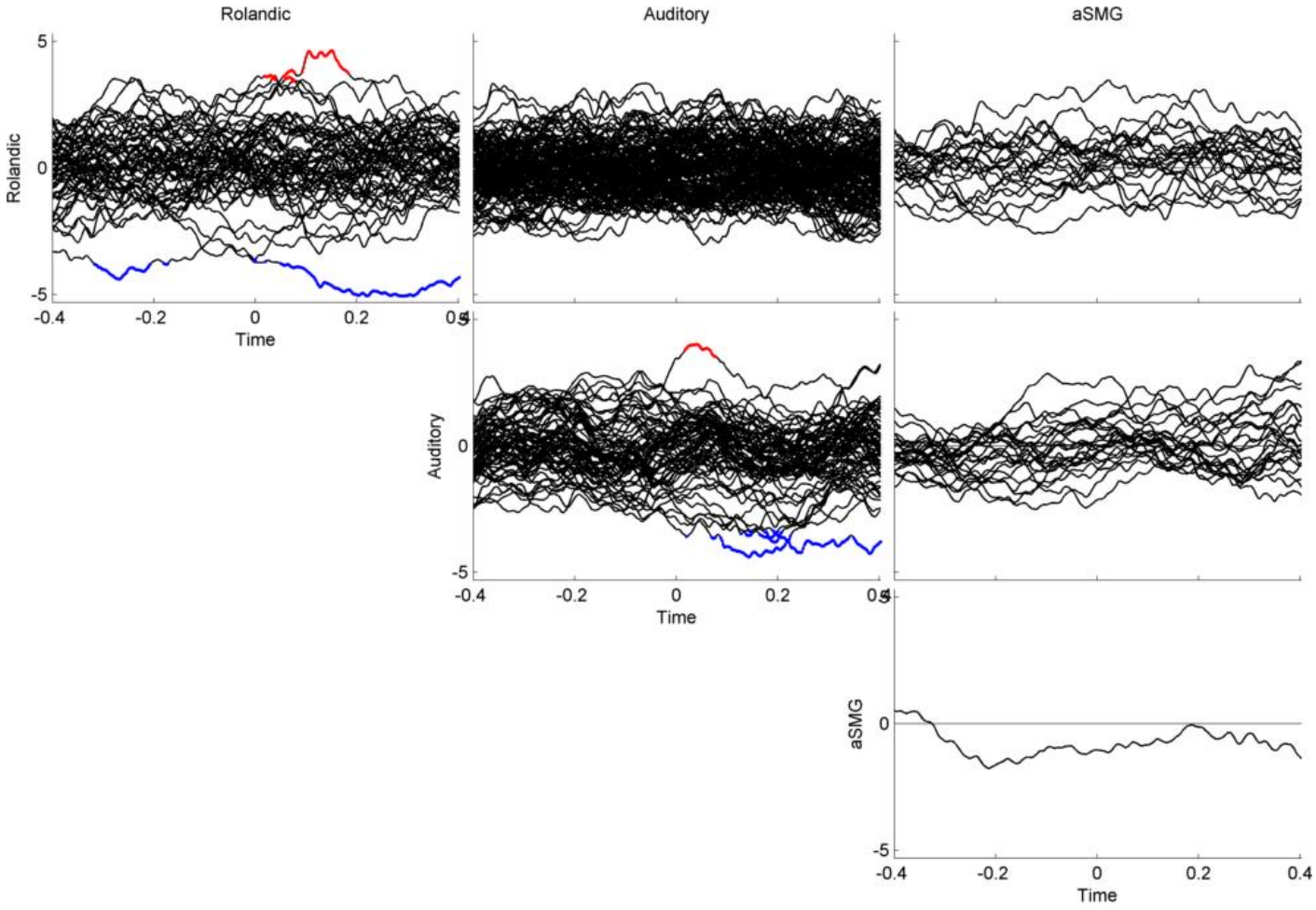
Subject C


Rolandic
Auditory
Rolandic
Auditory
Time
Time
5
4
3
2
1
0
-1
-2
-3
-4
-5
-0.4
-0.3
-0.2
-0.1
0
0.1
0.2
0.3
0.4

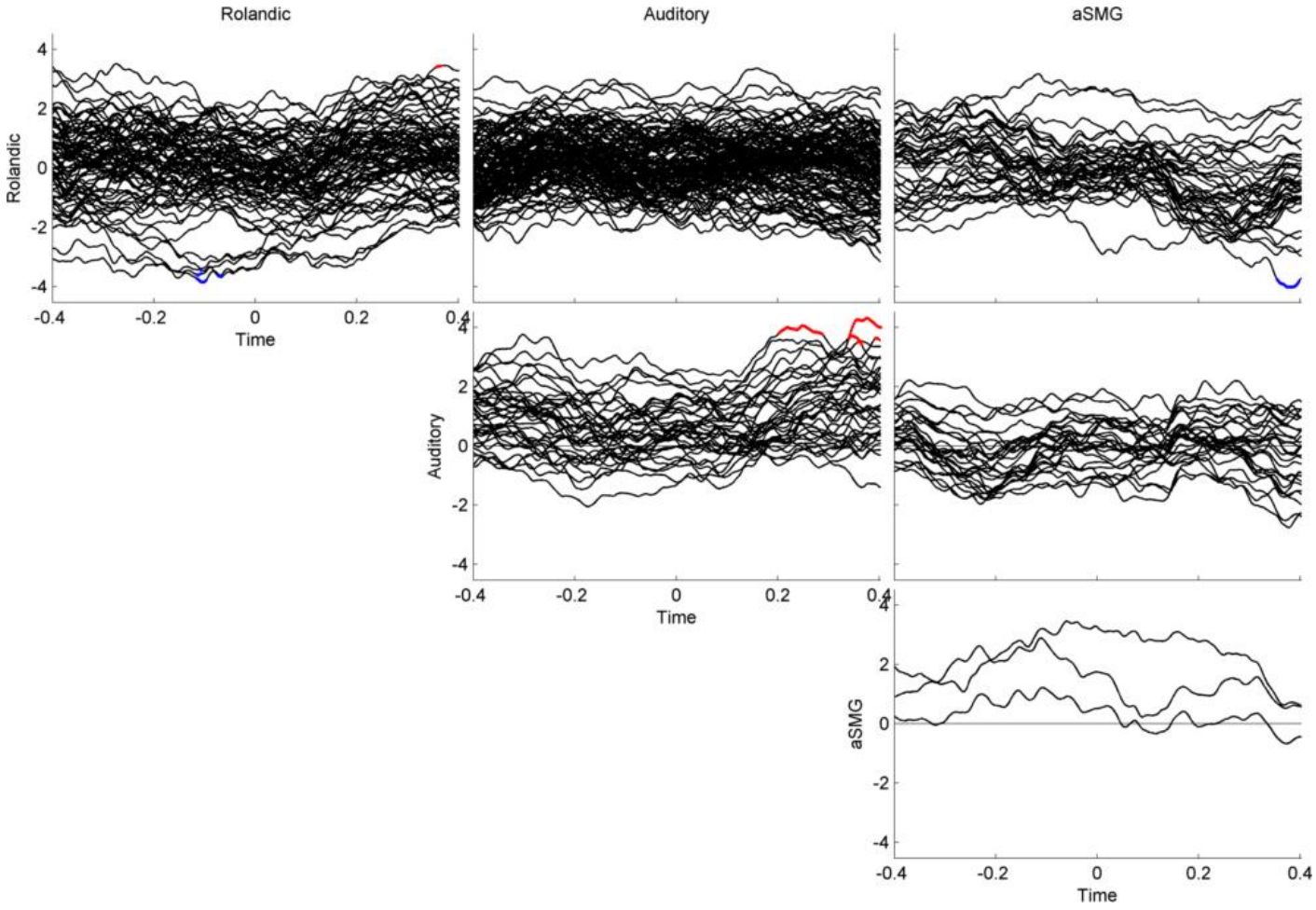
Subject F


Rolandic
Auditory
aSMG
Rolandic
Auditory
aSMG
Time
Time
Time
4
2
0
-2
-4
-0.4
-0.2
0
0.2
0.4

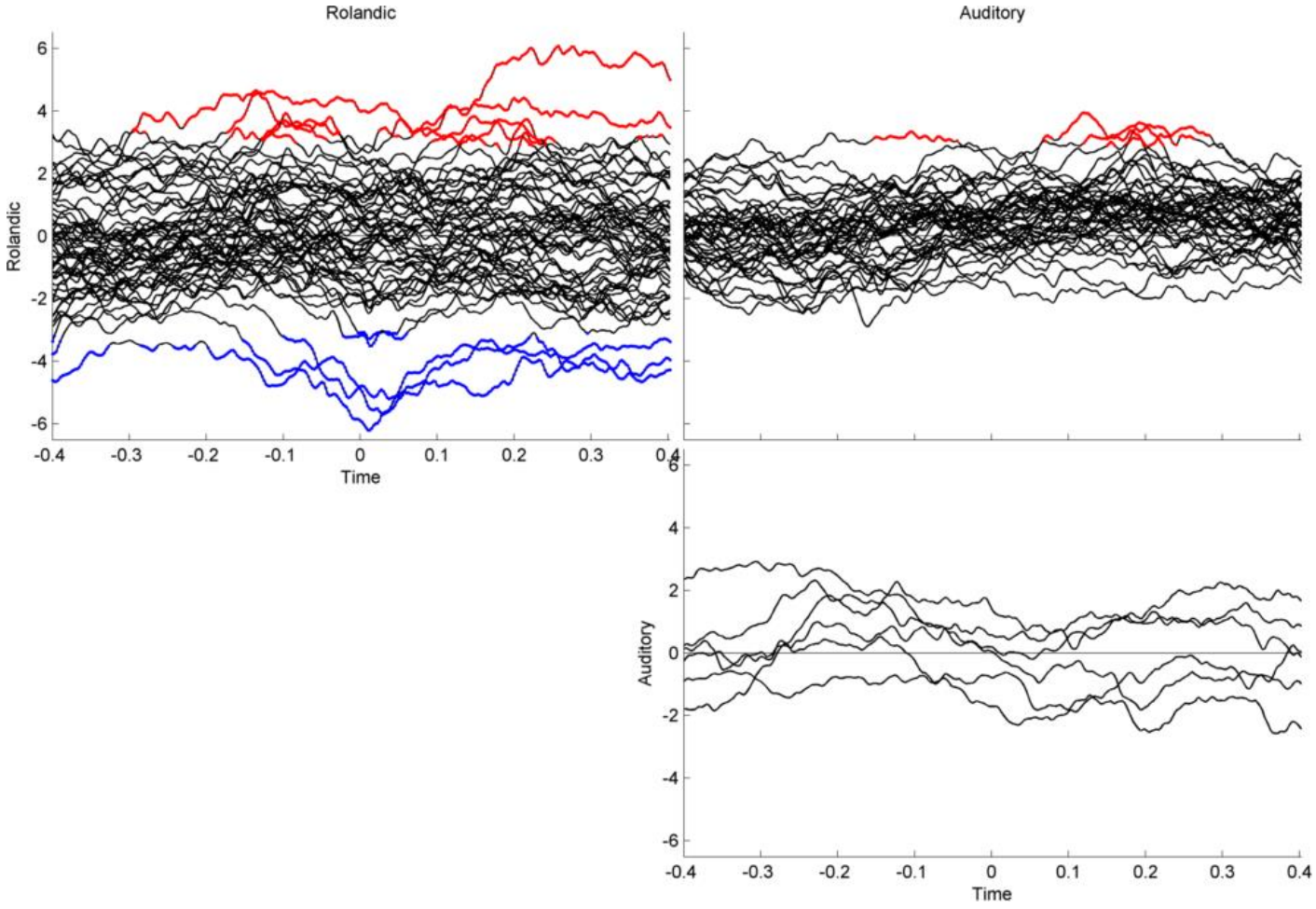
Subject G


Rolandic
Auditory
aSMG
Rolandic
Auditory
aSMG
5
0
-5
-0.4
-0.2
0
0.2
0.4
Time

Subject H


Rolandic
Auditory
aSMG
Rolandic
Auditory
aSMG
Time
Time
Time
4
2
0
-2
-4
-0.4
-0.2
0
0.2
0.4

Subject I


Rolandic
Auditory
Rolandic
Auditory
Time
Time
6
4
2
0
-2
-4
-6
-0.4
-0.3
-0.2
-0.1
0
0.1
0.2
0.3
0.4

CHAPTER V. CONCLUSION

This dissertation concerns the problem of inferring dynamic changes in functional integration related to a task. While functional and effective connectivity analyses have had success in the resting state paradigm, theoretical models of neural communication hypothesize that brain areas dynamically route information based on behavioral needs. With the availability of high temporal resolution data such as ECoG, it should be possible to probe functional networks on the timescale of behavior. We described a statistical framework for inferring changes in coupling that are tied to task dynamics and an application of the framework to an overt reading task. Here we summarize the contributions that this work makes to the field and discuss future directions of research.

1. Innovation and Impact

1.1 Dealing with the Curse of Dimensionality

One of the key issues in modern statistics is the "Curse of Dimensionality", that with increasing data, the degrees of freedom in statistical models tend to grow faster than the data itself. This classical problem is at the core of functional network inference, where the number of possible connections between nodes grows with the square of the number of nodes. With hardware technology allowing for ever more simultaneous neural recordings, functional connectivity inference technology is challenged to keep up. In this dissertation, we developed approaches to deal with the curse of dimensionality in the context of task-related functional connectivity inference.

It is a problem that the degrees of freedom grow faster than the data because degrees of freedom relate directly to the uncertainty in statistical estimates. In the context

of network inference, this means that the more nodes there are in a network, the less reliable the estimate of the network will be. It is common in the functional network literature to report inferred networks without a measure of statistical variability, making it difficult to distinguish true effects from noise. Hence the technique that we described in Chapter II to estimate the variability of networks represents a significant improvement over current practice in the field.

The problem of dimensionality can be mitigated by incorporating prior knowledge into the inference and restricting its scope. For example, choosing frequencies of interest in a coherence analysis based on prior knowledge about frequency bands that are likely to be involved in the task reduces the frequency dimension to a limited number of values. In addition, using theoretical distributions when appropriate, as we described in Chapter III, can eliminate unnecessary variance in the statistical estimators.

Another approach to dealing with dimensionality involves using aggregate network statistics to describe relevant features of the networks even if the details of the networks are imperfectly estimated, for example by using density to describe overall levels of connectivity. Network statistics typically have fewer degrees of freedom than raw networks and their degrees of freedom typically grow more slowly than the dimensions in the data, so in some cases they may be more statistically well-behaved in high-dimensional contexts (this may not be generally true, as discussed in Kolaczyk, 2009). Our technique to estimate uncertainty in aggregate network statistics, described in Chapter II, contributes a valuable tool to this approach.

Since the anatomical locations of the electrodes are typically known, the knowledge can be used to reduce the degrees of freedom by eliminating electrodes that are unlikely to be involved in the task and by grouping electrodes into regions. A barrier to progress in the field of functional network inference, especially for recording modalities with high temporal resolution where spatial averaging is inappropriate, has been the inability to match the spatial scale of inferred networks with the spatial scale of theoretical models of neural communication. By introducing canonical correlation and canonical coherence as coupling measures for functional connectivity analysis in Chapters II and III, we allow for the spatial scale of inferred networks to be larger than the spatial scale of the recordings. This can be used to infer networks at the level of functional areas, which are thought to modulate their connectivity dynamically based on the task (Fries, 2005; Kopell et al., 2010; Wang, 2010; Miller et al., 2012).

Our tools to deal with high-dimensional network inference, being modular and flexible, can be applied to a wide range of experimental paradigms and scientific questions.

1.2 The value of functional connectivity inference

Functional connectivity inference techniques, such as the framework described here, are especially useful for exploratory, descriptive analyses in the absence of strong prior knowledge about the nature of underlying coupling. They are therefore very flexible and applicable to a wide range of paradigms. However, because they do not incorporate generative models of the data or the mechanisms of communication between areas, they do not make inferences about the causes of the inferred connections. Rather, the

descriptive coupling statistics used in functional connectivity analysis are intended to detect possible signatures of communication, which require subsequent interpretation on the part of researchers.

The fact that functional connectivity does not use generative models of coupling mechanisms does not imply that functional connectivity cannot be tailored to different potential coupling mechanisms. For example, in Chapter III we describe how coherence and canonical coherence can outperform correlation and canonical correlation in the situation where the underlying connectivity is inherently rhythmic, as the communication through coherence hypothesis proposes (Fries, 2005; Fries et al., 2007; Fries, 2009). Coupling statistics that are well-matched to the underlying coupling will in general be able to detect coupling better than poorly matched statistics of the same complexity.

1.3 Contributions to speech production research

Chapter IV applies the techniques developed in Chapters II and III to an overt reading paradigm. Since speech research has historically been limited to data with low temporal resolution, using ECoG data has the potential to offer new insights into the dynamics of speech processing. In order to ground the analysis in familiar metrics, we describe the functional connectivity results in the context of first-order features of the data: ERPs, spectrograms, and broadband power averages.

The ERPs averaged over electrodes in rolandic cortex were consistent with prior work on the readiness potential, specifically the late-phase readiness potential over primary motor cortex (Shibasaki and Hallett, 2006). Over auditory cortex, the ERPs varied significantly across electrodes. This result has been observed before (Flinker et al.,

2010), and may be related to the fine-grained functional organization of higher order auditory cortex.

The first-order spectral features of the data were also consistent with previous research, showing a beta depression (Pfurtscheller and Lopes da Silva, 1999) and a broadband power increase (Edwards et al., 2010). The beta depression occurred 113 ms before speech onset over rolandic cortex and 41 ms after speech onset over auditory cortex, averaged over all subjects. The broadband increase occurred later, with an average onset of 52 ms before speech onset for rolandic cortex and 92 ms after speech onset for auditory cortex. The rolandic broadband latency accords well with predictions from the DIVA model, which proposed a 40 ms latency from primary motor cortex activation to EMG activity (Guenther et al., 2006). The DIVA model has an auditory latency of 20 ms: the mismatch here is likely due to the coverage of auditory electrodes in this study, which did not include primary auditory cortex.

In terms of networks, two subjects (D and E) showed interesting network dynamics. Subject E had dense correlation-based connectivity between rolandic electrodes that was positive within electrodes over the precentral gyrus and negative between the precentral and postcentral gyri. The precentral connections appeared at about the same time as alpha coherence-based connections, about 150 ms before speech onset, and the pre- to postcentral connections appeared at about the same time as negative beta connections, right at speech onset. The negative beta connections were likely related to a loss of beta power and suggest that the baseline interval, during silences, contained coherent beta activity across the central sulcus. This is consistent with a role of beta in

motor preparation and attention (Engel and Fries, 2010; Wang, 2010). The increase in alpha coherence was not accompanied by an increase in alpha power, and may point to a common thalamic driver (Hughes, 2005; Jones et al., 2009). Subject D had similar alpha-frequency effects, and a different beta-frequency loss of coherence, between rolandic and auditory cortex, potentially pointing to rolandic-auditory coherence during baseline. While these results only occurred in 2 of 9 subjects, it is notable that these two subjects had a relatively large number of trials and good spatial coverage of rolandic cortex. Hence a more structured experimental paradigm with more trials may uncover more robust effects.

2. Future Directions

2.1 Statistics

In many situations, the number of statistical inferences grows as the dimensionality increases. This is the case with the functional network analyses described here, which estimate the edge weight for each edge and apply a statistical hypothesis test to each. This raises the problem of multiple testing, which we correct using the False Discovery Rate procedure, FDR (Benjamini and Hochberg, 1995; Kramer et al., 2009). FDR, however, is sensitive to model misspecification, a problem that becomes increasingly noticeable with large numbers of hypothesis tests. We dealt with this potential problem in Chapter III by improving the statistical model introduced in Chapter II, including more sources of known variance. Another approach, the empirical FDR (Efron, 2010), uses the data itself to define a null hypothesis, under the assumption that most of the hypothesis tests should fall under the null. This approach is a part of a family

of techniques developed largely for large-scale genetics work, and could be very useful for functional connectivity analysis[1].

2.2 Speech

Speech is difficult to study. It involves a large number of brain areas both cortical and subcortical, all interacting dynamically to produce coordinated behavior at a fine temporal scale. This requires recording modalities that have fine temporal and spatial resolutions in addition to wide spatial coverage. No current recording modalities have all of these features, so speech researchers must find ways to compromise and target parts of the system. The fact that speech can only be studied in humans further limits the recording modalities that are available for scientific use. While ECoG has good temporal resolution, adequate spatial resolution and spatial scale, and good signal-to-noise ratios, the fact that it is limited to patients undergoing a stressful clinical procedure limits the amount of time that researchers can spend with the patients and the quality and amount of the data collected. In this context, it is essential to be able to optimize the statistical techniques being applied to the data to maximize statistical power.

The methodologies developed in this dissertation are designed for specifically this situation. While the network results were noisy, there is reason to believe that with a more structured experimental paradigm and more trials, the interesting network results from Subjects D and E may replicate to more subjects.

---

[1] I tried this, but ran into problems because the connections in my data set do not easily separate into empirical null distribution vs everything else. That is, I get empty networks. I interpret this to mean that binary networks are not supported by the data. It would be interesting to try in a context with more statistical power.

In addition, other statistical techniques could narrow down the sources of the effects observed in Chapter IV. For example, the alpha coherence over rolandic cortex during speech was suggestive of all-to-all connectivity, which could be better detected using a region-level coherence statistic such as global coherence (Cimenser et al., 2011). This type of strong region-level coherence may be indicative of a common thalamic driver, and an approach such as dynamic causal modeling (Stephan et al., 2010) could be used to test a generative model based on this hypothesis against other potential models. Developing generative models is the logical next step after the descriptive analytics have been used to develop hypotheses.

DIVA is one such generative model, and it could be tested and improved using ECoG experiments. The specific features of ECoG data open up possibilities to study aspects of speech that have previously been difficult to study. For example, almost all ECoG patients have coverage of primary motor cortex, primary somatosensory cortex, and the superior temporal gyrus. These areas are all important components of the feedback subsystems in the DIVA model, meaning that ECoG data could be particularly suited to testing the model's predictions of speech network responses during auditory or somatosensory perturbation paradigms (Guenther et al., 2006).

BIBLIOGRAPHY

Achard, S., Salvador, R., Whitcher, B., Suckling, J., and Bullmore, E. (2006). A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs. *Journal of neuroscience* 26**,** 63-72.

Allen, E.A., Damaraju, E., Plis, S.M., Erhardt, E.B., Eichele, T., and Calhoun, V.D. (2012). Tracking Whole-Brain Connectivity Dynamics in the Resting State. *Cerebral Cortex*. doi:10.1093/cercor/bhs352.

Anderson, T.W. (2003). *An introduction to multivariate statistical analysis.* Hoboken, N.J.: Wiley-Interscience.

Aoki, F., Fetz, E., Shupe, L., Lettich, E., and Ojemann, G. (2001). Changes in power and coherence of brain activity in human sensorimotor cortex during performance of visuomotor tasks. *Biosystems* 63**,** 89-99.

Arroyo, S., Lesser, R.P., Gordon, B., Uematsu, S., Jackson, D., and Webber, R. (1993). Functional significance of the mu rhythm of human cortex: an electrophysiologic study with subdural electrodes. *Electroencephalography and clinical neurophysiology* 87**,** 76-87.

Arvesen, J.N. (1969). Jackknifing U-statistics. *The Annals of Mathematical Statistics***,** 2076-2100.

Astolfi, L., Cincotti, F., Mattia, D., De Vico Fallani, F., Salinari, S., Marciani, M., Witte, H., and Babiloni, F. (2009). "Study of the time-varying cortical connectivity changes during the attempt of foot movements by spinal cord injured and healthy subjects", in: *IEEE Engineering in Medicine and Biology Society Annual Conference EMBS*: IEEE), 2208-2211.

Baker, S.N. (2007). Oscillatory interactions between sensorimotor cortex and the periphery. *Current Opinion in Neurobiology* 17**,** 649-655. doi:10.1016/j.conb.2008.01.007.

Bassett, D.S., Meyer-Lindenberg, A., Achard, S., Duke, T., and Bullmore, E. (2006). Adaptive reconfiguration of fractal small-world human brain functional networks. *Proceedings of the National Academy of Sciences* 103**,** 19518-19523.

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* 57**,** 289-300.

Bettus, G., Ranjeva, J.-P., Wendling, F., Bénar, C.G., Confort-Gouny, S., Régis, J., Chauvel, P., Cozzone, P.J., Lemieux, L., and Bartolomei, F. (2011). Interictal

functional connectivity of human epileptic networks assessed by intracerebral EEG and BOLD signal fluctuations. *PLoS ONE* 6**,** e20071.

Biswal, B., Yetkin, F.Z., Haughton, V.M., and Hyde, J.S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magnetic resonance in medicine* 34**,** 537-541.

Bokil, H., Purpura, K., Schoffelen, J.M., Thomson, D., and Mitra, P. (2007). Comparing spectra and coherences for groups of unequal size. *Journal of neuroscience methods* 159**,** 337-345. doi:S0165-0270(06)00328-1

Bolstad, A., Van Veen, B.D., and Nowak, R. (2011). Causal Network Inference Via Group Sparse Regularization. *IEEE Transactions on Signal Processing* 59**,** 2628-2641. doi:10.1109/TSP.2011.2129515.

Börgers, C., and Kopell, N.J. (2008). Gamma oscillations and stimulus selection. *Neural computation* 20**,** 383-414. doi:10.1162/neco.2007.07-06-289.

Bostock, M., Ogievetsky, V., and Heer, J. (2011). D3 Data-Driven Documents. *IEEE Transactions on Visualization and Computer Graphics* 17**,** 2301-2309. doi:10.1109/tvcg.2011.185.

Bouchard, K.E., Mesgarani, N., Johnson, K., and Chang, E.F. (2013). Functional organization of human sensorimotor cortex for speech articulation. *Nature* 495**,** 327-332. doi:10.1038/nature11911.

Breshears, J.D., Gaona, C.M., Roland, J.L., Sharma, M., Bundy, D.T., Shimony, J.S., Rashid, S., Eisenman, L.N., Hogan, R.E., and Snyder, A.Z. (2012). Mapping sensorimotor cortex using slow cortical potential resting-state networks while awake and under anesthesia. *Neurosurgery* 71**,** 305.

Brillinger, D.R. (1981). *Time series : data analysis and theory.* San Francisco: Holden-Day.

Brillinger, D.R. (2001). *Time series : data analysis and theory.* Philadelphia: Society for Industrial and Applied Mathematics.

Bullmore, E., and Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience* 10**,** 186-198. doi:10.1038/nrn2575.

Burns, S.P., Sritharan, D., Jouny, C., Bergey, G., Crone, N., Anderson, W.S., and Sarma, S.V. (2012). A network analysis of the dynamics of seizure. *IEEE Engineering in Medicine and Biology Society Conference Proceedings* 2012**,** 4684-4687. doi:10.1109/EMBC.2012.6347012.

Buschman, T.J., and Miller, E.K. (2007). Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science* 315**,** 1860-1862. doi:10.1126/science.1138071.

Buzsaki, G., Anastassiou, C.A., and Koch, C. (2012). The origin of extracellular fields and currents--EEG, ECoG, LFP and spikes. *Nature Reviews Neuroscience* 13**,** 407-420. doi:10.1038/nrn3241.

Buzsaki, G., and Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science* 304**,** 1926-1929. doi:10.1126/science.1099745.

Calhoun, V.D., Kiehl, K.A., and Pearlson, G.D. (2008). Modulation of temporally coherent brain networks estimated using ICA at rest and during cognitive tasks. *Human Brain Mapping* 29**,** 828-838.

Canolty, R.T., Ganguly, K., Kennerley, S.W., Cadieu, C.F., Koepsell, K., Wallis, J.D., and Carmena, J.M. (2010). Oscillatory phase coupling coordinates anatomically dispersed functional cell assemblies. *Proceedings of the National Academy of Sciences of the United States of America* 107**,** 17356-17361. doi:10.1073/pnas.1008306107.

Casella, G., and Berger, R.L. (1990). *Statistical inference.* Duxbury Press Belmont, CA.

Chang, C., and Glover, G.H. (2010). Time-frequency dynamics of resting-state brain connectivity measured with fMRI. *Neuroimage* 50**,** 81-98. doi:10.1016/j.neuroimage.2009.12.011.

Chang, S.E., Horwitz, B., Ostuni, J., Reynolds, R., and Ludlow, C.L. (2011). Evidence of left inferior frontal-premotor structural and functional connectivity deficits in adults who stutter. *Cerebral Cortex* 21**,** 2507-2518. doi:10.1093/cercor/bhr028.

Cimenser, A., Purdon, P.L., Pierce, E.T., Walsh, J.L., Salazar-Gomez, A.F., Harrell, P.G., Tavares-Stoeckel, C., Habeeb, K., and Brown, E.N. (2011). Tracking brain states under general anesthesia by using global coherence analysis. *Proceedings of the National Academy of Sciences of the United States of America* 108**,** 8832-8837. doi:10.1073/pnas.1017041108.

Cohen, M.R., and Kohn, A. (2011). Measuring and interpreting neuronal correlations. *Nature Neuroscience* 14**,** 811-819. doi:10.1038/nn.2842.

Cordes, D., Haughton, V.M., Arfanakis, K., Carew, J.D., Turski, P.A., Moritz, C.H., Quigley, M.A., and Meyerand, M.E. (2001). Frequencies contributing to functional connectivity in the cerebral cortex in "resting-state" data. *American Journal of Neuroradiology* 22**,** 1326-1333.

Crone, N.E., Sinai, A., and Korzeniewska, A. (2006). High-frequency gamma oscillations and human brain mapping with electrocorticography. *Progress in brain research* 159**,** 275-295. doi:S0079-6123(06)59019-3

Deecke, L., Engel, M., Lang, W., and Kornhuber, H.H. (1986). Bereitschaftspotential preceding speech after holding breath. *Experimental Brain Research* 65**,** 219-223.

Dempster, A.P. (1966). "Estimation in Multivariate Analysis," in *Multivariate analysis,* ed. P.R. Krishnaiah. (New York: Academic Press), pp. 315-334.

Derambure, P., Bourriez, J., Defebvre, L., Cassim, F., Josien, E., Duhamel, A., Destee, A., and Guieu, J. (1997). Abnormal Cortical Activation During Planning of Voluntary Movement in Patients with Epilepsy with Focal Motor Seizures: Event ‐ Related Desynchronization Study of Electroencephalographic mu Rhythm. *Epilepsia* 38**,** 655-662.

Dudoit, S., and Laan, M.J.V.D. (2008). *Multiple testing procedures with applications to genomics.* New York: Springer.

Edwards, E., Nagarajan, S.S., Dalal, S.S., Canolty, R.T., Kirsch, H.E., Barbaro, N.M., and Knight, R.T. (2010). Spatiotemporal imaging of cortical activation during verb generation and picture naming. *Neuroimage* 50**,** 291-301. doi:10.1016/j.neuroimage.2009.12.035.

Efron, B. (2010). *Large-scale inference : empirical Bayes methods for estimation, testing, and prediction.* Cambridge ; New York: Cambridge University Press.

Efron, B., and Tibshirani, R. (1993). *An introduction to the bootstrap.* New York: Chapman & Hall.

Eguiluz, V.M., Chialvo, D.R., Cecchi, G.A., Baliki, M., and Apkarian, A.V. (2005). Scale-free brain functional networks. *Physical review letters* 94**,** 018102.

Eickhoff, S.B., Heim, S., Zilles, K., and Amunts, K. (2009). A systems perspective on the effective connectivity of overt speech production. *Philosophical transactions of the Royal Society of London Series A-Mathematical physical and engineering sciences* 367**,** 2399-2421. doi:10.1098/rsta.2008.0287.

Engel, A.K., and Fries, P. (2010). Beta-band oscillations—signalling the status quo? *Current Opinion in Neurobiology* 20**,** 156-165. doi:10.1016/j.conb.2010.02.015.

Engel, J., Kuhl, D.E., Phelps, M.E., and Mazziotta, J.C. (1982). Interictal cerebral glucose metabolism in partial epilepsy and its relation to EEG changes. *Annals of neurology* 12**,** 510-517.

Enochson, L.D., and Goodman, N.R. (1965). "Gaussian approximations to the distribution of sample coherence". DTIC Document).

Esposito, F., Bertolino, A., Scarabino, T., Latorre, V., Blasi, G., Popolizio, T., Tedeschi, G., Cirillo, S., Goebel, R., and Di Salle, F. (2006). Independent component model of the default-mode brain function: Assessing the impact of active thinking. *Brain research bulletin* 70**,** 263-269.

Ferrarini, L., Veer, I.M., Baerends, E., Van Tol, M.J., Renken, R.J., Van Der Wee, N.J., Veltman, D.J., Aleman, A., Zitman, F.G., Penninx, B.W., Van Buchem, M.A., Reiber, J.H., Rombouts, S.A., and Milles, J. (2009). Hierarchical functional modularity in the resting-state human brain. *Human Brain Mapping* 30**,** 2220-2231. doi:10.1002/hbm.20663.

Fingelkurts, A.A., Fingelkurts, A.A., and Kähkönen, S. (2005). Functional connectivity in the brain--is it an elusive concept? *Neuroscience and biobehavioral reviews* 28**,** 827-836. doi:10.1016/j.neubiorev.2004.10.009.

Fisher, R.A. (1915). Frequency Distribution of the Values of the Correlation Coefficient in Samples from an Indefinitely Large Population. *Biometrika* 10**,** 507-521. doi:10.2307/2331838.

Flinker, A., Chang, E.F., Barbaro, N.M., Berger, M.S., and Knight, R.T. (2011). Sub-centimeter language organization in the human temporal lobe. *Brain and Language* 117**,** 103-109. doi:10.1016/j.bandl.2010.09.009.

Flinker, A., Chang, E.F., Kirsch, H.E., Barbaro, N.M., Crone, N.E., and Knight, R.T. (2010). Single-trial speech suppression of auditory cortex activity in humans. *Journal of Neuroscience* 30**,** 16643-16650. doi:10.1523/JNEUROSCI.1809-10.2010.

Fox, M.D., Corbetta, M., Snyder, A.Z., Vincent, J.L., and Raichle, M.E. (2006). Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems. *Proceedings of the National Academy of Sciences* 103**,** 10046-10051.

Fox, M.D., and Raichle, M.E. (2007). Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nature Reviews Neuroscience* 8**,** 700-711.

Fransson, P. (2006). How default is the default mode of brain function?: Further evidence from intrinsic BOLD signal fluctuations. *Neuropsychologia* 44**,** 2836-2845.

Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends in cognitive sciences* 9**,** 474-480. doi:10.1016/j.tics.2005.08.011.

Fries, P. (2009). Neuronal Gamma-Band Synchronization as a Fundamental Process in Cortical Computation. *Annual Review of Neuroscience* 32**,** 209-224. doi:10.1146/annurev.neuro.051508.135603.

Fries, P., Nikolić, D., and Singer, W. (2007). The gamma cycle. *Trends in neurosciences* 30**,** 309-316. doi:10.1016/j.tins.2007.05.005.

Friston, K.J. (1994). Functional and effective connectivity in neuroimaging: A synthesis. *Human Brain Mapping* 2**,** 56-78. doi:10.1002/hbm.460020107.

Friston, K.J. (2011). Functional and effective connectivity: a review. *Brain Connectivity* 1**,** 13-36. doi:10.1089/brain.2011.0008.

Ginestet, C.E., Nichols, T.E., Bullmore, E.T., and Simmons, A. (2011). Brain Network Analysis: Separating Cost from Topology Using Cost-Integration. *PLoS ONE* 6**,** e21570. doi:10.1371/journal.pone.0021570.t001.

Golfinopoulos, E., Tourville, J.A., Bohland, J.W., Ghosh, S.S., Nieto-Castanon, A., and Guenther, F.H. (2011). fMRI investigation of unexpected somatosensory feedback perturbation during speech. *Neuroimage* 55**,** 1324-1338. doi:10.1016/j.neuroimage.2010.12.065.

Golfinopoulos, E., Tourville, J.A., and Guenther, F.H. (2010). The integration of large-scale neural network modeling and functional brain imaging in speech motor control. *Neuroimage* 52**,** 862-874. doi:10.1016/j.neuroimage.2009.10.023.

Greenblatt, R.E., Pflieger, M.E., and Ossadtchi, A.E. (2012). Connectivity measures applied to human brain electrophysiological data. *Journal of Neuroscience Methods* 207**,** 1-16. doi:10.1016/j.jneumeth.2012.02.025.

Greicius, M.D., and Menon, V. (2004). Default-mode activity during a passive sensory task: uncoupled from deactivation but impacting activation. *Journal of cognitive neuroscience* 16**,** 1484-1492.

Guenther, F.H., Ghosh, S.S., and Tourville, J.A. (2006). Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain and Language* 96**,** 280-301. doi:10.1016/j.bandl.2005.06.001.

Handwerker, D.A., Roopchansingh, V., Gonzalez-Castillo, J., and Bandettini, P.A. (2012). Periodic changes in fMRI connectivity. *Neuroimage* 63**,** 1712-1719. doi:10.1016/j.neuroimage.2012.06.078.

Hastie, T., Tibshirani, R., and Friedman, J.H. (2009). *The elements of statistical learning : data mining, inference, and prediction.* New York, NY: Springer.

He, B.J., Snyder, A.Z., Zempel, J.M., Smyth, M.D., and Raichle, M.E. (2008). Electrophysiological correlates of the brain's intrinsic large-scale functional architecture. *Proceedings of the National Academy of Sciences* 105**,** 16039-16044.

He, B.J., Zempel, J.M., Snyder, A.Z., and Raichle, M.E. (2010). The temporal structures and functional significance of scale-free brain activity. *Neuron* 66**,** 353-369. doi:10.1016/j.neuron.2010.04.020.

Hermes, D., Miller, K.J., Noordmans, H.J., Vansteensel, M.J., and Ramsey, N.F. (2010). Automated electrocorticographic electrode localization on individually rendered brain surfaces. *Journal of neuroscience methods* 185**,** 293-298.

Hickok, G., and Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience* 8**,** 393-402. doi:10.1038/nrn2113.

Hill, B.M. (1975). A Simple General Approach to Inference About the Tail of a Distribution. *The Annals of Statistics* 3**,** 1163-1174. doi:10.1214/aos/1176343247.

Hiltunen, T., Kantola, J., Elseoud, A.A., Lepola, P., Suominen, K., Starck, T., Nikkinen, J., Remes, J., Tervonen, O., and Palva, S. (2014). Infra-slow EEG fluctuations are correlated with resting-state network dynamics in fMRI. *Journal of neuroscience* 34**,** 356-362.

Hinkley, D.V. (1978). Improving the jackknife with special reference to correlation estimation. *Biometrika* 65**,** 13-21.

Hughes, S.W. (2005). Thalamic Mechanisms of EEG Alpha Rhythms and Their Pathological Implications. *The Neuroscientist* 11**,** 357-372. doi:10.1177/1073858405277450.

Hutchison, R.M., Womelsdorf, T., Allen, E.A., Bandettini, P.A., Calhoun, V.D., Corbetta, M., Della Penna, S., Duyn, J.H., Glover, G.H., Gonzalez-Castillo, J., Handwerker, D.A., Keilholz, S., Kiviniemi, V., Leopold, D.A., De Pasquale, F., Sporns, O., Walter, M., and Chang, C. (2013). Dynamic functional connectivity: promise, issues, and interpretations. *Neuroimage* 80**,** 360-378. doi:10.1016/j.neuroimage.2013.05.079.

James, A.T. (1964). Distributions of Matrix Variates and Latent Roots Derived from Normal Samples. *The Annals of Mathematical Statistics* 35**,** 475-501. doi:10.2307/2238504.

Jewett, D.L., and Williston, J.S. (1971). Auditory-evoked far fields averaged from the scalp of humans. *Brain* 94**,** 681-696.

Jobard, G., Crivello, F., and Tzourio-Mazoyer, N. (2003). Evaluation of the dual route theory of reading: a metanalysis of 35 neuroimaging studies. *Neuroimage* 20**,** 693-712.

Jones, S.R., Pritchett, D.L., Sikora, M.A., Stufflebeam, S.M., Hamalainen, M., and Moore, C.I. (2009). Quantitative analysis and biophysically realistic neural modeling of the MEG mu rhythm: rhythmogenesis and modulation of sensory-evoked responses. *Journal of Neurophysiology* 102**,** 3554-3572. doi:10.1152/jn.00535.2009.

Kilpatrick, L.A., Zald, D.H., Pardo, J.V., and Cahill, L.F. (2006). Sex-related differences in amygdala functional connectivity during resting conditions. *Neuroimage* 30**,** 452-461. doi:10.1016/j.neuroimage.2005.09.065.

Ko, A.L., Darvas, F., Poliakov, A., Ojemann, J., and Sorensen, L.B. (2011). Quasi-periodic fluctuations in default mode network electrophysiology. *Journal of neuroscience* 31**,** 11728-11732.

Kolaczyk, E.D. (2009). *Statistical analysis of network data : methods and models.* New York: Springer.

Koopmans, L.H. (1995). *The spectral analysis of time series.* San Diego: Academic Press.

Kopell, N., Kramer, M.A., Malerba, P., and Whittington, M.A. (2010). Are Different Rhythms Good for Different Functions? *Frontiers in Human Neuroscience* 4. doi:10.3389/fnhum.2010.00187.

Kornhuber, H., and Deecke, L. (1964). Hirnpotentialanderungen beim Menschen vor und nach Willkurbewegungen, dargestellt mit Magnetband-Speicherung und Ruckwartsanalyse. *Pflugers Archiv fur die gesamte physiologie des menschen und der tiere* 281**,** 52.

Korzeniewska, A., Franaszczuk, P.J., Crainiceanu, C.M., Kuś, R., and Crone, N.E. (2011). Dynamics of large-scale cortical interactions at high gamma frequencies during word production: event related causality (ERC) analysis of human electrocorticography (ECoG). *NeuroImage* 56**,** 2218-2237. doi:10.1016/j.neuroimage.2011.03.030.

Kramer, M.A., and Cash, S.S. (2012). Epilepsy as a disorder of cortical network organization. *Neuroscientist* 18**,** 360-372. doi:10.1177/1073858411422754.

Kramer, M.A., Eden, U.T., Cash, S.S., and Kolaczyk, E.D. (2009). Network inference with confidence from multivariate time series. *Physical Review E* 79**,** 061916. doi:10.1103/PhysRevE.79.061916.

Kramer, M.A., Eden, U.T., Kolaczyk, E.D., Zepeda, R., Eskandar, E.N., and Cash, S.S. (2010). Coalescence and fragmentation of cortical networks during focal seizures. *Journal of neuroscience* 30**,** 10076-10085.

Kramer, M.A., Eden, U.T., Lepage, K.Q., Kolaczyk, E.D., Bianchi, M.T., and Cash, S.S. (2011). Emergence of persistent networks in long-term intracranial EEG recordings. *Journal of neuroscience* 31**,** 15757-15767.

Kramer, M.A., Kolaczyk, E.D., and Kirsch, H.E. (2008). Emergent network topology at seizure onset in humans. *Epilepsy Research* 79**,** 173-186. doi:10.1016/j.eplepsyres.2008.02.002.

Kubanek, J., Brunner, P., Gunduz, A., Poeppel, D., and Schalk, G. (2013). The tracking of speech envelope in the human cortex. *PLoS One* 8**,** e53398. doi:10.1371/journal.pone.0053398.

Kuhlman, W.N. (1978). Functional topography of the human mu rhythm. *Electroencephalography and clinical neurophysiology* 44**,** 83-93.

Lawley, D.N. (1959). Tests of Significance in Canonical Analysis. *Biometrika* 46**,** 59-66. doi:10.2307/2332808.

Leonardi, N., and Van De Ville, D. (2015). On spurious and real fluctuations of dynamic functional connectivity during rest. *Neuroimage* 104**,** 430-436. doi:10.1016/j.neuroimage.2014.09.007.

Leuthardt, E.C., Pei, X.M., Breshears, J., Gaona, C., Sharma, M., Freudenberg, Z., Barbour, D., and Schalk, G. (2012). Temporal evolution of gamma activity in human cortex during an overt and covert word repetition task. *Frontiers in Human Neuroscience* 6**,** 99. doi:10.3389/fnhum.2012.00099.

Lindén, H., Tetzlaff, T., Potjans, T.C., Pettersen, K.H., Grün, S., Diesmann, M., and Einevoll, G.T. (2011). Modeling the spatial reach of the LFP. *Neuron* 72**,** 859-872.

Lisker, L., and Abramson, A.S. (1967). Some effects of context on voice onset time in English stops. *Language and Speech* 10**,** 1-28.

Lotte, F., Brumberg, J.S., Brunner, P., Gunduz, A., Ritaccio, A.L., Guan, C., and Schalk, G. (2015). Electrocorticographic representations of segmental features in continuous speech. *Frontiers in Human Neuroscience* 9**,** 97. doi:10.3389/fnhum.2015.00097.

Magri, C., Mazzoni, A., Logothetis, N.K., and Panzeri, S. (2012). Optimal band separation of extracellular field potentials. *Journal of Neuroscience Methods* 210**,** 66-78. doi:10.1016/j.jneumeth.2011.11.005.

Manning, J.R., Jacobs, J., Fried, I., and Kahana, M.J. (2009). Broadband Shifts in Local Field Potential Power Spectra Are Correlated with Single-Neuron Spiking in Humans. *Journal of Neuroscience* 29**,** 13613-13620. doi:10.1523/JNEUROSCI.2041-09.2009.

Mardia, K.V., Kent, J.T., and Bibby, J.M. (1979). *Multivariate analysis.* New York: Academic Press.

Maris, E., and Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG-data. *Journal of neuroscience methods* 164**,** 177-190.

Marsden, J., Werhahn, K., Ashby, P., Rothwell, J., Noachtar, S., and Brown, P. (2000). Organization of cortical activities related to movement in humans. *Journal of neuroscience* 20**,** 2307-2314.

Matsumoto, R. (2004). Functional connectivity in the human language system: a cortico-cortical evoked potential study. *Brain* 127**,** 2316-2330. doi:10.1093/brain/awh246.

Mcfarland, D.J., Miner, L.A., Vaughan, T.M., and Wolpaw, J.R. (2000). Mu and beta rhythm topographies during motor imagery and actual movements. *Brain topography* 12**,** 177-186.

Meunier, D., Achard, S., Morcom, A., and Bullmore, E. (2009). Age-related changes in modular organization of human brain functional networks. *Neuroimage* 44**,** 715-723. doi:10.1016/j.neuroimage.2008.09.062.

Micheloyannis, S., Pachou, E., Stam, C.J., Vourkas, M., Erimaki, S., and Tsirka, V. (2006). Using graph theoretical analysis of multi channel EEG to evaluate the neural efficiency hypothesis. *Neuroscience Letters* 402**,** 273-277. doi:10.1016/j.neulet.2006.04.006.

Miller, K.J., Foster, B.L., and Honey, C.J. (2012). Does rhythmic entrainment represent a generalized mechanism for organizing computation in the brain? *Frontiers in computational neuroscience* 6**,** 1-3. doi:10.3389/fncom.2012.00085.

Miller, K.J., Honey, C.J., Hermes, D., Rao, R.P., Dennijs, M., and Ojemann, J.G. (2014). Broadband changes in the cortical surface potential track activation of functionally diverse neuronal populations. *Neuroimage* 85 Pt 2**,** 711-720. doi:10.1016/j.neuroimage.2013.08.070.

Miller, K.J., Leuthardt, E.C., Schalk, G., Rao, R.P., Anderson, N.R., Moran, D.W., Miller, J.W., and Ojemann, J.G. (2007). Spectral changes in cortical surface potentials during motor movement. *Journal of Neuroscience* 27**,** 2424-2432. doi:10.1523/JNEUROSCI.3886-06.2007.

Miller, K.J., Sorensen, L.B., Ojemann, J.G., and Den Nijs, M. (2009a). Power-law scaling in the brain surface electric potential. *PLoS Computational Biology* 5**,** e1000609. doi:10.1371/journal.pcbi.1000609.

Miller, K.J., Zanos, S., Fetz, E.E., Den Nijs, M., and Ojemann, J.G. (2009b). Decoupling the cortical power spectrum reveals real-time representation of individual finger movements in humans. *Journal of Neuroscience* 29**,** 3132-3137. doi:10.1523/JNEUROSCI.5506-08.2009.

Mitra, P., and Bokil, H. (2008). *Observed brain dynamics.* Oxford ; New York: Oxford University Press.

Murias, M., Swanson, J.M., and Srinivasan, R. (2007). Functional connectivity of frontal cortex in healthy and ADHD children reflected in EEG coherence. *Cerebral Cortex* 17**,** 1788-1799. doi:10.1093/cercor/bhl089.

Nieto-Castanon, A., Ghosh, S.S., Tourville, J.A., and Guenther, F.H. (2003). Region of interest based analysis of functional imaging data. *Neuroimage* 19**,** 1303-1316. doi:10.1016/S1053-8119(03)00188-5.

Ohara, S., Mima, T., Baba, K., Ikeda, A., Kunieda, T., Matsumoto, R., Yamamoto, J., Matsuhashi, M., Nagamine, T., and Hirasawa, K. (2001). Increased synchronization of cortical oscillatory activities between human supplementary motor and primary sensorimotor areas during voluntary movements. *Journal of neuroscience* 21**,** 9377-9386.

Ojemann, G.A., Ojemann, J., and Ramsey, N.F. (2013). Relation between functional magnetic resonance imaging (fMRI) and single neuron, local field potential (LFP) and electrocorticography (ECoG) activity in human cortex. *Frontiers in Human Neuroscience* 7**,** 34. doi:10.3389/fnhum.2013.00034.

Pachou, E., Vourkas, M., Simos, P., Smit, D., Stam, C.J., Tsirka, V., and Micheloyannis, S. (2008). Working memory in schizophrenia: an EEG study using power spectrum and coherence analysis to estimate cortical activation and network behavior. *Brain topography* 21**,** 128-137.

Peeva, M.G., Guenther, F.H., Tourville, J.A., Nieto-Castanon, A., Anton, J.L., Nazarian, B., and Alario, F.X. (2010). Distinct representations of phonemes, syllables, and supra-syllabic sequences in the speech production network. *Neuroimage* 50**,** 626-638. doi:10.1016/j.neuroimage.2009.12.065.

Pei, X., Leuthardt, E.C., Gaona, C.M., Brunner, P., Wolpaw, J.R., and Schalk, G. (2011). Spatiotemporal dynamics of electrocorticographic high gamma activity during overt and covert word repetition. *NeuroImage* 54**,** 2960-2972. doi:10.1016/j.neuroimage.2010.10.029.

Pereda, E., Quiroga, R.Q., and Bhattacharya, J. (2005). Nonlinear multivariate analysis of neurophysiological signals. *Progress in Neurobiology* 77**,** 1-37. doi:10.1016/j.pneurobio.2005.10.003.

Pfurtscheller, G., and Lopes Da Silva, F.H. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical neurophysiology* 110**,** 1842-1857. doi:10.1016/S1388-2457(99)00141-8.

Pickands, J., Iii (1975). Statistical Inference Using Extreme Order Statistics. *The Annals of Statistics* 3**,** 119-131. doi:10.1214/aos/1176343003.

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical neurophysiology* 118**,** 2128-2148.

Ponten, S.C., Bartolomei, F., and Stam, C.J. (2007). Small-world networks and epilepsy: graph theoretical analysis of intracerebrally recorded mesial temporal lobe seizures. *Clinical neurophysiology* 118**,** 918-927. doi:10.1016/j.clinph.2006.12.002.

Price, C.J. (2012). A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *NeuroImage* 62**,** 816-847. doi:10.1016/j.neuroimage.2012.04.062.

Quenouille, M.H. (1956). Notes on bias in estimation. *Biometrika***,** 353-360.

Raichle, M.E., Macleod, A.M., Snyder, A.Z., Powers, W.J., Gusnard, D.A., and Shulman, G.L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences* 98**,** 676-682.

Ray, S., and Maunsell, J.H. (2011). Different origins of gamma rhythm and high-gamma activity in macaque visual cortex. *PLoS Biology* 9**,** e1000610. doi:10.1371/journal.pbio.1000610.

Ritaccio, A., Boatman-Reich, D., Brunner, P., Cervenka, M.C., Cole, A.J., Crone, N., Duckrow, R., Korzeniewska, A., Litt, B., Miller, K.J., Moran, D.W., Parvizi, J., Viventi, J., Williams, J., and Schalk, G. (2011). Proceedings of the Second International Workshop on Advances in Electrocorticography. *Epilepsy & Behavior* 22**,** 641-650. doi:10.1016/j.yebeh.2011.09.028.

Rubino, D., Robbins, K.A., and Hatsopoulos, N.G. (2006). Propagating waves mediate information transfer in the motor cortex. *Nature Neuroscience* 9**,** 1549-1557. doi:10.1038/nn1802.

Rubinov, M., and Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. *Neuroimage* 52**,** 1059-1069. doi:10.1016/j.neuroimage.2009.10.003.

Sakoglu, U., Pearlson, G.D., Kiehl, K.A., Wang, Y.M., Michael, A.M., and Calhoun, V.D. (2010). A method for evaluating dynamic functional network connectivity and task-modulation: application to schizophrenia. *Magnetic Resonance Materials in Physics, Biology, and Medicine* 23**,** 351-366. doi:10.1007/s10334-010-0197-8.

Salvador, R., Suckling, J., Coleman, M.R., Pickard, J.D., Menon, D., and Bullmore, E. (2005). Neurophysiological architecture of functional magnetic resonance images of human brain. *Cerebral Cortex* 15**,** 1332-1342. doi:10.1093/cercor/bhi016.

Schalk, G., Mcfarland, D.J., Hinterberger, T., Birbaumer, N., and Wolpaw, J.R. (2004). BCI2000: A General-Purpose Brain-Computer Interface (BCI) System. *IEEE transactions on bio-medical engineering* 51**,** 1034-1043. doi:10.1109/TBME.2004.827072.

Scheffer-Teixeira, R., Belchior, H., Leao, R.N., Ribeiro, S., and Tort, A.B. (2013). On high-frequency field oscillations (>100 Hz) and the spectral leakage of spiking activity. *Journal of Neuroscience* 33**,** 1535-1539. doi:10.1523/JNEUROSCI.4217-12.2013.

Schindler, K., Leung, H., Elger, C.E., and Lehnertz, K. (2007). Assessing seizure dynamics by analysing the correlation structure of multichannel intracranial EEG. *Brain* 130**,** 65-77. doi:10.1093/brain/awl304.

Schnitzler, A., and Gross, J. (2005). Normal and pathological oscillatory communication in the brain. *Nature Reviews Neuroscience* 6**,** 285-296. doi:10.1038/nrn1650.

Schreier, P.J., and Scharf, L.L. (2010). *Statistical signal processing of complex-valued data : the theory of improper and noncircular signals.* Cambridge: Cambridge University Press.

Sehatpour, P., Molholm, S., Schwartz, T.H., Mahoney, J.R., Mehta, A.D., Javitt, D.C., Stanton, P.K., and Foxe, J.J. (2008). A human intracranial study of long-range oscillatory coherence across a frontal–occipital–hippocampal brain network during visual object processing. *Proceedings of the National Academy of Sciences* 105**,** 4399-4404.

Shah, A.S., Bressler, S.L., Knuth, K.H., Ding, M., Mehta, A.D., Ulbert, I., and Schroeder, C.E. (2004). Neural dynamics and the fundamental mechanisms of event-related brain potentials. *Cerebral Cortex* 14**,** 476-483. doi:10.1093/cercor/bhh009.

Share, D.L. (2008). On the Anglocentricities of current reading research and practice: the perils of overreliance on an" outlier" orthography. *Psychological bulletin* 134**,** 584.

Shehzad, Z., Kelly, A.M., Reiss, P.T., Gee, D.G., Gotimer, K., Uddin, L.Q., Lee, S.H., Margulies, D.S., Roy, A.K., Biswal, B.B., Petkova, E., Castellanos, F.X., and Milham, M.P. (2009). The resting brain: unconstrained yet reliable. *Cerebral Cortex* 19**,** 2209-2229. doi:10.1093/cercor/bhn256.

Shibasaki, H., and Hallett, M. (2006). What is the Bereitschaftspotential? *Clinical neurophysiology* 117**,** 2341-2356.

Siegel, M., Donner, T.H., and Engel, A.K. (2012). Spectral fingerprints of large-scale neuronal interactions. *Nature Reviews Neuroscience* 13**,** 121-134.

Simonyan, K., Ostuni, J., Ludlow, C.L., and Horwitz, B. (2009). Functional but not structural networks of the human laryngeal motor cortex show left hemispheric lateralization during syllable but not breathing production. *Journal of Neuroscience* 29**,** 14912-14923. doi:10.1523/JNEUROSCI.4897-09.2009.

Smith, R.L. (1987). Estimating Tails of Probability Distributions. *The Annals of Statistics* 15**,** 1174-1207. doi:10.1214/aos/1176350499.

Smith, S.M., Fox, P.T., Miller, K.L., Glahn, D.C., Fox, P.M., Mackay, C.E., Filippini, N., Watkins, K.E., Toro, R., and Laird, A.R. (2009). Correspondence of the brain's functional architecture during activation and rest. *Proceedings of the National Academy of Sciences* 106**,** 13040-13045.

Sporns, O., Chialvo, D.R., Kaiser, M., and Hilgetag, C.C. (2004). Organization, development and function of complex brain networks. *Trends in cognitive sciences* 8**,** 418-425. doi:10.1016/j.tics.2004.07.008.

Srinivas, K.V., Jain, R., Saurav, S., and Sikdar, S.K. (2007). Small-world network topology of hippocampal neuronal network is lost, in an in vitro glutamate injury model of epilepsy. *European Journal of Neuroscience* 25**,** 3276-3286. doi:10.1111/j.1460-9568.2007.05559.x.

Stam, C.J. (2004). Functional connectivity patterns of human magnetoencephalographic recordings: a 'small-world' network? *Neuroscience Letters* 355**,** 25-28. doi:10.1016/j.neulet.2003.10.063.

Stam, C.J., Jones, B.F., Nolte, G., Breakspear, M., and Scheltens, P. (2007). Small-world networks and functional connectivity in Alzheimer's disease. *Cerebral Cortex* 17**,** 92-99. doi:10.1093/cercor/bhj127.

Stam, C.J., and Van Dijk, B.W. (2002). Synchronization likelihood: an unbiased measure of generalized synchronization in multivariate data sets. *Physica D-Nonlinear Phenomena* 163**,** 236-251. doi:10.1016/S0167-2789(01)00386-4.

Stephan, K.E., Penny, W.D., Moran, R.J., Den Ouden, H.E., Daunizeau, J., and Friston, K.J. (2010). Ten simple rules for dynamic causal modeling. *Neuroimage* 49**,** 3099-3109.

Stephen, E.P., Lepage, K.Q., Eden, U.T., Brunner, P., Schalk, G., Brumberg, J.S., Guenther, F.H., and Kramer, M.A. (2014). Assessing dynamics, spatial scale, and uncertainty in task-related brain network analyses. *Frontiers in computational neuroscience* 8. doi:10.3389/fncom.2014.00031.

Strang, G. (2003). *Introduction to Linear Algebra.* Wellesley: Wellesley-Cambridge Press.

Supekar, K., Menon, V., Rubin, D., Musen, M., and Greicius, M.D. (2008). Network analysis of intrinsic functional brain connectivity in Alzheimer's disease. *PLoS Computational Biology* 4**,** e1000100. doi:10.1371/journal.pcbi.1000100.

Szurhaj, W., Derambure, P., Labyt, E., Cassim, F., Bourriez, J.L., Isnard, J., Guieu, J.D., and Mauguiere, F. (2003). Basic mechanisms of central rhythms reactivity to preparation and execution of a voluntary movement: a stereoelectroencephalographic study. *Clinical neurophysiology* 114**,** 107-119. doi:10.1016/s1388-2457(02)00333-4.

Thompson, B. (1990). Finding a Correction for the Sampling Error in Multivariate Measures of Relationship: A Monte Carlo Study. *Educational and Psychological Measurement* 50**,** 15-31. doi:10.1177/0013164490501003.

Thomson, D., and Chave, A. (1991). "Jackknifed error estimates for spectra, coherences, and transfer functions. Advances in spectrum analysis and array processing," in *Advances in spectrum analysis and array processing,* ed. S. Haykin. (Englewood Cliffs, New Jersey: Prentice Hall), 58-113.

Thomson, D.J. (1982). Spectrum estimation and harmonic analysis. *Proceedings of the IEEE* 70**,** 1055-1096. doi:10.1109/proc.1982.12433.

Thomson, D.J., Lanzerotti, L.J., Vernon, F.L., Lessard, M.R., and Smith, L.T.P. (2007). Solar modal structure of the engineering environment. *Proceedings of the Ieee* 95**,** 1085-1132. doi:Doi 10.1109/Jproc.2007.894712.

Toro, C., Deuschl, G., Thatcher, R., Sato, S., Kufta, C., and Hallett, M. (1994). Event-related desynchronization and movement-related cortical potentials on the ECoG and EEG. *Electroencephalography and Clinical Neurophysiology* 93**,** 380-389.

Tourville, J.A., Reilly, K.J., and Guenther, F.H. (2008). Neural mechanisms underlying auditory feedback control of speech. *Neuroimage* 39**,** 1429-1443. doi:10.1016/j.neuroimage.2007.09.054.

Towle, V.L., Yoon, H.A., Castelle, M., Edgar, J.C., Biassou, N.M., Frim, D.M., Spire, J.P., and Kohrman, M.H. (2008). ECoG gamma activity during a language task: differentiating expressive and receptive speech areas. *Brain* 131**,** 2013-2027. doi:10.1093/brain/awn147.

Truccolo, W.A., Ding, M., Knuth, K.H., Nakamura, R., and Bressler, S.L. (2002). Trial-to-trial variability of cortical evoked responses: implications for the analysis of functional connectivity. *Clinical neurophysiology* 113**,** 206-226. doi:10.1016/s1388-2457(01)00739-8.

Tukey, J.W. (1958). "Bias and confidence in not-quite large samples", in: *The Annals of Mathematical Statistics*: INST MATHEMATICAL STATISTICS IMS BUSINESS OFFICE-SUITE 7, 3401 INVESTMENT BLVD, HAYWARD, CA 94545), 614-614.

Van Den Heuvel, M.P., and Pol, H.E.H. (2010). Exploring the brain network: a review on resting-state fMRI functional connectivity. *European Neuropsychopharmacology* 20**,** 519-534.

Wang, X.J. (2010). Neurophysiological and computational principles of cortical rhythms in cognition. *Physiological Reviews* 90**,** 1195-1268. doi:10.1152/physrev.00035.2008.

Weiss, S., and Mueller, H.M. (2012). "Too Many betas do not Spoil the Broth": The Role of Beta Brain Oscillations in Language Processing. *Frontiers in Psychology* 3. doi:10.3389/fpsyg.2012.00201.

Wohlert, A.B. (1993). Event-related brain potentials preceding speech and nonspeech oral movements of varying complexity. *Journal of Speech and Hearing Research* 36**,** 897-905.

Womelsdorf, T., Schoffelen, J.M., Oostenveld, R., Singer, W., Desimone, R., Engel, A.K., and Fries, P. (2007). Modulation of Neuronal Interactions Through Neuronal Synchronization. *Science* 316**,** 1609-1612. doi:10.1126/science.1139597.

Wood, C.C., Spencer, D.D., Allison, T., Mccarthy, G., Williamson, P.D., and Goff, W.R. (1988). Localization of human sensorimotor cortex during surgery by cortical surface recording of somatosensory evoked potentials. *Journal of Neurosurgery* 68**,** 99-111. doi:10.3171/jns.1988.68.1.0099.

Yeo, B.T., Krienen, F.M., Sepulcre, J., Sabuncu, M.R., Lashkari, D., Hollinshead, M., Roffman, J.L., Smoller, J.W., Zollei, L., Polimeni, J.R., Fischl, B., Liu, H., and Buckner, R.L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology* 106**,** 1125-1165. doi:10.1152/jn.00338.2011.

Zaveri, H.P., Duckrow, R.B., and Spencer, S.S. (2009). Concerning the observation of an electrical potential at a distance from an intracranial electrode contact. *Clinical neurophysiology* 120**,** 1873-1875. doi:10.1016/j.clinph.2009.08.001.

CURRICULUM VITAE