

Modeling the fast-timescale network dynamics that underlie complex thought

A Thesis
Submitted to the Faculty
in partial fulfillment of the requirements for the
degree of
Doctor of Philosophy
in
Psychological and Brain Sciences
by Lucy L. W. Owen

Guarini School of Graduate and Advanced Studies
Dartmouth College
Hanover, New Hampshire
October, 2021

Examining Committee:

Jeremy R. Manning, Ph.D., Chair

Richard F. Betzel, Ph.D.

Emily S. Finn, Ph.D.

James V. Haxby, Ph.D.

F. Jon Kull, Ph.D.
Dean of the Guarini School of Graduate and Advanced Studies

ABSTRACT

The human brain is a complex system, organized on multiple scales of space and time, and functions in both a local and distributed manner. Modeling something as inherently distributed as complex thought holds the potential to expand our understanding of this multi-tiered organization in the brain as well as the properties of brain activity at each of those levels.

However, one of the many limitations in our pursuit in this type of modeling is the resolution of our data. Here, I first describe a model that attempts to resolve the discrepancy of high temporal and high spatial resolutions, by leveraging the redundancy of information across people, across tasks, and in an individual’s brain. We use a multi-subject brain model to capture millimeter-scale spatial resolutions and millisecond-scale temporal resolutions throughout the entire brain of each patient. I then describe how we test and validate this method, and also show that our approach generalizes across people and tasks.

Additionally, in order to better understand the dynamic interactions between the brain structures that underlie our thoughts, the tools and data we use must reflect the complexity of this system. I next describe the development and application of a model that explores dynamic higher-order network interactions in brain data collected at varying levels of engagement. Our method provides a scalable model that estimates high-order dynamic correlations in timeseries data. I explore these interactions that support complex cognition by using brain data collected with a naturalistic paradigm and investigating different levels of engagement.

Overall, this work focuses on modeling brain activity and understanding how our brains support cognition. It has involved understanding the network dynamics of the brain and has centrally focused on modeling fast-timescale whole-brain networks with the goal of better understanding these complex systems.

ACKNOWLEDGMENTS

Most importantly, I want to thank my advisor, Jeremy Manning. I honestly doubt there has ever been a more supportive and understanding advisor. Jeremy, thank you for your never-ending insights, your enthusiasm for science, and your always perfectly chosen gif/meme/party-parrot on slack.

I would also like to thank former and current lab members. I particularly want to thank Andrew Heusser for his guidance. Even though we only overlapped briefly, his help and friendship meant so much to me and set me up for success in my PhD.

I also want to thank my friends, collaborators, and colleagues at Dartmouth. For many reasons, it was not the easiest 5 years, but this community is resilient and I'm grateful for the support.

And lastly, I will forever be grateful to my family for their bottomless support, encouragement, and love. I truly would be nothing without you.

TABLE OF CONTENTS

ABSTRACT	ii
ACKNOWLEDGMENTS	iii
LIST OF FIGURES	vi
LIST OF APPENDICES	viii
CHAPTER	
1 Introduction	1
1.1 What do we need to model the brain?	1
1.1.1 Data resolution	2
1.1.2 Scalability	3
1.1.3 Naturalistic tasks	4
1.2 What else can we learn about the brain?	6
1.3 Organization of this thesis	6
2 A Gaussian process model of human electrocorticographic data	8
2.1 Abstract	8
2.2 Introduction	8
2.3 Approach	10
2.4 Results	16
2.5 Discussion	27
2.6 Conclusion	30
3 High-level cognition during story listening is reflected in high-order dynamic correlations in neural activity patterns	31
3.1 Abstract	31
3.2 Introduction	32
3.3 Approach	35
3.4 Results	50
3.5 Discussion	59
4 Discussion	63

4.0.1	Summary	63
4.0.2	High-resolution spatiotemporal recordings	63
4.0.3	Higher-order interactions	64
4.0.4	Future directions	65
4.0.5	Concluding remarks	67
	APPENDICES	68
	BIBLIOGRAPHY	87

LIST OF FIGURES

FIGURE

2.1	Methods overview	12
2.2	Reconstruction accuracy across all electrodes in two ECoG datasets	19
2.3	Electrode sampling density by location	20
2.4	Reconstruction accuracy across all electrodes in two ECoG datasets for each frequency band	22
2.5	Most informative recording locations	24
2.6	Most informative recording locations by frequency band	26
3.1	Neural patterns	33
3.2	Estimating dynamic high-order correlations	36
3.3	Examples of kernel functions	37
3.4	Decoding analysis pipeline	48
3.5	Recovering known dynamic first-order correlations from synthetic data	51
3.6	Recovery of simulated first-order and second-order dynamic correlations	53
3.7	Across-participant timepoint decoding accuracy varies with correlation order and cognitive engagement	55
3.8	Statistical summary of decoding accuracies for different neural features	57
3.9	Top terms associated with the most strongly correlated nodes at each order	58
3.10	Proposed high-order network dynamics underlying high-level cognition during story listening	60
A1	Reconstruction accuracy across all electrodes in two ECoG datasets, broken down by experiment	69
A2	Reconstruction accuracy for Dataset 2, Experiments 1 and 2	70
A3	Reconstruction accuracy across all electrodes in Dataset 2	71
A4	Reconstruction accuracy across all electrodes in two ECoG datasets for power in each frequency band	72
A5	Most informative recording locations by frequency band	73
A6	Reconstruction accuracy versus within-subject data features for two ECoG datasets	74
A7	Temporal stability of estimated patient-specific correlation matrices	75
A8	Stability of full-brain correlation matrices across patients	76
B1	Across-participant timepoint decoding accuracy varies with correlation order and cognitive engagement across kernels	78

B2	Across-participant timepoint decoding accuracy varies with correlation order and cognitive engagement across widths	79
B3	Recovery heatmap of simulated first-order and second-order dynamic correlations across time and widths	80
B4	Recovery heatmap of simulated first-order and second-order dynamic correlations across features and widths	81
B5	Top terms associated with the most strongly correlated nodes at each order, for the <i>intact</i> experimental condition	82
B6	Top terms associated with the most strongly correlated nodes at each order, for the <i>paragraph</i> experimental condition	83
B7	Top terms associated with the most strongly correlated nodes at each order, for the <i>word</i> experimental condition	84
B8	Top terms associated with the most strongly correlated nodes at each order, for the <i>rest</i> experimental condition	85
B9	Cumulative percent variance explained as a function of the number of principle components	86

LIST OF APPENDICES

A Supplementary information for: A Gaussian process model of human electrocorticographic data	68
B Supplementary information for: High-level cognition during story listening is reflected in high-order dynamic correlations in neural activity patterns	77

CHAPTER 1

Introduction

Our brains are incredibly complex systems and are organized on multiple scales of space and time. Information can be carried in any number of ways, from individual neurons to brain regions, which can function in both local and distributed manners. Additionally, different subsystems can operate on their own specific functions while simultaneously participating in larger networks and supporting a variety of functions. With all this complexity, how can we ever begin to understand the brain?

Learning about this intriguingly and maddeningly complicated structure, fundamental to our existence, is the motivation behind this work. The goal of this thesis is to explore how our brains support complex cognition by modeling patterns in brain activity. I have worked on the development of novel methods and tools which attempt to target this complexity while attempting to maintain the restraint needed to generalize. I have also worked with brain data that specifically attempts to harness these naturally complex patterns in thought. The following thesis contains my research focusing on modeling fast-timescale whole-brain networks that underlie complex thought, with the goal of better understanding how our brains support cognition.

1.1 What do we need to model the brain?

Complex thought requires integration of information and coordination throughout the brain. These interactions between brain regions, or networks, happen at fast timescales. To better understand complex cognition, we wish to formalize these fast and dynamic interactions, mathematically.

However, there are many facets to consider in our pursuit to model these interactions. One major limitation is the resolution of our data, and the discrepancy that currently exists between high temporal and high spatial resolutions. Another is modeling constraints

and computational power needed to capture these network interactions. And finally, using suitably relevant brain data collected from naturalistic tasks.

1.1.1 Data resolution

In other model systems, such as zebrafish (Randlett et al., 2015; Betzel, 2020) and *C. elegans* (Nguyen et al., 2016; Betzel and Vértes, 2021), high spatiotemporal resolution recordings have transformed our understanding of neural systems. While techniques exist to measure spatial or temporal properties of these networks in humans, there remains a fundamental trade-off in non-invasive human neuroimaging pertaining to spatial and temporal resolution (Sejnowski et al., 2014). For example, recording techniques like scalp electroencephalography (EEG), and magnetoencephalography (MEG) produce high temporal resolution brain recordings, but at poor spatial resolutions. Conversely, functional magnetic resonance imaging (fMRI) produce high spatial resolution brain recordings, but at poor temporal resolutions.

1.1.1.1 Non-invasive attempts to resolve this discrepancy

One increasingly common approach has been to improve the spatial resolution of MEG or scalp EEG data through beamforming (Sarvas, 1987; Snyder, 1991; Baillet et al., 2001; Hillebrand et al., 2005). Beamforming uses the sensor locations as well as detailed brain conductance models to localize signals from deep in the brain. However, one serious challenge is that the magnetic fields produced by external sources are substantially stronger than those produced by brain signals originating from these deep brain sources. Thus obtaining adequate signal quality often requires averaging the measured responses over tens to hundreds of responses or trials (e.g., see review by Hillebrand et al., 2005), which significantly impacts experimental design and the use of naturalistic stimuli.

A different approach to obtaining high spatiotemporal resolution neural data has been to collect fMRI and EEG data together. However, the signal quality of both recordings suffers during simultaneous fMRI-EEG (e.g., see review by Huster et al., 2012). Also, like beamforming, experimental design is an important consideration. Where fMRI experiments often attempt to maximize the number of post-stimulus samples by locking the events to the regularly spaced image acquisition times (TRs), EEG experiments typically use jittered stimulus presentation times to maximize the difference between brain signal to external noise.

1.1.1.2 Electrocorticography

Unfortunately, the “gold standard” for sampling at high temporal resolutions from precise locations is invasive. These recordings are taken either from electrodes implanted directly onto the cortical surface (electrocorticography; ECoG) or into deep brain structures (intracranial EEG; iEEG) in patients prior to epilepsy surgery. These data yield ideal spatial and temporal resolutions. However, these techniques come at the cost of substantially reduced brain coverage. Moreover, the locations of the electrodes are determined clinically, and not based on research needs.

Due to these limitations, clinicians must optimize their diagnosis based on the resolution constraints imposed by non-invasive recording techniques and the coverage constraints imposed by the intracranial recordings. Therefore, the development of a new diagnostic tool to infer brain patterns at high temporal and spatial resolution from just a few measurements is not just important for patient outcomes, but is critical for better understanding the dynamic interactions in the human brain.

1.1.2 Scalability

In addition to improved data resolution, quantifying the dynamics of neural networks is fundamental to gaining insights into information processing in the brain. To better understand how the brain operates, the tools we use must reflect the complexity of this system. The challenge is that these dynamic network patterns cannot be directly observed from conventional brain recordings, but rather they must be inferred using computational methods applied to the brain recordings.

1.1.2.1 Higher-order correlations

Recent work has shown that cognition is mediated by dynamic interactions between brain structures (Grossberg, 1988; Friston, 2000; Sporns and Honey, 2006; Turk-Browne, 2013; Demertzi et al., 2019; Solomon et al., 2019; Lurie et al., 2018; Preti et al., 2017; Zou et al., 2019; Mack et al., 2017; Bressler and Kelso, 2001; McIntosh, 2000). In particular, Simony et al. (2016), showed that dynamic first-order correlations (i.e., pairwise correlations between brain structures that change from moment to moment) change according to ongoing cognitive processes. Extending this idea beyond first-order interactions, higher-order dynamic interactions may be better at characterizing the complexity of the brain: a system that operates on multiple scales of space and time, and that can be decomposed into subcomponents at many scales, and also the interactions between those subcomponents.

The complexity of this system can be difficult to conceptualize and even more difficult to quantify. One approach is treating the brain as a dynamic connected graph. Each node is a distinct element, which can be a specific brain structure or even a network, and each edge reflects the interrelationships between the corresponding nodes. Higher-order interactions, then, reflect homologous patterns of these interrelationships.

In other words, if the changing patterns of correlations between two regions, A and B , are similar to those between two other regions, C and D , this would be reflected in the second-order correlations between $(A-B)$ and $(C-D)$. In this way, second-order correlations identify similarities and differences between subgraphs of the brain's connectome. Analogously, third-order correlations reflect homologies between second-order correlations—i.e., homologous patterns of homologous interactions between brain regions. More generally, higher-order correlations reflect homologies between patterns of lower-order correlations.

1.1.2.2 Computational challenges

Higher-order correlations may hold further insights into the interactions required for higher cognitive processing. Although advances in network neuroscience have soared, and recent work has used graph theory and topology to characterize how brain networks reconfigure during cognition (Bassett et al., 2006; Zheng et al., 2019; McIntosh and Jirsa, 2019; Toker and Sommer, 2019; Sizemore et al., 2018; Reimann et al., 2017; Betzel et al., 2019; Zamani Esfahlani et al., 2020; Faskowitz et al., 2020), this work had not been extended to interactions beyond the second order. This has been due, in part, to computational limitations. To calculate higher order correlations naively, one would square the number of features at each order, and the computations would grow exponentially. This becomes computationally intractable, even with just a few features, very quickly. Moreover, as increasingly large datasets and databases (Van Essen et al., 2013) are being generated, all with many subjects, and each with large numbers of timepoints and features, there are major computational challenges associated with data-driven analyses. Finding ways to scalably quantify the complexity of these interactions is critical in the pursuit of understanding how our brains support cognition.

1.1.3 Naturalistic tasks

Investigating the networks underlying time-varying, higher-order phenomena such as cognition requires particular consideration of the evaluated brain data. Specifically, the task

used during data collection and the richness of the response. Tasks that require multimodal integration at varying spatial and temporal scales may better approximate natural complex thought.

Naturalistic stimuli, like stories and movies, are complex and exhibit rich structure at a variety of timescales. In comparison to simple and targeted paradigms, which seek to limit and control the sampled space, naturalistic paradigms attempt to broaden that sampled space. They attempt to evoke more ecologically-valid brain responses while still remaining in the confines of the lab. Their use has advanced our understanding of human cognition (e.g., Hasson et al., 2004; Bartels and Zeki, 2005) and underlying neural dynamics (e.g., Sonkusare et al., 2019; Nastase et al., 2020).

Although naturalistic paradigms are less controlled, certain aspects of data quality are improved. For example, participants have been shown to be more attentive and engaged with naturalistic stimuli (Vanderwal et al., 2019) which may provide a better replacement for resting-state in which study goals are open ended (Finn, 2021). Sustaining attention is always an important consideration in task design, but is particularly crucial for populations where attention is already a challenge (Eickhoff et al., 2020; Cantlon and Li, 2013). Naturalistic stimuli also avoid the inherent confounds from repetitive stimulus presentation (Hasson et al., 2010).

1.1.3.1 Intersubject reliability

Moreover, naturalistic stimuli have been shown to elicit more reliable brain activity (Hasson et al., 2004, 2010; Haxby et al., 2020a), in comparison to controlled experiments where the same stimuli in the same conditions are repeated multiple times. With this reliability across participants, researchers can evaluate the similarities in task-evoked brain responses (Haxby et al., 2020b; Hasson et al., 2004, 2010; Haxby et al., 2011) and study how complex representations are encoded in brain activity (Naselaris et al., 2011; Hamilton and Huth, 2018). Fluctuations in narrative engagement and thereby patterns of brain coactivation, have also been shown to predict later recall (Song et al., 2021). (For comprehensive reviews on naturalistic tasks and intersubject correlation see Sonkusare et al., 2019; Nastase et al., 2020)

1.1.3.2 Simony et al. (2016)

The fMRI dataset collected by Simony et al. (2016) is particularly suited to better understand the complexity underlying cognition. The dataset is comprised of four experimental

conditions that exposed participants to stimuli that varied systematically in how cognitively engaging they were. The intact experimental condition had participants listen to an audio recording of a 10-minute story. The scrambled experimental conditions had participants listen to a temporally scrambled version of the story, where either the paragraphs or words occurred out of order. Finally, in a rest experimental condition, participants lay in the scanner with no overt stimulus. This [public dataset](#) provides a unique way of understanding the different brain activity dynamics at varying levels of cognitive processing and engagement.

1.2 What else can we learn about the brain?

The choices we make when modeling brain activity are fundamental to what we can learn from the model. For example, what is the scope of the model? From ion channels opening and closing to networks of brain structures interacting, the brain operates on vastly different scales of time and space. Given this the multi-scale and modular nature of the brain, it is critical to understand at what scale(s) we are modeling (Bassett and Gazzaniga, 2011; Bassett et al., 2018).

What about the assumptions of the model? How generalizable or specific do we want it to be (Scheinost et al., 2019)? We can also think about the fit of a model. Too snug, it won't generalize. Too loose, and it won't specify. And once we make the assumptions, what does it mean when we can make accurate predictions of held-out data (Bassett et al., 2018)? What other properties of brain activity can we begin to elucidate? These are just some of the questions that I plan to explore in my dissertation.

1.3 Organization of this thesis

The goal of my thesis will be to better understand complex cognition by mathematically formalizing the fast and dynamic interactions in the brain. To do this, I will address the data resolution issue in Chapter 2 by providing a multi-subject brain model that captures millisecond-resolution dynamics but at similar spatial resolutions to fMRI. Chapter 3 will address computation tractability by providing a scalable model that quantifies complexity in brain data. Moreover, as motivated in Section 1.1.3, I will explore these interactions that support complex cognition by using brain data collected with a naturalistic paradigm and investigating varying levels of engagement.

In addition to these modeling goals, Chapter 4 will discuss what *else* we can learn about properties of the brain. Specifically, I will examine redundancy in brain activity, and how

consistent it is across people and across tasks. I will also explore the level of interactions throughout and across our brains that support our thoughts.

CHAPTER 2

A Gaussian process model of human electrocorticographic data

This chapter was published as: Lucy L.W. Owen, Tudor A. Muntianu, Andrew C. Heusser, Patrick Daly, Katherine Scangos, and Jeremy R. Manning, entitled "A Gaussian process model of human electrocorticographic data." *Cerebral Cortex*, Volume 30, Issue 10, October 2020, Pages 5333–5345. <https://doi.org/10.1093/cercor/bhaa115>.

2.1 Abstract

We present a model-based method for inferring full-brain neural activity at millimeter-scale spatial resolutions and millisecond-scale temporal resolutions using standard human intracranial recordings. Our approach makes the simplifying assumptions that different people's brains exhibit similar correlational structure, and that activity and correlation patterns vary smoothly over space. One can then ask, for an arbitrary individual's brain: given recordings from a limited set of locations in that individual's brain, along with the observed spatial correlations learned from other people's recordings, how much can be inferred about ongoing activity at *other* locations throughout that individual's brain? We show that our approach generalizes across people and tasks, thereby providing a person- and task-general means of inferring high spatiotemporal resolution full-brain neural dynamics from standard low-density intracranial recordings.

2.2 Introduction

Modern human brain recording techniques are fraught with compromise (Sejnowski et al., 2014). Commonly used approaches include functional magnetic resonance imaging

(fMRI), scalp electroencephalography (EEG), and magnetoencephalography (MEG). For each of these techniques, neuroscientists and electrophysiologists must choose to optimize spatial resolution at the cost of temporal resolution (e.g., as in fMRI) or temporal resolution at the cost of spatial resolution (e.g., as in EEG and MEG). A less widely used approach (due to requiring work with neurosurgical patients) is to record from electrodes implanted directly onto the cortical surface (electrocorticography; ECoG) or into deep brain structures (intracranial EEG; iEEG). However, these intracranial approaches also require compromise: the high spatiotemporal resolution of intracranial recordings comes at the cost of substantially reduced brain coverage, since safety considerations limit the number of electrodes one may implant in a given patient's brain. Further, the locations of implanted electrodes are determined by clinical, rather than research, needs.

An increasingly popular approach is to improve the effective spatial resolution of MEG or scalp EEG data by using a geometric approach called *beamforming* to solve the biomagnetic or bioelectrical inverse problem (Sarvas, 1987). This approach entails using detailed brain conductance models (often informed by high spatial resolution anatomical MRI images) along with the known sensor placements (localized precisely in 3D space) to reconstruct brain signals originating from theoretical point sources deep in the brain (and far from the sensors). Traditional beamforming approaches must overcome two obstacles. First, the inverse problem beamforming seeks to solve has infinitely many solutions. Researchers have made progress towards constraining the solution space by assuming that signal-generating sources are localized on a regularly spaced grid spanning the brain and that individual sources are small relative to their distances to the sensors (Snyder, 1991; Baillet et al., 2001; Hillebrand et al., 2005). The second, and in some ways much more serious, obstacle is that the magnetic fields produced by external (noise) sources are substantially stronger than those produced by the neuronal changes being sought (i.e., at deep structures, as measured by sensors at the scalp). This means that obtaining adequate signal quality often requires averaging the measured responses over tens to hundreds of responses or trials (e.g., see review by Hillebrand et al., 2005).

Another approach to obtaining high spatiotemporal resolution neural data has been to collect fMRI and EEG data simultaneously. Simultaneous fMRI-EEG has the potential to balance the high spatial resolution of fMRI with the high temporal resolution of scalp EEG, thereby, in theory, providing the best of both worlds. In practice, however, the signal quality of both recordings suffers substantially when the two techniques are applied simultaneously (e.g., see review by Huster et al., 2012). In addition, the experimental designs that are ideally suited to each technique individually are somewhat at odds. For example,

fMRI experiments often lock stimulus presentation events to the regularly spaced image acquisition time (TR), which maximizes the number of post-stimulus samples. By contrast, EEG experiments typically employ jittered stimulus presentation times to maximize the experimentalist’s ability to distinguish electrical brain activity from external noise sources such as from 60 Hz alternating current power sources.

The current “gold standard” for precisely localizing signals and sampling at high temporal resolution is to take (ECoG or iEEG) recordings from implanted electrodes (but from a limited set of locations in any given brain). This begs the following question: what can we infer about the activity exhibited by the rest of a person’s brain, given what we learn from the limited intracranial recordings we have from their brain and additional recordings taken from *other* people’s brains? Here we develop an approach, which we call *SuperEEG*¹, based on Gaussian process regression (Rasmussen, 2006). SuperEEG entails using data from multiple people to estimate activity patterns at arbitrary locations in each person’s brain (i.e., independent of their electrode placements). We test our SuperEEG approach using two large datasets of intracranial recordings (Sederberg et al., 2003, 2007b,a; Manning et al., 2011, 2012; Ezzyat et al., 2017; Horak et al., 2017; Kragel et al., 2017; Kucewicz et al., 2017; Lin et al., 2017; Solomon et al., 2018; Weidemann et al., 2019; Ezzyat et al., 2018; Kucewicz et al., 2018). We show that the SuperEEG algorithm recovers signals well from electrodes that were held out of the training dataset. We also examine the factors that influence how accurately activity may be estimated (recovered), which may have implications for electrode design and placement in neurosurgical applications.

2.3 Approach

The SuperEEG approach to inferring high temporal resolution full-brain activity patterns is outlined and summarized in Figure 2.1. We describe (in this section) and evaluate (in *Results*) our approach using two large previously collected datasets comprising multi-session intracranial recordings. Dataset 1 comprises multi-session recordings taken from 6876 electrodes implanted in the brains of 88 epilepsy patients (Sederberg et al., 2003, 2007b,a; Manning et al., 2011, 2012). Each recording session lasted from 0.2–3 h (total recording time: 0.3–14.2 h; Fig. S6E). During each recording session, the patients participated in a free recall list learning task, which lasted for up to approximately 1 h. In addition, the

¹The term “SuperEEG” was coined by Robert J. Sawyer in his popular science fiction novel *The Terminal Experiment* (Sawyer, 1995). SuperEEG is a fictional technology that measures ongoing neural activity throughout the entire living human brain at arbitrarily high spatiotemporal resolution.

recordings included “buffer” time (the length varied by patient) before and after each experimental session, during which the patients went about their regular hospital activities (confined to their hospital room, and primarily in bed). These additional activities included interactions with medical staff and family, watching television, reading, and other similar activities. For the purposes of the Dataset 1 analyses presented here, we aggregated all data across each recording session, including recordings taken during the main experimental task as well as during non-experimental time. We used Dataset 1 to develop our main SuperEEG approach, and to examine the extent to which SuperEEG might be able to generate task-general predictions. Dataset 2 comprised multi-session recordings from 14860 electrodes implanted in the brains of 131 epilepsy patients (Ezzyat et al., 2017; Horak et al., 2017; Kragel et al., 2017; Kucewicz et al., 2017; Lin et al., 2017; Solomon et al., 2018; Weidemann et al., 2019; Ezzyat et al., 2018; Kucewicz et al., 2018). Each recording session lasted from 0.4–2.2 h (total recording time: 0.4–6.6 h; Fig. S6K). Whereas Dataset 1 included recordings taken as the patients participated in a variety of activities, Dataset 2 included recordings taken as each patient performed each of two specific experimental memory tasks: a random word list free recall task (Experiment 1) and a categorized word list free recall task (Experiment 2). We used Dataset 2 to further examine the ability of SuperEEG to generalize its predictions within versus across tasks. Figure S6 provides additional information about both datasets.

We first applied fourth order Butterworth notch filters to remove 60 Hz (± 0.5 Hz) line noise from every recording (from every electrode). Next, we downsampled the recordings (regardless of the original samplerate) to 250 Hz. This downsampling step served to both normalize for differences in sampling rates across patients and to ease the computational burden of our subsequent analyses. We then excluded any electrodes that showed putative epileptiform activity. Specifically, we excluded from further analysis any electrode that exhibited a maximum kurtosis of 10 or greater across all of that patient’s recording sessions. We also excluded any patients with fewer than 2 electrodes that passed this criteria, as the SuperEEG algorithm requires measuring correlations between 2 or more electrodes from each patient. For Dataset 1, this yielded clean recordings from 4168 electrodes implanted throughout the brains of 67 patients (Fig. 2.1A, colored dots); for Dataset 2, this yielded clean recordings from 5023 electrodes implanted in 78 patients. Each individual patient contributed electrodes from a limited set of brain locations, which we localized in a common space (MNI152; Grabner et al., 2006); an example Dataset 1 patient’s 54 electrodes that survived the kurtosis thresholding procedure are highlighted in black and red (Fig. 2.1A).

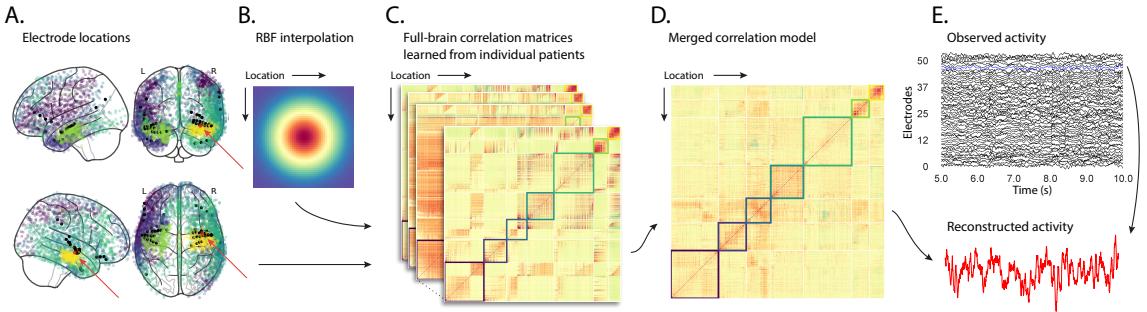


Figure 2.1: Methods overview. **A. Electrode locations.** Each dot reflects the location of a single electrode implanted in the brain of a Dataset 1 patient. A held-out recording location from one patient is indicated in red, and the patient’s remaining electrodes are indicated in black. The electrodes from the remaining patients are colored by k -means cluster (computed using the full-brain correlation model shown in Panel D). **B. Radial basis function kernel.** Each electrode contributed by the patient (black) weights on the full set of locations under consideration (all dots in Panel A, defined as \bar{R} in the text). The weights fall off with positional distance (in MNI152 space) according to an RBF. **C. Per-patient correlation matrices.** After computing the pairwise correlations between the recordings from each patient’s electrodes, we use RBF-weighted averages to estimate correlations between all locations in \bar{R} . We obtain an estimated full-brain correlation matrix using each patient’s data. **D. Merged correlation model.** We combine the per-patient correlation matrices (Panel C) to obtain a single full-brain correlation model that captures information contributed by every patient. Here we have sorted the rows and columns to reflect k -means clustering labels (using $k=7$; Yeo et al., 2011), whereby we grouped locations based on their correlations with the rest of the brain (i.e., rows of the matrix displayed in the panel). The boundaries denote the cluster groups. The rows and columns of Panel C have been sorted using the Panel D-derived cluster labels. **E. Reconstructing activity throughout the brain.** Given the observed recordings from the given patient (shown in black; held-out recording is shown in blue), along with a full-brain correlation model (Panel D), we use Equation 2.12 to reconstruct the most probable activity at the held-out location (red).

The recording from a given electrode is maximally informative about the activity of the neural tissue immediately surrounding its recording surface. However, brain regions that are distant from the recording surface of the electrode also contribute to the recording, albeit (*ceteris paribus*) to a much lesser extent. One mechanism underlying these contributions is volume conduction. The precise rate of falloff due to volume conduction (i.e., how much a small volume of brain tissue at location x contributes to the recording from an electrode at location η) depends on the size of the recording surface, the electrode's impedance, and the conductance profile of the volume of brain between x and η . As an approximation of this intuition, we place a Gaussian radial basis function (RBF) at the location η of each electrode's recording surface (Fig. 2.1B). We use the values of the RBF at any brain location x as a rough estimate of how much structures around x contributed to the recording from location η :

$$\text{rbf}(x|\eta, \lambda) = \exp \left\{ -\frac{\|x - \eta\|^2}{\lambda} \right\}, \quad (2.1)$$

where the width variable λ is a parameter of the algorithm (which may in principle be set according to location-specific tissue conductance profiles) that governs the level of spatial smoothing. In choosing λ for the analyses presented here, we sought to maximize spatial resolution (which implies a small value of λ) while also maximizing the algorithm's ability to generalize to any location throughout the brain, including those without dense electrode coverage (which implies a large value of λ). Here we set $\lambda = 20$, guided in part by our prior related work (Manning et al., 2014, 2018), and in part by examining the brain coverage with non-zero weights achieved by placing RBFs at each electrode location in Dataset 1 and taking the sum (across all electrodes) at each voxel in a 4 mm^3 MNI brain. (We then held λ fixed for our analyses of Dataset 2.) We note that this value could in theory be further optimized, e.g., using cross validation or a formal model (e.g., Manning et al., 2018).

A second mechanism whereby a given region x can contribute to the recording at η is through (direct or indirect) anatomical connections between structures near x and η . Although anatomical and functional correlations can differ markedly (e.g., Adachi et al., 2012; Honey et al., 2009; ni et al., 2014), we use temporal correlations in the data to estimate these anatomical connections (Becker et al., 2018). Let \bar{R} be the set of locations at which we wish to estimate local field potentials, and let $R_s \subseteq \bar{R}$ be set of locations at which we observe local field potentials from patient s (excluding the electrodes that did not pass the kurtosis test described above). In the analyses below we define $\bar{R} = \cup_{s=1}^S R_s$. We

can calculate the expected inter-electrode correlation matrix for patient s , where $C_{s,k}(i,j)$ is the correlation between the time series of voltages for electrodes i and j from subject s during session k , using:

$$\bar{C}_s = r \left(\frac{1}{n} \left(\sum_{k=1}^n z(C_{s,k}) \right) \right), \text{ where} \quad (2.2)$$

$$z(r) = \frac{\log(1+r) - \log(1-r)}{2} \text{ is the Fisher } z\text{-transformation and} \quad (2.3)$$

$$z^{-1}(z) = r(z) = \frac{\exp(2z) - 1}{\exp(2z) + 1} \text{ is its inverse.} \quad (2.4)$$

Next, we use Equation 2.1 to construct a number of to-be-estimated locations by number of patient electrode locations weight matrix, W_s . Specifically, W_s approximates how informative the recordings at each location in R_s are in reconstructing activity at each location in \bar{R} , where the contributions fall off with an RBF according to the distances between the corresponding locations:

$$W_s(i,j) = \text{rbf}(i|j, \lambda). \quad (2.5)$$

Given this weight matrix, W_s , and the observed inter-electrode correlation matrix for patient s , \bar{C}_s , we can estimate the correlation matrix for all locations in \bar{R} (\hat{C}_s ; Fig. 2.1C) using:

$$\hat{N}_s(x, y) = \sum_{i=1}^{|R_s|} \sum_{j=1}^{i-1} W(x, i) \cdot W(y, j) \cdot z(\bar{C}_s(i, j)) \quad (2.6)$$

$$\hat{D}_s(x, y) = \sum_{i=1}^{|R_s|} \sum_{j=1}^{i-1} W(x, i) \cdot W(y, j). \quad (2.7)$$

$$\hat{C}_s = r \left(\frac{\hat{N}_s}{\hat{D}_s} \right). \quad (2.8)$$

After estimating the numerator (\hat{N}_s) and denominator (\hat{D}_s) placeholders for each \hat{C}_s , we aggregate these estimates across the S patients to obtain a single expected full-brain corre-

lation matrix (\hat{K} ; Fig. 2.1D):

$$\hat{K} = \text{r} \left(\frac{\sum_{s=1}^S \hat{N}_s}{\sum_{s=1}^S \hat{D}_s} \right). \quad (2.9)$$

Intuitively, the numerators capture the general structures of the patient-specific estimates of full-brain correlations, and the denominators account for which locations were near the implanted electrodes in each patient. To obtain \hat{K} , we compute a weighted average across the estimated patient-specific full-brain correlation matrices, where patients with observed electrodes near a particular set of locations in \hat{K} contribute more to the estimate.

Having used the multi-patient data to estimate a full-brain correlation matrix at the set of locations in \bar{R} that we wish to know about, we next use \hat{K} to estimate activity patterns everywhere in \bar{R} , given observations at only a subset of locations in \bar{R} (Fig. 2.1E).

Let α_s be the set of indices of patient s 's electrode locations in \bar{R} (i.e., the locations in R_s), and let β_s be the set of indices of all other locations in \bar{R} . In other words, β_s reflects the locations in \bar{R} where we did not observe a recording for patient s (these are the recording locations we will want to fill in using SuperEEG). We can sub-divide \hat{K} as follows:

$$\hat{K}_{\beta_s, \alpha_s} = \hat{K}(\beta_s, \alpha_s), \text{ and} \quad (2.10)$$

$$\hat{K}_{\alpha_s, \alpha_s} = \hat{K}(\alpha_s, \alpha_s). \quad (2.11)$$

Here $\hat{K}_{\beta_s, \alpha_s}$ represents the correlations between the “unknown” activity at the locations indexed by β_s and the observed activity at the locations indexed by α_s , and $\hat{K}_{\alpha_s, \alpha_s}$ represents the correlations between the observed recordings (at the locations indexed by α_s).

Let Y_{s,k,α_s} be the number-of-timepoints (T) by $|\alpha_s|$ matrix of (observed) voltages from the electrodes in α_s during session k from patient s . Then we can estimate the voltage from patient s 's k^{th} session at the locations in β_s as follows (Rasmussen, 2006):

$$\hat{Y}_{s,k,\beta_s} = ((\hat{K}_{\beta_s, \alpha_s} \cdot \hat{K}_{\alpha_s, \alpha_s}^{-1}) \cdot Y_{s,k,\alpha_s}^T)^T. \quad (2.12)$$

This equation is the foundation of the SuperEEG algorithm. Whereas we observe recordings only at the locations indexed by α_s , Equation 2.12 allows us to estimate the recordings at all locations indexed by β_s , which we can define *a priori* to include any locations we wish, throughout the brain. This yields estimates of the time-varying voltages at *every* lo-

cation in \bar{R} , provided that we define \bar{R} in advance to include the union of all of the locations in R_s and all of the locations at which we wish to estimate recordings (i.e., a timeseries of voltages).

We designed our approach to be agnostic to electrode impedances, as electrodes that do not exist do not have impedances. Therefore our algorithm recovers voltages in standard deviation (z -scored) units rather than attempting to recover absolute voltages. (This property reflects the fact that $\hat{K}_{\beta_s, \alpha_s}$ and $\hat{K}_{\alpha_s, \alpha_s}$ are correlation matrices rather than covariance matrices.) Also, we note that Equation 2.12 requires computing a T by T matrix, which can become computationally expensive when T is very large (e.g., for the Dataset 1 patient with the longest recording time, $T = 12,786,750$; also see Fig. S6, Panels E and K). However, because Equation 2.12 is time invariant, we may compute Y_{s,k,β_s} in a piecewise manner by filling in Y_{s,k,β_s} one row at a time (using the corresponding samples from Y_{s,k,α_s}).

The SuperEEG algorithm described above and in Figure 2.1 allows us to estimate, up to a constant scaling factor, local field potentials (LFPs) for each patient at all arbitrarily chosen locations in the set \bar{R} , *even if we did not record that patient’s brain at all of those locations*. We next turn to an evaluation of the accuracy of those estimates.

2.4 Results

We used a cross-validation approach to test the accuracy with which the SuperEEG algorithm reconstructs activity throughout the brain. For each patient in turn, we estimated full-brain correlation matrices (Eqn. 2.9) using data from all of the *other* patients. This step ensured that the data we were reconstructing could not also be used to estimate the between-location correlations that drove the reconstructions via Equation 2.12 (otherwise the analysis would be circular). For that held-out patient, we held out each electrode in turn. We used Equation 2.12 to reconstruct activity at the held-out electrode location, using the correlation matrix learned from all other patients’ data as \hat{K} , and using activity recorded from the other electrodes from the held-out patient as Y_{s,k,α_s} . (For analyses examining the stability of our estimates of \hat{K} across time and patients, see Figs. S7 and S8, respectively). We then asked: how closely did each of the SuperEEG-estimated recordings at those electrodes match the observed recordings from those electrodes (i.e., how closely did the estimated \hat{Y}_{s,k,β_s} match the observed Y_{s,k,β_s})?

We used this general approach to quantify the algorithm’s performance across the full dataset. For each held-out electrode, from each held-out patient in turn, we computed the average correlation (across recording sessions) between the SuperEEG-reconstructed volt-

age traces and the observed voltage traces from that electrode. For this analysis we set \bar{R} to be the union of all electrode locations across all patients. This yielded a single correlation coefficient for each electrode location in \bar{R} , reflecting how well the SuperEEG algorithm was able to recover the recording at that location by incorporating data across patients (black histogram in Fig. 2.2A, map in Fig. 2.2C). The observed distribution of correlations was centered well above zero (mean: $r = 0.51$; t -test comparing mean of distribution of z -transformed average patient correlation coefficients to 0: $t(66) = 23.55, p < 10^{-10}$), indicating that the SuperEEG algorithm recovers held-out activity patterns substantially better than random guessing.

Next, we compared the quality of these across-participant reconstructions (i.e., computed using a correlation model learned from other patients' data) to reconstructions generated using a correlation model trained using the in-patient's data. In other words, for this within-patient benchmark analysis we estimated \hat{C}_s (Eqn. 2.8) for each patient in turn, using recordings from all of that patient's electrodes except at the location we were reconstructing. These within-patient reconstructions serve as an estimate of how well data from all of the other electrodes from that single patient may be used to estimate held-out data from the same patient. This allows us to ask how much information about the activity at a given electrode might be inferred through (a) volume conductance or other sources of "leakage" from activity patterns measured from the patient's other electrodes and (b) across-electrode correlations learned from that single patient. As shown in Figure 2.2A (gray histogram), the distribution of within-patient correlations was centered well above zero (mean: $r = 0.32$; t -test comparing mean of distribution of z -transformed average patient correlation coefficients to 0: $t(66) = 15.16, p < 10^{-10}$). However, the across-patient correlations were substantially higher (t -test comparing average z -transformed within versus across patient electrode correlations: $t(66) = 9.17, p < 10^{-10}$). This is an especially conservative test, given that the across-patient SuperEEG reconstructions exclude (from the correlation matrix estimates) all data from the patient whose data is being reconstructed. We repeated each of these analyses on a second independent dataset and found similar results (Fig. 2.2B, D; within versus across reconstruction accuracy: $t(77) = 11.25, p < 10^{-10}$). We also replicated this result separately for each of the two experiments from Dataset 2 (Fig. S3). This overall finding, that reconstructions of held-out data using correlation models learned from *other* patient's data yield higher reconstruction accuracy than correlation models learned from the patient whose data is being reconstructed, has two important implications. First, it implies that distant electrodes provide additional predictive power to the data reconstructions beyond the information contained solely in nearby electrodes. This follows from the

fact that each patient’s grid, strip, and depth electrodes are implanted in a unique set of locations, so for any given electrode the closest electrodes in the full dataset tend to come from the same patient. Second, it implies that the spatial correlations learned using the SuperEEG algorithm are, to some extent, similar across people.

The recordings we analyzed from Dataset 1 comprised data collected as the patients performed a variety of (largely idiosyncratic) tasks throughout each day’s recording session. That we observed reliable reconstructions across patients suggests that the spatial correlations derived from the SuperEEG algorithm are, to some extent, similar across tasks. We tested this finding more directly using Dataset 2. In Dataset 2, the recordings were limited to times when each patient was participating in one of two experiments. Experiment 1 is a random-word list free recall task; Experiment 2 is a categorized list free recall task (24 patients participated in both). We wondered whether a correlation model learned from data from one experiment might yield good predictions of data from the other experiment. Further, we wondered about the extent to which it might be beneficial or harmful to combine data across tasks.

To test the task-specificity of the SuperEEG-derived correlation models, we restricted the dataset to the 24 patients that participated in both experiments and repeated the above within- and across-patient cross validation procedures separately for Experiment 1 and Experiment 2 data from Dataset 2. We then compared the reconstruction accuracies for held-out electrodes, for models trained within versus across the two experiments, or combining across both experiments (Fig. S1). In every case we found that across-patient models trained using data from all other patients out-performed within-patient models trained on data only from the subject contributing the given electrode ($ts(23) > 6.50, ps < 10^{-5}$). All reconstruction accuracies also reliably exceeded chance performance ($ts(23) > 8.00, ps < 10^{-8}$). Average reconstruction accuracy was highest for the across-patient models limited to data from the same experiment (mean accuracy: $r = 0.68$); next-highest for the models that combined data across both experiments (mean accuracy: $r = 0.61$); and lowest for models trained across tasks (mean accuracy: $r = 0.47$). This pattern of results also held for each of the Dataset 2 experiments individually (Fig. S2). Taken together, these results indicate that there are reliable commonalities in the spatial correlations of full-brain activity across tasks, but that there are also reliable differences in these spatial correlations across tasks. Whereas reconstruction accuracy benefits from incorporating data from other patients, reconstruction accuracy is highest when constrained to within-task data, or data that includes a variety of tasks (e.g., Dataset 1, or combining across the two Dataset 2 experiments).

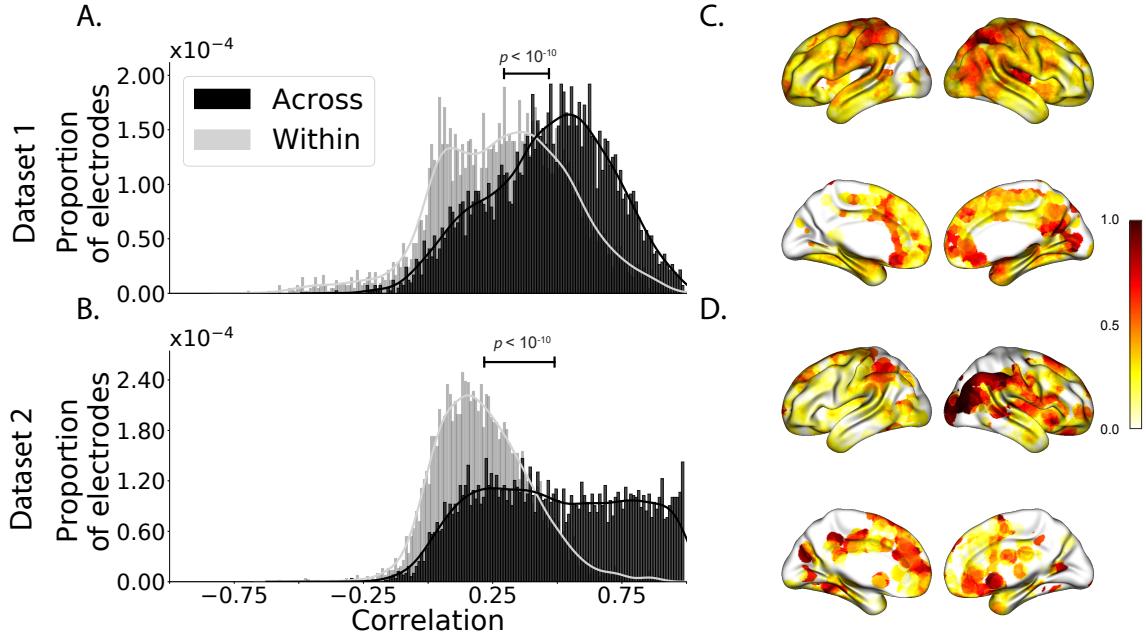


Figure 2.2: Reconstruction accuracy across all electrodes in two ECoG datasets. A. Distributions of correlations between observed versus reconstructed activity by electrode, for Dataset 1. The across-patient distribution (black) reflects reconstruction accuracy (correlation) using a correlation model learned from all but one patient's data, and then applied to that held-out patient's data. The within-patient distribution (gray) reflects performance using a correlation model learned from the same patient who contributed the to-be-reconstructed electrode. **B. Distributions of correlations for Dataset 2.** This panel is in the same format as Panel A, but reflects results obtained from Dataset 2. The histograms aggregate data across both Dataset 2 experiments; for results broken down by experiment see Figures S2 and S3. **C.–D. Reconstruction accuracy by location.** The colors denote the average across-session correlations, using the across-patient correlation model, between the observed and reconstructed activity at the given electrode location projected to the cortical surface (Combrisson et al., 2019). Panel C displays the map for Dataset 1 and Panel D displays the map for Dataset 2.

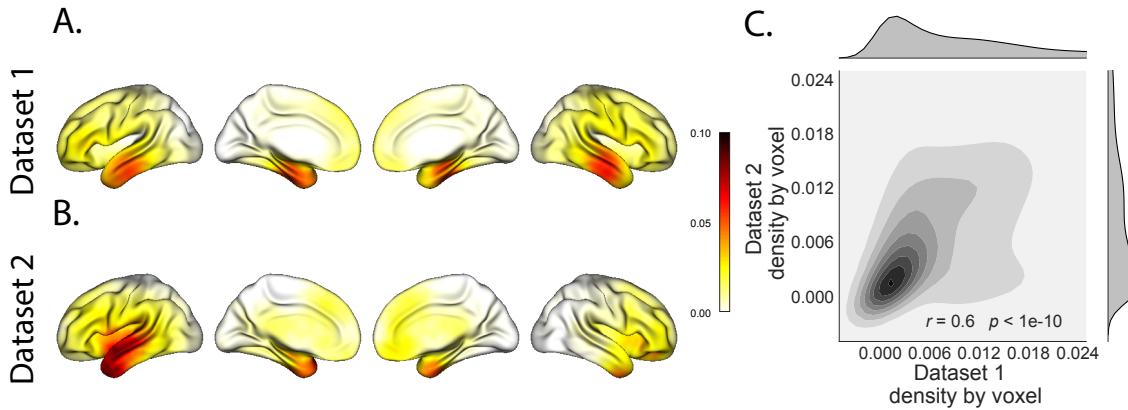


Figure 2.3: Electrode sampling density by location. **A. Electrode sampling density by voxel location in Dataset 1.** Each voxel is colored by the proportion of total electrodes in the dataset that are located within 20 MNI units of the given voxel. **B. Electrode sampling density by voxel location in Dataset 2.** This panel displays the sampling density map for Dataset 2, in the same format as Panel A. **C. Correspondence in sampling density by voxel location across Datasets 1 and 2.** The two-dimensional histogram displays the per-voxel sampling densities in the two Datasets, and the one-dimensional histograms display the proportions of voxels in each dataset with the given density value. The correlation reported in the panel is across voxels in the 4 mm^3 MNI152 brain.

Although both datasets we examined provide good full-brain coverage (when considering data from every patient), electrodes were not sampled uniformly throughout the brain. For example, in our patient population, electrodes are more likely to be implanted in regions like the medial temporal lobe (MTL), and are rarely implanted in occipital cortex (Fig. 2.3A, B). Separately for each dataset, for each voxel in the 4 mm^3 voxel MNI152 brain, we computed the proportion of electrodes in the dataset that were contained within a 20 MNI unit radius sphere centered on that voxel. We defined the *density* at that location as this proportion. Across Datasets 1 and 2, the electrode placement densities were similar (correlation by voxel: $r = 0.6, p < 10^{-10}$). We wondered whether regions with good coverage might be associated with better reconstruction accuracy. For example, Figures 2.2C and D indicate that some electrodes in the MTL (which tends to be relatively densely sampled) have relatively high reconstruction accuracy, and occipital electrodes (which tends to be relatively sparsely sampled) tend to have relatively low reconstruction accuracy. To test whether this held more generally across the entire brain, for each dataset we computed the electrode placement density for each electrode from each patient (using the proportion of *other* patients' electrodes within 20 MNI units of the given electrode). We then correlated these density values with the across-patient reconstruction accuracies

for each electrode. We found no reliable correlation between reconstruction accuracy and density for Dataset 1 ($r = 0.05, p = 0.70$) and a reliable negative correlation for Dataset 2 ($r = -0.21, p = 0.05$). This suggests that the reconstruction accuracies we observed are *not* driven solely by sampling density, but rather may also reflect higher order properties of neural dynamics such as functional correlations between distant voxels (Betzel et al., 2017).

Prior work in humans and animals has shown that the spatial profile of the local field potential differs by frequency band (e.g., with respect to volume conductance properties and contribution to the local field potential; Buzsaki et al., 2012; Fries et al., 2007; Crone et al., 2011). For example, lower frequency components of the local field potential tend to have higher power and extend further in space than high-frequency components (e.g., Miller et al., 2007; Manning et al., 2009). We wondered whether the reconstructions we observed might be differently weighting or considering the contributions of activity at different frequency bands. We therefore examined a range of frequency bands (δ : 2–4 Hz; θ : 4–8 Hz; α : 8–12 Hz; β : 12–30 Hz; γ_L : 30–60 Hz; and γ_H : 60–100 Hz), along with a measure of broadband (BB) power. We used second-order Butterworth bandpass filters to compute the activity patterns within each narrow frequency band. We defined broadband power as the mean height of a linear robust regression fit in log-log space to the order 4 Morelet wavelet-computed power spectrum at 50 log-spaced frequencies from 2–100 Hz (Manning et al., 2009). We then repeated our within-subject and across-subject cross-validated reconstruction accuracy tests (analogous to Fig. 2.2) separately for each frequency band (Fig. 2.4). (We also carried out a similar analysis on the Hilbert transform-computed spectral power within each narrow band; see Fig. S4.) Across both datasets, we found that our approach is best at reconstructing patterns of broadband activity (right-most bars in Figs. 2.4A and D), a correlate of population firing rate (Manning et al., 2009). We also achieved good reconstruction accuracy within each narrow frequency band (Figs. 2.4 and S4). Activity at lower frequencies (δ, θ, α , and β) tended to be reconstructed better than high-frequency patterns (γ_L and γ_H), with reconstruction accuracy peaking in the θ band. Overall, these results indicate that our approach is able to accurately recover information within the 2–100 Hz range.

A basic assumption of our approach (and of most prior ECoG work) is that electrode recordings are most informative about the neural activity near the recording surface of the electrode. But if we consider that activity patterns throughout the brain are meaningfully correlated, are there particular implantation locations that, if recorded from a given patient’s brain, yield especially high reconstruction accuracies throughout the rest of their brain? For

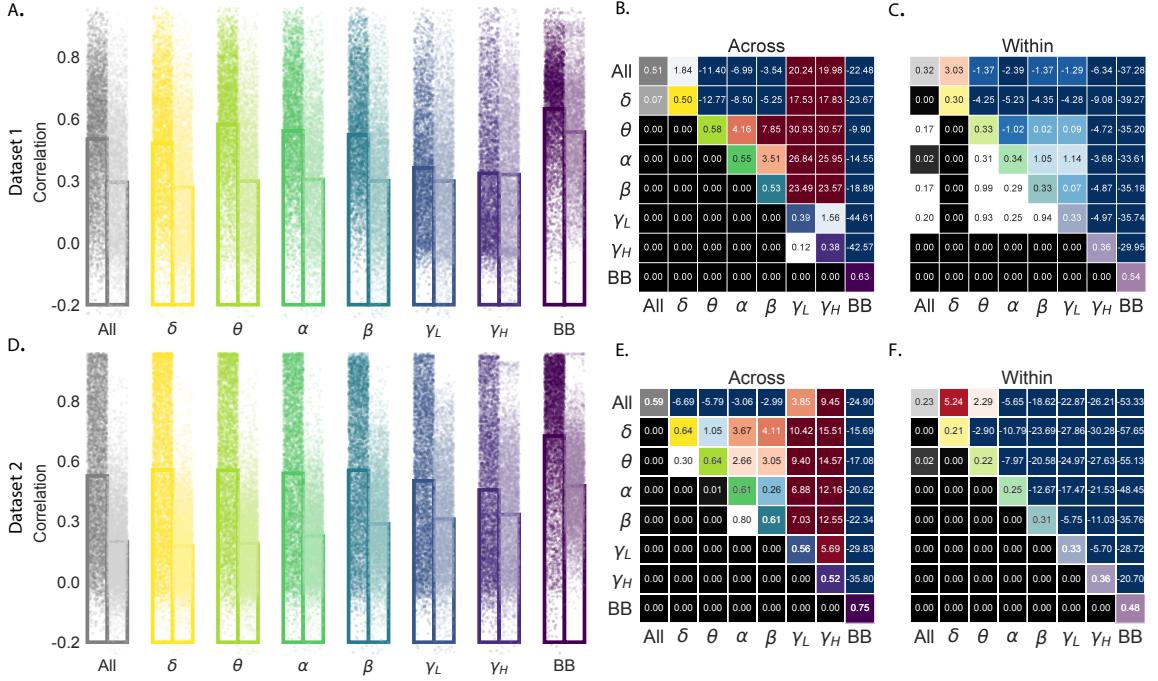


Figure 2.4: Reconstruction accuracy across all electrodes in two ECoG datasets for each frequency band. A. Distributions of correlations between observed versus reconstructed activity by electrode for each frequency band in Dataset 1. Each color denotes a different frequency band. Within each color group, the darker dots and bar on the left display the distribution (and mean) across-patient reconstruction accuracies (analogous to the black histograms in Fig. 2.2). The lighter dots and bar on the right display the distribution (and mean) within-patient reconstruction accuracies (analogous to the gray histograms in Fig. 2.2). Each dot indicates the reconstruction accuracy for one electrode in the dataset. To facilitate visual comparison with the frequency-specific results, the leftmost bars (gray) re-plot the histograms in Figure 2.2A. **B. Statistical summary of across-patient reconstruction accuracy by electrode for each frequency band in Dataset 1.** In the upper triangles of each map, warmer colors (positive t -values) indicate that the reconstruction accuracy for the frequency band in the given row was greater (via a two-tailed paired-sample t -test) than for the frequency band in the given column. Cooler colors (negative t -values) indicate that reconstruction accuracy for the frequency band in the given row was lower than for the frequency band in the given column. The lower triangles of each map denote the corresponding p -values for the t -tests. The diagonal entries display the average reconstruction accuracy within each frequency band. **C. Statistical summary of within-patient reconstruction accuracy by electrode for each frequency band in Dataset 1.** This panel displays the within-patient statistical summary, in the same format as Panel B. **D. Distributions of correlations between observed versus reconstructed activity by electrode, for each frequency band in Dataset 2.** This panel displays reconstruction accuracy distributions for each frequency band for Dataset 2. **E.–F. Statistical summaries of across-patient and within-patient reconstruction accuracy by electrode for each frequency band in Dataset 2.** These panels are in the same as Panels B and C, but display results from Dataset 2.

example, one might hypothesize that brain structures that are heavily interconnected with many other structures could be more informative about full-brain activity patterns than comparatively isolated structures. To test this hypothesis, we computed the average reconstruction accuracy across all of each patient’s electrodes (using our across-patients cross validation test; black histograms in Fig. 2.2A and B). We first labeled each patient’s electrodes, in each dataset, with the average reconstruction accuracy for that patient. In other words, we assigned every electrode from each patient the same value, reflecting how well the activity patterns for that patient were reconstructed. Next, for each voxel in the 4 mm³ MNI brain, we computed the average value across any electrode (from any patient) that came within 20 MNI units of that voxel’s center. This yielded an *information score* for each voxel, reflecting the (weighted) average reconstruction accuracy across any patients with electrodes near each voxel, where the averages were weighted to reflect patients who had more electrodes implanted near that location. We created a single map of these information scores for each dataset, highlighting regions that are especially informative about activity in *other* brain areas (Figs. 2.5A and B). Despite task and patient differences across the two datasets, we nonetheless found that the information score maps from both datasets were correlated (voxelwise correlation between information scores across the two datasets: $r = 0.18, p < 10^{-10}$). Our finding that there were some commonalities between the two datasets’ information score maps lends support to the notion that different brain areas are (reliably) differently informative about full-brain activity patterns. We also examined the intersection between the top 10% most informative voxels across the two datasets (gray areas in Fig. 2.5C, networks shown in Fig. 2.6A, top row). Supporting the notion that structures that are highly interconnected with the rest of the brain are most informative about full-brain activity patterns, the intersecting set of voxels with the highest information scores included major portions of the dorsal attention network (e.g., inferior parietal lobe, precuneus, inferior temporal gyrus, thalamus, and striatum) as well as some portions of the default mode network (e.g., angular gyrus) that are highly interconnected with a large proportion of the brain’s gray matter (e.g., Tomasi and Volkow, 2011).

We also wondered whether the map of information scores might vary as a function of the spectral components of the activity patterns under consideration. We computed analogous maps of information scores for each individual frequency band. Across Datasets 1 and 2 (with the exception of α -band activity), we observed reliable positive correlations between the voxelwise maps of information scores (δ : $r = 0.09, p < 10^{-57}$; θ : $r = 0.24, p < 10^{-60}$; α : $r = -0.03, p < 0.001$; β : $r = 0.02, p = 0.0011$; γ_L : $r = 0.1, p < 10^{-67}$; γ_H : $r = 0.03, p < 10^{-7}$; broadband: $r = 0.21, p < 10^{-297}$).

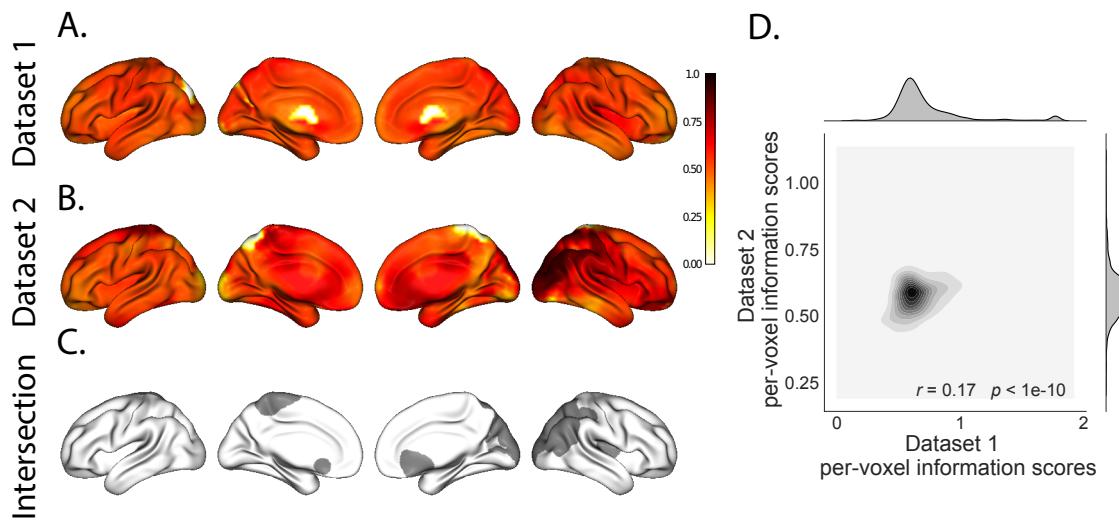


Figure 2.5: Most informative recording locations. **A. Dataset 1 information scores by voxel.** The voxel colors reflect the weighted average reconstruction accuracy across all electrodes from any patients with at least one electrode within 20 MNI units of the given voxel. **B. Dataset 2 information scores by voxel.** This panel is in the same format as Panel A. **C. Intersection.** Gray areas indicate the intersections between the top 10% most informative voxels in each map and projected onto the cortical surface (Combrisson et al., 2019). **D. Correspondence in information scores by voxel across Datasets 1 and 2.** The correlation reported in the Panel is between the per-voxel information scores across Datasets 1 and 2.

To gain additional insight into which regions were most informative about full-brain activity patterns at different frequency bands, we next computed (for each frequency band) the intersection of the top 10% highest information scores across the maps for Datasets 1 and 2 (analogous to our approach in Fig. 2.5C). This yielded a single map of the (reliably) most informative locations, for each frequency band we examined. We then carried out *post hoc* analyses on each of these maps to characterize the underlying structural and functional properties of each set of regions we identified as being particularly informative about one or more types of neural pattern (Figs. 2.6 and S5).

A growing body of neuroscientific research is concerned with characterizing the *parcellations* of anatomical and functional brain networks (for review see Zalesky et al., 2010; Arslan et al., 2018). The dominant approaches entail obtaining a full-brain connectivity matrix using either diffusion tensor imaging to identify the brain’s network of white matter connections, or functional connectivity (typically applied to resting state data) to correlate the patterns of activity exhibited by different brain structures. One can then apply graph theoretic approaches to assign each brain structure (typically a single fMRI voxel) to one or more networks (for review see Bullmore and Sporns, 2009). The result is a set of distinct (or partially overlapping) brain “networks” that may be further examined to elucidate their potential functional role. We overlaid a well-cited seven-network parcellation map identified by Yeo et al. (2011) onto the maps of brain locations that were most informative about each type of neural pattern. For each of these information maps, we computed the proportion of voxels in the most informative brain regions that belonged to each of the seven networks identified by Yeo et al. (2011); Figure 2.6D. We found that the regions we identified as being most informative about different neural patterns varied markedly with respect to which functional networks they belonged to (Fig. 2.6A, B).

The variability we observed in the frequency-specific information score maps is consistent with the notion that there is no “universal” brain region that reflects all types of activity patterns throughout the rest of the brain. Rather, each region’s activity patterns appear to be characterized by different spectral profiles, and the ability to infer full-brain activity patterns at a particular frequency band depends on the structural and functional connectome specific to that frequency band (Fig. 2.6E). We wondered how the maps we found might fit in with prior work. To this end, in addition to examining the anatomical profiles of each map, we used Neurosynth (Rubin et al., 2017) to identify (using meta analyses of the neuroimaging literature) the top five most common terms associated with each frequency-specific map (Fig. 2.6C). We found that δ patterns across the brain were best predicted by regions of ventromedial prefrontal cortex, striatum, and thalamus (yellow). These regions

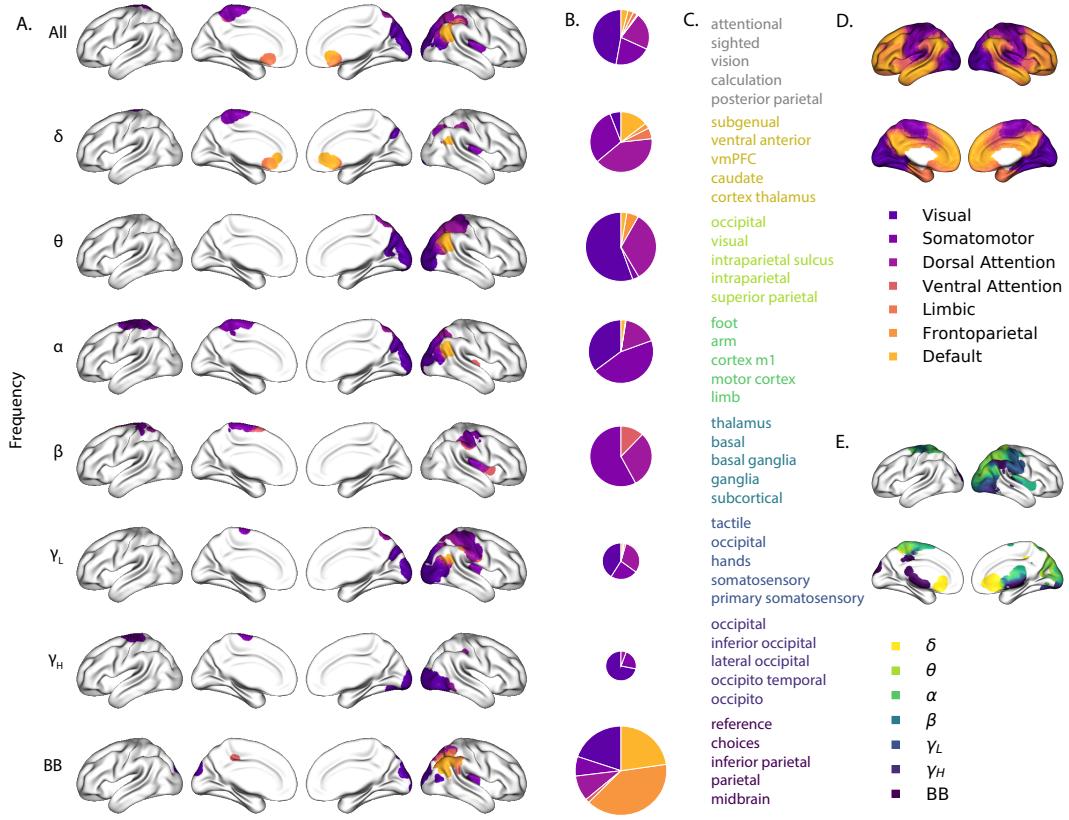


Figure 2.6: Most informative recording locations by frequency band. **A. Intersections between information score maps by frequency band.** The regions indicated in each row depict the intersection between the top 10% most informative locations across Datasets 1 and 2. **B. Network memberships of the most informative brain regions.** The pie charts display the proportions of voxels in each region that belong to the seven networks identified by Yeo et al. (2011). The relative sizes of the charts for each frequency band reflect the average across-subject reconstruction accuracies (Figs. 2.4A, D). The voxels in Panel A are colored according to the same network memberships. **C. Neurosynth terms associated with the most informative brain regions, by frequency band.** The lists in each row display the top five neurosynth terms (Rubin et al., 2017) decoded for each region. **D. Network parcellation map and legend.** The parcellation defined by Yeo et al. (2011) is displayed on the inflated brain maps. The colors and network labels serve as a legend for Panels A and B. **E. Combined map of the most informative brain regions.** The map displays the union of the most informative maps in Panel A, colored by frequency band. The labels also serve as a legend for Panel C.

are also implicated in modulating δ oscillations during sleep, and are heavily interconnected with cortex (e.g., Amzica and Steriade, 1998). The brain areas most informative about full-brain θ patterns were occipital and parietal regions associated with visual processing and visual attention (light green). Prior work has implicated θ oscillations in these areas in periodic sampling of visual attention (e.g., Busch and VanRullen, 2010). We found that full-brain α patterns were best predicted by motor areas (dark green), which also exhibit α band changes during voluntary movements (e.g., Jurkiewicz et al., 2006). Striatum and thalamus (teal) were most informative about full-brain β patterns. Prior work has implicated striatal β activity in sensory and motor processing (Feingold et al., 2015) and thalamic β activity has been implicated in modulating widespread β patterns across neocortex (Sherman et al., 2016). Somatosensory areas (dark blue) were most informative about full-brain γ_L patterns. Prior work has implicated somatosensory γ_L in somatosensory processing and motor planning (Ihara et al., 2003). Occipital cortex (purple) was most informative about full-brain γ_H patterns. Occipital γ_H has also been linked with visual processing and reading (Wu et al., 2011) and the transmission of visual representations from low-order to higher-order visual areas (Matsumoto et al., 2013). Full-brain broadband patterns were best predicted by inferior parietal cortex precuneus (maroon). Functional neuroimaging BOLD responses (Simony et al., 2016) and broadband ECoG patterns (Honey et al., 2012) in these default-mode hubs have been implicated in processing context-dependent representations that unfold over long timescales.

2.5 Discussion

Are our brain's networks static or dynamic? And to what extent are the network properties of our brains stable across people and tasks? One body of work suggests that our brain's *functional* networks are dynamic (e.g., Manning et al., 2018; Owen et al., 2021), person-specific (e.g., Finn et al., 2015), and task-specific (e.g., Turk-Browne, 2013). In contrast, although the gross anatomical structure of our brains changes meaningfully over the course of years as our brains develop, on the timescales of typical neuroimaging experiments (i.e., hours to days) our anatomical networks are largely stable (e.g., Casey et al., 2000). Further, many aspects of brain anatomy, including white matter structure, are largely preserved across people (e.g., Talairach and Tournoux, 1988; Jahanshad et al., 2013; Mori et al., 2008). There are several possible means of reconciling this apparent inconsistency between dynamic person- and task-specific functional networks versus stable anatomical networks. For example, relatively small magnitude anatomical differences

across people may be reflected in reliable functional connectivity differences. Along these lines, one recent study found that diffusion tensor imaging (DTI) structural data is similar across people, but may be used to predict person-specific resting state functional connectivity data (Becker et al., 2018). Similarly, other work indicates that task-specific functional activations may be predicted by resting state functional connectivity data (Cole et al., 2016; Tavor et al., 2016). Another (potentially complementary) possibility is that our functional networks are constrained by anatomy, but nevertheless exhibit (potentially rapid) task-dependent changes (e.g., Sporns and Betzel, 2016). This prior work differs from ours in a number of ways. For example, fMRI data has substantially higher spatial resolution than (raw) ECoG data, and fMRI data have nearly complete spatial overlap across participants whereas ECoG data have minimal spatial overlap across participants. Nevertheless, our work draws inspiration from those studies in that we also attempt to estimate held-out activity patterns across people and tasks.

Here we have taken a model-based approach to studying whether high spatiotemporal resolution activity patterns throughout the human brain may be explained by a static connectome model that is shared across people and tasks. Specifically, we trained a model to take in recordings from a subset of brain locations, and then predicted activity patterns during the same interval, but at *other* locations that were held out from the model. Our model, based on Gaussian process regression, was built on three general hypotheses about the nature of the correlational structure of neural activity (each of which we tested). First, we hypothesized that functional correlations are stable over time and across tasks. We found that, although aspects of the patients' functional correlations were stable across tasks, we achieved better reconstruction accuracy when we trained the model on within-task data. This suggests that our general approach could be extended to better model across-task changes, e.g., following Cole et al. (2016); Tavor et al. (2016); and others. Second, we hypothesized that some of the correlational structure of people's brain activity is similar across individuals. Consistent with this hypothesis, our model explained each patient's data best when trained using data from *other* patients— even when compared models trained within-patient. Third, we resolved ambiguities in the data by hypothesizing that neural activity from nearby sources tends to be similar, all else being equal. This hypothesis was supported through our finding that all of the models we trained that incorporated this spatial smoothness assumption predicted held-out data well above chance.

Another important finding is that SuperEEG-based reconstructions accurately recover activity patterns at a broad range of frequencies (as well as broadband patterns). However, brain networks differed in how informative they were about activity within each frequency

band. Prior work has largely treated region-specific narrowband and broadband activity as an indicator that activity at those frequency ranges reflects that the given region is representing or supporting a particular function. Our work suggests a complementary interpretation that when we observe a particular neural pattern in a particular brain region, it may instead (or in addition) reflect how that region is transmitting information to the rest of the brain via signalling at the given frequency range.

One potential limitation of our approach is that it does not provide a natural means of estimating the precise timing of single-neuron action potentials. Prior work has shown that gamma band and broadband activity in the LFP may be used to estimate the firing rates of neurons that underly the population contributing to the LFP (Miller et al., 2008; Manning et al., 2009; Jacobs et al., 2010; Crone et al., 2011). Because SuperEEG reconstructs LFPs throughout the brain, one could in principle use broadband power in the reconstructed signals to estimate the corresponding firing rates (though not the timings of individual action potentials). We found that we were able to reconstruct full-brain patterns of broadband power well (Fig. 2.4).

A second potential limitation of our approach is that it relies on ECoG data from epilepsy patients. Recent work comparing functional correlations in epilepsy patients (measured using ECoG) and healthy individuals (measured using fMRI) suggests that there are gross similarities between these populations (e.g., Reddy et al., 2018; Kucyi et al., 2018). Nevertheless, because all of the patients we examined have drug-resistant epilepsy, it remains uncertain how generally the findings reported here might apply more broadly to the population at large (e.g., non-clinical populations).

Beyond providing a means of estimating ongoing activity throughout the brain using already-implanted electrodes, our work also has implications how to optimize electrode placements in neurosurgical evaluations. Electrodes are typically implanted to maximize coverage of suspected epileptogenic tissue. However, our findings suggest that this approach might be improved upon. Specifically, one could leverage not only the non-invasive recordings taken during an initial monitoring period (as is currently done routinely), but also recordings collected from *other* patients. We could then ask: given what we learn from other patients' data (and potentially from the scalp EEG recordings of this new patient), where should we place a fixed number of electrodes to maximize our ability to map seizure foci? As shown in Figures 2.5, 2.6, and S5, recordings from different regions vary with respect to how informative they are about different narrowband and broadband full-brain activity patterns.

By providing a means of reconstructing full-brain activity patterns, the SuperEEG ap-

proach maps ECoG recordings from different patients into a common neural space, despite that different patients' electrodes were implanted in different locations. This feature of our approach enables across-patient ECoG studies, analogous to across-subject fMRI studies (e.g., Haxby et al., 2001; Norman et al., 2006; Haxby et al., 2011). Whereas the focus of this manuscript is to specifically evaluate which aspects of neural activity patterns SuperEEG recovers well (or poorly), in parallel work we are training across-patient classifiers by leveraging the common neural spaces obtained by applying SuperEEG to multi-patient ECoG data. For example, we have shown that SuperEEG-derived activity patterns may be used to accurately predict psychiatric conditions such as depression (Scangos et al., 2021). Analogous approaches could in principle be used to develop improved brain-computer interfaces and/or to carry out other analyses that would benefit from high spatiotemporal resolution full-brain data in individuals, projected into a common ECoG space across people.

2.6 Conclusion

Over the past several decades, neuroscientists have begun to leverage the strikingly profound mathematical structure underlying the brain's complexity to infer how our brains carry out computations to support our thoughts, actions, and physiological processes. Whereas traditional beamforming techniques rely on geometric source-localization of signals measured at the scalp, here we propose an alternative approach that leverages the rich correlational structure of two large datasets of human intracranial recordings. In doing so, we are one step closer to observing, and perhaps someday understanding, the full spatiotemporal structure of human neural activity.

CHAPTER 3

High-level cognition during story listening is reflected in high-order dynamic correlations in neural activity patterns

This chapter was published as: Lucy L.W. Owen, Thomas H. Chang, and Jeremy R. Manning, entitled "High-level cognition during story listening is reflected in high-order dynamic correlations in neural activity patterns." Nature Communications, in press. <https://www.biorxiv.org/content/10.1101/763821v3>.

3.1 Abstract

Our thoughts arise from coordinated patterns of interactions between brain structures that change with our ongoing experiences. High-order dynamic correlations in neural activity patterns reflect different subgraphs of the brain's functional connectome that display homologous lower-level dynamic correlations. Here we test the hypothesis that high-level cognition is reflected in high-order dynamic correlations in brain activity patterns. We develop an approach to estimating high-order dynamic correlations in timeseries data, and we apply the approach to neuroimaging data collected as human participants either listen to a ten-minute story or listen to a temporally scrambled version of the story. We train across-participant pattern classifiers to decode (in held-out data) when in the session each neural activity snapshot was collected. We find that classifiers trained to decode from high-order dynamic correlations yield the best performance on data collected as participants listened to the (unscrambled) story. By contrast, classifiers trained to decode data from scrambled versions of the story yielded the best performance when they were trained using first-order dynamic correlations or non-correlational activity patterns. We suggest that as our thoughts

become more complex, they are reflected in higher-order patterns of dynamic network interactions throughout the brain.

3.2 Introduction

A central goal in cognitive neuroscience is to elucidate the neural code: i.e., the mapping between (a) mental states or cognitive representations and (b) neural activity patterns. One means of testing models of the neural code is to ask how accurately that model is able to “translate” neural activity patterns into known (or hypothesized) mental states or cognitive representations Haxby et al. (2001); Norman et al. (2006); Tong and Pratte (2012); Mitchell et al. (2008); Kamitani and Tong (2005); Nishimoto et al. (2011); Pereira et al. (2018); Huth et al. (2012, 2016). Training decoding models on different types of neural features (Fig. 3.1a) can also help to elucidate which specific aspects of neural activity patterns are informative about cognition and, by extension, which types of neural activity patterns might compose the neural code. For example, prior work has used region of interest analyses to estimate the anatomical locations of specific neural representations Etzel et al. (2009), or to compare the relative contributions to the neural code of multivariate activity patterns versus dynamic correlations between neural activity patterns Manning et al. (2018); Fong et al. (2019). An emerging theme in this literature is that cognition is mediated by dynamic interactions between brain structures Grossberg (1988); Friston (2000); Sporns and Honey (2006); Bassett et al. (2006); Turk-Browne (2013); Demertzi et al. (2019); Solomon et al. (2019); Lurie et al. (2018); Preti et al. (2017); Zou et al. (2019); Mack et al. (2017); Bressler and Kelso (2001); McIntosh (2000).

Studies of the neural code to date have primarily focused on univariate or multivariate neural patterns Norman et al. (2006), or (more recently) on patterns of dynamic first-order correlations (i.e., interactions between pairs of brain structures Manning et al. (2018); Fong et al. (2019); Lurie et al. (2018); Preti et al. (2017); Zou et al. (2019); Demertzi et al. (2019)). What might the future of this line of work hold? For example, is the neural code implemented through higher-order interactions between brain structures Reimann et al. (2017)? Second-order correlations reflect homologous patterns of correlation. In other words, if the dynamic patterns of correlations between two regions, A and B , are similar to those between two other regions, C and D , this would be reflected in the second-order correlations between $(A-B)$ and $(C-D)$. In this way, second-order correlations identify similarities and differences between subgraphs of the brain’s connectome. Analogously, third-order correlations reflect homologies between second-order correlations—i.e., homologous

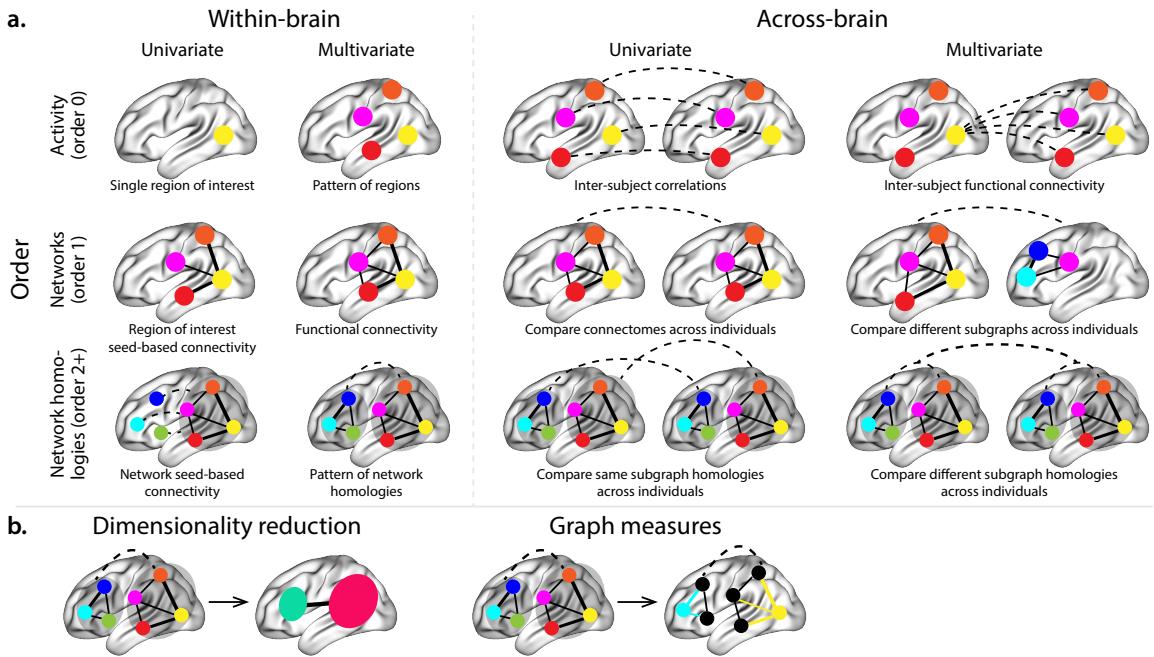


Figure 3.1: Neural patterns. a. A space of neural features. Within-brain analyses are carried out within a single brain, whereas across-brain analyses compare neural patterns across two or more individuals' brains. Univariate analyses characterize the activities of individual units (e.g., nodes, small networks, hierarchies of networks, etc.), whereas multivariate analyses characterize the patterns of activity across units. Order 0 patterns involve individual nodes; order 1 patterns involve node-node interactions; order 2 (and higher) patterns relate to interactions between homologous networks. Each of these patterns may be static (e.g., averaging over time) or dynamic. **b. Summarizing neural patterns.** To efficiently compute with complex neural patterns, it can be useful to characterize the patterns using summary measures. Dimensionality reduction algorithms project the patterns onto lower-dimensional spaces whose dimensions reflect weighted combinations or non-linear transformations of the dimensions in the original space. Graph measures characterize each unit's participation in its associated network.

patterns of homologous interactions between brain regions. More generally, higher-order correlations reflect homologies between patterns of lower-order correlations. We can then ask: which “orders” of interaction are most reflective of high-level cognitive processes?

One reason one might expect to see homologous networks in a dataset is related to the notion that network dynamics reflect ongoing neural computations or cognitive processing Beaty et al. (2016). If the nodes in two brain networks are interacting (within each network) in similar ways then, according to our characterization of network dynamics, we refer to the similarities between those patterns of interaction as higher-order correlations. When higher-order correlations are themselves changing over time, we can also attempt to capture and characterize those high-order dynamics.

Another central question pertains to the extent to which the neural code is carried by activity patterns that directly reflect ongoing cognition Haxby et al. (2001); Norman et al. (2006), versus the dynamic properties of the network structure itself, independent of specific activity patterns in any given set of regions Bassett et al. (2006). For example, graph measures such as centrality and degree Bullmore and Sporns (2009) may be used to estimate how a given brain structure is “communicating” with other structures, independently of the specific neural representations carried by those structures. If one considers a brain region’s position in the network (e.g., its eigenvector centrality) as a dynamic property, one can compare how the positions of different regions are correlated, and/or how those patterns of correlations change over time. We can also compute higher-order patterns in these correlations to characterize homologous subgraphs in the connectome that display similar changes in their constituent brain structures’ interactions with the rest of the brain.

To gain insights into the above aspects of the neural code, we developed a computational framework for estimating dynamic high-order correlations in timeseries data. This framework provides an important advance, in that it enables us to examine patterns of higher-order correlations that are computationally intractable to estimate via conventional methods. Given a multivariate timeseries, our framework provides timepoint-by-timepoint estimates of the first-order correlations, second-order correlations, and so on. Our approach combines a kernel-based method for computing dynamic correlations in timeseries data with a dimensionality reduction step (Fig. 3.1b) that projects the resulting dynamic correlations into a low-dimensional space. We explored two dimensionality reduction approaches: principle components analysis Pearson (1901) (PCA), which preserves an approximately invertible transformation back to the original data McIntosh and Jirsa (2019); Toker and Sommer (2019); Gonzalez-Castillo et al. (2019), and a second non-invertible algorithm for computing dynamic patterns in eigenvector centrality Landau (1895). This latter ap-

proach characterizes correlations between each feature dimension’s relative position in the network (at each moment in time) in favor of the specific activity histories of different features Betzel et al. (2019); Sizemore et al. (2018); Reimann et al. (2017).

We validated our approach using synthetic data where the underlying correlations were known. We then applied our framework to a neuroimaging dataset collected as participants listened to either an audio recording of a ten-minute story, listened to a temporally scrambled version of the story, or underwent a resting state scan Simony et al. (2016). Temporal scrambling has been used in a growing number of studies, largely by Uri Hasson’s group, to identify brain regions that are sensitive to higher-order and longer-timescale information (e.g., cross-sensory integration, rich narrative meaning, complex situations, etc.) versus regions that are primarily sensitive to low-order (e.g., sensory) information. For example, Hasson et al. (2008) argues that when brain areas are sensitive to fine versus coarse temporal scrambling, this indicates that they are “higher order” in the sense that they process contextual information pertaining to further-away timepoints. By contrast, low-level regions, such as primary sensory cortices, do not meaningfully change their responses (after correcting for presentation order) even when the stimulus is scrambled at fine timescales.

We used a subset of the story listening and rest data to train across-participant classifiers to decode listening times (of groups of participants) using a blend of neural features (comprising neural activity patterns, as well as different orders of dynamic correlations between those patterns that were inferred using our computational framework). We found that both the PCA-based and eigenvector centrality-based approaches yielded neural patterns that could be used to decode accurately (i.e., well above chance). Both approaches also yielded the best decoding accuracy for data collected during (intact) story listening when high-order (PCA: second-order; eigenvector centrality: fourth-order) dynamic correlation patterns were included as features. When we trained classifiers on the scrambled stories or resting state data, only (relatively) lower-order dynamic patterns were informative to the decoders. Taken together, our results indicate that high-level cognition is supported by high-order dynamic patterns of communication between brain structures.

3.3 Approach

Our general approach to efficiently estimating high-order dynamic correlations comprises four general steps (Fig. 3.2). First, we derive a kernel-based approach to computing dynamic pairwise correlations in a T (timepoints) by K (features) multivariate timeseries, \mathbf{X}_0 . This yields a T by $\mathcal{O}(K^2)$ matrix of dynamic correlations, \mathbf{Y}_1 , where each row comprises

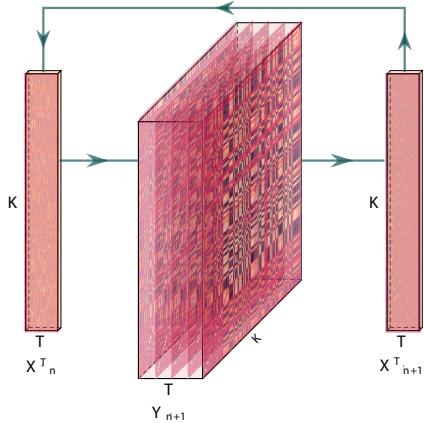


Figure 3.2: Estimating dynamic high-order correlations. Given a T by K matrix of multivariate timeseries data, \mathbf{X}_n (where $n \in \mathbb{N}, n \geq 0$), we use Equation 3.4 to compute a timeseries of K by K correlation matrices, \mathbf{Y}_{n+1} . We then approximate \mathbf{Y}_{n+1} with the T by K matrix \mathbf{X}_{n+1} . This process may be repeated to scalably estimate iteratively higher-order correlations in the data. Note that the transposes of \mathbf{X}_n and \mathbf{X}_{n+1} are displayed in the figure for compactness.

the upper triangle and diagonal of the correlation matrix at a single timepoint, reshaped into a row vector (this reshaped vector is $\left(\frac{K^2-K}{2} + K\right)$ -dimensional). Second, we apply a dimensionality reduction step to project the matrix of dynamic correlations back onto a K -dimensional space. This yields a T by K matrix, \mathbf{X}_1 , that reflects an approximation of the dynamic correlations reflected in the original data. Third, we use repeated applications of the kernel-based dynamic correlation step to \mathbf{X}_n and the dimensionality reduction step to the resulting \mathbf{Y}_{n+1} to estimate high-order dynamic correlations. Each application of these steps to a T by K time series \mathbf{X}_n yields a T by K matrix, \mathbf{X}_{n+1} , that reflects the dynamic correlations between the columns of \mathbf{X}_n . In this way, we refer to n as the order of the timeseries, where \mathbf{X}_0 (order 0) denotes the original data and \mathbf{X}_n denotes (approximated) n^{th} -order dynamic correlations between the columns of \mathbf{X}_0 . Finally, we use a cross-validation-based decoding approach to evaluate how well information contained in a given order (or weighted mixture of orders) may be used to decode relevant cognitive states. If including a given \mathbf{X}_n in the feature set yields higher classification accuracy on held-out data, we interpret this as evidence that the given cognitive states are reflected in patterns of n^{th} -order correlations.

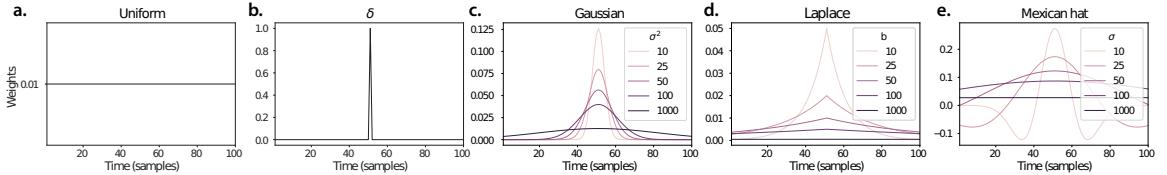


Figure 3.3: Examples of kernel functions. Each panel displays per-timepoint weights for a kernel centered at $t = 50$, evaluated at 100 timepoints ($\tau \in [1, \dots, 100]$). **a. Uniform kernel.** The weights are timepoint-invariant; observations at all timepoints are weighted equally, and do not change as a function of τ . This is a special case kernel function that reduces dynamic correlations to static correlations. **b. Dirac δ kernel.** Only the observation at timepoint t is given a non-zero weight (of 1). **c. Gaussian kernels.** Each kernel's weights fall off in time according to a Gaussian probability density function centered on time t . Weights derived using several different example width parameters (σ^2) are displayed. **d. Laplace kernels.** Each kernel's weights fall off in time according to a Laplace probability density function centered on time t . Weights derived using several different example width parameters (b) are displayed. **e. Mexican hat (Ricker wavelet) kernels.** Each kernel's weights fall off in time according to a Ricker wavelet centered on time t . This function highlights the contrasts between local versus surrounding activity patterns in estimating dynamic correlations. Weights derived using several different example width parameters (σ) are displayed.

Kernel-based approach for computing dynamic correlations

Given a T by K matrix of observations, \mathbf{X} , we can compute the (static) Pearson's correlation between any pair of columns, $\mathbf{X}(\cdot, i)$ and $\mathbf{X}(\cdot, j)$ using Pearson (1901):

$$\text{corr}(\mathbf{X}(\cdot, i), \mathbf{X}(\cdot, j)) = \frac{\sum_{t=1}^T (\mathbf{X}(t, i) - \bar{\mathbf{X}}(\cdot, i)) (\mathbf{X}(t, j) - \bar{\mathbf{X}}(\cdot, j))}{\sqrt{\sum_{t=1}^T \sigma_{\mathbf{X}(\cdot, i)}^2 \sigma_{\mathbf{X}(\cdot, j)}^2}}, \text{ where} \quad (3.1)$$

$$\bar{\mathbf{X}}(\cdot, k) = \frac{1}{T} \sum_{t=1}^T \mathbf{X}(t, k), \text{ and} \quad (3.2)$$

$$\sigma_{\mathbf{X}(\cdot, k)}^2 = \frac{1}{T} \sum_{t=1}^T (\mathbf{X}(t, k) - \bar{\mathbf{X}}(\cdot, k))^2 \quad (3.3)$$

We can generalize this formula to compute time-varying correlations by incorporating a kernel function that takes a time t as input, and returns how much the observed data at each timepoint $\tau \in [-\infty, \infty]$ contributes to the estimated instantaneous correlation Allen et al. (2012) at time t (Fig. 3.3).

Given a kernel function $\kappa_t(\cdot)$ for timepoint t , evaluated at timepoints $\tau \in [1, \dots, T]$, we can update the static correlation formula in Equation 3.1 to estimate the instantaneous

correlation at timepoint t :

$$\text{timecorr}_{\kappa_t}(\mathbf{X}(\cdot, i), \mathbf{X}(\cdot, j)) = \frac{\sum_{\tau=1}^T (\mathbf{X}(\tau, i) - \tilde{\mathbf{X}}_{\kappa_t}(\cdot, i)) (\mathbf{X}(\tau, j) - \tilde{\mathbf{X}}_{\kappa_t}(\cdot, j))}{\sqrt{\sum_{\tau=1}^T \tilde{\sigma}_{\kappa_t}^2(\mathbf{X}(\cdot, i)) \tilde{\sigma}_{\kappa_t}^2(\mathbf{X}(\cdot, j))}}, \text{ where} \quad (3.4)$$

$$\tilde{\mathbf{X}}_{\kappa_t}(\cdot, k) = \sum_{\tau=1}^T \kappa_t(\tau) \mathbf{X}(\tau, k), \quad (3.5)$$

$$\tilde{\sigma}_{\kappa_t}^2(\mathbf{X}(\cdot, k)) = \sum_{\tau=1}^T (\mathbf{X}(\tau, k) - \tilde{\mathbf{X}}_{\kappa_t}(\cdot, k))^2. \quad (3.6)$$

Here $\text{timecorr}_{\kappa_t}(\mathbf{X}(\cdot, i), \mathbf{X}(\cdot, j))$ reflects the correlation at time t between columns i and j of \mathbf{X} , estimated using the kernel κ_t . We evaluate Equation 3.4 in turn for each pair of columns in \mathbf{X} and for kernels centered on each timepoint in the timeseries, respectively, to obtain a T by K by K timeseries of dynamic correlations, \mathbf{Y} . For convenience, we then reshape the upper triangles and diagonals of each timepoint's symmetric correlation matrix into a row vector to obtain an equivalent T by $\binom{K^2-K}{2} + K$ matrix.

Dynamic inter-subject functional connectivity (DISFC)

Equation 3.4 provides a means of taking a single observation matrix, \mathbf{X}_n and estimating the dynamic correlations from moment to moment, \mathbf{Y}_{n+1} . Suppose that one has access to a set of multiple observation matrices that reflect the same phenomenon. For example, one might collect neuroimaging data from several experimental participants, as each participant performs the same task (or sequence of tasks). Let $\mathbf{X}_n^1, \mathbf{X}_n^2, \dots, \mathbf{X}_n^P$ reflect the T by K observation matrices ($n = 0$) or reduced correlation matrices ($n > 0$) for each of P participants in an experiment. We can use inter-subject functional connectivity Simony et al. (2016); Simony and Chang (2020) (ISFC) to compute the stimulus-driven correlations reflected in the multi-participant dataset at a given timepoint t using:

$$\bar{\mathbf{C}}(t) = M \left(R \left(\frac{1}{2P} \sum_{p=1}^P Z(\mathbf{Y}_{n+1}^p(t))^\top + Z(\mathbf{Y}_{n+1}^p(t)) \right) \right), \quad (3.7)$$

where M extracts and vectorizes the upper triangle and diagonal of a symmetric matrix, Z is the Fisher z -transformation Zar (2010):

$$Z(r) = \frac{\log(1+r) - \log(1-r)}{2}, \quad (3.8)$$

R is the inverse of Z :

$$R(z) = \frac{\exp(2z-1)}{\exp(2z+1)}, \quad (3.9)$$

and $\mathbf{Y}_{n+1}^p(t)$ denotes the correlation matrix at timepoint t (Eqn. 3.4) between each column of \mathbf{X}_n^p and each column of the average $\bar{\mathbf{X}}_n$ from all other participants, $\bar{\mathbf{X}}_n^{\setminus p}$:

$$\bar{\mathbf{X}}_n^{\setminus p} = \frac{1}{P-1} \sum_{q \in \setminus p} \mathbf{X}_n^q, \quad (3.10)$$

where $\setminus p$ denotes the set of all participants other than participant p . In this way, the T by $\left(\frac{K^2-K}{2} + K\right)$ DISFC matrix $\bar{\mathbf{C}}$ provides a time-varying extension of the ISFC approach developed by Simony et al. (2016).

Low-dimensional representations of dynamic correlations

Given a T by $\left(\frac{K^2-K}{2} + K\right)$ matrix of n^{th} -order dynamic correlations, \mathbf{Y}_n , we propose two general approaches to computing a T by K low-dimensional representation of those correlations, \mathbf{X}_n . The first approach uses dimensionality reduction algorithms to project \mathbf{Y}_n onto a K -dimensional space. The second approach uses graph measures to characterize the relative positions of each feature ($k \in [1, \dots, K]$) in the network defined by the correlation matrix at each timepoint.

Dimensionality reduction-based approaches to computing \mathbf{X}_n

The modern toolkit of dimensionality reduction algorithms include Principal Components Analysis Pearson (1901) (PCA), Probabilistic PCA Tipping and Bishop (1999) (PPCA), Exploratory Factor Analysis Spearman (1904) (EFA), Independent Components Analysis Jutten and Herault (1991); Comon et al. (1991) (ICA), t -Stochastic Neighbor Embedding van der Maaten and Hinton (2008) (t -SNE), Uniform Manifold Approximation and Projection McInnes et al. (2018) (UMAP), non-negative matrix factorization Lee and Seung (1999) (NMF), Topographic Factor Analysis Manning et al. (2014) (TFA), Hierarchi-

cal Topographic Factor analysis Manning et al. (2018) (HTFA), Topographic Latent Source Analysis Gershman et al. (2011) (TLSA), dictionary learning Mairal et al. (2009b,a), and deep auto-encoders Hinton and Salakhutdinov (2006), among others. While complete characterizations of each of these algorithms is beyond the scope of the present manuscript, the general intuition driving these approaches is to compute the T by K matrix, \mathbf{X} , that is closest to the original T by J matrix, \mathbf{Y} , where (typically) $K \ll J$. The different approaches place different constraints on what properties \mathbf{X} must satisfy and which aspects of the data are compared (and how) in order to optimize how well \mathbf{X} approximates \mathbf{Y} .

Applying dimensionality reduction algorithms to \mathbf{Y} yields an \mathbf{X} whose columns reflect weighted combinations (or nonlinear transformations) of the original columns of \mathbf{Y} . This has two main consequences. First, with each repeated dimensionality reduction, the resulting \mathbf{X}_n has lower and lower fidelity (with respect to what the “true” \mathbf{Y}_n might have looked like without using dimensionality reduction to maintain tractability). In other words, computing \mathbf{X}_n is a lossy operation. Second, whereas each column of \mathbf{Y}_n may be mapped directly onto specific pairs of columns of \mathbf{X}_{n-1} , the columns of \mathbf{X}_n reflect weighted combinations and/or nonlinear transformations of the columns of \mathbf{Y}_n . Many dimensionality reduction algorithms are invertible (or approximately invertible). However, attempting to map a given \mathbf{X}_n back onto the original feature space of \mathbf{X}_0 will usually require $\mathcal{O}(TK^{2^n})$ space and therefore becomes intractable as n or K grow large.

Graph measure approaches to computing \mathbf{X}_n

The above dimensionality reduction approaches to approximating a given \mathbf{Y}_n with a lower-dimensional \mathbf{X}_n preserve a (potentially recombined and transformed) mapping back to the original data in \mathbf{X}_0 . We also explore graph measures that instead characterize each feature’s relative position in the broader network of interactions and connections. To illustrate the distinction between the two general approaches we explore, suppose a network comprises nodes A and B , along with several other nodes. If A and B exhibit uncorrelated activity patterns, then by definition the functional connection (correlation) between them will be close to 0. However, if A and B each interact with other nodes in similar ways, we might attempt to capture those similarities between A ’s and B ’s interactions with those other members of the network.

In general, graph measures take as input a matrix of interactions (e.g., using the above notation, a K by K correlation matrix or binarized correlation matrix reconstituted from a single timepoint’s row of \mathbf{Y}), and return as output a set of K measures describing how each node (feature) sits within that correlation matrix with respect to the rest of the popula-

tion. Widely used measures include betweenness centrality (the proportion of shortest paths between each pair of nodes in the population that involves the given node in question Newman (2005); Opsahl et al. (2010); Barthélemy (2004); Geisberger et al. (2008); Freeman (1977)); diversity and dissimilarity (characterizations of how differently connected a given node is from others in the population Rao (1982); Lin (2009); Ricotta and Szeidl (2006)); eigenvector centrality and pagerank centrality (measures of how influential a given node is within the broader network Newman (2008); Bonacich (2007); Lohmann et al. (2010); Halu et al. (2013)); transfer entropy and flow coefficients (a measure of how much information is flowing from a given node to other nodes in the network Honey et al. (2007); Schreiber (2000)); k -coreness centrality (a measure of the connectivity of a node within its local subgraph Alvarez-Hamelin et al. (2005); Christakis and Fowler (2010)); within-module degree (a measure of how many connections a node has to its close neighbors in the network Rubinov and Sporns (2010)); participation coefficient (a measure of the diversity of a node's connections to different subgraphs in the network Rubinov and Sporns (2010)); and subgraph centrality (a measure of a node's participation in all of the network's subgraphs Estrada and Rodríguez-Velázquez (2005)); among others.

For a given graph measure, $\eta : \mathbb{R}^{K \times K} \rightarrow \mathbb{R}^K$, we can use η to transform each row of \mathbf{Y}_n in a way that characterizes the corresponding graph properties of each column. This results in a new T by K matrix, \mathbf{X}_n , that reflects how the features reflected in the columns of \mathbf{X}_{n-1} participate in the network during each timepoint (row).

Dynamic higher-order correlations

Because \mathbf{X}_n has the same shape as the original data \mathbf{X}_0 , approximating \mathbf{Y}_n with a lower-dimensional \mathbf{X}_n enables us to estimate high-order dynamic correlations in a scalable way. Given a T by K input matrix, the output of Equation 3.4 requires $\mathcal{O}(TK^2)$ space to store. Repeated applications of Equation 3.4 (i.e., computing dynamic correlations between the columns of the outputted dynamic correlation matrix) each require exponentially more space; in general the n^{th} -order dynamic correlations of a T by K timeseries occupies $\mathcal{O}(TK^{2^n})$ space. However, when we approximate or summarize the output of Equation 3.4 with a T by K matrix (as described above), it becomes feasible to compute even very high-order correlations in high-dimensional data. Specifically, approximating the n^{th} -order dynamic correlations of a T by K timeseries requires only $\mathcal{O}(TK^2)$ additional space— the same as would be required to compute first-order dynamic correlations. In other words, the space required to store $n + 1$ multivariate timeseries reflecting up to n^{th} order correlations

in the original data scales linearly with n using our approach (Fig. 3.2).

Data

We examined two types of data: synthetic data and human functional neuroimaging data. We constructed and leveraged the synthetic data to evaluate our general approach Thompson et al. (2018). Specifically, we tested how well Equation 3.4 could be used to recover known dynamic correlations using different choices of kernel (κ ; Fig. 3.3), for each of several synthetic datasets that exhibited different temporal properties. We also simulated higher-order correlations and tested how well Equation 3.4 could recover these correlations using the best kernel from the previous synthetic data analyses. We then applied our approach to a functional neuroimaging dataset to test the hypothesis that ongoing cognitive processing is reflected in high-order dynamic correlations. We used an across-participant classification test to estimate whether dynamic correlations of different orders contain information about which timepoint in a story participants were listening to.

Synthetic data: simulating dynamic first-order correlations

We constructed a total of 400 different multivariate timeseries, collectively reflecting a total of 4 qualitatively different patterns of dynamic first-order correlations (i.e., 100 datasets reflecting each type of dynamic pattern). Each timeseries comprised 50 features (dimensions) that varied over 300 timepoints. The observations at each timepoint were drawn from a zero-mean multivariate Gaussian distribution with a covariance matrix defined for each timepoint as described below. We drew the observations at each timepoint independently from the draws at all other timepoints; in other words, for each observation $s_t \sim \mathcal{N}(\mathbf{0}, \Sigma_t)$ at timepoint t , $p(s_t) = p(s_t | s_{\setminus t})$.

Constant. We generated data with stable underlying correlations to evaluate how Equation 3.4 characterized correlation “dynamics” when the ground truth correlations were static. We constructed 100 multivariate timeseries whose observations were each drawn from a single (stable) Gaussian distribution. For each dataset (indexed by m), we constructed a random covariance matrix, Σ_m :

$$\Sigma_m = \mathbf{C}\mathbf{C}^\top, \text{ where} \quad (3.11)$$

$$\mathbf{C}(i, j) \sim \mathcal{N}(0, 1), \text{ and where} \quad (3.12)$$

$i, j \in [1, 2, \dots, 50]$. In other words, all of the observations (for each of the 300 timepoints) within each dataset were drawn from a multivariate Gaussian distribution with the same covariance matrix, and the 100 datasets each used a different covariance matrix.

Random. We generated a second set of 100 synthetic datasets whose observations at each timepoint were drawn from a Gaussian distribution with a new randomly constructed (using Eqn. 3.11) covariance matrix. Because each timepoint’s covariance matrix was drawn independently from the covariance matrices for all other timepoints, these datasets provided a test of reconstruction accuracy in the absence of any meaningful underlying temporal structure in the dynamic correlations underlying the data.

Ramping. We generated a third set of 100 synthetic datasets whose underlying correlations changed gradually over time. For each dataset, we constructed two “anchor” covariance matrices using Equation 3.11, Σ_{start} and Σ_{end} . For each of the 300 timepoints in each dataset, we drew the observations from a multivariate Gaussian distribution whose covariance matrix at each timepoint $t \in [0, \dots, 299]$ was given by

$$\Sigma_t = \left(1 - \frac{t}{299}\right) \Sigma_{\text{start}} + \frac{t}{299} \Sigma_{\text{end}}. \quad (3.13)$$

The gradually changing correlations underlying these datasets allow us to evaluate the recovery of dynamic correlations when each timepoint’s correlation matrix is unique (as in the random datasets), but where the correlation dynamics are structured and exhibit first-order autocorrelations (as in the constant datasets).

Event. We generated a fourth set of 100 synthetic datasets whose underlying correlation matrices exhibited prolonged intervals of stability, interspersed with abrupt changes. For each dataset, we used Equation 3.11 to generate 5 random covariance matrices. We constructed a timeseries where each set of 60 consecutive samples was drawn from a Gaussian with the same covariance matrix. These datasets were intended to simulate a system that exhibits periods of stability punctuated by occasional abrupt state changes.

Synthetic data: simulating dynamic high-order correlations

We developed an iterative procedure for constructing timeseries data that exhibits known dynamic high-order correlations. The procedure builds on our approach to generating dynamic first-order correlations. Essentially, once we generate a timeseries with known first-

order correlations, we can use the known first-order correlations as a template to generate a new timeseries of second-order correlations. In turn, we can generate a timeseries of third-order correlations from the second-order correlations, and so on. In general, we can generate order n correlations given a timeseries of order $n - 1$ correlations, for any $n > 1$. Finally, given the order n timeseries, we can reverse the preceding process to generate an order $n - 1$ timeseries, an order $n - 2$ order timeseries, and so on, until we obtain an order 0 timeseries of simulated data that reflects the chosen high-order dynamics.

The central mathematical operation in our procedure is the Kronecker product (\otimes). The Kronecker product of a $K \times K$ matrix, m_1 , with itself (i.e., $m_1 \otimes m_1$) produces a new $K^2 \times K^2$ matrix, m_2 whose entries reflect a scaled tiling of the entries in m_1 . If these tilings (scaled copies of m_1) are indexed by row and column, then the tile in the i^{th} row and j^{th} column contains the entries of m_1 , multiplied by $m_1(i, j)$. Following this pattern, the Kronecker product $m_2 \otimes m_2$ yields the $K^4 \times K^4$ matrix m_3 whose tiles are scaled copies of m_2 . In general, repeated applications of the Kronecker self-product may be used to generate $m_{n+1} = m_n \otimes m_n$ for $n > 1$, where m_{n+1} is a $K^{2^n} \times K^{2^n}$ matrix. After generating a first-order timeseries of dynamic correlations (see Synthetic data: simulating dynamic first-order correlations), we use this procedure (applied independently at each timepoint) to transform it into a timeseries of n^{th} -order correlations. When m_{n+1} is generated in this way, the temporal structure of the full timeseries (i.e., constant, random, ramping, event) is preserved, since changes in the original first-order timeseries are also reflected in the scaled tilings of itself that comprise the higher-order matrices.

Given a timeseries of n^{th} -order correlations, we then need to work “backwards” in order to generate the order-zero timeseries. If the n^{th} -order correlation matrix at a given timepoint is m_n , then we can generate an order $n - 1$ correlation matrix (for $n > 1$) by taking a draw from $\mathcal{N}(0, m_n)$ and reshaping the resulting vector to have square dimensions. To force the resulting matrix to be symmetric, we remove its lower triangle, and replace the lower triangle with (a reflected version of) its upper triangle. Intuitively, the re-shaped matrix will look like a noisy (but symmetric) version of the template matrix, m_{n-1} . (When $n = 1$, no re-shaping is needed; the resulting K -dimensional vector may be used as the observation at the given timepoint.) After independently drawing each timepoint’s order $n - 1$ correlation matrix from that timepoint’s order n correlation matrix, this process can be applied repeatedly until $n = 0$. This results in a K -dimensional timeseries of T observations containing the specified high-order correlations at orders 1 through n . Following our approach to generating synthetic data exhibiting known first-order correlations, we constructed a total of 400 additional multivariate timeseries, collectively reflecting a total of 4 qualitatively dif-

ferent patterns of dynamic correlations (i.e., 100 datasets reflecting each type of dynamic pattern: constant, random, ramping, and event). Each timeseries comprised 10 zero-order features (dimensions) that varied over 300 timepoints. After applying our dynamic correlation estimation procedure, this yielded a 100-dimensional timeseries of first-order features that could then be used to estimate dynamic second-order correlations. (We chose to use $K = 10$ zero-order features for our higher order simulations in order to put the accuracy computations displayed in Figs. 3.5 and 3.6 on a roughly even footing.)

Functional neuroimaging data collected during story listening

We examined an fMRI dataset collected by Simony et al. (2016) that the authors have made publicly available at arks.princeton.edu/ark:/88435/dsp015d86p269k. The dataset comprises neuroimaging data collected as participants listened to an audio recording of a story (intact condition; 36 participants), listened to temporally scrambled recordings of the same story (17 participants in the paragraph-scrambled condition listened to the paragraphs in a randomized order and 36 in the word-scrambled condition listened to the words in a randomized order), or lay resting with their eyes open in the scanner (rest condition; 36 participants). Full neuroimaging details may be found in the original paper for which the data were collected Simony et al. (2016). Procedures were approved by the Princeton University Committee on Activities Involving Human Subjects, and by the Western Institutional Review Board (Puyallup, WA). All subjects were native English speakers with normal hearing and provided written informed consent.

Hierarchical topographic factor analysis (HTFA). Following our prior related work, we used HTFA Manning et al. (2018) to derive a compact representation of the neuroimaging data. In brief, this approach approximates the timeseries of voxel activations (44,415 voxels) using a much smaller number of radial basis function (RBF) nodes (in this case, 700 nodes, as determined by an optimization procedure Manning et al. (2018)). This provides a convenient representation for examining full-brain network dynamics. All of the analyses we carried out on the neuroimaging dataset were performed in this lower-dimensional space. In other words, each participant's data matrix, X_0 , was a number-of-timepoints by 700 matrix of HTFA-derived factor weights (where the row and column labels were matched across participants). Code for carrying out HTFA on fMRI data may be found as part of the BrainIAK toolbox Capota et al. (2017), which may be downloaded at brainiak.org.

Temporal decoding

We sought to identify neural patterns that reflected participants' ongoing cognitive processing of incoming stimulus information. As reviewed by Simony et al. (2016), one way of homing in on these stimulus-driven neural patterns is to compare activity patterns across individuals (e.g., using ISFC analyses). In particular, neural patterns will be similar across individuals to the extent that the neural patterns under consideration are stimulus-driven, and to the extent that the corresponding cognitive representations are reflected in similar spatial patterns across people Simony and Chang (2020). Following this logic, we used an across-participant temporal decoding test developed by Manning et al. (2018) to assess the degree to which different neural patterns reflected ongoing stimulus-driven cognitive processing across people (Fig. 3.4). The approach entails using a subset of the data to train a classifier to decode stimulus timepoints (i.e., moments in the story participants listened to) from neural patterns. We use decoding (forward inference) accuracy on held-out data, from held-out participants, as a proxy for the extent to which the inputted neural patterns reflected stimulus-driven cognitive processing in a similar way across individuals.

Forward inference and decoding accuracy

We used an across-participant correlation-based classifier to decode which stimulus timepoint matched each timepoint's neural pattern(Fig. 3.4). We first divided the participants into two groups: a template group, $\mathcal{G}_{\text{template}}$ (i.e., training data), and a to-be-decoded group, $\mathcal{G}_{\text{decode}}$ (i.e., test data). We used Equation 3.7 to compute a DISFC matrix for each group ($\bar{\mathbf{C}}_{\text{template}}$ and $\bar{\mathbf{C}}_{\text{decode}}$, respectively). We then correlated the rows of $\bar{\mathbf{C}}_{\text{template}}$ and $\bar{\mathbf{C}}_{\text{decode}}$ to form a number-of-timepoints by number-of-timepoints decoding matrix, Λ . In this way, the rows of Λ reflected timepoints from the template group, while the columns reflected timepoints from the to-be-decoded group. We used Λ to assign temporal labels to each row $\bar{\mathbf{C}}_{\text{decode}}$ using the row of $\bar{\mathbf{C}}_{\text{template}}$ with which it was most highly correlated. We then repeated this decoding procedure, but using $\mathcal{G}_{\text{decode}}$ as the template group and $\mathcal{G}_{\text{template}}$ as the to-be-decoded group. Given the true timepoint labels (for each group), we defined the decoding accuracy as the average proportion of correctly decoded timepoints, across both groups. We defined the relative decoding accuracy as the difference between the decoding accuracy and chance accuracy (i.e., $\frac{1}{T}$).

Feature weighting and testing

We sought to examine which types of neural features (i.e., activations, first-order dynamic correlations, and higher-order dynamic correlations) were informative to the temporal decoders. Using the notation above, these features correspond to \mathbf{X}_0 , \mathbf{X}_1 , \mathbf{X}_2 , \mathbf{X}_3 , and so on.

One challenge to fairly evaluating high-order correlations is that if the kernel used in Equation 3.4 is wider than a single timepoint, each repeated application of the equation will result in further temporal blur. Because our primary assessment metric is temporal decoding accuracy, this unfairly biases against detecting meaningful signal in higher-order correlations (relative to lower-order correlations). We attempted to mitigate temporal blur in estimating each \mathbf{X}_n by using a Dirac δ function kernel (which places all of its mass over a single timepoint; Fig. 3.3b, 3.4a) to compute each lower-order correlation ($\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_{n-1}$). We then used a new (potentially wider, as described below) kernel to compute \mathbf{X}_n from \mathbf{X}_{n-1} . In this way, temporal blurring was applied only in the last step of computing \mathbf{X}_n . We note that, because each \mathbf{X}_n is a low-dimensional representation of the corresponding \mathbf{Y}_n , the higher-order correlations we estimated reflect true correlations in the data with lower-fidelity than estimates of lower-order correlations. Therefore, even after correcting for temporal blurring, our approach is still biased against finding meaningful signal in higher-order correlations.

After computing each $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_{n-1}$ for each participant, we divided participants into two equally sized groups (± 1 for odd numbers of participants): $\mathcal{G}_{\text{train}}$ and $\mathcal{G}_{\text{test}}$. We then further subdivided $\mathcal{G}_{\text{train}}$ into $\mathcal{G}_{\text{train}_1}$ and $\mathcal{G}_{\text{train}_2}$. We then computed Λ (temporal correlation) matrices for each type of neural feature, using $\mathcal{G}_{\text{train}_1}$ and $\mathcal{G}_{\text{train}_2}$. This resulted in $n + 1$ Λ matrices (one for the original timeseries of neural activations, and one for each of n orders of dynamic correlations). Our objective was to find a set of weights for each of these Λ matrices such that the weighted average of the $n + 1$ matrices yielded the highest decoding accuracy. We used quasi-Newton gradient ascent Nocedal and Wright (2006), using decoding accuracy (for $\mathcal{G}_{\text{train}_1}$ and $\mathcal{G}_{\text{train}_2}$) as the objective function to be maximized, to find an optimal set of training data-derived weights, $\phi_{0,1,\dots,n}$, where $\sum_{i=0}^n \phi_i = 1$ and where $\phi_i \geq 0 \forall i \in [0, 1, \dots, n]$.

After estimating an optimal set of weights, we computed a new set of $n + 1$ Λ matrices correlating the DISFC patterns from $\mathcal{G}_{\text{train}}$ and $\mathcal{G}_{\text{test}}$ at each timepoint. We use the resulting decoding accuracy of $\mathcal{G}_{\text{test}}$ timepoints (using the weights in $\phi_{0,1,\dots,n}$ to average the Λ matrices) to estimate how informative the set of neural features containing up to n^{th} order correlations were.

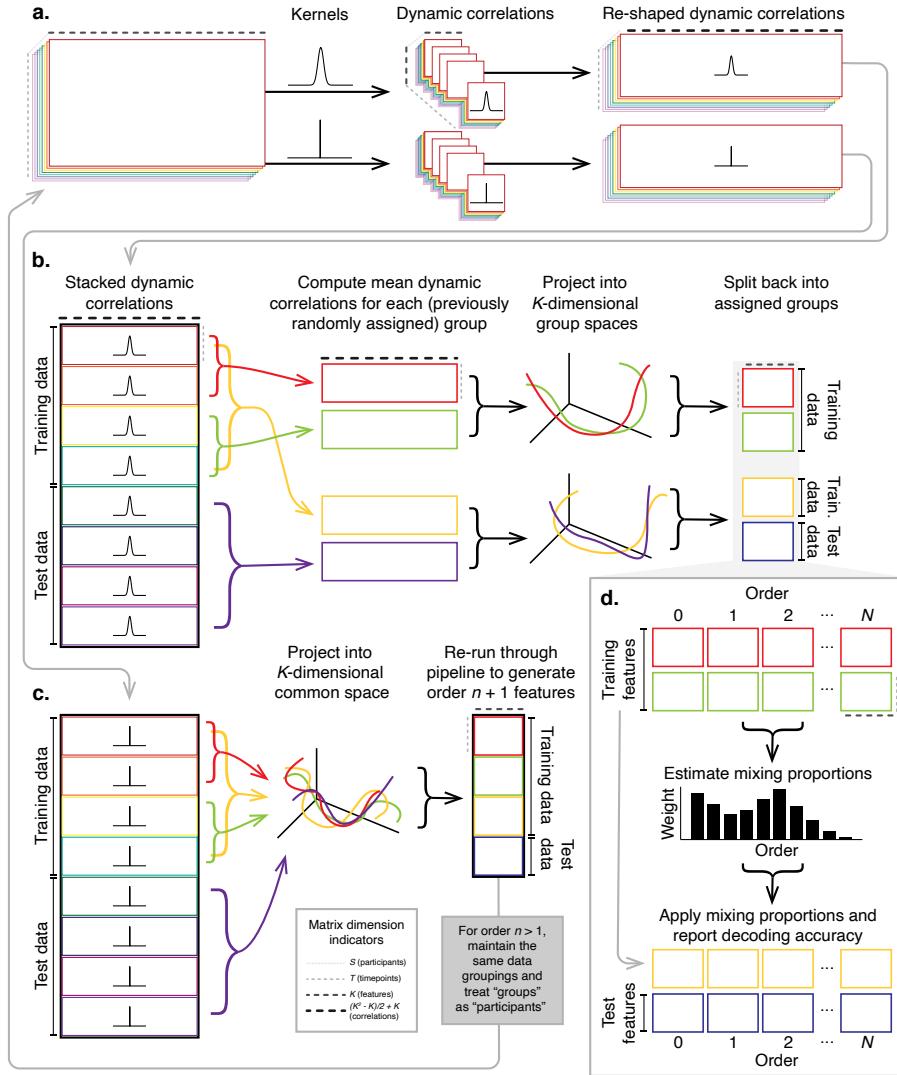


Figure 3.4: Decoding analysis pipeline. **a. Computing dynamic correlations from time-series data.** Given a timeseries of observations as a $T \times K$ matrix (or a set of S such matrices), we use Equation 3.4 to compute each participant's DISFC (relative to other participants in the training or test sub-group, as appropriate). We repeat this process twice—once using the analysis kernel (shown here as a Gaussian in the upper row of the panel), and once using a δ function kernel (lower row of the panel). **b. Projecting dynamic correlations into a lower-dimensional space.** We project the training and test data into K -dimensional spaces to create compact representations of dynamic correlations at the given order (estimated using the analysis kernel). **c. Kernel trick.** We project the dynamic correlations computed using a δ function kernel into a common K -dimensional space. These low-dimensional embeddings are fed back through the analysis pipeline in order to compute features at the next-highest order. **d. Decoding analysis.** We split the training data into two equal groups, and optimize the feature weights (i.e., dynamic correlations at each order) to maximize decoding accuracy. We then apply the trained classifier to the (held-out) test data.

We used a permutation-based procedure to form stable estimates of decoding accuracy for each set of neural features. In particular, we computed the decoding accuracy for each of 10 random group assignments of $\mathcal{G}_{\text{train}}$ and $\mathcal{G}_{\text{test}}$. We report the mean accuracy (along with 95% confidence intervals) for each set of neural features.

Identifying robust decoding results

The temporal decoding procedure we use to estimate which neural features support ongoing cognitive processing is governed by several parameters. In particular, Equation 3.4 requires defining a kernel function, which can take on different shapes and widths. For a fixed set of neural features, each of these parameters can yield different decoding accuracies. Further, the best decoding accuracy for a given timepoint may be reliably achieved by one set of parameters, whereas the best decoding accuracy for another timepoint might be reliably achieved by a different set of parameters, and the best decoding accuracy across all timepoints might be reliably achieved by still another different set of parameters. Rather than attempting to maximize decoding accuracy, we sought to discover the trends in the data that were robust to classifier parameters choices. Specifically, we sought to characterize how decoding accuracy varied (under different experimental conditions) as a function of which neural features were considered.

To identify decoding results that were robust to specific classifier parameter choices, we repeated our decoding analyses after substituting into Equation 3.4 each of a variety of kernel shapes and widths. We examined Gaussian (Fig. 3.3c), Laplace (Fig. 3.3d), and Mexican Hat (Fig. 3.3e) kernels, each with widths of 5, 10, 20, and 50 samples. We then report the average decoding accuracies across all of these parameter choices. This enabled us to (partially) factor out performance characteristics that were parameter-dependent, within the set of parameters we examined.

Reverse inference

The dynamic patterns we examined comprise high-dimensional correlation patterns at each timepoint. To help interpret the resulting patterns in the context of other studies, we created summary maps by computing the across-timepoint average pairwise correlations at each order of analysis (first order, second order, etc.). We selected the 10 strongest (absolute value) correlations at each order. Each correlation is between the dynamic activity patterns (or patterns of dynamic high-order correlations) measured at two RBF nodes (see Hierarchical Topographic Factor Analysis). Therefore, the 10 strongest correlations in-

volved up to 20 RBF nodes. Each RBF defines a spatial function whose activations range from 0 to 1. We constructed a map of RBF components that denoted the endpoints of the 10 strongest correlations (we set each RBF to have a maximum value of 1). We then carried out a meta analysis using Neurosynth Rubin et al. (2017) to identify the 10 terms most commonly associated with the given map. This resulted in a set of 10 terms associated with the average dynamic correlation patterns at each order.

3.4 Results

We sought to understand whether high-level cognition is reflected in dynamic patterns of high-order correlations. To that end, we developed a computational framework for estimating the dynamics of stimulus-driven high-order correlations in multivariate timeseries data (see Dynamic inter-subject functional connectivity (DISFC) and Dynamic higher-order correlations). We evaluated the efficacy of this framework at recovering known patterns in several synthetic datasets (see Synthetic data: simulating dynamic first-order correlations and Synthetic data: simulating dynamic higher-order correlations). We then applied the framework to a public fMRI dataset collected as participants listened to an auditorily presented story, listened to a temporally scrambled version of the story, or underwent a resting state scan (see Functional neuroimaging data collected during story listening). We used the relative decoding accuracies of classifiers trained on different sets of neural features to estimate which types of features reflected ongoing cognitive processing.

Recovering known dynamic first-order correlations

We generated synthetic datasets that differed in how the underlying first-order correlations changed over time. For each dataset, we applied Equation 3.4 with a variety of kernel shapes and widths. We assessed how well the true underlying correlations at each timepoint matched the recovered correlations (Fig. 3.5). For every kernel and dataset we tested, our approach recovered the correlation dynamics we embedded into the data. However, the quality of these recoveries varied across different synthetic datasets in a kernel-dependent way.

In general, wide monotonic kernel shapes (Laplace, Gaussian), and wider kernels (within a shape), performed best when the correlations varied gradually from moment-to-moment (Figs. 3.5a, c, and d). In the extreme, as the rate of change in correlations approaches 0 (Fig. 3.5a), an infinitely wide kernel would exactly recover the Pearson’s

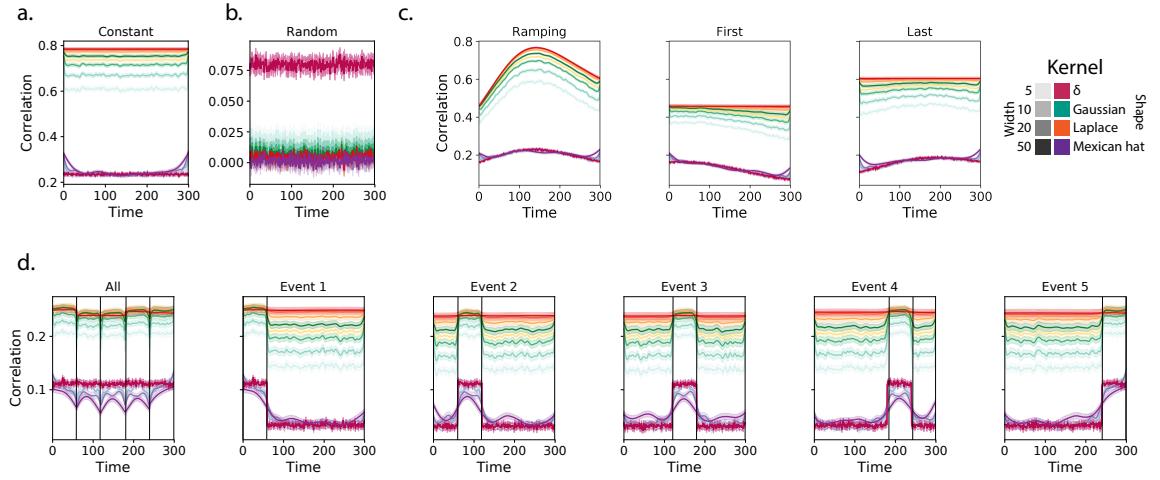


Figure 3.5: Recovering known dynamic first-order correlations from synthetic data.

Each panel displays the average correlations between the vectorized upper triangles of the recovered correlation matrix at each timepoint and either the true underlying correlation at each timepoint or a reference correlation matrix. (The averages are taken across 100 different randomly generated synthetic datasets of each given category, each with $K = 50$ features and $T = 300$ timepoints.) Error ribbons denote 95% confidence intervals of the mean (taken across datasets). Different colors denote different kernel shapes, and the shading within each color family denotes the kernel width parameter. For a complete description of each synthetic dataset, see Synthetic data: simulating dynamic first-order correlations. **a. Constant correlations.** These datasets have a stable (unchanging) underlying correlation matrix. **b. Random correlations.** These datasets are generated using a new independently drawn correlation matrix at each new timepoint. **c. Ramping correlations.** These datasets are generated by smoothly varying the underlying correlations between the randomly drawn correlation matrices at the first and last timepoints. The left panel displays the correlations between the recovered dynamic correlations and the underlying ground truth correlations. The middle panel compares the recovered correlations with the first timepoint's correlation matrix. The right panel compares the recovered correlations with the last timepoint's correlation matrix. **d. Event-based correlations.** These datasets are each generated using five randomly drawn correlation matrices that each remain stable for a fifth of the total timecourse. The left panel displays the correlations between the recovered dynamic correlations and the underlying ground truth correlations. The right panels compare the recovered correlations with the correlation matrices unique to each event. The vertical lines denote event boundaries. Source data are provided as a Source Data file.

correlation (e.g., compare Eqns. 3.1 and 3.4).

When the correlation dynamics were unstructured in time (Fig. 3.5b), a Dirac δ kernel (infinitely narrow) performed best. This is because, when every timepoint’s correlations are independent of the correlations at every other timepoint, averaging data over time dilutes the available signal. Following a similar pattern, holding kernel shape fixed, narrower kernel parameters better recovered randomly varying correlations.

Recovering known dynamic higher-order correlations

Following our approach to evaluating our ability to recover known dynamic first-order correlations from synthetic data, we generated an analogous second set of synthetic datasets that we designed to exhibit known dynamic first-order and second-order correlations (see Synthetic data: simulating dynamic higher-order correlations). We generated a total of 400 datasets (100 datasets for each category) that varied in how the first-order and second-order correlations changed over time. We then repeatedly applied Equation 3.4 using the overall best-performing kernel from our first-order tests (a Laplace kernel with a width of 20; Fig. 3.5) to assess how closely the recovered dynamic correlations matched the dynamic correlations we had embedded into the datasets.

Overall, we found that we could reliably recover both first-order and second-order correlations from the synthetic data (Fig. 3.6). When the correlations were stable for longer intervals, or changed gradually (constant, ramping, and event datasets), recovery performance was relatively high, and we were better able to recover dynamic first-order correlations than second-order correlations. This is because errors in our estimation procedure at lower orders necessarily propagate to higher orders (since lower-order correlations are used to estimate higher-order correlations). Conversely, when the correlations were particularly unstable (random datasets), we better recovered second-order correlations. This is because noise in our data generation procedure propagates from higher orders to lower orders (see Synthetic data: simulating dynamic high-order correlations).

We also examined the impact of the data duration (Fig. S3) and complexity (number of zero-order features; Fig. S4) on our ability to accurately recover ground truth first-order and second-order dynamic correlations. In general, we found that our approach better recovers ground truth dynamic correlations from longer duration timeseries data. We also found that our approach tends to best recover data generated using fewer zero-order features (i.e., lower complexity), although this tendency was not strictly monotonic. Further, because our data generation procedure requires $\mathcal{O}(K^4)$ memory to generate a second-order

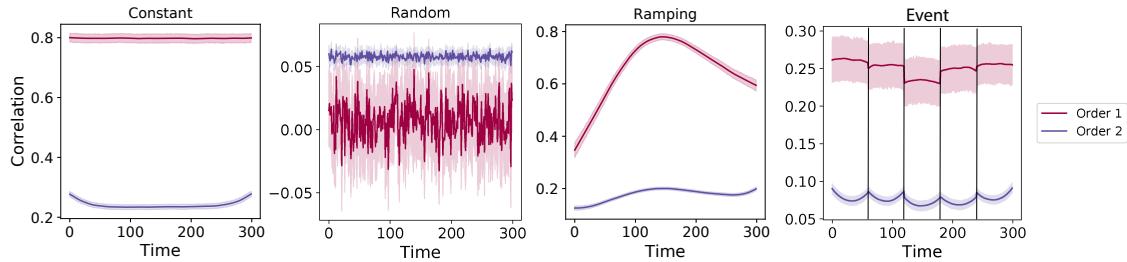


Figure 3.6: Recovery of simulated first-order and second-order dynamic correlations. Each panel displays the average correlations between the vectorized upper triangles of the recovered first-order and second-order correlation matrices and the true (simulated) first-order and second order correlation matrices at each timepoint and for each synthetic dataset. (The averages are taken across 100 different randomly generated synthetic datasets of each given category, each with $K = 10$ features and $T = 300$ timepoints.) Error ribbons denote 95% confidence intervals of the mean (taken across datasets). For a complete description of each synthetic dataset, see Synthetic data: simulating dynamic higher-order correlations. All estimates represented in this figure were computed using a Laplace kernel (width = 20). **Constant.** These datasets have stable (unchanging) underlying second-order correlation matrices. **Random.** These datasets are generated using a new independently drawn second-order correlation matrix at each timepoint. **Ramping.** These datasets are generated by smoothly varying the underlying second-order correlations between the randomly drawn correlation matrices at the first and last timepoints. **Event.** These datasets are each generated using five randomly drawn second-order correlation matrices that each remain stable for a fifth of the total timecourse. The vertical lines denote event boundaries. Note that the “dips” and “ramps” at the boundaries of sharp transitions (e.g., the beginning and ends of the “constant” and “ramping” datasets, and at the event boundaries of the “event” datasets) are finite-sample effects that reflect the reduced numbers of samples that may be used to accurately estimate correlations at sharp boundaries. Source data are provided as a Source Data file.

timeseries with K zero-order features, we were not able to fully explore how the number of zero-order features affects recovery accuracy as the number of features gets larger (e.g., as it approaches the number of features present in the fMRI data we examine below). Although we were not able to formally test this to our satisfaction, we expect that accurately estimating dynamic high-order correlations would require data with many more zero-order features than we were able to simulate. Our reasoning is that high-order correlations necessarily involve larger numbers of lower-order features, so achieving adequate “resolution” high-order timeseries might require many low-order features.

Taken together, our explorations using synthetic data indicated that we are able to partially, but not perfectly, recover ground truth dynamic first-order and second-order correlations. This suggests that our modeling approach provides a meaningful (if noisy) estimate of high-order correlations. We next turned to analyses of human fMRI data to examine whether the recovered dynamics might reflect the dynamics of human cognition during a naturalistic story-listening task.

Cognitively relevant dynamic high-order correlations in fMRI data

We used across-participant temporal decoders to identify cognitively relevant neural patterns in fMRI data (see Forward inference and decoding accuracy). The dataset we examined Simony et al. (2016) comprised four experimental conditions that exposed participants to stimuli that varied systematically in how cognitively engaging they were. The intact experimental condition (intact) had participants listen to an audio recording of a 10-minute story. The paragraph-scrambled experimental condition (paragraph) had participants listen to a temporally scrambled version of the story, where the paragraphs occurred out of order (but where the same total set of paragraphs were presented over the full listening interval). All participants in this condition experienced the scrambled paragraphs in the same order. The word-scrambled experimental condition (word) had participants listen to a temporally scrambled version of the story where the words in the story occurred in a random order. All participants in the word condition experienced the scrambled words in the same order. Finally, in a rest experimental condition (rest), participants lay in the scanner with no overt stimulus, with their eyes open (blinking as needed). This public dataset provided a convenient means of testing our hypothesis that different levels of cognitive processing and engagement are reflected in different orders of brain activity dynamics.

In brief, we computed timeseries of dynamic high-order correlations that were similar across participants in each of two randomly assigned groups: a training group and a test

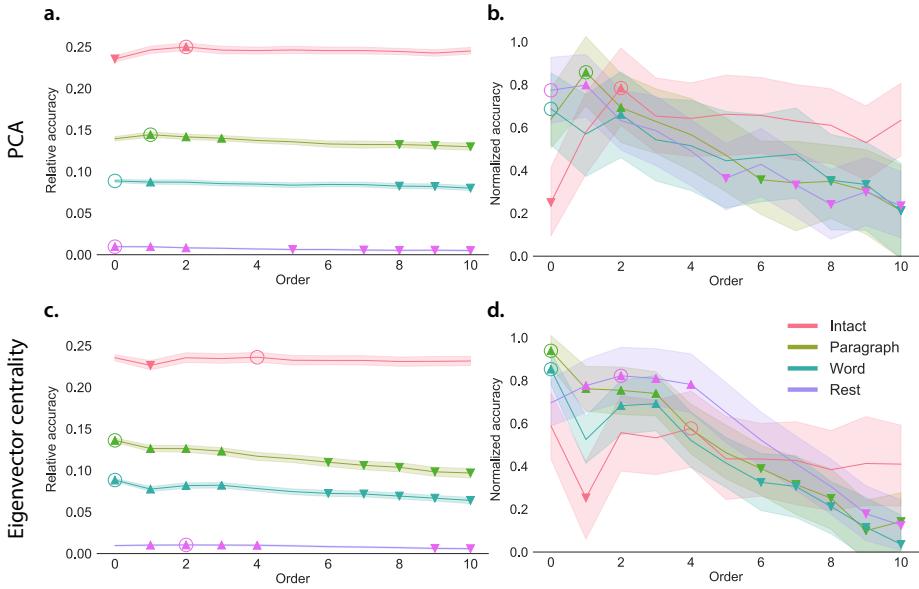


Figure 3.7: Across-participant timepoint decoding accuracy varies with correlation order and cognitive engagement. **a. Decoding accuracy as a function of order: PCA.** “Order” (*x*-axis) refers to the maximum order of dynamic correlations that were available to the classifiers (see Feature weighting and testing). The reported across-participant decoding accuracies are averaged over all kernel shapes and widths (see Identifying robust decoding results). The *y*-values are displayed relative to chance accuracy (intact: $\frac{1}{300}$; paragraph: $\frac{1}{272}$; word: $\frac{1}{300}$; rest: $\frac{1}{400}$; these chance accuracies were subtracted from the observed accuracies to obtain the relative accuracies reported on the *y*-axis). The error ribbons denote 95% confidence intervals of the means across cross-validation folds (i.e., random assignments of participants to the training and test sets). The colors denote the experimental condition. Arrows denote sets of features that yielded reliably higher (upward facing) or lower (downward facing) decoding accuracy than the mean of all other features (via a two-tailed *t*-test, thresholded at $p < 0.05$). Figure 3.8 displays additional comparisons between the decoding accuracies achieved using different sets of neural features. The circled values represent the maximum decoding accuracy within each experimental condition. **b. Normalized timepoint decoding accuracy as a function of order: PCA.** This panel displays the same results as Panel a, but here each curve has been normalized to have a maximum value of 1 and a minimum value of 0 (including the upper and lower bounds of the respective 95% confidence intervals of the mean). Panels a and b used PCA to project each high-dimensional pattern of dynamic correlations onto a lower-dimensional space. **c. Timepoint decoding accuracy as a function of order: eigenvector centrality.** This panel is in the same format as Panel a, but here eigenvector centrality has been used to project the high-dimensional patterns of dynamic correlations onto a lower-dimensional space. **d. Normalized timepoint decoding accuracy as a function of order: eigenvector centrality.** This panel is in the same format as Panel b, but here eigenvector centrality has been used to project the high-dimensional patterns of dynamic correlations onto a lower-dimensional space. See Figures S1 and S2 for decoding results broken down by kernel shape and width, respectively. Source data are provided as a Source Data file.

group. We then trained classifiers on the training group’s data to match each sample from the test group with a stimulus timepoint. Each classifier comprised a weighted blend of neural patterns that reflected up to n^{th} -order dynamic correlations (see Feature weighting and testing). We repeated this process for $n \in \{0, 1, 2, \dots, 10\}$. Our examinations of synthetic data suggested that none of the kernels we examined were “universal” in the sense of optimally recovering underlying correlations regardless of the temporal structure of those correlations. We found a similar pattern in the (real) fMRI data, whereby different kernels yielded different decoding accuracies, but no single kernel emerged as the clear “best.” In our analyses of neural data, we therefore averaged our decoding results over a variety of kernel shapes and widths in order to identify results that were robust to specific kernel parameters (see Identifying robust decoding results).

Our approach to estimating dynamic high-order correlations entails mapping the high-dimensional feature space of correlations (represented by a T by $\mathcal{O}(K^2)$ matrix) onto a lower-dimensional feature space (represented by a T by K matrix). We carried out two sets of analyses that differed in how this mapping was computed. The first set of analyses used PCA to find a low-dimensional embedding of the original dynamic correlation matrices (Fig. 3.7a,b). The second set of analyses characterized correlations in dynamics of each feature’s eigenvector centrality, but did not preserve the underlying activity dynamics (Fig. 3.7c,d).

Both sets of temporal decoding analyses yielded qualitatively similar results for the auditory (non-rest) conditions of the experiment (Fig. 3.7: pink, green, and teal lines; Fig. 3.8: three leftmost columns). The highest decoding accuracy for participants who listened to the intact (unscrambled) story was achieved using high-order dynamic correlations (PCA: second-order; eigenvector-centrality: fourth-order). Scrambled versions of the story were best decoded by lower-order correlations (PCA/paragraph: first-order; PCA/word: order zero; eigenvector centrality/paragraph: order zero; eigenvector centrality/word: order zero). The two sets of analyses yielded different decoding results on resting state data (Fig. 3.7: purple lines; Fig. 3.8: rightmost column). We note that, while the resting state times could be decoded reliably, the accuracies were only very slightly above chance. We speculate that the decoders might have picked up on attentional drift, boredom, or tiredness; we hypothesize that these all increased throughout the resting state scan. The decoders might be picking up on aspects of these loosely defined cognitive states that are common across individuals. The PCA-based approach achieved the highest resting state decoding accuracy using order zero features (non-correlational, activation-based), whereas the eigenvector centrality-based approach achieved the highest resting state decoding accuracy using

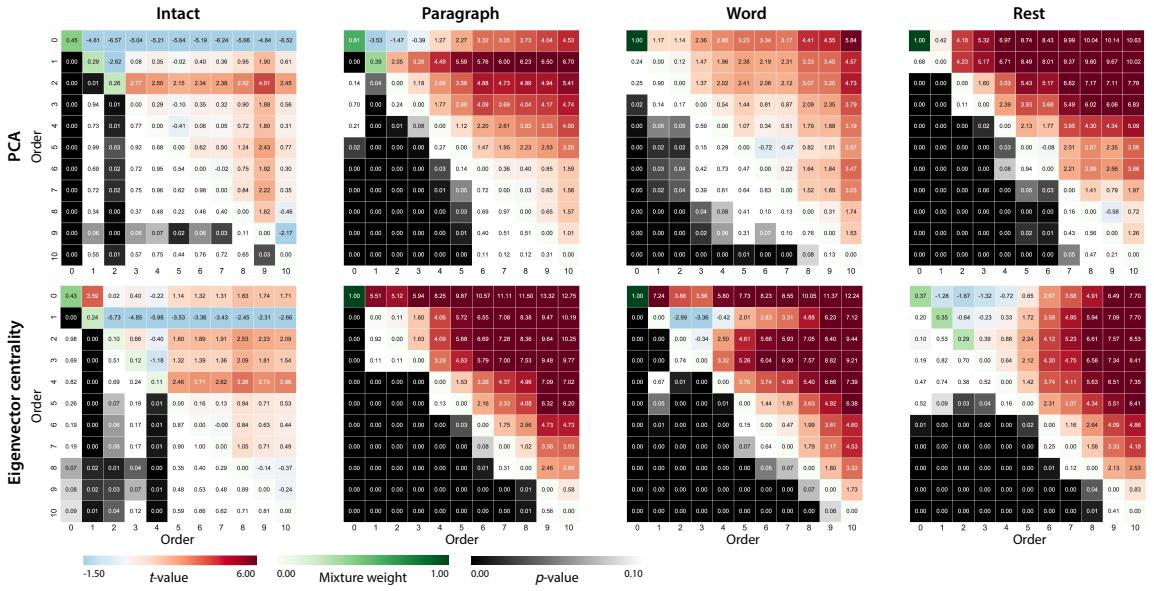


Figure 3.8: Statistical summary of decoding accuracies for different neural features.

Each column of matrices displays decoding results for one experimental condition (intact, paragraph, word, and rest). We considered dynamic activity patterns (order 0) and dynamic correlations at different orders (order > 0). We used two-tailed t -tests to compare the distributions of decoding accuracies obtained using each pair of features. The distributions for each feature reflect the set of average decoding accuracies (across all kernel parameters), obtained for each random assignment of training and test groups. In the upper triangles of each matrix, warmer colors (positive t -values) indicate that the neural feature indicated in the given row yielded higher accuracy than the feature indicated in the given column. Cooler colors (negative t -values) indicate that the feature in the given row yielded lower decoding accuracy than the feature in the given column. The lower triangles of each map denote the corresponding p -values for the t -tests. The diagonal entries display the relative average optimized weight given to each type of feature in a decoder that included all feature types (see Feature weighting and testing). Source data are provided as a Source Data file.

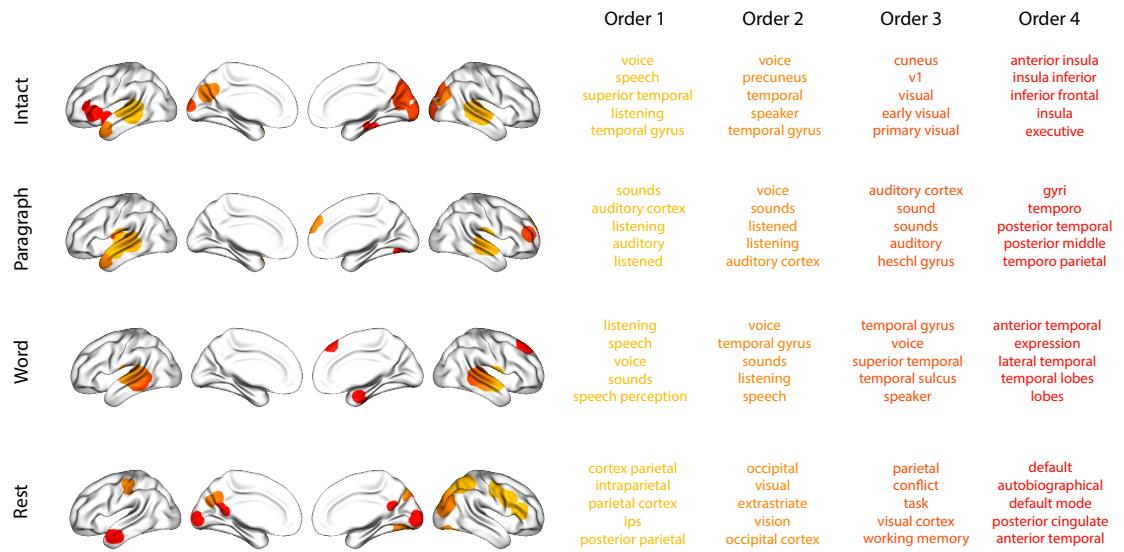


Figure 3.9: Top terms associated with the most strongly correlated nodes at each order. Each color corresponds to one order of inter-subject functional correlations. To calculate the dynamic correlations, eigenvector centrality has been used to project the high-dimensional patterns of dynamic correlations onto a lower-dimensional space at each previous order, which allows us to map the brain regions at each order by retaining the features of the original space. The inflated brain plots display the locations of the endpoints of the 10 strongest (absolute value) correlations at each order, thresholded at 0.999, and projected onto the cortical surface Combrisson et al. (2019). The lists of terms on the right display the top five Neurosynth terms Rubin et al. (2017) decoded from the corresponding brain maps for each order. Each row displays data from a different experimental condition. Additional maps and their corresponding Neurosynth terms may be found in the Supplementary materials (intact: Fig. S5; paragraph: Fig. S6; word: Fig. S7; rest: Fig. S8). Source data are provided as a Source Data file.

second-order correlations. Taken together, these analyses indicate that high-level cognitive processing (while listening to the intact story) is reflected in the dynamics of high-order correlations in brain activity, whereas lower-level cognitive processing (while listening to scrambled versions of the story that lack rich meaning) is reflected in the dynamics of lower-order correlations and non-correlational activity dynamics. Further, these patterns are associated both with the underlying activity patterns (characterized using PCA) and also with the changing relative positions that different brain areas occupy in their associated networks (characterized using eigenvector centrality).

Having established that patterns of high-order correlations are informative to decoders, we next wondered which specific networks of brain regions contributed most to these pat-

terns. As a representative example, we selected the kernel parameters that yielded decoding accuracies that were the most strongly correlated (across conditions and orders) with the average accuracies across all of the kernel parameters we examined. Using Figure 3.7c as a template, the best-matching kernel was a Laplace kernel with a width of 50 (Fig. 3.3d; also see Fig. S9). We used this kernel to compute a single K by K n^{th} -order DISFC matrix for each experimental condition. We then used Neurosynth Rubin et al. (2017) to compute the terms most highly associated with the most strongly correlated pairs of regions in each of these matrices (Fig. 3.9; see Reverse inference).

For all of the story listening conditions (intact, paragraph, and word; top three rows of Fig. 3.9), we found that first- and second-order correlations were most strongly associated with auditory and speech processing areas. During intact story listening, third-order correlations reflected integration with visual areas, and fourth-order correlations reflected integration with areas associated with high-level cognition and cognitive control, such as the ventrolateral prefrontal cortex. However, when participants listened to temporally scrambled stories, these higher-order correlations instead involved interactions with additional regions associated with speech and semantic processing (second and third rows of Fig. 3.9). By contrast, we found a much different set of patterns in the resting state data (Fig. 3.9, bottom row). First-order resting state correlations were most strongly associated with regions involved in counting and numerical understanding. Second-order resting state correlations were strongest in visual areas; third-order correlations were strongest in task-positive areas; and fourth-order correlations were strongest in regions associated with autobiographical and episodic memory. We carried out analogous analyses to create maps (and decode the top associated Neurosynth terms) for up to fifteenth-order correlations (Figs. S5, S6, S7, and S8). Of note, examining fifteenth-order correlations between 700 nodes using conventional methods would have required storing roughly $\frac{700^{2 \times 15}}{2} \approx 1.13 \times 10^{85}$ floating point numbers—assuming single-precision (32 bits each), this would require roughly 32 times as many bits as there are molecules in the known universe! Although these fifteenth-order correlations do appear (visually) to have some well-formed structure, we provide this latter example primarily as a demonstration of the efficiency and scalability of our approach.

3.5 Discussion

We tested the hypothesis that high-level cognition is reflected in high-order brain network dynamics Solomon et al. (2019); Reimann et al. (2017). We examined high-order network dynamics in functional neuroimaging data collected during a story listening experiment.

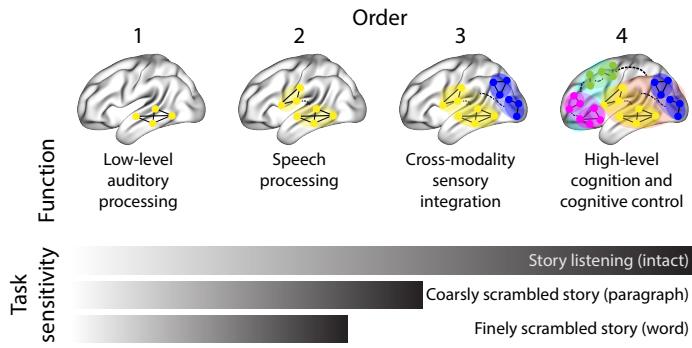


Figure 3.10: Proposed high-order network dynamics underlying high-level cognition during story listening. Schematic depicts higher orders of network interactions supporting higher-level aspects of cognitive processing. When tasks evoke richer, deeper, and/or higher-level processing, this is reflected in higher-order network interactions.

When participants listened to an auditory recording of the story, participants exhibited similar high-order brain network dynamics. By contrast, when participants instead listened to temporally scrambled recordings of the story, only lower-order brain network dynamics were similar across participants. Our results indicate that higher orders of network interactions support higher-level aspects of cognitive processing (Fig. 3.10).

The notion that cognition is reflected in (and possibly mediated by) patterns of first-order network dynamics has been suggested by or proposed in myriad empirical studies and reviews Demertzi et al. (2019); Turk-Browne (2013); Lurie et al. (2018); Fong et al. (2019); Park et al. (2018); Preti et al. (2017); Manning et al. (2018); Roy et al. (2019); Liégeois et al. (2019); Zou et al. (2019); Chang and Glover (2010); Gonzalez-Castillo et al. (2019); McIntosh (2000); Bressler and Kelso (2001). Our study extends this line of work by finding cognitively relevant higher-order network dynamics that reflect ongoing cognition. Our findings also complement other work that uses graph theory and topology to characterize how brain networks reconfigure during cognition Bassett et al. (2006); Zheng et al. (2019); McIntosh and Jirsa (2019); Toker and Sommer (2019); Sizemore et al. (2018); Reimann et al. (2017); Betzel et al. (2019).

An open question not addressed by our study pertains to how different structures integrate incoming information with different time constants. For example, one line of work suggests that the cortical surface comprises a structured map such that nearby brain structures process incoming information at similar timescales. Low-level sensory areas integrate information relatively quickly, whereas higher-level regions integrate information relatively slowly Baldassano et al. (2017); Hasson et al. (2008, 2015); Honey et al. (2012); Lerner et al. (2011, 2014); Chien and Honey (2019). A similar hierarchy appears to play a role in

predicting future events Lee et al. (2020). Other related work in human and mouse brains indicates that the temporal response profile of a given brain structure may relate to how strongly connected that structure is with other brain areas Fallon et al. (2020). Further study is needed to understand the role of temporal integration at different scales of network interaction, and across different anatomical structures. Importantly, our analyses do not speak to the physiological basis of higher-order dynamics, and could reflect nonlinearities, chaotic patterns, non-stationarities, and/or multistability, etc. However, our decoding analyses do indicate that higher-order dynamics are consistent across individuals, and therefore unlikely to reflect non-stimulus-driven dynamics that are unlikely to be similar across individuals.

One limitation of our approach relates to how noise propagates in our estimation procedure. Specifically, our procedure for estimating high-order dynamic correlations depends on estimates of lower-order dynamic correlations. This means that our measures of which higher-order patterns are reliable and stable across experimental conditions are partially confounded with the stability of lower-order patterns. Prior work suggests that the stability of what we refer to here as first-order dynamics likely varies across the experimental conditions we examined Simony et al. (2016). Therefore a caveat to our claim that richer stimuli evoke more stable higher-order dynamics is that our approach assumes that those high-order dynamics reflect relations or interactions between lower-order features.

Another potential limitation of our approach relates to recent work suggesting that the brain undergoes rapid state changes, for example across event boundaries Baldassano et al. (2017). Shappell et al. (2019) used hidden semi-Markov models to estimate state-specific network dynamics Vidaurre et al. (2018). Our general approach might be extended by considering putative state transitions. For example, rather than weighting all timepoints using a similar kernel (Eqn. 3.4), the kernel function could adapt on a timepoint-by-timepoint basis such that only timepoints determined to be in the same “state” were given non-zero weight.

Identifying high-order network dynamics associated with high-level cognition required several important methods advances. First, we used kernel-based dynamic correlations to extended the notion of (static) inter-subject functional connectivity Simony et al. (2016) to a dynamic measure of inter-subject functional connectivity (DISFC) that does not rely on sliding windows Manning et al. (2018), and that may be computed at individual timepoints. This allowed us to precisely characterize stimulus-evoked network dynamics that were similar across individuals. Second, we developed a computational framework for efficiently and scalably estimating high-order dynamic correlations. Our approach uses

dimensionality reduction algorithms and graph measures to obtain low-dimensional embeddings of patterns of network dynamics. Third, we developed an analysis framework for identifying robust decoding results by carrying out our analyses using a range of parameter values and identifying which results were robust to specific parameter choices. By showing that high-level cognition is reflected in high-order network dynamics, we have elucidated the next step on the path towards understanding the neural basis of cognition.

CHAPTER 4

Discussion

4.0.1 Summary

Modeling something as inherently distributed as complex thought holds the potential to expand our understanding of this multi-tiered organization in the brain as well as the properties of brain activity at each of those levels. The content of this thesis has focused on modeling fast-timescale whole-brain networks, with the goal of better understanding these complex systems. The preceding chapters describe the novel methods and computational tools I have helped develop, demonstrate their applications, and expound on the different properties of brain activity they capture. Here, I will summarize key takeaways from these chapters and highlight the potential for ongoing work.

4.0.2 High-resolution spatiotemporal recordings

To better understand complex cognition, we wish to mathematically formalize the fast and dynamic interactions in the brain. However, one of the many limitations in our pursuit in this type of modeling is the resolution of our data. Chapter 2 resolves the discrepancy of high temporal and high spatial resolutions, by leveraging the redundancy of information across people, across tasks, and in an individual’s brain.

The results described in Chapter 2 explore how much can be inferred about activity patterns throughout the brain given recordings from a limited set of brain regions. Our method relies on electrocorticographic (ECoG) recordings from electrodes that are implanted in the brains of patients with drug-resistant epilepsy. The approach uses a machine learning technique, Gaussian Process regression (Rasmussen, 2006), which leverages data from different patients’ brains to build up a model of how activity in different brain regions is related. We then use these relations to infer activity patterns throughout the entire brain of each patient at millimeter-scale spatial resolutions and millisecond-scale temporal resolutions.

We tested and validated our method using two of the largest ECoG dataset available. Moreover, in a followup study led by Scangos et al. (2021), we used our method to create whole brain reconstructions to explore biomarkers of depression. Not only did this work test and validate our method once again, and on yet another ECoG dataset, this study showed that these predictions can also provide diagnostic insights relevant to mental health, such as anxiety and depression.

Our work has found that activity throughout a person’s brain can be inferred using recordings from a limited number of locations from that person’s brain in conjunction with recordings from other people’s brains. These findings have deep implications for the structure of brain data. The first implication is that the correlational structure of different people’s brain data is largely preserved across individuals and across tasks. Additionally, insights about inter-dependency, or redundancy, in brain data have been elucidated by this work. Redundancy has, in fact, been *harnessed* to make these predictions. Specifically, there must be some degree of predictability in activity patterns are across the brain, across people, and across tasks, in order for these predictions to work.

4.0.3 Higher-order interactions

In order to better understand the dynamic interactions between brain structures that underlie our thoughts, the tools and data we use must reflect the complexity of this system. Chapter 3 describes the development and application of a model that explores dynamic higher order correlations in brain data collected using a naturalistic paradigm at varying levels of engagement. This chapter is centered around modeling the high temporal resolution network dynamics of complex thought.

Our method estimates high-order dynamic correlations in timeseries data in a computationally tractable way. Using this model, we examine which aspects of brain activity, and at which order of complexity, underlie human cognition. Specifically, we applied our approach in Chapter 3 to the neuroimaging dataset described in Chapter 1.1.3.2. Data was collected as human participants either listen to a ten-minute story or listen to a temporally scrambled version of the story. By training across-participant pattern classifiers to decode (in held-out data) when in the session each neural activity snapshot was collected, we could better understand the complexity of neural activity patterns that were informative, and therefore important for higher order cognition.

We report that classifiers trained to decode from high-order dynamic correlations yielded the best performance on data collected as participants listened to the (unscram-

bled) story. By contrast, classifiers trained to decode data from scrambled versions of the story yielded the best performance when they were trained using first-order dynamic correlations or non-correlational activity patterns. The results of this study suggest that more engaging tasks may be represented by higher-order network dynamics than less engaging tasks.

This work has examined how interactivity of brain patterns in fMRI data varies as a function of the cognitive richness of the stimulus. With this project, we explored whether high-order dynamic interactions between brain structures support our thoughts. We found that more cognitively rich experiences evoke more higher-order network dynamics

4.0.4 Future directions

Bridging these two lines of research, one future direction will be to explore whole-brain, high-spatial resolution networks using these models. By doing this, we could gain insights into neural patterns at high spatiotemporal resolution by exploring functional connectivity and tracking information flow throughout the brain. However, to fully utilize the whole-brain, high-spatial resolution network models, there are several avenues we could explore to improve the whole-brain ECoG predictions. More broadly, improvements to the predictions could provide an opportunity for applying the most exciting timeseries and network dynamic analysis tools being developed in the fMRI literature to much faster timescale ECoG data.

Currently, our method to predict whole-brain ECoG activity has several underlying assumptions that can be improved. Our model assumes uniform spatial blurring with a stable covariance across time and people, and no time lags. However, work from our lab and others suggest that functional networks are not only dynamic (e.g., Manning et al., 2018; Owen et al., 2021), they are person-specific (e.g., Finn et al., 2015), and task-specific (e.g., Turk-Browne, 2013). There are several ways we can extend our approach to make it more flexible and potentially improve these ECoG predictions.

One way to refine the covariance model would be to incorporate functional and structural MRI data to estimate Bayesian priors over the covariance matrices inferred. Our covariance model currently assumes that in the absence of other information, nearby brain locations will behave similarly, and that this similarity falls off as distance increases. Using resting-state connectivity data acquired from The Human Connectome Project (Van Essen et al., 2013), we could functionally estimate correlated brain regions. Additionally, we could constrain our model based on anatomical connections predicted using Diffusion Ten-

sor Imaging. Furthermore, we could improve these prediction more precisely for individual patients by adding specific information from different recording sources about a single patient, like their own DTI data, and by expanding the model to incorporate other non-invasive measures such as EEG.

Another underlying assumption of our current covariance model is stability across time. To improve our predictions, we could develop our model to allow us to quantify dynamic connectivity and network synchrony at a millisecond timescale. One potential refinement to our static model would be to allow for time lags in correlation values relative to white matter path length, which could also be derived from DTI data.

With our model, we could also investigate electrode locations that are particularly informative to these estimates. Our current validation technique uses a jack-knife approach in which we hold out recordings for each electrode and compare the actual activity to the predicted activity we reconstruct. From this, we acquire a correlation value that reflects our predictive accuracy at each location. Expanding on this approach, we could investigate electrode locations that are particularly informative to these estimates. Specifically, by holding out every electrode in turn as well as one additional electrode, and using this subset of data, we could compare predicted and actual activity and investigate if the location of the held-out location was particularly informative to the predictive accuracy. The resulting map would yield predictive accuracy by pairwise location. Mapping predictive accuracy of each recording location would not only be informative for clinicians in optimizing electrode implantation, but would also have direct implications of region to region functional connectivity.

In tandem with these model improvements, we could also track information flow in our whole brain ECoG predictions. By using our method to track the flow of neural activity, we could not only better characterize seizure propagation, but we could more accurately localize the epileptic focus for resection. Critically, with current diagnostic techniques epilepsy surgery has only a 60-80% success rate. Using the resection location and surgery outcome (success or failure), we could validate the predictions made by our procedure. Additionally, we could use Neurosynth Rubin et al. (2017) to characterize specific brain region function in the path of the seizure location and correlate it with seizure manifestation. By comparing our prediction with neurologists estimates and resection locations, we could provide evidence of diagnostic relevance. However, beyond pertaining to clinical treatment, this technique may also be used to track the flow of information in the brain. Specifically, how the complex sensory information is synchronized and integrated over time and across many brain regions. Having a clinically validated technique to track the flow of information

would be crucial in better understanding these fast-timescale network interactions.

4.0.5 Concluding remarks

This thesis has explored the network dynamics of the brain and has centrally focused on modeling fast-timescale whole-brain networks with the goal of better understanding these complex systems. In this process, I have developed methods and new computational tools. However, these tools can also be leveraged to answer a wide range of questions. It is my hope that others will find these tools useful to their own work.

APPENDIX A

Supplementary information for: A Gaussian process model of human electrocorticographic data

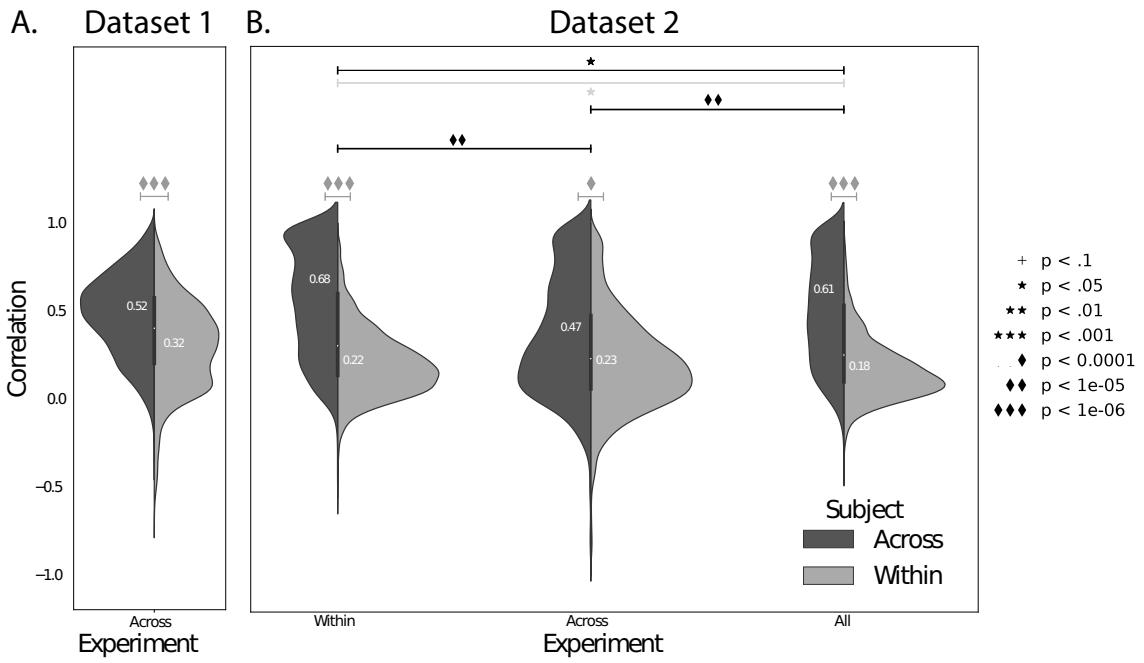


Figure A1: Reconstruction accuracy across all electrodes in two ECoG datasets, broken down by experiment. **A. Distributions of correlations between observed versus reconstructed activity by electrode, for Dataset 1.** The across-patient distribution (black) reflects reconstruction accuracy (correlation) using a correlation model learned from all but one patient’s data, and then applied to that held-out patient’s data. The within-patient distribution (gray) reflects performance using a correlation model learned from the same patient who contributed the to-be-reconstructed electrode. The split violin plot displays the same distributions shown in Figure 2A. The white numbers within each half denote the means of each distribution. **B. Distributions of correlations for Dataset 2.** The split violin plots are in the same format as those in Panel A. The leftmost plot (“Within”) displays the same distributions shown in Figure 2B. All reconstructions reflected in the distribution were carried out using a model trained and tested using data from the same experiment. The middle plots (“Across”) reflect reconstructions trained and tested using data from different experiments. The rightmost plot (“All”) reflect reconstructions obtained using models trained and tested on data from both experiments. The black distributions reflect models trained and tested across patients (analogous to the black histograms in Figure 2) and the gray distributions reflect models trained and tested within patient (analogous to the gray histograms in Figure 2). The dark gray significance bars reflect paired-sample t -tests comparing the (z -transformed) reconstruction accuracy for each electrode obtained within versus across patients. The black significance bars reflect paired-sample t -tests comparing the across-patient reconstruction accuracies across datasets. The light gray significance bars reflect paired-sample t -tests comparing the within-patient reconstruction accuracies across datasets. The symbols denote the corresponding p -values of those statistical tests.

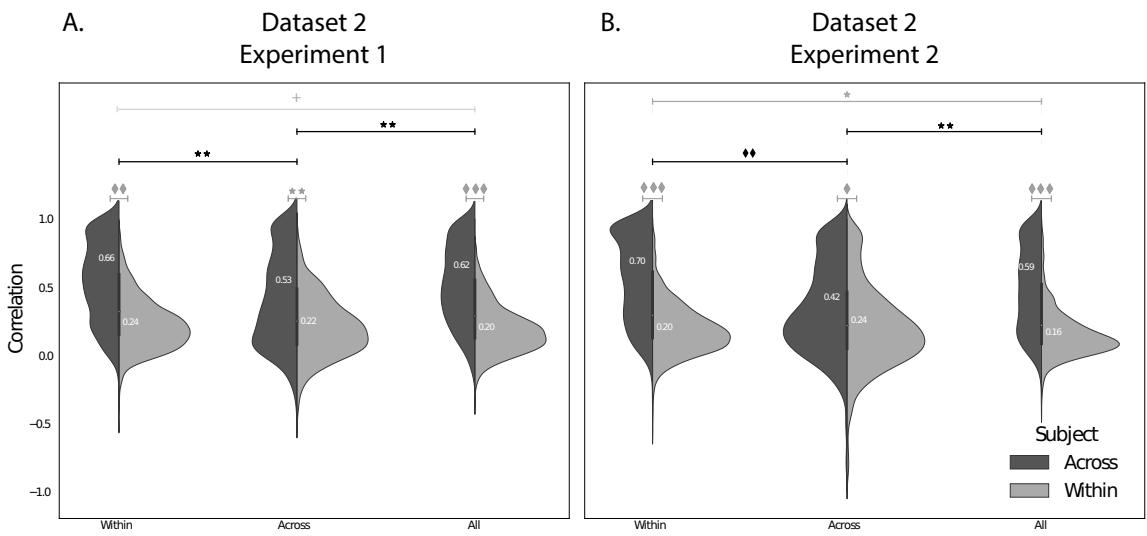


Figure A2: Reconstruction accuracy for Dataset 2, Experiments 1 and 2. **A. Distributions of correlations between observed versus reconstructed activity by electrode, for Dataset 2 (Experiment 1).** The plots are in the same format as Figure A1B, but reflect data only from Experiment 1 in Dataset 2. **A. Distributions of correlations between observed versus reconstructed activity by electrode, for Dataset 2 (Experiment 2).** The plots are in the same format as Panel A, but reflect data only from Experiment 2 in Dataset 2.

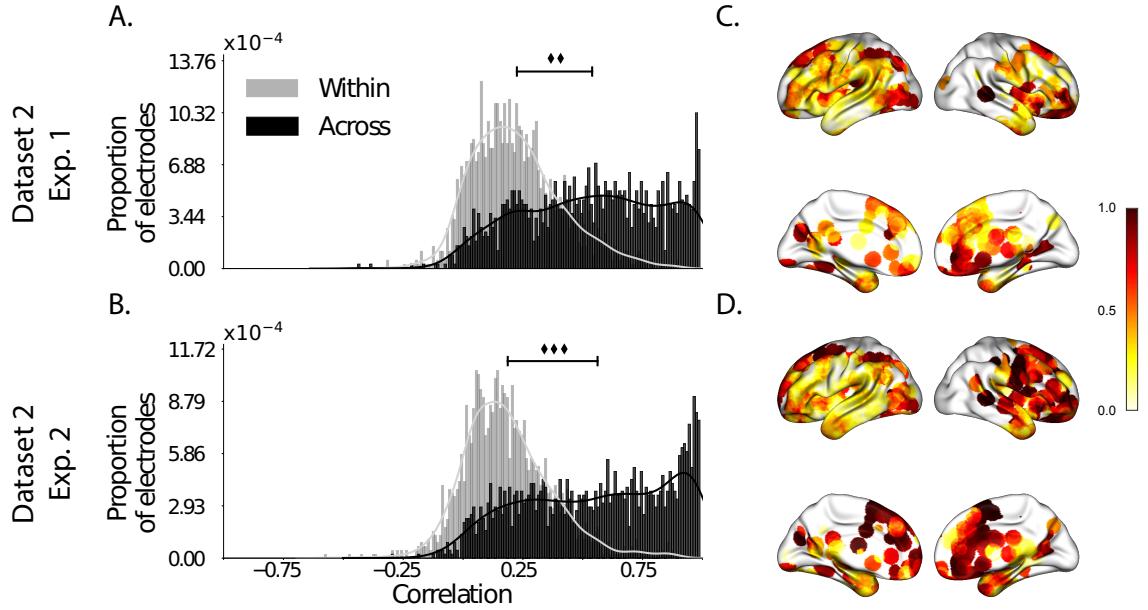


Figure A3: Reconstruction accuracy across all electrodes in Dataset 2. This figure is analogous to Figure 2 in the main text, but displays data for Dataset 2 broken down by experiment. **A. Distributions of correlations between observed versus reconstructed activity by electrode for Experiment 1.** The across-patient distribution (black) reflects reconstruction accuracy (correlation) using a correlation model learned from all but one patient's data, and then applied to that held-out patient's data. The within-patient distribution (gray) reflects performance using a correlation model learned from the same patient who contributed the to-be-reconstructed electrode. **B. Distributions of correlations for Experiment 2.** This panel is in the same format as Panel A, but reflects results obtained from Experiment 2. **C.–D. Reconstruction accuracy by location.** The colors denote the average across-session correlations, using the across-patient correlation model, between the observed and reconstructed activity at the given electrode location projected to the cortical surface (Combrisson et al., 2019). Panel C displays the map for Experiment 1 and Panel D displays the map for Experiment 2.

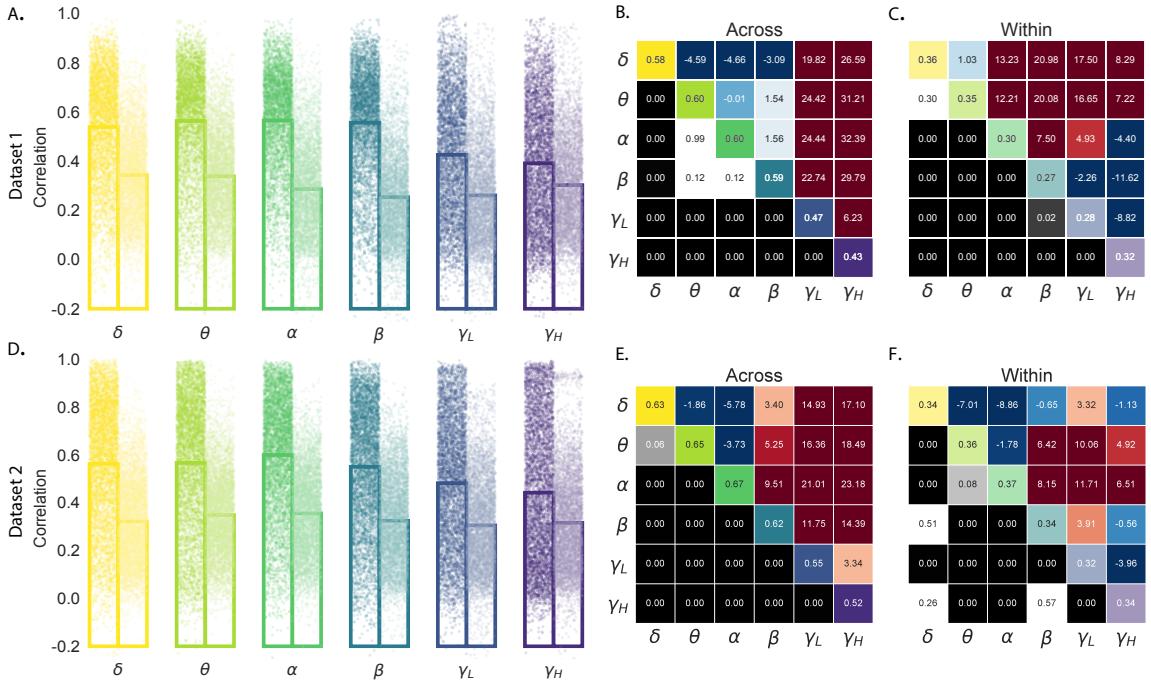


Figure A4: Reconstruction accuracy across all electrodes in two ECoG datasets for power in each frequency band. This figure is analogous to Figure 3 in the main text, but displays the reconstruction accuracies of the *power* in each frequency band, rather than the bandpass-filtered voltages, as in Figure 3. **A. Distributions of correlations between observed versus reconstructed power by electrode for each frequency band in Dataset 1.** Each color denotes a different frequency band. Within each color group, the darker dots and bar on the left display the distribution (and mean) across-patient reconstruction accuracies (analogous to the black histograms in Fig. 2). The lighter dots and bar on the right display the distribution (and mean) within-patient reconstruction accuracies (analogous to the gray histograms in Fig. 2). Each dot indicates the reconstruction accuracy for one electrode in the dataset. **B. Statistical summary of across-patient reconstruction accuracy by electrode for each frequency band in Dataset 1.** In the upper triangles of each map, warmer colors (positive *t*-values) indicate that the reconstruction accuracy for the frequency band in the given row was greater (via a two-tailed paired-sample *t*-test) than for the frequency band in the given column. Cooler colors (negative *t*-values) indicate that reconstruction accuracy for the frequency band in the given row was lower than for the frequency band in the given column. The lower triangles of each map denote the corresponding *p*-values for the *t*-tests. The diagonal entries display the average reconstruction accuracy within each frequency band. **C. Statistical summary of within-patient reconstruction accuracy by electrode for each frequency band in Dataset 1.** This panel displays the within-patient statistical summary, in the same format as Panel B. **D. Distributions of correlations between observed versus reconstructed activity by electrode, for each frequency band in Dataset 2.** This panel displays reconstruction accuracy distributions for each frequency band for Dataset 2. **E–F. Statistical summaries of across-patient and within-patient reconstruction accuracy by electrode for each frequency band in Dataset 2.** These panels are in the same as Panels B and C, but display results from Dataset 2.

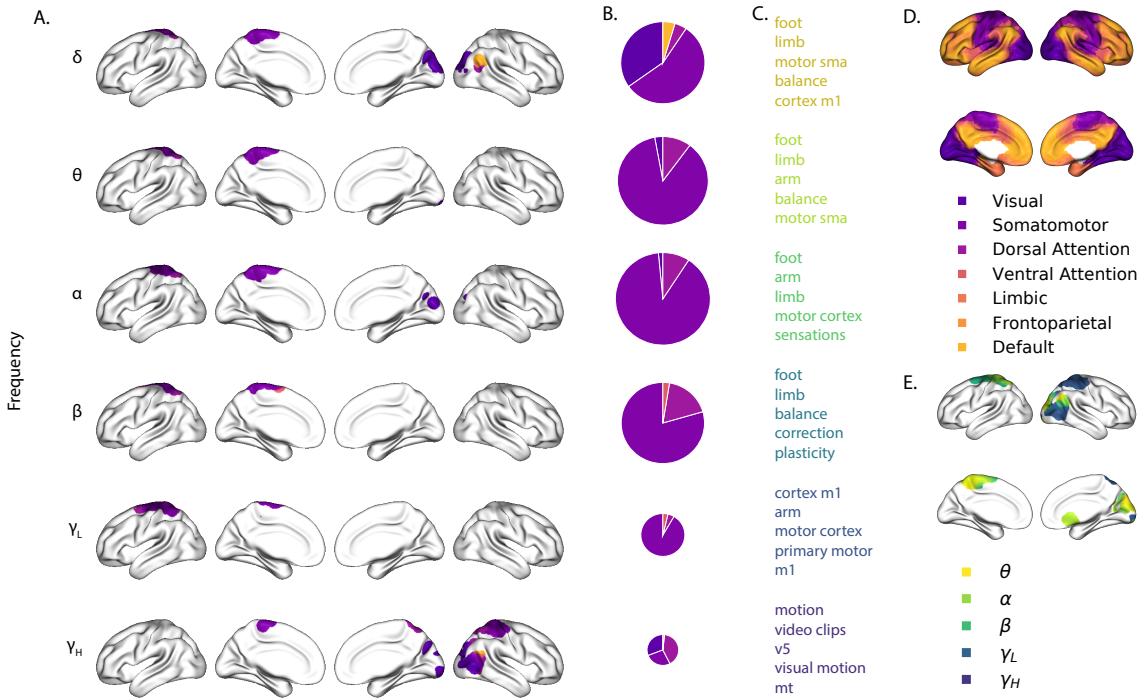


Figure A5: Most informative recording locations by frequency band. This figure is analogous to Figure 6 in the main text, but displays the brain regions that were most informative about the *power* in each frequency band (whereas Fig. 6 displays the regions that were most informative about the bandpass-filtered voltages within each band). **A. Intersections between information score maps by frequency band.** The regions indicated in each row depict the intersection between the top 10% most informative locations across Datasets 1 and 2. **B. Network memberships of the most informative brain regions.** The pie charts display the proportions of voxels in each region that belong to the seven networks identified by Yeo et al. (2011). The relative sizes of the charts for each frequency band reflect the average across-subject reconstruction accuracies (Figs. 3A, D). The voxels in Panel A are colored according to the same network memberships. **C. Neurosynth terms associated with the most informative brain regions, by frequency band.** The lists in each row display the top five neurosynth terms (Rubin et al., 2017) decoded for each region. **D. Network parcellation map and legend.** The parcellation defined by Yeo et al. (2011) is displayed on the inflated brain maps. The colors and network labels serve as a legend for Panels A and B. **E. Combined map of the most informative brain regions.** The map displays the union of the most informative maps in Panel A, colored by frequency band. The labels also serve as a legend for Panel C.

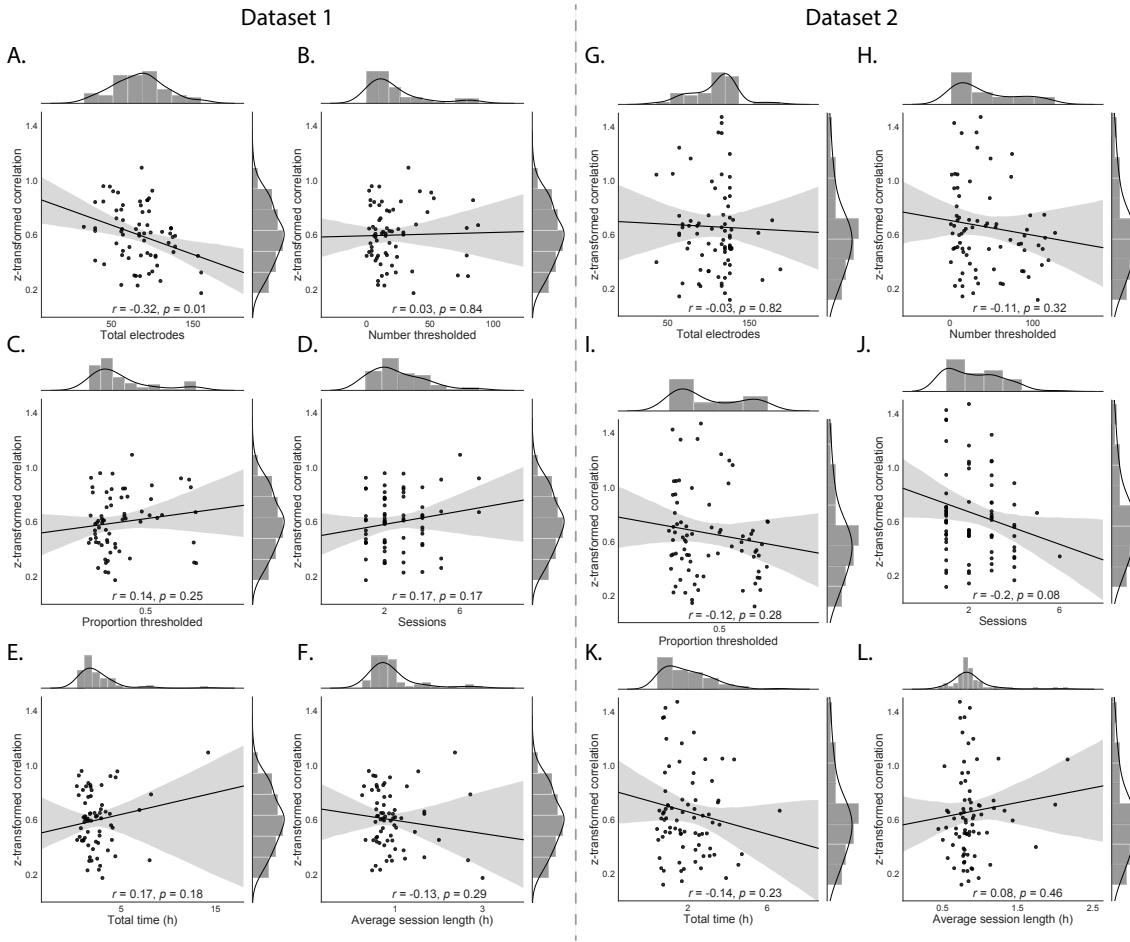


Figure A6: Reconstruction accuracy versus within-subject data features for two ECoG datasets. The individual dots in each panel reflect data from one patient. The least-squares linear regression lines (with shaded 95% confidence intervals) and the correlations reported in each panel are between the data shown on the x - and y -axes of the panel. The y -axis in each panel denotes the average z -transformed correlation between the observed and predicted voltage timeseries at each electrode (across all of the patient's electrodes). The cross-validated predictions were obtained using the across-patient model (Fig. 2A and B, black histograms). **A–F. Dataset 1.** **A. Total electrodes.** The x -coordinates of each dot display the total number of electrodes implanted in each patient's brain. **B. Number thresholded.** The x -coordinates of each dot display the number of implanted electrodes that survived the kurtosis-based filtering procedure (see Approach). **C. Proportion thresholded.** The x -coordinates of each dot display the proportion of each patient's implanted electrodes (relative to the total number) that survived the kurtosis-based filtering procedure. **D. Sessions.** The x -coordinates of each dot display the number of distinct recording sessions contributed by each patient. **E. Total time.** The x -coordinates of each dot display the total duration (in hours) of each patient's recordings, across all of their recording sessions. **F. Average session length.** The x -coordinates of each dot display average duration (in hours) of the patient's recording sessions. **G–L. Features from Dataset 2.** These panels display analogous information to Panels A–F, but for Dataset 2 patients.

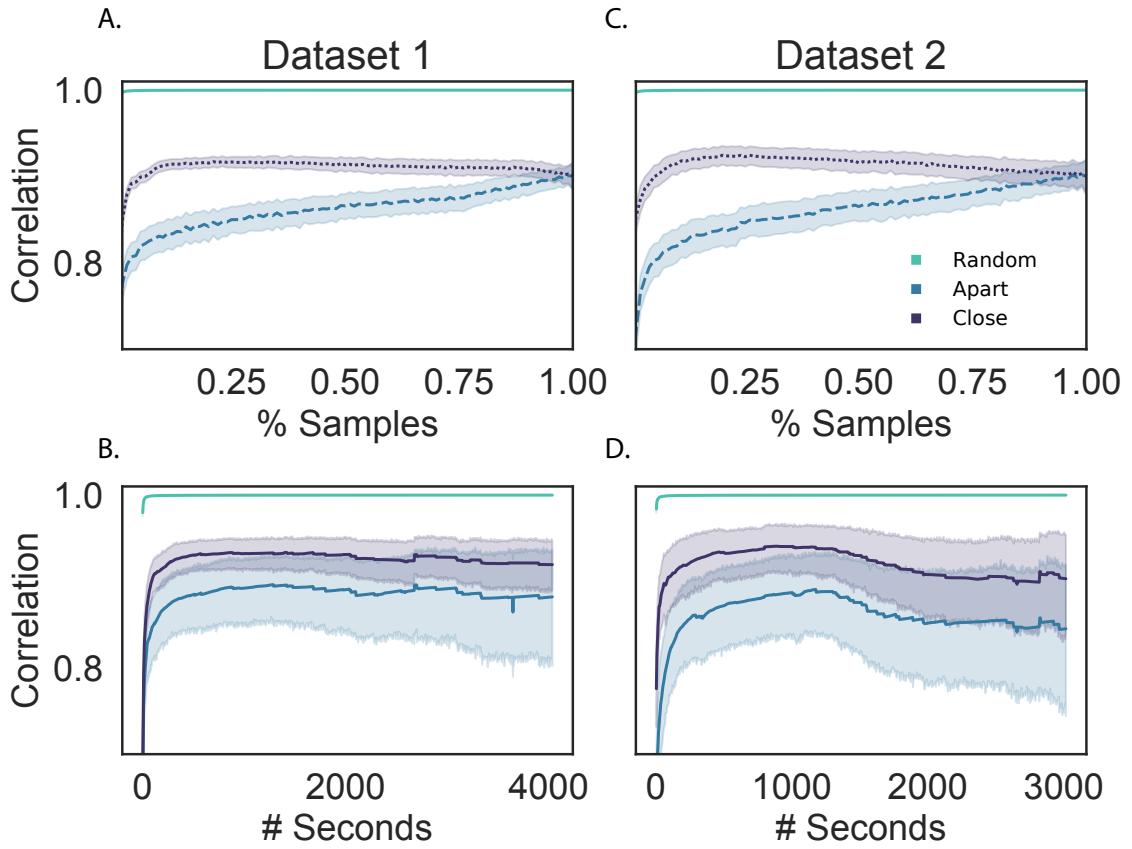


Figure A7: Temporal stability of estimated patient-specific correlation matrices. For each patient s , we computed their patient-specific correlation matrix (C_s) using two different subsets of their recorded data. We correlated the upper triangles of these estimates of C_s to obtain a single correlation for each participant, reflecting how stable (in time) the estimates were. We computed this stability measure, for each patient, as a function of how much data was used to estimate each C_s (expressed as a proportion of the patient’s total dataset, as in Panels **A** and **C**; or as the total duration of the patient’s recordings, as in Panels **B** and **D**). We also varied how the timepoints that went into each estimate were related. Specifically, we first concatenated the data from each patient’s recording sessions to construct a single data timeseries (ignoring session boundaries) for each patient. We then constructed two estimates of C_s by drawing each patient’s data from two equally sized sets of t timepoints chosen from their multi-session data matrix. These timepoints were drawn either at *random* without replacement (teal lines); from equally sized timespans at the start and end of the multi-session data matrix (*apart*, blue lines); or from two equally sized timespans just prior to and just after the data midpoint (*close*, purple lines). For the random condition of this analysis, as we increased t , we ensured that all of the timepoints included in the analysis for smaller values of t were also included in the analyses for larger values of t . For example, the randomly drawn timepoints for $t = 2000$ included the same timepoints drawn for $t = 1000$, plus an additional 1000 new timepoints. Panels **A** and **B** display the resulting correlations for Dataset 1, and panels **C** and **D** display the correlations for Dataset 2. Error ribbons in all panels denote 95% confidence intervals (computed across patients).

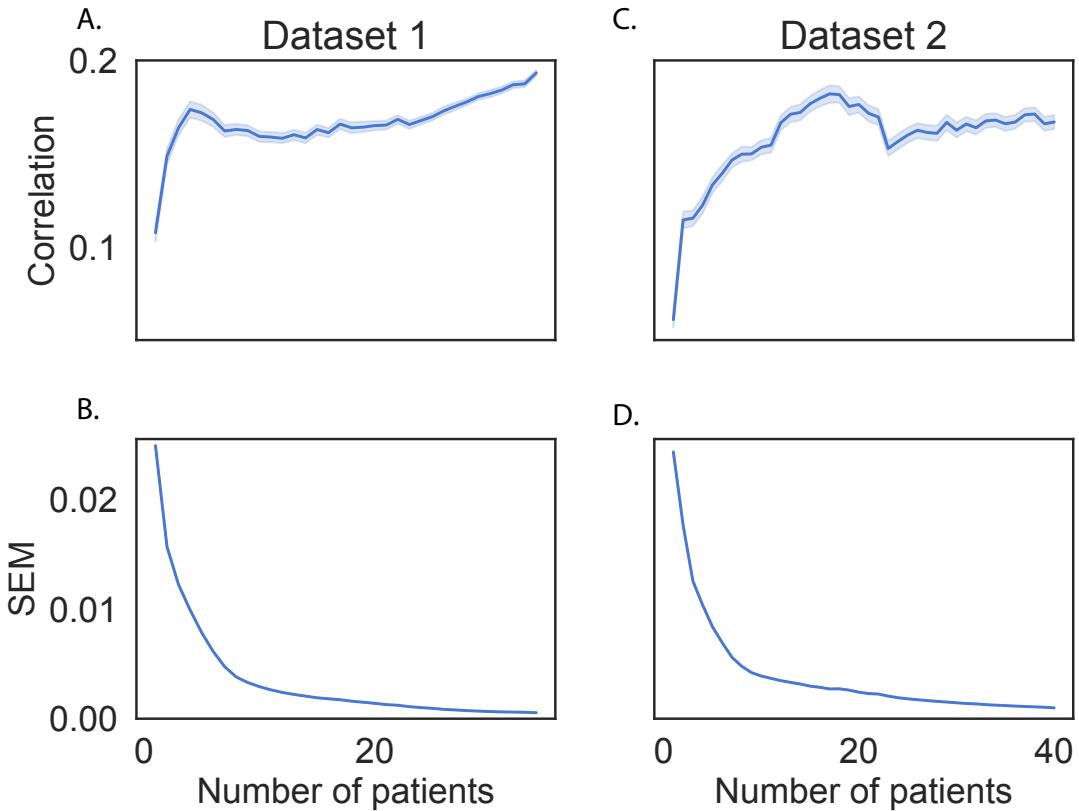


Figure A8: Stability of full-brain correlation matrices across patients. We estimated full-brain correlation matrices (\hat{K}) using different subsets of patients from each dataset. We explored how stable these estimates were as a function of how many patients were used to compute \hat{K} . For each sample size n (number of patients), we drew two non-overlapping sets of n patients (without replacement) from the full set of patients in each dataset. We then estimated \hat{K} using the two sets of n patients. We computed the correlation between the upper triangles of these matrices as a measure of how similar the matrices were. We repeated this procedure 500 times for each value of n (ranging from 1 up to $\frac{1}{2}$ of the total number of patients in the given dataset); this yielded a distribution of correlation coefficients for each value of n . Panels A and C display the mean correlations as a function of n for Datasets 1 and 2, respectively. The error ribbons denote 95% confidence intervals (computed across iterations). Panels B and D display the standard error of the mean (SEM) of these distributions of correlations as a function of n , for Datasets 1 and 2, respectively. The error ribbons denote bootstrap-derived 95% confidence intervals.

APPENDIX B

Supplementary information for: High-level cognition during story listening is reflected in high-order dynamic correlations in neural activity patterns

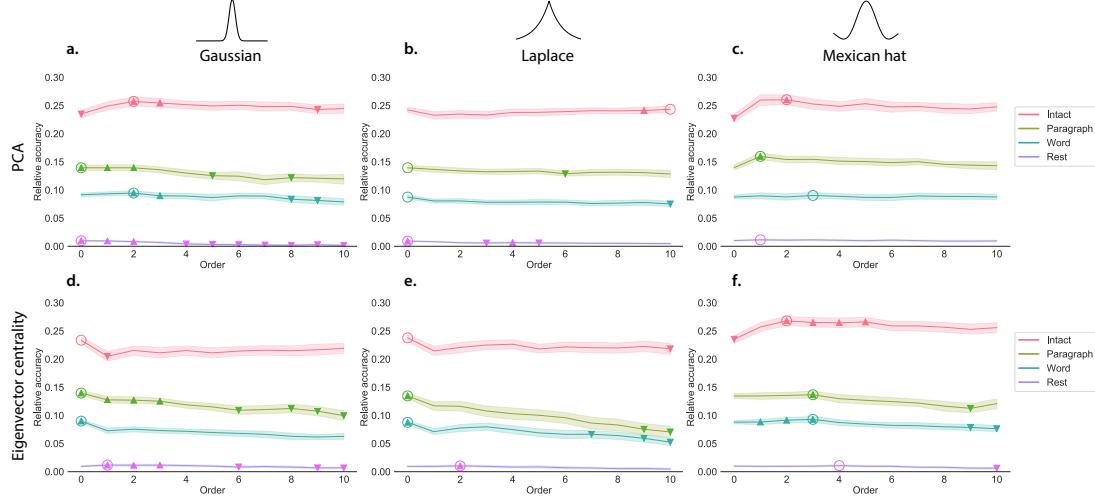


Figure B1: Across-participant timepoint decoding accuracy varies with correlation order and cognitive engagement across kernels. a.-c. Timepoint decoding accuracy as a function of order: PCA. *Order* (*x*-axis) refers to the maximum order of dynamic correlations that were available to the classifiers (see *Feature weighting and testing*). The reported across-participant decoding accuracies for **a. Gaussian**, **b. Laplace**, and **c. Mexican hat** kernel shapes are averaged over all widths (see *Identifying robust decoding results*). The *y*-values are displayed relative to chance accuracy (intact: $\frac{1}{300}$; paragraph: $\frac{1}{272}$; word: $\frac{1}{300}$; rest: $\frac{1}{400}$). The error ribbons denote 95% confidence intervals of the mean across cross-validation folds (i.e., random assignments of participants to the training and test sets). The colors denote the experimental condition. Arrows denote sets of features that yielded reliably higher (upward facing) or lower (downward facing) decoding accuracy than the mean of all other features (via a two-tailed *t*-test, thresholded at $p < 0.05$; exact values are reported in the Source Data file). The circled values represent the maximum decoding accuracy within each experimental condition. Panels a.-c. used PCA to project each high-dimensional pattern of dynamic correlations onto a lower-dimensional space. **d.-f. Timepoint decoding accuracy as a function of order: eigenvector centrality.** These panels are in the same format as Panel a.-c., but here eigenvector centrality has been used to project the high-dimensional patterns of dynamic correlations into lower-dimensional spaces. Source data are provided as a Source Data file.

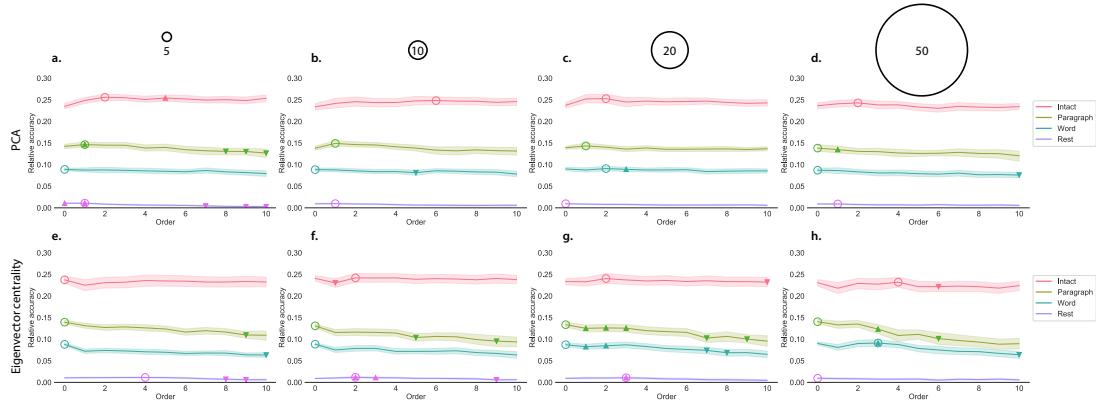


Figure B2: Across-participant timepoint decoding accuracy varies with correlation order and cognitive engagement across widths. **a.-d. Timepoint decoding accuracy as a function of order: PCA.** *Order* (*x*-axis) refers to the maximum order of dynamic correlations that were available to the classifiers (see *Feature weighting and testing*). The reported across-participant decoding accuracies for kernel widths of **a. 5**, **b. 10**, **c. 20**, and **d. 50** are averaged over all kernel shapes (see *Identifying robust decoding results*). The *y*-values are displayed relative to chance accuracy (intact: $\frac{1}{300}$; paragraph: $\frac{1}{272}$; word: $\frac{1}{300}$; rest: $\frac{1}{400}$). The error ribbons denote 95% confidence intervals of the mean across cross-validation folds (i.e., random assignments of participants to the training and test sets). The colors denote the experimental condition. Arrows denote sets of features that yielded reliably higher (upward facing) or lower (downward facing) decoding accuracy than the mean of all other features (via a two-tailed *t*-test, thresholded at $p < 0.05$; exact values are reported in the Source Data file). The circled values represent the maximum decoding accuracy within each experimental condition. Panels a.-d. used PCA to project each high-dimensional pattern of dynamic correlations onto a lower-dimensional space. **e.-h. Timepoint decoding accuracy as a function of order: eigenvector centrality.** These panels are in the same format as Panel a.-d., but here eigenvector centrality has been used to project the high-dimensional patterns of dynamic correlations into lower-dimensional spaces. Source data are provided as a Source Data file.

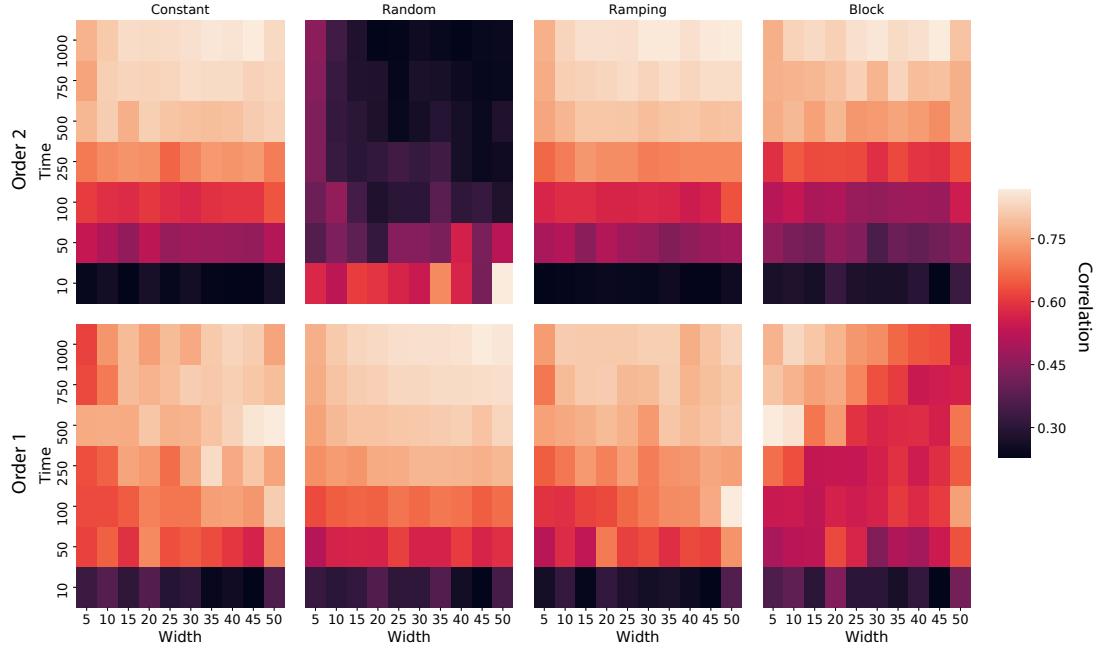


Figure B3: Recovery heatmap of simulated first-order and second-order dynamic correlations across time and widths. Each panel displays a heatmap of the average correlations between the vectorized upper triangles of the recovered first-order and second-order correlation matrices and the true (simulated) first-order and second order correlation matrices. The averages are taken across several randomly generated synthetic datasets for each timeseries pattern (constant, random, ramping, and data; see *Synthetic data: simulating dynamic higher-order correlations*). The x -axes of each heatmap denote varying Laplace-shaped kernel widths (see Figs. B1 and B2), and the y -axes of each heatmap denote varying durations (in samples) of the synthetic datasets. A total of 10 synthetic datasets were generated for each duration (row) and timeseries pattern (panel). Source data are provided as a Source Data file.

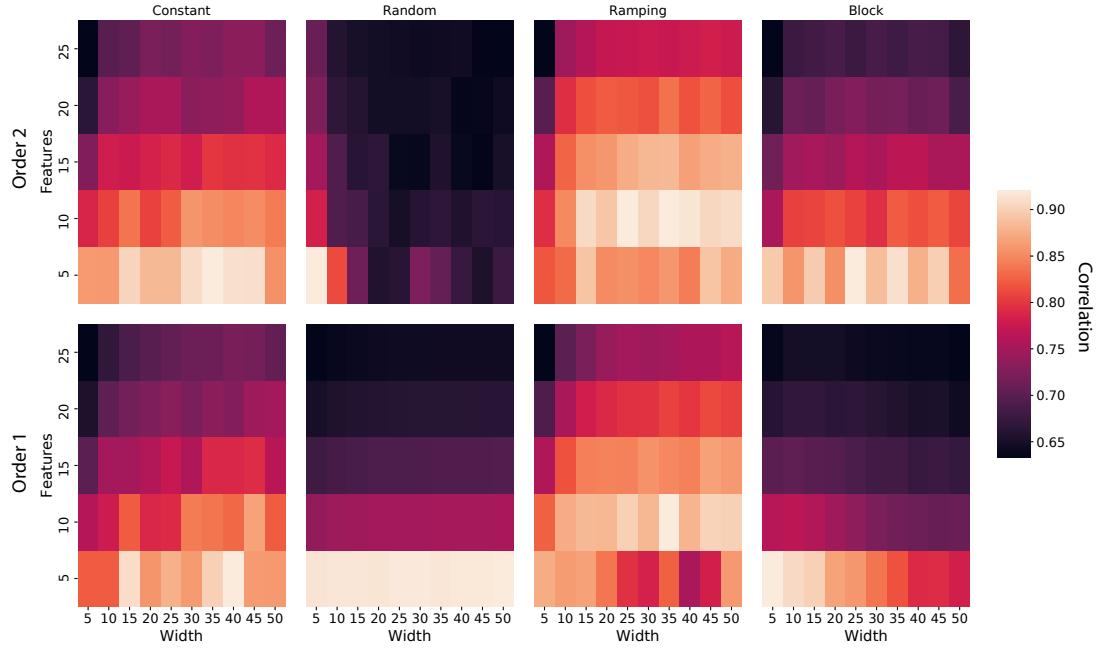
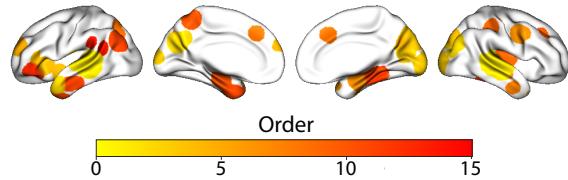
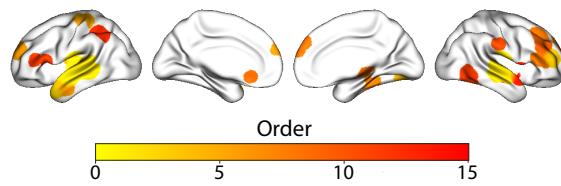


Figure B4: Recovery heatmap of simulated first-order and second-order dynamic correlations across features and widths. Each panel displays a heatmap of the average correlations between the vectorized upper triangles of the recovered first-order and second-order correlation matrices and the true (simulated) first-order and second order correlation matrices. The averages are taken across several randomly generated synthetic datasets for each timeseries pattern (constant, random, ramping, and data; see *Synthetic data: simulating dynamic higher-order correlations*). The x -axes of each heatmap denote varying Laplace-shaped kernel widths (see Figs. B1 and B2), and the y -axes of each heatmap denote varying numbers of features (K) used in each of the synthetic datasets. A total of 10 synthetic datasets were generated for each number of features (row) and timeseries pattern (panel). Source data are provided as a Source Data file.



Order 1	Order 2	Order 3	Order 4	Order 5
voice speech superior temporal listening temporal gyrus sts speech perception spoken superior temporal sulcus	voice precuneus temporal speaker temporal gyrus spoken heard sounds comprehension speech perception	cuneus v1 visual early visual primary visual lingual gyrus occipital lingual visual cortex integrate	anterior insula insula inferior inferior frontal insula executive inferior frontal pole frontal gyrus asd	integrate spatially anterior hippocampus occipital recollection orbital hippocampal visual stream visual cortex visual
Order 6	Order 7	Order 8	Order 9	Order 10
putamen face ffa ffa subsequent fusiform face images selective basal ganglia ganglia basal	hippocampal hippocampus amygdala hippocampus anterior hippocampus medial temporal parahippocampal mtl temporal lobe parahippocampal cortex lobe	face recognition recognition face fusiform face ffa faces fusiform gyrus hippocampal medial temporal	monitoring error conflict task frontal cortex read errors insula anterior color demands	posterior insula insular extrastriate insular cortex lateral occipital letter occipital visual painful insula
Order 11	Order 12	Order 13	Order 14	Order 15
memory encoding subsequent hippocampus hippocampal associative memory lobe mtl mtl retrieval episodic	angular gyrus angular anterior temporal temporal lobes default default mode network dmns self referential dmn default network	parahippocampal navigation objects parahippocampal cortex parahippocampal gyrus anterior hippocampus encoding memory encoding fusiform face ffa	inferior frontal inferior words pseudowords frontal gyrus semantic reading chinese word orthographic	motor task sensorimotor cortex contralateral hand sensorimotor finger primary spatially tapping index finger

Figure B5: Top terms associated with the most strongly correlated nodes at each order, for the *intact* experimental condition. Each color corresponds to one order of inter-subject functional correlations. The inflated brain plots display the locations of the endpoints of the 10 strongest (absolute value) correlations at each order, projected onto the cortical surface (Combrisson et al., 2019). The lists of terms display the top 10 Neurosynth terms (Rubin et al., 2017) decoded from the corresponding brain maps for each order. (Also see Fig. 6, top row, in the main text.) Source data are provided as a Source Data file.

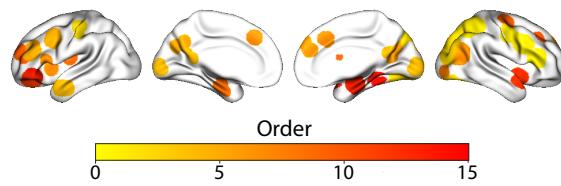


Order 1	Order 2	Order 3	Order 4	Order 5
sounds auditory cortex listening auditory listened auditory stimuli speech primary auditory sound acoustic	voice precuneus temporal speaker temporal gyrus spoken heard sounds comprehension speech perception	auditory cortex sound sounds auditory heschl gyrus heschl acoustic gyrus stg planum listened	gyri temporo posterior temporal posterior middle temporo parietal visual stream secondary somatosensory insular verb temporal gyri	occipital extrastriate visual reaching objects arm dlpfc incorrect occipital gyrus index finger
Order 6	Order 7	Order 8	Order 9	Order 10
anterior temporal mental states lateral temporal tom pole temporal pole comprehension virtual mind tom theory of mind	fusiform object memory encoding fusiform gyrus fusiform gyrus venral visual faces reaching parahippocampal gyrus objects	ventrolateral fusiform gyrus inferior frontal inferior fusiform intentions pfc cognitive impairment semantic mild cognitive	chinese reading naming judgment letter english speakers outcomes ventral striatum nucleus accumbens	prefrontal prefrontal cortex dorsolateral prefrontal cortex precuneus precuneus posterior medial dorsolateral precuneus dorsomedial prefrontal medial prefrontal
Order 11	Order 12	Order 13	Order 14	Order 15
tactile secondary somatosensory precuneus somatosensory cortex pain temporo parietal primary somatosensory somatosensory somatosensory cortices precuneus posterior	ifg gyrus ifg viewing angular gyrus v5 angular motion social interaction visual temporo parietal	inferior parietal parietal lobule working memory ipl working anterior intraparietal parietal cortex frontoparietal fronto parietal posterior parietal	musical anterior superior superior temporal listening spoken music comprehension voice speech frontal operculum	tom mind tom temporal medial temporal sulcus dorsomedial prefrontal theory mind sts dorsomedial medial prefrontal

Figure B6: Top terms associated with the most strongly correlated nodes at each order, for the *paragraph* experimental condition. This figure is in the same format as Figure B5, but displays results for the paragraph-scrambled story listening condition. (Also see Fig. 6, second row, in the main text.) Source data are provided as a Source Data file.



Figure B7: Top terms associated with the most strongly correlated nodes at each order, for the *word* experimental condition. This figure is in the same format as Figure B5, but displays results for the word-scrambled story listening condition. (Also see Fig. 6, third row, in the main text.) Source data are provided as a Source Data file.



Order 1	Order 2	Order 3	Order 4	Order 5
cortex parietal intraparietal parietal cortex ips posterior parietal parietal intraparietal sulcus symbolic superior parietal prefrontal parietal	occipital visual estrastiate vision occipital cortex visual cortex visual stream motion v1 mt	parietal conflict task visual cortex working memory supplementary motor working arithmetic supplementary anterior temporal	default autobiographical default mode posterior cingulate anterior temporal dmn early visual mental imagery primary visual autobiographical memory	dorsolateral frontal middle frontal motion mt verbal memory visual motion dorsolateral prefrontal vision
Order 6	Order 7	Order 8	Order 9	Order 10
locations navigation medial medial prefrontal orienting precuneus memory retrieval location retrieval dorsomedial	insula anterior error anterior insula monitoring insula conflict frontal cortex insular taste errors	occipital precuneus letter inferior occipital visual cortex precuneus occipital cortex lateral occipital spatially sulcus	posterior insula temporal lobe mtl amygdala responses lobe mtl insula lobe mild cognitive insular cognitive impairment	inferior occipital anterior prefrontal face selective face ffa dlpfc ffa fusiform face dorsolateral cortex dlpfc
Order 11	Order 12	Order 13	Order 14	Order 15
hippocampal hippocampus anterior hippocampus lobe amygdala hippocampus parahippocampal gyrus mtl medial temporal temporal lobe lobe mtl	food ofc cortex ofc orbitofrontal cortex nucleus accumbens accumbens orbitofrontal value orbital hypothalamus	insula taste insular insular cortex unpleasant pleasant amygdala amygdala insula posterior insula pain	amygdala hippocampus hippocampal hippocampus anterior hippocampus medial temporal amygdala parahippocampal navigation unpleasant episodic	s1 striatal outcomes monetary memory encoding accumbens mesolimbic nucleus accumbens frontal money

Figure B8: Top terms associated with the most strongly correlated nodes at each order, for the *rest* experimental condition. This figure is in the same format as Figure B5, but displays results for the resting state condition. (Also see Fig. 6, bottom row, in the main text.) Source data are provided as a Source Data file.

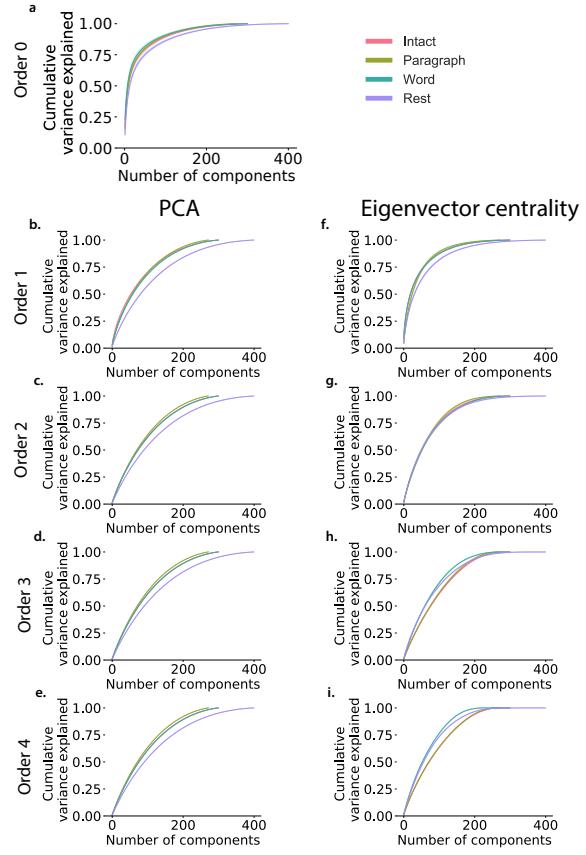


Figure B9: Cumulative percent variance explained as a function of the number of principle components. *Order* refers to the order of the dynamic correlations calculated. Principle components analysis was performed, and reduced independently for each subject. Maximum number of components varies with the total time for each condition (intact: 300; paragraph: 272; word: 300; rest: 400). **a. Cumulative percent variance as a function of number of components for Order 0.** PCA was performed on the raw activity patterns (Order 0). **b.-e. Cumulative percent variance as a function of number of components for Orders 1.-4.:** PCA Dynamic correlation were calculated for orders 1-4 using PCA to project each high-dimensional pattern of dynamic correlations onto a lower-dimensional space. **f.-i. Cumulative percent variance as a function of number of components for Orders 1.-4.: eigenvector centrality.** These panels are in the same format as Panel b.-e., but here eigenvector centrality has been used to project the high-dimensional patterns of dynamic correlations onto a lower-dimensional space. Source data are provided as a Source Data file.

BIBLIOGRAPHY

- Adachi, Y., Osada, T., Sporns, O., Watanabe, T., Matsui, T., Miyamoto, K., and Miyashita, Y. (2012). Functional connectivity between anatomically unconnected areas is shaped by collective network-level effects in the macaque cortex. *Cerebral Cortex*, 22(7):1586–1592.
- 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*, 24(3):663–676.
- Alvarez-Hamelin, I., Dall’Asta, L., Barrat, A., and Vespignani, A. (2005). k -corr decomposition: a tool for the visualization of large scale networks. *arXiv*, page cs/0504107v2.
- Amzica, F. and Steriade, M. (1998). Electrophysiological correlates of sleep delta waves. *Electroencephalography and Clinical Neurophysiology*, 107(2):69–83.
- Arslan, S., Ktena, S. I., Makropoulos, A., Robinson, E. C., Rueckert, D., and Parisot, S. (2018). Human brain mapping: a systematic comparison of parcellation methods for the human cerebral cortex. *NeuroImage*, 170(15):5–30.
- Baillet, S., Mosher, J. C., and Leahy, R. M. (2001). Electromagnetic brain mapping. *IEEE Signal Processing Magazine*, 18:14–30.
- Baldassano, C., Chen, J., Zadbood, A., Pillow, J. W., Hasson, U., and Norman, K. A. (2017). Discovering event structure in continuous narrative perception and memory. *Neuron*, 95(3):709–721.
- Bartels, A. and Zeki, S. (2005). The chronoarchitecture of the cerebral cortex. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 360(1456):733–750.
- Barthélemy, M. (2004). Betweenness centrality in large complex networks. *European Physical Journal B*, 38:163–168.
- Bassett, D., 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, USA*, 103(51):19518–19523.

- Bassett, D. S. and Gazzaniga, M. S. (2011). Understanding complexity in the human brain. *Trends in cognitive sciences*, 15(5):200–209.
- Bassett, D. S., Zurn, P., and Gold, J. I. (2018). On the nature and use of models in network neuroscience. *Nature reviews. Neuroscience*, 19(9):566–578.
- Beaty, R. E., Benedek, M., Silvia, P. J., and Schacter, D. L. (2016). Creative cognition and brain network dynamics. *Trends in Cognitive Sciences*, 20(2):87–95.
- Becker, C. O., Pequito, S., Pappas, G. J., Miller, M. B., and Bassett, S. T. G., and Preciado, V. M. (2018). Spectral mapping of brain functional connectivity from diffusion imaging. *Scientific Reports*, 8(1411):doi.org/10.1038/s41598-017-18769-x.
- Betzel, R. and Vértes, P. (2021). Organizing principles of the *c. elegans* contactome. *Cell Systems*, 12(7):689–691.
- Betzel, R. F. (2020). Organizing principles of whole-brain functional connectivity in zebrafish larvae. *Network Neuroscience*, 4(1):234–256.
- Betzel, R. F., Byrge, L., Esfahlani, F. Z., and Kennedy, D. P. (2019). Temporal fluctuations in the brain’s modular architecture during movie-watching. *bioRxiv*, doi.org/10.1101/750919.
- Betzel, R. F., Medaglia, J. D., Kahn, A. E., Soffer, J., Schonhaut, D. R., and Bassett, D. S. (2017). Inter-regional ECoG correlations predicted by communication dynamics, geometry, and correlated gene expression. *arXiv*, 1706.06088.
- Bonacich, P. (2007). Some unique properties of eigenvector centrality. *Social Networks*, 29(4):555–564.
- Bressler, S. L. and Kelso, J. A. S. (2001). Cortical coordination dynamics and cognition. *Trends in Cognitive Sciences*, 5(1):26–36.
- Bullmore, E. and Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, 10(3):186–198.
- Busch, N. A. and VanRullen, R. (2010). Spontaneous EEG oscillations reveal periodic sampling of visual attention. *Proceedings of the National Academy of Sciences, USA*, 107(37):16048–16053.
- Buzsaki, G., Anastassiou, C., and Koch, C. (2012). The origin of extracellular fields and currents—EEG, ECoG, lfp, and spikes. *Nature Reviews Neuroscience*, 13:407–419.
- Cantlon, J. F. and Li, R. (2013). Neural activity during natural viewing of sesame street statistically predicts test scores in early childhood. *PLOS Biology*, 11(1):e1001462–.
- Capota, M., Turek, J., Chen, P.-H., Zhu, X., Manning, J. R., Sundaram, N., Keller, B., Wang, Y., and Shin, Y. S. (2017). Brain imaging analysis kit.

- Casey, B. J., Giedd, J. N., and Thomas, K. M. (2000). Structural and functional brain development and its relation to cognitive development. *Biological Psychology*, 54(1–3):241–257.
- Chang, C. and Glover, G. H. (2010). Time-frequency dynamics of resting-state brain connectivity measured with fMRI. *NeuroImage*, 50:81–98.
- Chien, H.-Y. S. and Honey, C. J. (2019). Constructing and forgetting temporal context in the human cerebral cortex. *bioRxiv*, doi.org/10.1101/761593.
- Christakis, N. A. and Fowler, J. H. (2010). Social network sensors for early detection of contagious outbreaks. *PLoS One*, 5(9):e12948.
- Cole, M. W., Ito, T., Bassett, D. S., and Schultz, D. H. (2016). Activity flow over resting-state networks shapes cognitive task activations. *Nature Neuroscience*, 19(12):1718–1726.
- Combrisson, E., Vallat, R., O'Reilly, C., Jas, M., Pascarella, A., I Saive, A., Thiery, T., Meunier, D., Altukhov, D., Lajnef, T., Ruby, P., Guillot, A., and Jerbi, K. (2019). Visbrain: a multi-purpose GPU-accelerated open-source suite for multimodal brain data visualization. *Frontiers in Neuroinformatics*, 13(14):1–14.
- Comon, P., Jutten, C., and Herault, J. (1991). Blind separation of sources, part II: problems statement. *Signal Processing*, 24(1):11–20.
- Crone, N. E., Korzeniewska, A., and Franaszczuk, P. J. (2011). Cortical gamma responses: searching high and low. *International Journal of Psychophysiology*, 79(1):9–15.
- Demertzi, A., Tagliazucchi, E., Dehaene, S., Deco, G., Barttfeld, P., Raimondo, F., Martial, C., Fernández-Espejo, D., Rohaut, B., Voss, H. U., Schiff, N. D., Owen, A. M., Laureys, S., Naccache, L., and Sitt, J. D. (2019). Human consciousness is supported by dynamic complex patterns of brain signal coordination. *Science Advances*, 5(2):eaat7603.
- Eickhoff, S. B., Milham, M., and Vanderwal, T. (2020). Towards clinical applications of movie fmri. *NeuroImage*, 217:116860.
- Estrada, E. and Rodríguez-Velázquez, J. A. (2005). Subgraph centrality in complex networks. *Physical Review E*, 71(5):056103.
- Etzel, J. A., Gazzola, V., and Keysers, C. (2009). An introduction to anatomical ROI-based fMRI classification. *Brain Research*, 1281:114–125.
- Ezzyat, Y., Kragel, J. E., Burke, J. F., Levy, D. F., Lyalenko, A., Wanda, P., O'Sullivan, L., Hurley, K. B., Busygin, S., Pedisich, I., Sperling, M. R., Worrell, G. A., Kucewicz, M. T., Davis, K. A., Lucas, T. H., Inman, C. S., Lega, B. C., Jobst, B. C., Sheth, S. A., Zaghloul, K., Jutras, M. J., Stein, J. M., Das, S. R., Gorniak, R., Rizzuto, D. S., and Kahana, M. J. (2017). Direct brain stimulation modulates encoding states and memory performance in humans. *Current Biology*, 27:1–8.

- Ezzyat, Y., Wanda, P. A., Levy, D. F., Kadel, A., Aka, A., Pedisich, I., Sperling, M. R., Sharan, A. D., Lega, B. C., Burks, A., Gross, R. E., Inman, C. S., Jobst, B. C., Gorenstein, M. A., Davis, K. A., Worrell, G. A., Kucewicz, M. T., Stein, J. M., Gorniak, R., Das, S. R., Rizzuto, D. S., and Kahana, M. J. (2018). Closed-loop stimulation of temporal cortex rescues functional networks and improves memory. *Nature Communications*, 9(365):doi.org/10.1038/s41467-017-02753-0.
- Fallon, J., Ward, P. G. D., Parkes, L., and Oldham, S. (2020). Timescales of spontaneous fMRI fluctuations relate to structural connectivity in the brain. *Network Neuroscience*, 4(3):788–806.
- Faskowitz, J., Esfahlani, F. Z., Jo, Y., Sporns, O., and Betzel, R. F. (2020). Edge-centric functional network representations of human cerebral cortex reveal overlapping system-level architecture. *Nature Neuroscience*, 23(12):1644–1654.
- Feingold, J., Gibson, D. J., DePasquale, B., and Graybiel, A. M. (2015). Bursts of beta oscillation differentiate postperformance activity in the striatum and motor cortex of monkeys performing movement tasks. *Proceedings of the National Academy of Sciences, USA*, 112(44):13687–13692.
- Finn, E. S. (2021). Is it time to put rest to rest? *Trends in Cognitive Sciences*.
- Finn, E. S., Shen, X., Scheinost, D., Rosenberg, M. D., Huang, J., Chun, M. M., Papademetris, X., and Constable, R. T. (2015). Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. *Nature Neuroscience*, 18:1664–1671.
- Fong, A. H. C., Yoo, K., Rosenberg, M. D., Zhang, S., Li, C.-S. R., Scheinost, D., Constable, R. T., and Chun, M. M. (2019). Dynamic functional connectivity during task performance and rest predicts individual differences in attention across studies. *NeuroImage*, 188:14–25.
- Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry*, 40(1):35–41.
- Fries, P., Nikolić, D., and Singer, W. (2007). The gamma cycle. *Trends in Neurosciences*, 30(7):309–316.
- Friston, K. J. (2000). The labile brain. I. neuronal transients and nonlinear coupling. *Philosophical Transactions of the Royal Society of London*, 355B:215–236.
- Geisberger, R., Sanders, P., and Schultes, D. (2008). Better approximation of betweenness centrality. *Proceedings of the Meeting on Algorithm Engineering and Experiments*, pages 90–100.
- Gershman, S. J., Blei, D. M., Pereira, F., and Norman, K. A. (2011). A topographic latent source model for fMRI data. *NeuroImage*, 57:89–100.

- Gonzalez-Castillo, J., Caballero-Gaudes, C., Topolski, N., Handwerker, D. A., Pereira, F., and Bandettini, P. A. (2019). Imaging the spontaneous flow of thought: distinct periods of cognition contribute to dynamic functional connectivity during rest. *NeuroImage*, 202(116129).
- Grabner, G., Janke, A. L., Budge, M. M., Smith, D., Pruessner, J., and Collins, D. L. (2006). Symmetric atlasing and model based segmentation: an application to the hippocampus in older adults. *Medical Image Computing and Computer-Assisted Intervention*, 9:58–66.
- Grossberg, S. (1988). Nonlinear neural networks: principles, mechanisms, and architectures. *Neural Networks*, 1(1):17–61.
- Halu, A., Mondragón, R. J., Panzarasa, P., and Bianconi, G. (2013). Multiplex PageRank. *PLoS One*, 8(10):e78293.
- Hamilton, L. S. and Huth, A. G. (2018). The revolution will not be controlled: natural stimuli in speech neuroscience. *Language, cognition and neuroscience*, 35(5):573–582.
- Hasson, U., Chen, J., and Honey, C. J. (2015). Hierarchical process memory: memory as an integral component of information processing. *Trends in Cognitive Sciences*, 19(6):304–315.
- Hasson, U., Malach, R., and Heeger, D. J. (2010). Reliability of cortical activity during natural stimulation. *Trends in cognitive sciences*, 14(1):40–48.
- Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., and Malach, R. (2004). Intersubject synchronization of cortical activity during natural vision. *Science*, 303:1634–1640.
- Hasson, U., Yang, E., Vallines, I., Heeger, D. J., and Rubin, N. (2008). A hierarchy of temporal receptive windows in human cortex. *The Journal of Neuroscience*, 28(10):2539–2550.
- Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., and Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, 293:2425–2430.
- Haxby, J. V., Gobbini, M. I., and Nastase, S. A. (2020a). Naturalistic stimuli reveal a dominant role for agentic action in visual representation. *NeuroImage*, 216:116561.
- Haxby, J. V., Guntupalli, J. S., Connolly, A. C., Halchenko, Y. O., Conroy, B. R., Gobbini, M. I., Hanke, M., and Ramadge, P. J. (2011). A common, high-dimensional model of the representational space in human ventral temporal cortex. *Neuron*, 72:404–416.
- Haxby, J. V., Guntupalli, J. S., Nastase, S. A., and Ma, F. (2020b). Hyperalignment: modeling shared information encoded in idiosyncratic cortical topographies. *eLife*, 9:e56601.

- Hillebrand, A., Singh, K. D., Holliday, I. E., Furlong, P. L., and Barnes, G. R. (2005). A new approach to neuroimaging with magnetoencephalography. *Human Brain Mapping*, 25:199–211.
- Hinton, G. E. and Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science*, 313(5786):504–507.
- Honey, C. J., Kötter, R., Breakspear, M., and Sporns, O. (2007). Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *Proceedings of the National Academy of Sciences, USA*, 104(24):10240–10245.
- Honey, C. J., Sporns, O., Cammoun, L., Gigandet, X., Thiran, J. P., Meuli, R., and Hagmann, P. (2009). Predicting human resting-state functional connectivity from structural connectivity. *Proceedings of the National Academy of Sciences, USA*, 106(6):2035–2040.
- Honey, C. J., Thesen, T., Donner, T. H., Silbert, L. J., Carlson, C. E., Devinsky, O., Doyle, J. C., Rubin, N., Heeger, D. J., and Hasson, U. (2012). Slow cortical dynamics and the accumulation of information over long timescales. *Neuron*, 76:423–434.
- Horak, P. C., Meisenhelter, S., Song, Y., Testorf, M. E., Kahana, M. J., Viles, W. D., Bujarski, K. A., Connolly, A. C., Robbins, A. A., Sperling, M. R., Sharan, A. D., Worrell, G. A., Miller, L. R., Gross, R. E., Davis, K. A., Roberts, D. W., Lega, B., Sheth, S. A., Zaghloul, K. A., Stein, J. M., Das, S. R., Rizzuto, D. S., and Jobst, B. C. (2017). Interictal epileptiform discharges impair word recall in multiple brain areas. *Epilepsia*, 58(3):373–380.
- Huster, R. J., Debener, S., Eichele, T., and Herrmann, C. S. (2012). Methods for simultaneous EEG-fMRI: an introductory review. *The Journal of Neuroscience*, 32(18):6053–6060.
- Huth, A. G., de Heer, W. A., Griffiths, T. L., Theunissen, F. E., and Gallant, J. L. (2016). Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature*, 532:453–458.
- Huth, A. G., Nisimoto, S., Vu, A. T., and Gallant, J. L. (2012). A continuous semantic space describes the representation of thousands of object and action categories across the human brain. *Neuron*, 76(6):1210–1224.
- Ihara, A., Hirata, M., Yanagihara, K., Ninomiya, H., Imai, K., Ishii, R., Osaki, Y., Sakihara, K., Izumi, H., Imaoka, H., Kato, A., and Yoshimine, T. (2003). Neuromagnetic gamma-band activity in the primary and secondary somatosensory areas. *NeuroReport*, 14(2):273–277.
- Jacobs, J., Manning, J., and Kahana, M. J. (2010). Response to Miller: “broadband” versus “high gamma” electrocorticographic signals. *The Journal of Neuroscience*, 30(19).

- Jahanshad, N., Kochunov, P. V., Sprooten, E., Mandl, R. C., Nichols, T. E., Almasy, L., Blangero, J., Brouwer, R. M., Curran, J. E., de Zubiray, G. I., Duggirala, R., Fox, P. T., Hong, L. E., Landman, B. A., Martin, N. G., McMahon, K. L., Medland, S. E., Mitchell, B. D., Olvera, R. L., Peterson, C. P., Starr, J. M., Sussmann, J. E., Toga, A. W., Wardlaw, J. M., Wright, M. J., Pol, H. E. H., Pastin, M. E., McIntosh, A. M., Deary, I. J., Thompson, P. M., and Glahn, D. C. (2013). Multi-site genetic analysis of diffusion images and voxelwise heritability analysis: A pilot project of the enigma-dti working group. *NeuroImage*, 81(1):455–469.
- Jurkiewicz, M. T., Gaetz, W. C., Bostan, A. C., and Cheyne, D. (2006). Post-movement beta rebound is generated in motor cortex: evidence from neuromagnetic recordings. *NeuroImage*, 32(3):1281–1289.
- Jutten, C. and Herault, J. (1991). Blind separation of sources, part I: an adaptive algorithm based on neuromimetic architecture. *Signal Processing*, 24(1):1–10.
- Kamitani, Y. and Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nature Neuroscience*, 8:679–685.
- Kragel, J. E., Ezzyat, Y., Sperling, M. R., Gorniak, R., Worrell, G. A., Berry, B. M., Inman, C., Lin, J.-J., Davis, K. A., Das, S. R., Stein, J. M., Jobst, B. C., Zaghloul, K. A., Sheth, S. A., Rizzuto, D. S., and Kahana, M. J. (2017). Similar patterns of neural activity predict memory formation during encoding and retrieval. *NeuroImage*, 155:70–71.
- Kucewicz, M. T., Berry, B. M., Kremen, V., Brinkmann, B. H., Sperling, M. R., Jobst, B. C., Gross, R. E., Lega, B., Sheth, S. A., Stein, J. M., Das, S. R., Gorniak, R., Stead, S. M., Rizzuto, D. S., Kahana, M. J., and Worrell, G. A. (2017). Dissecting gamma frequency activity during human memory processing. *Brain*, 140:1337–1350.
- Kucewicz, M. T., Berry, B. M., Miller, L. R., Khadjevand, F., Ezzyat, Y., Stein, J. M., Kremen, V., Brinkmann, B. H., Wanda, P., Sperling, M. R., Gorniak, R., Davis, K. A., Jobst, B. C., Gross, R. E., Lega, B., Gompel, J. V., Stead, S. M., Rizzuto, D. S., Kahana, M. J., and Worrell, G. A. (2018). Evidence for verbal memory enhancement with electrical brain stimulation in the lateral temporal cortex. *Brain*, 141(4):971–978.
- Kucyi, A., Schrouff, J., Bickel, S., Foster, B. L., Shine, J. M., and Parvizi, J. (2018). Intracranial electrophysiology reveals reproducible intrinsic functional connectivity with human brain networks. *The Journal of Neuroscience*, 38(17):4230–4242.
- Landau, E. (1895). Zur relativen Wertbemessung der Turnierresultate. *Deutsches Wochenschach*, 11:366–369.
- Lee, C. S., Aly, M., and Baldassano, C. (2020). Anticipation of temporally structured events in the brain. *bioRxiv*, 10.1101/2020.10.14.338145.
- Lee, D. D. and Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature*, 401:788–791.

- Lerner, Y., Honey, C. J., Katkov, M., and Hasson, U. (2014). Temporal scaling of neural responses to compressed and dilated natural speech. *Journal of Neurophysiology*, 111:2433–2444.
- Lerner, Y., Honey, C. J., Silbert, L. J., and Hasson, U. (2011). Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. *The Journal of Neuroscience*, 31(8):2906–2915.
- Liégeois, R., Li, J., Kong, R., Orban, C., Van De Ville, D., Ge, T., Sabuncu, M. R., and Yeo, B. T. T. (2019). Resting brain dynamics at different timescales capture distinct aspects of human behavior. *Nature Communications*, 10(2317):1–9.
- Lin, J. (2009). Divergence measures based on the Shannon entropy. *IEEE Transactions on Information Theory*, 37(1):145–151.
- Lin, J.-J., Rugg, M. D., Das, S., Stein, J., Rizzuto, D. S., Kahana, M. J., and Lega, B. C. (2017). Theta band power increases in the posterior hippocampus predict successful episodic memory encoding in humans. *Hippocampus*, 27:1040–1053.
- Lohmann, G., Margulies, D. S., Horstmann, A., Pleger, B., Lepsien, J., Goldhahn, D., Schloegl, H., Stumvoll, M., Villringer, A., and Turner, R. (2010). Eigenvector centrality mapping for analyzing connectivity patterns in fMRI data of the human brain. *PLoS One*, 5(4):e10232.
- Lurie, D., Kessler, D., Bassett, D., Betzel, R., Breakspear, M., Keilholz, S., Kucyi, A., Liégeois, R., Lindquist, M., McIntosh, A., Poldrack, R., Shine, J., Thompson, W., Bielczyk, N., Douw, L., Kraft, D., Miller, R., Muthuraman, M., Pasquini, L., Razi, A., Vidaurre, D., Xie, H., and Calhoun, V. (2018). On the nature of time-varying functional connectivity in resting fMRI. *PsyArXiv*, doi.org/10.31234/osf.io/xtzre.
- Mack, M. L., Preston, A. R., and Love, B. C. (2017). Medial prefrontal cortex compresses concept representations through learning. *bioRxiv*, doi.org/10.1101/178145.
- Mairal, J., Ponce, J., Sapiro, G., Zisserman, A., and Bach, F. R. (2009a). Supervised dictionary learning. *Advances in Neural Information Processing Systems*, pages 1033–1040.
- Mairal, J. B., Bach, F., Ponce, J., and Sapiro, G. (2009b). Online dictionary learning for sparse coding. *Proceedings of the International Conference on Machine Learning*, pages 689–696.
- Manning, J. R., Jacobs, J., Fried, I., and Kahana, M. J. (2009). Broadband shifts in LFP power spectra are correlated with single-neuron spiking in humans. *The Journal of Neuroscience*, 29(43):13613–13620.
- Manning, J. R., Polyn, S. M., Baltuch, G., Litt, B., and Kahana, M. J. (2011). Oscillatory patterns in temporal lobe reveal context reinstatement during memory search. *Proceedings of the National Academy of Sciences, USA*, 108(31):12893–12897.

- Manning, J. R., Ranganath, R., Norman, K. A., and Blei, D. M. (2014). Topographic factor analysis: a Bayesian model for inferring brain networks from neural data. *PLoS One*, 9(5):e94914.
- Manning, J. R., Sperling, M. R., Sharan, A., Rosenberg, E. A., and Kahana, M. J. (2012). Spontaneously reactivated patterns in frontal and temporal lobe predict semantic clustering during memory search. *The Journal of Neuroscience*, 32(26):8871–8878.
- Manning, J. R., Zhu, X., Willke, T. L., Ranganath, R., Stachenfeld, K., Hasson, U., Blei, D. M., and Norman, K. A. (2018). A probabilistic approach to discovering dynamic full-brain functional connectivity patterns. *NeuroImage*, 180:243–252.
- Matsumoto, J. Y., Stead, M., Kucewicz, M. T., Matsumoto, A. J., Peters, P. A., Brinkmann, B. H., Danstrom, J. C., Goerss, S. J., Marsh, W. R., Meyer, F. B., et al. (2013). Network oscillations modulate interictal epileptiform spike rate during human memory. *Brain*, 136(8).
- McInnes, L., Healy, J., and Melville, J. (2018). UMAP: uniform manifold approximation and projection for dimension reduction. *arXiv*, 1802(03426).
- McIntosh, A. R. (2000). Towards a network theory of cognition. *Neural Networks*, 13(8–9):861–870.
- McIntosh, A. R. and Jirsa, V. K. (2019). The hidden repertoire of brain dynamics and dysfunction. *Network Neuroscience*, doi.org/10.1162/netn_a_00107.
- Miller, K. J., Leuthardt, E. C., Schalk, G., Rao, R. P. N., Anderson, N. R., Moran, D. W., Miller, J. W., and Ojemann, J. G. (2007). Spectral changes in cortical surface potentials during motor movement. *The Journal of Neuroscience*, 27:2424–2432.
- Miller, K. J., Shenoy, P., den Nijs, M., Sorensen, L. B., Rao, R. P. N., and Ojemann, J. G. (2008). Beyond the gamma band: the role of high-frequency features in movement classification. *IEEE Transactions on Biomedical Engineering*, 55(5):1634–1637.
- Mitchell, T. M., Shinkareva, S. V., Carlson, A., Chang, K. M., Malave, V. L., Mason, R. A., and Just, M. A. (2008). Predicting human brain activity associated with the meanings of nouns. *Science*, 320(5880):1191.
- Mori, S., Oishi, K., Jiang, H., Jiang, L., Li, X., Akhter, K., Hua, K., Faria, A. V., Mahmood, A., Woods, R., Toga, A. W., Pike, G. B., Neto, P. R., Evans, A., Zhang, J., Huang, H., Miller, M. I., van Zijl, P., and Mazziotta, J. (2008). Stereotaxic white matter atlas based on diffusion tensor imaging in an icbm template. *NeuroImage*, 40(2):570–582.
- Naselaris, T., Kay, K. N., Nishimoto, S., and Gallant, J. L. (2011). Encoding and decoding in fmri. *NeuroImage*, 56(2):400–410.
- Nastase, S. A., Goldstein, A., and Hasson, U. (2020). Keep it real: rethinking the primacy of experimental control in cognitive neuroscience. *NeuroImage*, 222:117254.

- Newman, M. E. J. (2005). A measure of betweenness centrality based on random walks. *Social Networks*, 27:39–54.
- Newman, M. E. J. (2008). The mathematics of networks. *The New Palgrave Encyclopedia of Economics*, 2:1–12.
- Nguyen, J. P., Shipley, F. B., Linder, A. N., Plummer, G. S., Liu, M., Setru, S. U., Shaevitz, J. W., and Leifer, A. M. (2016). Whole-brain calcium imaging with cellular resolution in freely behaving *caenorhabditis elegans*. *Proceedings of the National Academy of Sciences, USA*, 113(8):E1074–E1081.
- ni, J. G., van den Heuvel, M. P., Avena-Koenigsberger, A., de Mendizabal, N. V., Betzel, R. F., Griffa, A., Hagmann, P., Corominas-Murtra, B., Thiran, J.-P., and Sporns, O. (2014). Resting-brain functional connectivity predicted by analytic measures of network communication. *Proceedings of the National Academy of Sciences, USA*, 111(2):833–838.
- Nishimoto, S., Vu, A. T., Naselaris, T., Benjamini, Y., Yu, B., and Gallant, J. L. (2011). Reconstructing visual experience from brain activity evoked by natural movies. *Current Biology*, 21:1–6.
- Nocedal, J. and Wright, S. J. (2006). *Numerical optimization*. Springer, New York, NY.
- Norman, K. A., Polyn, S. M., Detre, G. J., and Haxby, J. V. (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences*, 10(9):424–430.
- Opsahl, T., Agneessens, F., and Skvoretz, J. (2010). Node centrality in weighted networks: generalizing degree and shortest paths. *Social Networks*, 32:245–251.
- Owen, L. L. W., Chang, T. H., and Manning, J. R. (2021). High-level cognition during story listening is reflected in high-order dynamic correlations in neural activity patterns. *Nature Communications*, 12(1):5728.
- Park, H.-J., Friston, K. J., Pae, C., Park, B., and Razi, A. (2018). Dynamic effective connectivity in resting state fMRI. *NeuroImage*, 180:594–608.
- Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2:559–572.
- Pereira, F., Lou, B., Pritchett, B., Ritter, S., Gershman, S. J., Kanwisher, N., Botvinick, M., and Fedorenko, E. (2018). Toward a universal decoder of linguistic meaning from brain activation. *Nature Communications*, 9(963):1–13.
- Preti, M. G., Bolton, T. A. W., and Van De Ville, D. (2017). The dynamic functional connectome: state-of-the-art and perspectives. *NeuroImage*, 160:41–54.

- Randlett, O., Wee, C. L., Naumann, E. A., Nnaemeka, O., Schoppik, D., Fitzgerald, J. E., Portugues, R., Lacoste, A. M. B., Riegler, C., Engert, F., and Schier, A. F. (2015). Whole-brain activity mapping onto a zebrafish brain atlas. *Nature Methods*, 12(11):1039–1046.
- Rao, C. R. (1982). Diversity and dissimilarity coefficients: a unified approach. *Theoretical Population Biology*, 21(1):24–43.
- Rasmussen, C. E. (2006). *Gaussian processes for machine learning*. MIT Press.
- Reddy, P. G., Betzel, R. F., Khambhati, A. N., Shah, P., Kini, L., Litt, B., Lucas, T. H., Davis, K. A., and Bassett, D. S. (2018). Genetic and neuroanatomical support for functional brain network dynamics in epilepsy authors genetic and neuroanatomical support for functional brain network dynamics in epilepsy. *arXiv*, 1809.03934.
- Reimann, M. W., Nolte, M., Scolamiero, M., Turner, K., Perin, R., Chindemi, G., Dłotko, P., Levi, R., Hess, K., and Markram, H. (2017). Cliques of neurons bound into cavities provide a missing link between structure and function. *Frontiers in Computational Neuroscience*, 11(48):1–16.
- Ricotta, C. and Szeidl, L. (2006). Towards a unifying approach to diversity measures: bridging the gap between the Shannon entropy and Rao's quadratic index. *Theoretical Population Biology*, 70(3):237–243.
- Roy, D. S., Park, Y.-G., Ogawa, S. K., Cho, J. H., Choi, H., Kamensky, L., Martin, J., Chung, K., and Tonegawa, S. (2019). Brain-wide mapping of contextual fear memory engram ensembles supports the dispersed engram complex hypothesis. *bioRxiv*, doi.org/10.1101/668483.
- Rubin, T. N., Kyoejo, O., Gorgolewski, K. J., Jones, M. N., Poldrack, R. A., and Yarkoni, T. (2017). Decoding brain activity using a large-scale probabilistic functional-anatomical atlas of human cognition. *PLoS Computational Biology*, 13(10):e1005649.
- Rubinov, M. and Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. *NeuroImage*, 52:1059–1069.
- Sarvas, J. (1987). Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. *Physics in Medicine and Biology*, 32(1):11–22.
- Sawyer, R. J. (1995). *The terminal experiment*. Harper Prism.
- Scangos, K. W., Khambhati, A. N., Daly, P. M., Owen, L. W., Manning, J. R., Ambrose, J. B., Austin, E., Dawes, H. E., Krystal, A. D., and Chang, E. F. (2021). Distributed subnetworks of depression defined by direct intracranial neurophysiology. *Frontiers in Human Neuroscience*, 15:561.
- Scheinost, D., Noble, S., Horien, C., Greene, A. S., Lake, E. M., Salehi, M., Gao, S., Shen, X., O'Connor, D., Barron, D. S., Yip, S. W., Rosenberg, M. D., and Constable, R. T.

- (2019). Ten simple rules for predictive modeling of individual differences in neuroimaging. *NeuroImage*, 193:35–45.
- Schreiber, T. (2000). Measuring information transfer. *Physical Review Letters*, 85(2):461–464.
- Sederberg, P. B., Kahana, M. J., Howard, M. W., Donner, E. J., and Madsen, J. R. (2003). Theta and gamma oscillations during encoding predict subsequent recall. *The Journal of Neuroscience*, 23(34):10809–10814.
- Sederberg, P. B., Schulze-Bonhage, A., Madsen, J. R., Bromfield, E. B., Litt, B., Brandt, A., and Kahana, M. J. (2007a). Gamma oscillations distinguish true from false memories. *Psychological Science*, 18(11):927–932.
- Sederberg, P. B., Schulze-Bonhage, A., Madsen, J. R., Bromfield, E. B., McCarthy, D. C., Brandt, A., Tully, M. S., and Kahana, M. J. (2007b). Hippocampal and neocortical gamma oscillations predict memory formation in humans. *Cerebral Cortex*, 17(5):1190–1196.
- Sejnowski, T. J., Churchland, P. S., and Movshon, J. A. (2014). Putting big data to good use in neuroscience. *Nature Neuroscience*, 17:1440–1441.
- Shappell, H., Caffo, B. S., Pekar, J. J., and Lindquist, M. A. (2019). Improved state change estimation in dynamic functional connectivity using hidden semi-Markov models. *NeuroImage*, 191:243–257.
- Sherman, M. A., Lee, S., Law, R., Haegens, S., Thorn, C. A., Hämäläinen, M. S., Moore, C. I., and Jones, S. R. (2016). Neural mechanisms of transient neocortical beta rhythms: converging evidence from humans, computational modeling, monkeys, and mice. *Proceedings of the National Academy of Sciences, USA*, 113(33):E4885–E4894.
- Simony, E. and Chang, C. (2020). Analysis of stimulus-induced brain dynamics during naturalistic paradigms. *NeuroImage*, 216:116461.
- Simony, E., Honey, C. J., Chen, J., and Hasson, U. (2016). Dynamic reconfiguration of the default mode network during narrative comprehension. *Nature Communications*, 7(12141):1–13.
- Sizemore, A. E., Giusti, C., Kahn, A., Vettel, J. M., Betzel, R. F., and Bassett, D. S. (2018). Cliques and cavities in the human connectome. *Journal of Computational Neuroscience*, 44(1):115–145.
- Snyder, A. Z. (1991). Dipole source localization in the study of EP generators: a critique. *Electroencephalography and Clinical Neurophysiology*, 80(4):321–325.
- Solomon, E. A., Gross, R., Lega, B., Sperling, M. R., Worrell, G., Sheth, S. A., Zaghloul, K. A., Jobst, B. C., Stein, J. M., Das, S., Gorniak, R., Inman, C., Seger, S., Kragel,

- J. E., Rizzuto, D. S., and Kahana, M. J. (2018). MTL functional connectivity predicts stimulation-induced theta power. *Nature Communications*, In press.
- Solomon, S. H., Medaglia, J. D., and Thompson-Schill, S. L. (2019). Implementing a concept network model. *Behavior Research Methods*, 51:1717–1736.
- Song, H., Finn, E. S., and Rosenberg, M. D. (2021). Neural signatures of attentional engagement during narratives and its consequences for event memory. *Proceedings of the National Academy of Sciences*, 118(33).
- Sonkusare, S., Breakspear, M., and Guo, C. (2019). Naturalistic stimuli in neuroscience: Critically acclaimed. *Trends in Cognitive Sciences*, 23(8):699–714.
- Spearman, C. (1904). General intelligence, objectively determined and measured. *American Journal of Psychology*, 15:201–292.
- Sporns, O. and Betzel, R. F. (2016). Modular brain networks. *Annual Review of Psychology*, 67:613–640.
- Sporns, O. and Honey, C. J. (2006). Small worlds inside big brains. *Proceedings of the National Academy of Sciences, USA*, 103(51):19219–19220.
- Talairach, J. and Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain*. Verlag, Stuttgart, Germany.
- Tavor, I., Jones, O. P., Mars, R. B., Smith, S. M., Behrens, T. E., and Jbabdi, S. (2016). Task-free MRI predicts individual differences in brain activity during task performance. *Science*, 352(6282):216–220.
- Thompson, W. H., Richter, C. G., Plavén-Sigray, P., and Fransson, P. (2018). Simulations to benchmark time-varying connectivity methods for fMRI. *PLoS Computational Biology*, 14(5):e1006196.
- Tipping, M. E. and Bishop, C. M. (1999). Probabilistic principal component analysis. *Journal of Royal Statistical Society, Series B*, 61(3):611–622.
- Toker, D. and Sommer, F. T. (2019). Information integration in large brain networks. *PLoS Computational Biology*, 15(2):e1006807.
- Tomasi, D. and Volkow, N. D. (2011). Association between functional connectivity hubs and brain networks. *Cerebral Cortex*, 21:2003–2013.
- Tong, F. and Pratte, M. S. (2012). Decoding patterns of human brain activity. *Annual Review of Psychology*, 63:483–509.
- Turk-Browne, N. B. (2013). Functional interactions as big data in the human brain. *Science*, 342:580–584.

- van der Maaten, L. J. P. and Hinton, G. E. (2008). Visualizing high-dimensional data using t-SNE. *Journal of Machine Learning Research*, 9:2579–2605.
- Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E. J., Yacoub, E., and Ugurbil, K. (2013). The wu-minn human connectome project: An overview. *NeuroImage*, 80:62–79.
- Vanderwal, T., Eilbott, J., and Castellanos, F. X. (2019). Movies in the magnet: Naturalistic paradigms in developmental functional neuroimaging. *Developmental cognitive neuroscience*, 36:100600–100600.
- Vidaurre, D., Abeysuriya, R., Becker, R., Quinn, A. J., Alfaro-Almagro, F., Smith, S. M., and Woolrich, M. W. (2018). Discovering dynamic brain networks from big data in rest and task. *NeuroImage*, 180:646–656.
- Weidemann, C. T., Kragel, J. E., Lega, B. C., Worrell, G. A., Sperling, M. R., Sharan, A. D., Jobst, B. C., Khadjevand, F., Davis, K. A., Wanda, P. A., Kadel, A., Rizzuto, D. S., and Kahana, M. J. (2019). Neural activity reveals interactions between episodic and semantic memory systems during retrieval. *Journal of Experimental Psychology: General*, 148(1):1–12.
- Wu, H. C., Nagasawa, T., Brown, E. C., Juhasz, C., Rothermel, R., Hoechstetter, K., Shah, A., Mittal, S., Fuerst, D., Sood, S., and Asano, E. (2011). Gamma-oscillations modulated by picture naming and word reading: intracranial recording in epileptic patients. *Clinical Neurophysiology*, 122(10):1929–1942.
- Yeo, B. T. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., Roffman, J. L., Smoller, J. W., Zollei, L., Polimieni, 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(3):1125–1165.
- Zalesky, A., Fornito, A., Harding, I. H., Cocchi, L., Yücel, M., Pantelis, C., and Bullmore, E. T. (2010). Whole-brain anatomical networks: does the choice of nodes matter? *NeuroImage*, 50(3):970–983.
- Zamani Esfahlani, F., Bertolero, M. A., Bassett, D. S., and Betzel, R. F. (2020). Space-independent community and hub structure of functional brain networks. *NeuroImage*, 211:116612.
- Zar, J. H. (2010). *Biostatistical analysis*. Prentice-Hall.
- Zheng, M., Allard, A., Hagmann, P., and Serrano, M. . A. (2019). Geometric renormalization unravels self-similarity of the multiscale human connectome. *arXiv*, 1904.11793.
- Zou, Y., Donner, R. V., Marwan, N., Donges, J. F., and Kurths, J. (2019). Complex network approaches to nonlinear time series analysis. *Physics Reports*, 787:1–97.