

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

3 Lucy L. W. Owen¹, Thomas H. Chang^{1,2}, and Jeremy R. Manning^{1,†}

¹Department of Psychological and Brain Sciences,
Dartmouth College, Hanover, NH

²Amazon.com, Seattle, WA

†Address correspondence to jeremy.r.manning@dartmouth.edu

4 October 23, 2020

5 **Abstract**

6 Our thoughts arise from coordinated patterns of interactions between brain structures that change with
7 our ongoing experiences. High-order dynamic correlations in neural activity patterns reflect different
8 subgraphs of the brain’s connectome that display homologous lower-level dynamic correlations. We tested
9 the hypothesis that high-level cognition is reflected in high-order dynamic correlations in brain activity
10 patterns. We developed an approach to estimating high-order dynamic correlations in timeseries data, and
11 we applied the approach to neuroimaging data collected as human participants either listened to a ten-
12 minute story, listened to a temporally scrambled version of the story, or underwent a resting state scan. We
13 trained across-participant pattern classifiers to decode (in held-out data) when in the session each neural
14 activity snapshot was collected. We found that classifiers trained to decode from high-order dynamic
15 correlations yielded the best performance on data collected as participants listened to the (unscrambled)
16 story. By contrast, classifiers trained to decode data from scrambled versions of the story or during
17 the resting state scan yielded the best performance when they were trained using first-order dynamic
18 correlations or non-correlational activity patterns. We suggest that as our thoughts become more complex,
19 they are reflected in higher-order patterns of dynamic network interactions throughout the brain.

20 **Introduction**

21 A central goal in cognitive neuroscience is to elucidate the *neural code*: the mapping between (a) mental
22 states or cognitive representations and (b) neural activity patterns. One means of testing models of the
23 neural code is to ask how accurately that model is able to “translate” neural activity patterns into known
24 (or hypothesized) mental states or cognitive representations (e.g., Haxby et al., 2001; Huth et al., 2016, 2012;
25 Kamitani & Tong, 2005; Mitchell et al., 2008; Nishimoto et al., 2011; Norman et al., 2006; Pereira et al.,
26 2018; Tong & Pratte, 2012). Training decoding models on different types of neural features (Fig. 1a) can
27 also help to elucidate which specific aspects of neural activity patterns are informative about cognition—
28 and, by extension, which types of neural activity patterns might comprise the neural code. For example,
29 prior work has used region of interest analyses to estimate the anatomical locations of specific neural
30 representations (e.g., Etzel et al., 2009), or to compare the relative contributions to the neural code of

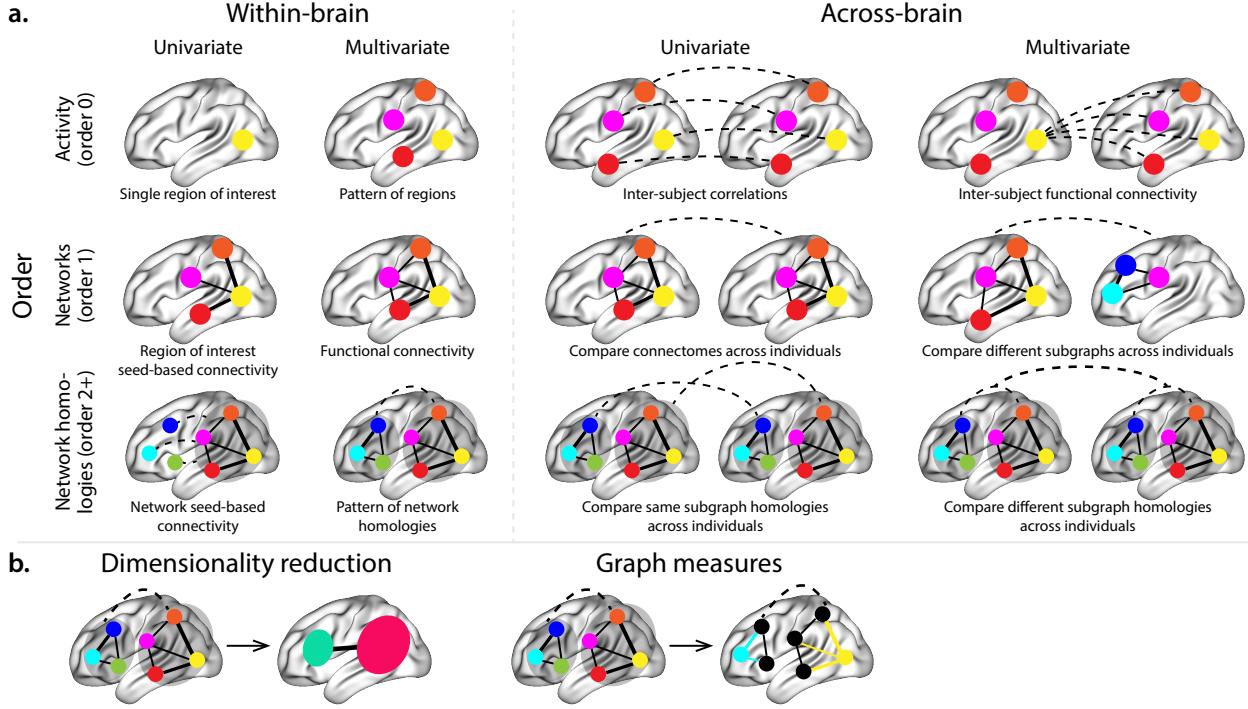


Figure 1: Neural patterns. a. A space of neural features. Within-brain analyses are carried out within a single brain, whereas across-brain analyses compare neural patterns across two or more individuals' brains. Univariate analyses characterize the activities of individual units (e.g., nodes, small networks, hierarchies of networks, etc.), whereas multivariate analyses characterize the patterns of activities 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.

multivariate activity patterns versus dynamic correlations between neural activity patterns (e.g., Fong et al., 2019; Manning et al., 2018). An emerging theme in this literature is that cognition is mediated by dynamic interactions between brain structures (Bassett et al., 2006; Demertzi et al., 2019; Lurie et al., 2018; Mack et al., 2017; Preti et al., 2017; Solomon et al., 2019; Sporns & Honey, 2006; Turk-Browne, 2013; Zou et al., 2019). Studies of the neural code to date have primarily focused on univariate or multivariate neural patterns (for review see Norman et al., 2006), or (more recently) on patterns of dynamic first-order correlations (i.e., interactions between pairs of brain structures; Demertzi et al., 2019; Fong et al., 2019; Lurie et al., 2018; Manning et al., 2018; Preti et al., 2017; Zou 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 (e.g., see Reimann et al., 2017)? Second-order correlations reflect *homologous* patterns of correlation.

41 In other words, if the dynamic patterns of correlations between two regions, A and B , are similar to those
42 between two other regions, C and D , this would be reflected in the second-order correlations between ($A-B$)
43 and ($C-D$). In this way, second-order correlations identify similarities and differences between subgraphs
44 of the brain's connectome. Analogously, third-order correlations reflect homologies between second-order
45 correlations—i.e., homologous patterns of homologous interactions between brain regions. More generally,
46 higher-order correlations reflect homologies between patterns of lower-order correlations. We can then ask:
47 which “orders” of interaction are most reflective of high-level cognitive processes?

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

59 To gain insights into the above aspects of the neural code, we developed a computational framework
60 for estimating dynamic high-order correlations in timeseries data. This framework provides an important
61 advance, in that it enables us to examine patterns of higher-order correlations that are computationally
62 intractable to estimate via conventional methods. Given a multivariate timeseries, our framework provides
63 timepoint-by-timepoint estimates of the first-order correlations, second-order correlations, and so on. Our
64 approach combines a kernel-based method for computing dynamic correlations in timeseries data with a di-
65 mensionality reduction step (Fig. 1b) that projects the resulting dynamic correlations into a low-dimensional
66 space. We explored two dimensionality reduction approaches: principle components analysis (PCA; Pear-
67 son, 1901), which preserves an approximately invertible transformation back to the original data (e.g., this
68 follows related approaches taken by Gonzalez-Castillo et al., 2019; McIntosh & Jirsa, 2019; Toker & Som-
69 mer, 2019); and a second non-invertible algorithm that explored patterns in eigenvector centrality (Landau,
70 1895). This latter approach characterizes correlations between each feature dimension’s relative *position* in
71 the network in favor of the specific activity histories of different features (also see Betzel et al., 2019; Reimann
72 et al., 2017; Sizemore et al., 2018).

73 We validated our approach using synthetic data where the underlying correlations were known. We

74 then applied our framework to a neuroimaging dataset collected as participants listened to either an audio
75 recording of a ten-minute story, listened to a temporally scrambled version of the story, or underwent a
76 resting state scan (Simony et al., 2016). We used a subset of the data to train across-participant classifiers
77 to decode listening times using a blend of neural features (comprising neural activity patterns, as well as
78 different orders of dynamic correlations between those patterns that were inferred using our computational
79 framework). We found that both the PCA-based and eigenvector centrality-based approaches yielded neural
80 patterns that could be used to decode accurately (i.e., well above chance). Both approaches also yielded
81 the best decoding accuracy for data collected during (intact) story listening when high-order (PCA: second-
82 order; eigenvector centrality: fourth-order) dynamic correlation patterns were included as features. When
83 we trained classifiers on the scrambled stories or resting state data, only lower-order dynamic patterns were
84 informative to the decoders. Taken together, our results indicate that high-level cognition is supported by
85 high-order dynamic patterns of communication between brain structures.

86 Results

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

97 Recovering known dynamic correlations from synthetic data

98 Recovering dynamic first-order correlations

99 We generated synthetic datasets that differed in how the underlying first-order correlations changed over
100 time. For each dataset, we applied Equation 4 with a variety of kernel shapes and widths. We assessed how
101 well the true underlying correlations at each timepoint matched the recovered correlations (Fig. 2). For every
102 kernel and dataset we tested, our approach recovered the correlation dynamics we embedded into the data.

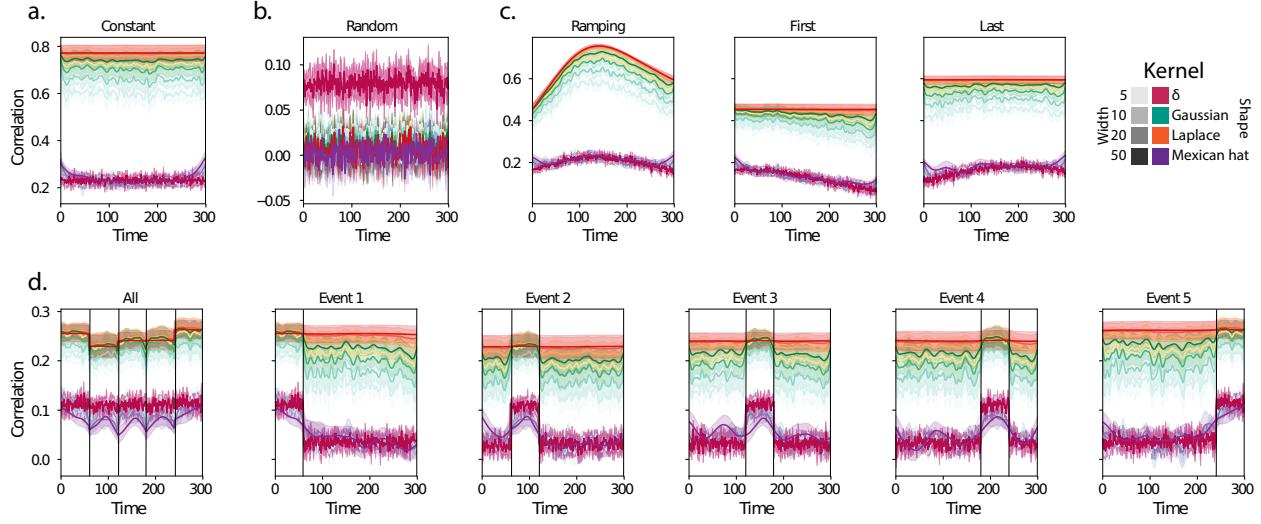


Figure 2: Recovering known dynamic first-order correlations from synthetic data. Each panel displays the average correlations between the vectorized upper triangles of the recovered correlation matrix at each timepoint and either the true underlying correlation at each timepoint or a reference correlation matrix. (The averages are taken across 10 different randomly generated synthetic datasets of the given category.) Error ribbons denote 95% confidence intervals (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.

103 However, the quality of these recoveries varied across different synthetic datasets in a kernel-dependent
 104 way.

105 In general, wide monotonic kernel shapes (Laplace, Gaussian), and wider kernels (within a shape),
 106 performed best when the correlations varied gradually from moment-to-moment (Figs. 2a, c, and d). In the
 107 extreme, as the rate of change in correlations approaches 0 (Fig. 2a), an infinitely wide kernel would exactly
 108 recover the Pearson's correlation (e.g., compare Eqns. 1 and 4).

109 When the correlation dynamics were unstructured in time (Fig. 2b), a Dirac δ kernel (infinitely narrow)
 110 performed best. This is because, when every timepoint's correlations are independent of the correlations at
 111 every other timepoint, averaging data over time dilutes the available signal. Following a similar pattern,
 112 holding kernel shape fixed, narrower kernel parameters better recovered randomly varying correlations.

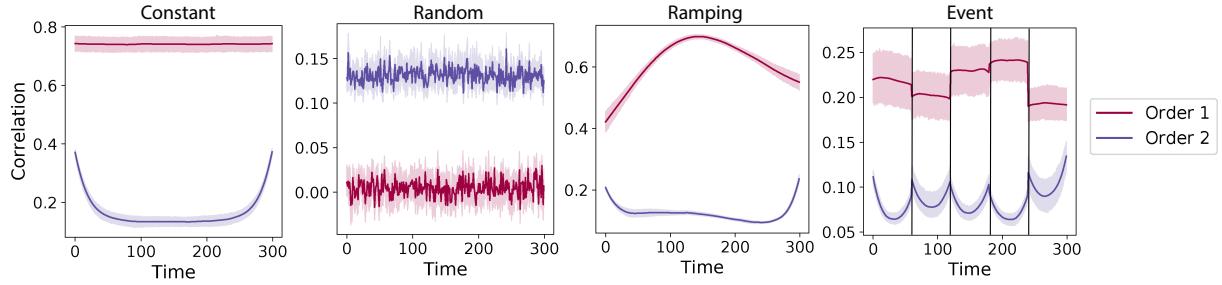


Figure 3: Recovery of simulated first-order and second-order dynamic correlations. Each panel displays the average correlations between the vectorized upper triangles of the recovered correlation matrix and the simulated correlation matrix at each timepoint and for each synthetic dataset. (The averages are taken across 10 different randomly generated synthetic datasets of the given category.) Error ribbons denote 95% confidence intervals (taken across datasets). For a complete description of each synthetic dataset, see *Synthetic data: simulating dynamic higher-order correlations*. All results in this figure used a Laplace kernel (width = 20). **a. Constant correlations.** These first and second order dynamic correlation were simulated from datasets have a stable (unchanging) underlying correlation matrix. **b. Random correlations.** These first and second order dynamic correlation were simulated from datasets are generated using a new independently drawn correlation matrix at each new timepoint. **c. Ramping correlations.** These first and second order dynamic correlation were simulated from datasets are generated by smoothly varying the underlying correlations between the randomly drawn correlation matrices at the first and last timepoints and displays the correlations between the recovered dynamic correlations and the simulalted correlation first and second order correlation matrices. **d. Event-based correlations.** These first and second order dynamic correlation were simulated from datasets are each generated using five randomly drawn correlation matrices that each remain stable for a fifth of the total timecourse.

113 Recovering dynamic higher-order correlations

114 Following our approach to evaluating our ability to recover known dynamic first-order correlations from
 115 synthetic data, we generated an analogous second set of synthetic datasets that we designed to exhibit
 116 known dynamic first-order *and* second-order correlations (see *Synthetic data: simulating dynamic higher-*
 117 *order correlations*). We generated a total of 40 datasets that varied in how the first-order and second-order
 118 correlations changed over time. We then repeatedly applied Equation 4 using the overall best-performing
 119 kernel from our first-order tests (a Laplace kernel with a width of 20; Fig. 2) to assess how closely the
 120 recovered dynamic correlations matched the dynamic correlations we had embedded into the datasets.

121 Overall, we found that we could reliably recover both first-order and second-order correlations from
 122 the synthetic data (Fig. 3). When the correlations were stable for longer intervals, or changed gradually
 123 (constant, ramping, and event datasets), recovery performance was relatively high, and we were better able
 124 to recover dynamic first-order correlations than second-order correlations. We expected that this would
 125 happen, given that errors in our estimation procedure at lower orders necessarily propagate to higher orders
 126 (since lower-order correlations are used to estimate higher-order correlations). Interestingly, we also found
 127 that when the correlations were particularly *unstable* (random datasets), we better recovered second-order
 128 correlations.

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

134 **Cognitively relevant dynamic high-order correlations in fMRI data**

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

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

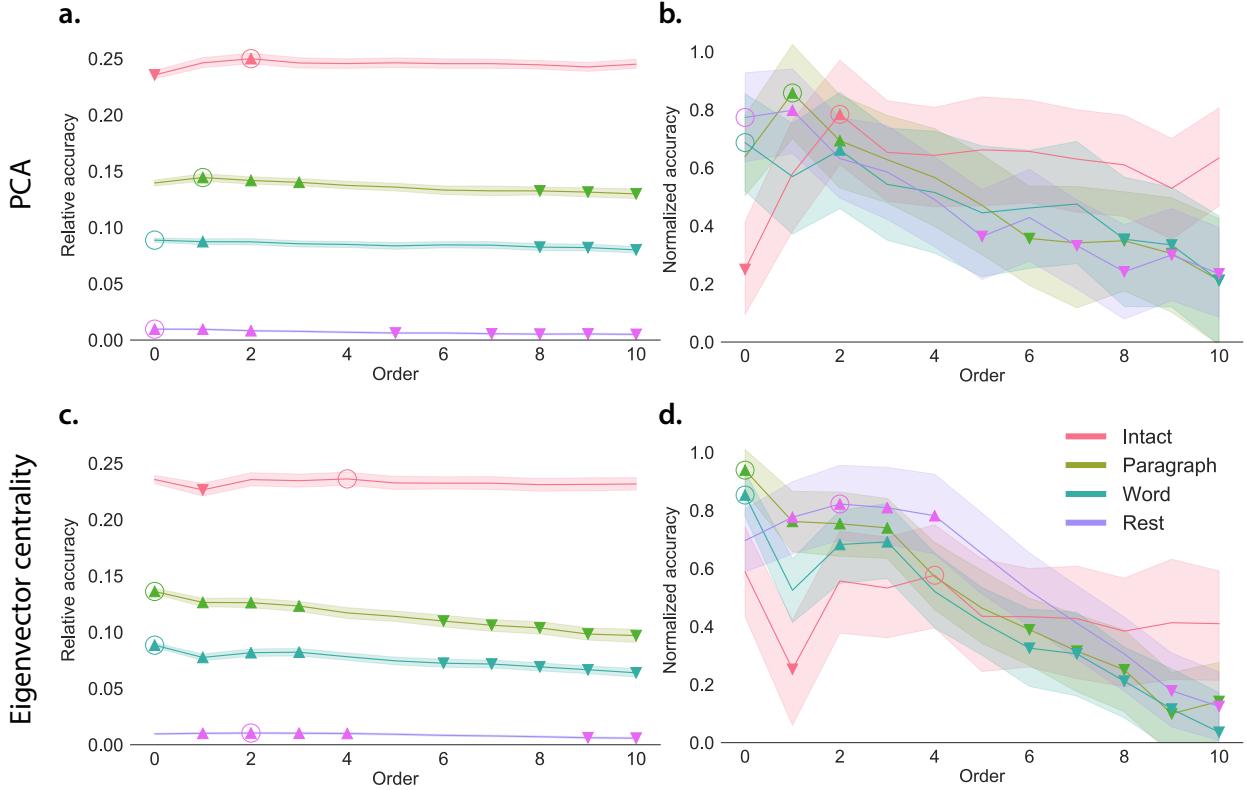


Figure 4: Across-participant timepoint decoding accuracy varies with correlation order and cognitive engagement. **a. Decoding accuracy as a function of order: PCA.** Order (*x*-axis) refers to the maximum order of dynamic correlations that were available to the classifiers (see *Feature weighting and 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: $\frac{1}{400}$). The error ribbons denote 95% confidence intervals 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 (upwards facing) or lower (downward facing) decoding accuracy than the mean of all other features (via a two-tailed test, thresholded at $p < 0.05$). Figure 5 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). 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.

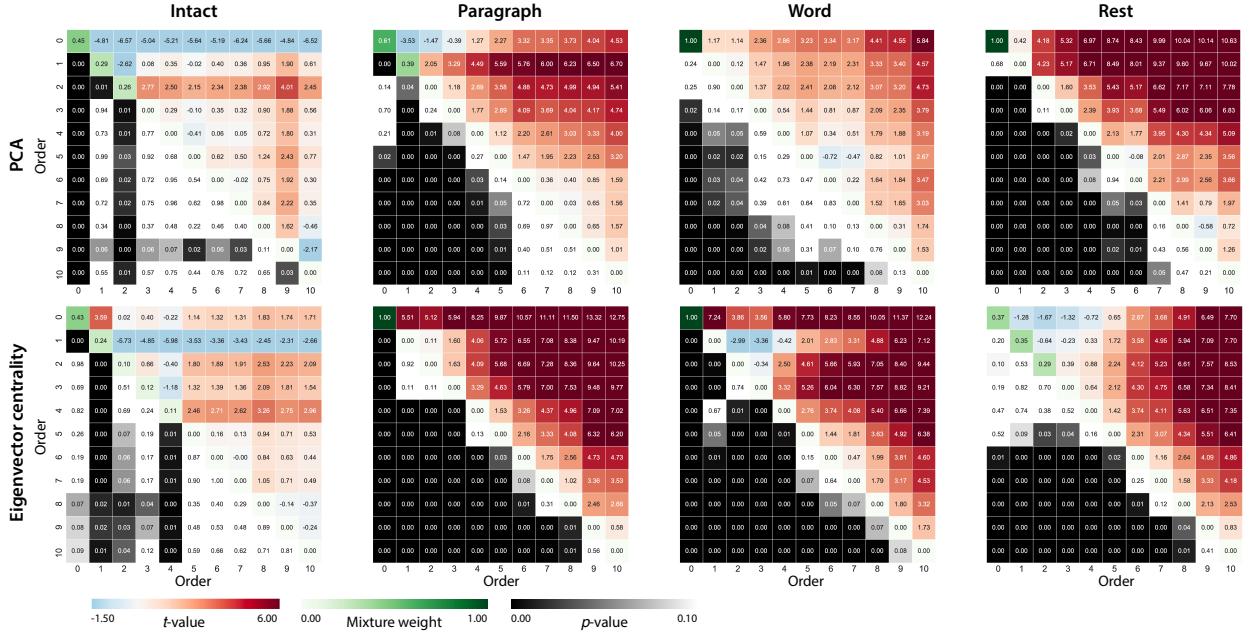


Figure 5: Statistical summary of decoding accuracies for different neural features. Each column displays decoding results for one experimental condition (intact, paragraph, word, and rest). We considered dynamic activity patterns (order 0) and dynamic correlations at different orders (order > 0). We 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 map, 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*).

Our approach to estimating dynamic high-order correlations entails mapping the high-dimensional feature space of correlations (represented by a T by $O(K^2)$ 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. 4a,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. 4c,d).

Both sets of temporal decoding analyses yielded qualitatively similar results for the auditory (non-rest) conditions of the experiment (Fig. 4: pink, green, and teal lines; Fig. 5: 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

172 of analyses yielded different decoding results on resting state data (Fig. 4: purple lines; Fig. 5: rightmost
173 column). We note that while the resting state times could be decoded reliably, the accuracies were only very
174 slightly above chance. We speculate that the decoders might have picked up on attentional drift, boredom,
175 or tiredness; we hypothesize that these all increased throughout the resting state scan. The decoders might
176 be picking up on aspects of these loosely defined cognitive states that are common across individuals. The
177 PCA-based approach achieved the highest resting state decoding accuracy using order zero features (non-
178 correlational, activation-based), whereas the eigenvector centrality-based approach achieved the highest
179 resting state decoding accuracy using second-order correlations. Taken together, these analyses indicate
180 that high-level cognitive processing (while listening to the intact story) is reflected in the dynamics of high-
181 order correlations in brain activity, whereas lower-level cognitive processing (while listening to scrambled
182 versions of the story that lack rich meaning) is reflected in the dynamics of lower-order correlations and
183 non-correlational activity dynamics. Further, these patterns are associated both with the underlying activity
184 patterns (characterized using PCA) and also with the changing relative positions that different brain areas
185 occupy in their associated networks (characterized using eigenvector centrality).

186 Having established that patterns of high-order correlations are informative to decoders, we next won-
187 dered which specific networks of brain regions contributed most to these patterns. As a representative
188 example, we selected the kernel parameters that yielded decoding accuracies that best matched the average
189 accuracies across all of the kernel parameters we examined. Using Figure 4c as a template, the best-matching
190 kernel was a Laplace kernel with a width of 50 (Fig. 9d). We used this kernel to compute a single K by K
191 n^{th} -order DISFC matrix for each experimental condition. We then used Neurosynth (Rubin et al., 2017) to
192 compute the terms most highly associated with the most strongly correlated pairs of regions in each of these
193 matrices (Fig. 6; see *Reverse inference*).

194 For all of the story listening conditions (intact, paragraph, and word), we found that first- and second-
195 order correlations were most strongly associated with auditory and speech processing areas. During
196 intact story listening, third-order correlations reflected integration with visual areas, and fourth-order
197 correlations reflected integration with areas associated with high-level cognition and cognitive control,
198 such as the ventrolateral prefrontal cortex. However, during listening to temporally scrambled stories,
199 these higher-order correlations instead involved interactions with additional regions associated with speech
200 and semantic processing. By contrast, we found a much different set of patterns in the resting state data.
201 First-order resting state correlations were most strongly associated with regions involved in counting and
202 numerical understanding. Second-order resting state correlations were strongest in visual areas; third-order
203 correlations were strongest in task-positive areas; and fourth-order correlations were strongest in regions
204 associated with autobiographical and episodic memory. We carried out analogous analyses to create maps

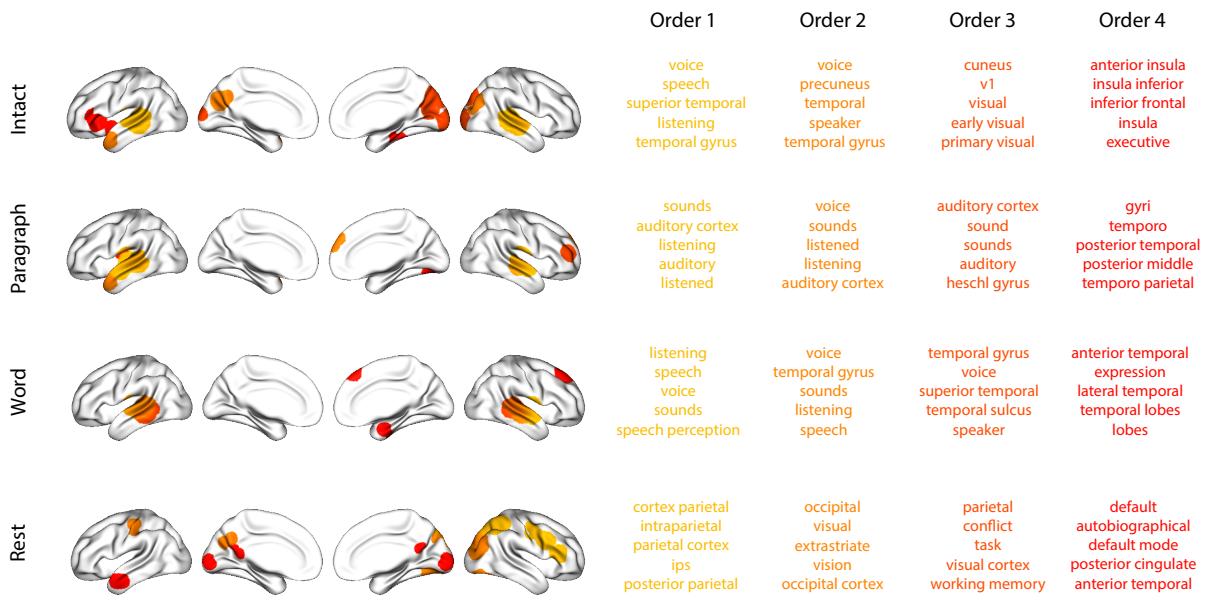


Figure 6: Top terms associated with the endpoints of the strongest correlations. Each color corresponds to one order of inter-subject functional correlations. To calculate the dynamic correlations, eigenvector centrality has been used to project the high-dimensional patterns of dynamic correlations onto 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. S1; paragraph: Fig. S2; word: Fig. S3; rest: Fig. S4).

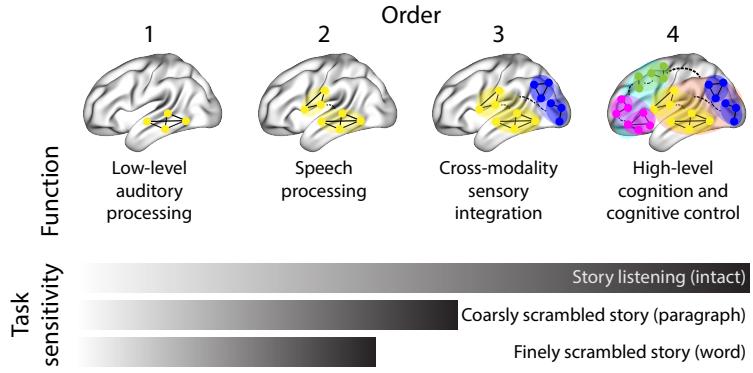


Figure 7: Proposed high-order network dynamics underlying high-level cognition during story listening. Schematic depicts higher orders of network interactions supporting higher-level aspects of cognitive processing. When tasks evoke richer, deeper, and/or higher-level processing, this is reflected in higher-order network interactions.

(and decode the top associated Neurosynth terms) for up to fifteenth-order correlations (Figs. S1, S2, S3, and S4). 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.

Discussion

We tested the hypothesis that high-level cognition is reflected in high-order brain network dynamics (e.g., see Reimann et al., 2017; Solomon et al., 2019). We examined high-order network dynamics in functional neuroimaging data collected during a story listening experiment. 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. 7).

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 (e.g., Chang & Glover, 2010; Demertzis et al., 2019; Fong et al., 2019; Gonzalez-Castillo et al., 2019; Liégeois et al., 2019; Lurie et al., 2018; Park et al., 2018; Preti et al., 2017; Roy et al., 2019; Turk-Browne, 2013; Zou et al., 2019). Our study extends this line of work by finding cognitively relevant *higher-order* network dynamics that reflect ongoing cognition. Our findings complement other work that uses graph theory and topology to characterize how

225 brain networks reconfigure during cognition (e.g., Bassett et al., 2006; Betzel et al., 2019; McIntosh & Jirsa,
226 2019; Reimann et al., 2017; Sizemore et al., 2018; Toker & Sommer, 2019; Zheng et al., 2019).

227 An open question not addressed by our study pertains to how different structures integrate incoming
228 information with different time constants. For example, one line of work suggests that the cortical surface
229 comprises a structured map such that nearby brain structures process incoming information at similar
230 timescales. Low-level sensory areas integrate information relatively quickly, whereas higher-level regions
231 integrate information relatively slowly (Baldassano et al., 2017; Chien & Honey, 2019; Hasson et al., 2015,
232 2008; Honey et al., 2012; Lerner et al., 2014, 2011). A similar hierarchy appears to play a role in predicting
233 future events (C. S. Lee et al., 2020). Other related work in human and mouse brains indicates that the
234 temporal response profile of a given brain structure may relate to how strongly connected that structure
235 is with other brain areas (Fallon et al., 2019). Further study is needed to understand the role of temporal
236 integration at different scales of network interaction, and across different anatomical structures. Import-
237 antly, our analyses do not speak to the physiological basis of higher-order dynamics, and could reflect
238 nonlinearities, chaotic patterns, non-stationarities, or multistability, etc. However, our decoding analyses
239 do indicate that higher-order dynamics are consistent across individuals, and therefore unlikely to be driven
240 by non-stimulus-driven dynamics which are unlikely to be similar across individuals.

241 Another potential limitation of our approach relates to recent work suggesting that the brain undergoes
242 rapid state changes, for example across event boundaries (e.g., Baldassano et al., 2017). Shappell et al.
243 (2019) used hidden semi-Markov models to estimate state-specific network dynamics (also see Vidaurre et
244 al., 2018). Our general approach might be extended by considering putative state transitions. For example,
245 rather than weighting all timepoints using a similar kernel (Eqn. 4), the kernel function could adapt on a
246 timepoint-by-timepoint basis such that only timepoints determined to be in the same “state” were given
247 non-zero weight.

248 Identifying high-order network dynamics associated with high-level cognition required several impor-
249 tant methods advances. First, we used kernel-based dynamic correlations to extended the notion of (static)
250 inter-subject functional connectivity (Simony et al., 2016) to a dynamic measure of inter-subject functional
251 connectivity (DISFC) that does not rely on sliding windows, and that may be computed at individual
252 timepoints. This allowed us to precisely characterize stimulus-evoked network dynamics that were similar
253 across individuals. Second, we developed a computational framework for efficiently and scalably estimat-
254 ing high-order dynamic correlations. Our approach uses dimensionality reduction algorithms and graph
255 measures to obtain low-dimensional embeddings of patterns of network dynamics. Third, we developed
256 an analysis framework for identifying robust decoding results by carrying out our analyses using a range
257 of parameter values and then identifying which results were robust to specific parameter choices.

258 **Concluding remarks**

259 The complex hierarchy of dynamic interactions that underlie our thoughts is perhaps the greatest mystery in
260 modern science. Methods for characterizing the dynamics of high-order correlations in neural data provide
261 a window into the neural basis of cognition. By showing that high-level cognition is reflected in high-order
262 network dynamics, we have elucidated the next step on the path towards understanding the neural basis
263 of cognition.

264 **Methods**

265 Our general approach to efficiently estimating high-order dynamic correlations comprises four general
266 steps (Fig. 8). First, we derive a kernel-based approach to computing dynamic pairwise correlations in
267 a T (timepoints) by K (features) multivariate timeseries, \mathbf{X}_0 . This yields a T by $O(K^2)$ matrix of dynamic
268 correlations, \mathbf{Y}_1 , where each row comprises the upper triangle and diagonal of the correlation matrix at
269 a single timepoint, reshaped into a row vector (this reshaped vector is $(\frac{K^2-K}{2} + K)$ -dimensional). Second,
270 we apply a dimensionality reduction step to project the matrix of dynamic correlations back onto a K -
271 dimensional space. This yields a T by K matrix, \mathbf{X}_1 , that reflects an approximation of the dynamic correlations
272 reflected in the original data. Third, we use repeated applications of the kernel-based dynamic correlation
273 step to \mathbf{X}_n and the dimensionality reduction step to the resulting \mathbf{Y}_{n+1} to estimate high-order dynamic
274 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
275 reflects the dynamic correlations between the columns of \mathbf{X}_n . In this way, we refer to n as the *order* of the
276 timeseries, where \mathbf{X}_0 (order 0) denotes the original data and \mathbf{X}_n denotes (approximated) n^{th} -order dynamic
277 correlations between the columns of \mathbf{X}_0 . Finally, we use a cross-validation-based decoding approach to
278 evaluate how well information contained in a given order (or weighted mixture of orders) may be used
279 to decode relevant cognitive states. If including a given \mathbf{X}_n in the feature set yields higher classification
280 accuracy on held-out data, we interpret this as evidence that the given cognitive states are reflected in
281 patterns of n^{th} -order correlations.

282 All of the code used to produce the figures and results in this manuscript, along with links to the
283 corresponding datasets, may be found at github.com/ContextLab/timecorr-paper. In addition, we have
284 released a Python toolbox for computing dynamic high-order correlations in timeseries data; our toolbox
285 may be found at timecorr.readthedocs.io.

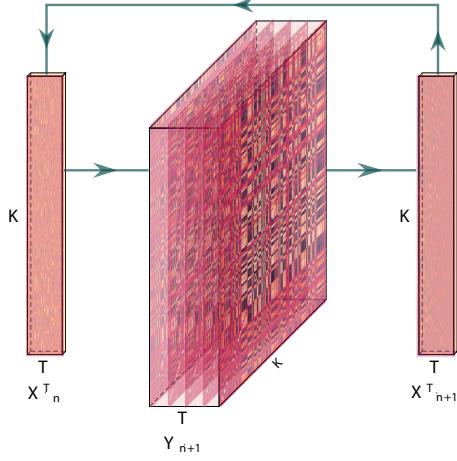


Figure 8: **Estimating dynamic high-order correlations.** Given a T by K matrix of multivariate timeseries data, \mathbf{X}_n (where $n \in \mathbb{N}, n \geq 0$), we use Equation 4 to compute a timeseries of K by K correlation matrices, \mathbf{Y}_{n+1} . We then approximate \mathbf{Y}_{n+1} with the 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.

286 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):

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

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

$$\sigma_{\mathbf{X}(\cdot, k)}^2 = \frac{1}{T} \sum_{t=1}^T (\mathbf{X}(t, k) - \bar{\mathbf{X}}(\cdot, k))^2 \quad (3)$$

- 287 We can generalize this formula to compute time-varying correlations by incorporating a *kernel function* that
 288 takes a time t as input, and returns how much the observed data at each timepoint $\tau \in [-\infty, \infty]$ contributes
 289 to the estimated instantaneous correlation at time t (Fig. 9; also see Allen et al., 2012, for a similar approach).

290

Given a kernel function $\kappa_t(\cdot)$ for timepoint t , evaluated at timepoints $\tau \in [1, \dots, T]$, we can update the

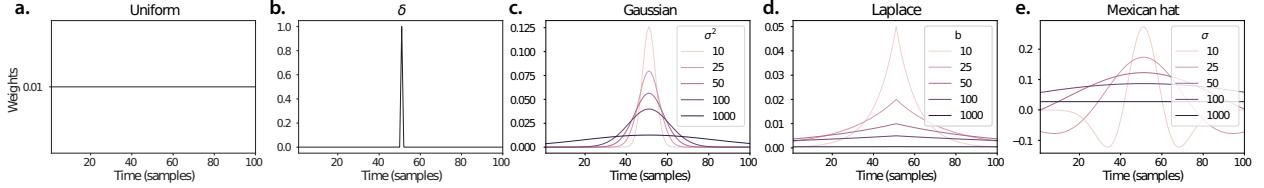


Figure 9: Examples of kernel functions. Each panel displays per-timepoint weights for a kernel centered at $t = 50$, evaluated at 100 timepoints ($\tau \in [1, \dots, 100]$). **a. Uniform kernel.** The weights are timepoint-invariant; observations at all timepoints are weighted equally, and do not change 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.

static correlation formula in Equation 1 to estimate the *instantaneous correlation* at timepoint t :

$$\text{timecorr}_{\kappa_t}(\mathbf{X}(\cdot, i), \mathbf{X}(\cdot, j)) = \frac{\sum_{\tau=1}^T (\mathbf{X}(\tau, i) - \tilde{\mathbf{X}}_{\kappa_t}(\cdot, i))(\mathbf{X}(\tau, j) - \tilde{\mathbf{X}}_{\kappa_t}(\cdot, j))}{\sqrt{\sum_{\tau=1}^T \tilde{\sigma}_{\kappa_t}^2(\mathbf{X}(\cdot, i)) \tilde{\sigma}_{\kappa_t}^2(\mathbf{X}(\cdot, j))}}, \text{ where} \quad (4)$$

$$\tilde{\mathbf{X}}_{\kappa_t}(\cdot, k) = \sum_{\tau=1}^T \kappa_t(\tau) \mathbf{X}(\tau, k), \quad (5)$$

$$\tilde{\sigma}_{\kappa_t}^2(\mathbf{X}(\cdot, k)) = \sum_{\tau=1}^T (\mathbf{X}(\tau, k) - \tilde{\mathbf{X}}_{\kappa_t}(\cdot, k))^2. \quad (6)$$

291 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
 292 the kernel κ_t . We evaluate Equation 4 in turn for each pair of columns in \mathbf{X} and for kernels centered on each
 293 timepoint in the timeseries, respectively, to obtain a T by K by K timeseries of dynamic correlations, \mathbf{Y} . For
 294 convenience, we then reshape the upper triangles and diagonals of each timepoint's symmetric correlation
 295 matrix into a row vector to obtain an equivalent T by $\left(\frac{K^2-K}{2} + K\right)$ matrix.

296 Dynamic inter-subject functional connectivity (DISFC)

Equation 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* (ISFC; Simony & Chang, 2020; Simony et al., 2016) to compute the stimulus-driven correlations reflected in the multi-participant dataset at a given timepoint t using:

$$\bar{\mathbf{C}}(t) = M \left(R \left(\frac{1}{2P} \sum_{p=1}^P Z(\mathbf{Y}_{n+1}^p(t))^\top + Z(\mathbf{Y}_{n+1}^p(t)) \right) \right), \quad (7)$$

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

$$Z(r) = \frac{\log(1+r) - \log(1-r)}{2}, \quad (8)$$

R is the inverse of Z :

$$R(z) = \frac{\exp(2z-1)}{\exp(2z+1)}, \quad (9)$$

and $\mathbf{Y}_{n+1}^p(t)$ denotes the correlation matrix at timepoint t (Eqn. 4) between each column of \mathbf{X}_n^p and each column of the average \mathbf{X}_n from all *other* participants, $\bar{\mathbf{X}}_n^p$:

$$\bar{\mathbf{X}}_n^p = \frac{1}{P-1} \sum_{q \in \setminus p} \mathbf{X}_n^q, \quad (10)$$

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

299 Low-dimensional representations of dynamic correlations

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

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

306 The modern toolkit of dimensionality reduction algorithms include Principal Components Analysis (PCA;
307 Pearson, 1901), Probabilistic PCA (PPCA; Tipping & Bishop, 1999), Exploratory Factor Analysis (EFA;

308 Spearman, 1904), Independent Components Analysis (ICA; Comon et al., 1991; Jutten & Herault, 1991),
309 *t*-Stochastic Neighbor Embedding (*t*-SNE; van der Maaten & Hinton, 2008), Uniform Manifold Approximation
310 and Projection (UMAP; McInnes et al., 2018), non-negative matrix factorization (NMF; D. D. Lee
311 & Seung, 1999), Topographic Factor Analysis (TFA; Manning et al., 2014), Hierarchical Topographic Factor
312 analysis (HTFA; Manning et al., 2018), Topographic Latent Source Analysis (TLSA; Gershman et al.,
313 2011), dictionary learning (J. Mairal et al., 2009; J. B. Mairal et al., 2009), and deep auto-encoders (Hinton
314 & Salakhutdinov, 2006), among others. While complete characterizations of each of these algorithms is
315 beyond the scope of the present manuscript, the general intuition driving these approaches is to compute
316 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
317 approaches place different constraints on what properties \mathbf{X} must satisfy and which aspects of the data are
318 compared (and how) in order to optimize how well \mathbf{X} approximates \mathbf{Y} .

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

328 **Graph measure approaches to computing \mathbf{X}_n**

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

337 In general, graph measures take as input a matrix of interactions (e.g., using the above notation, a K
338 by K correlation matrix or binarized correlation matrix reconstituted from a single timepoint’s row of \mathbf{Y}),

339 and return as output a set of K measures describing how each node (feature) sits within that correlation
340 matrix with respect to the rest of the population. Widely used measures include betweenness centrality (the
341 proportion of shortest paths between each pair of nodes in the population that involves the given node
342 in question; e.g., Barthélemy, 2004; Freeman, 1977; Geisberger et al., 2008; Newman, 2005; Opsahl et al.,
343 2010); diversity and dissimilarity (characterizations of how differently connected a given node is from others
344 in the population; e.g., Lin, 2009; Rao, 1982; Ricotta & Szeidl, 2006); eigenvector centrality and pagerank
345 centrality (measures of how influential a given node is within the broader network; e.g., Bonacich, 2007;
346 Halu et al., 2013; Lohmann et al., 2010; Newman, 2008); transfer entropy and flow coefficients (a measure of
347 how much information is flowing from a given node to other nodes in the network; e.g., Honey et al., 2007;
348 Schreiber, 2000); k -coreness centrality (a measure of the connectivity of a node within its local subgraph; e.g.,
349 Alvarez-Hamelin et al., 2005; Christakis & Fowler, 2010); within-module degree (a measure of how many
350 connections a node has to its close neighbors in the network; e.g., Rubinov & Sporns, 2010); participation
351 coefficient (a measure of the diversity of a node's connections to different subgraphs in the network; e.g.,
352 Rubinov & Sporns, 2010); and subgraph centrality (a measure of a node's participation in all of the network's
353 subgraphs; e.g., Estrada & Rodríguez-Velázquez, 2005); among others.

354 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
355 characterizes the corresponding graph properties of each column. This results in a new T by K matrix,
356 \mathbf{X}_n , that reflects how the features reflected in the columns of \mathbf{X}_{n-1} participate in the network during each
357 timepoint (row).

358 **Dynamic higher-order correlations**

359 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
360 enables us to estimate high-order dynamic correlations in a scalable way. Given a T by K input matrix, the
361 output of Equation 4 requires $O(TK^2)$ space to store. Repeated applications of Equation 4 (i.e., computing
362 dynamic correlations between the columns of the outputted dynamic correlation matrix) each require
363 exponentially more space; in general the n^{th} -order dynamic correlations of a T by K timeseries occupies
364 $O(TK^{2^n})$ space. However, when we approximate or summarize the output of Equation 4 with a T by K matrix
365 (as described above), it becomes feasible to compute even very high-order correlations in high-dimensional
366 data. Specifically, approximating the n^{th} -order dynamic correlations of a T by K timeseries requires only
367 $O(TK^2)$ additional space— the same as would be required to compute first-order dynamic correlations. In
368 other words, the space required to store $n + 1$ multivariate timeseries reflecting up to n^{th} order correlations
369 in the original data scales linearly with n using our approach (Fig. 8).

370 **Data**

371 We examined two types of data: synthetic data and human functional neuroimaging data. We constructed
372 and leveraged the synthetic data to evaluate our general approach (for a related validation approach see
373 Thompson et al., 2018). Specifically, we tested how well Equation 4 could be used to recover known dynamic
374 correlations using different choices of kernel (κ ; Fig. 9), for each of several synthetic datasets that exhibited
375 different temporal properties. We also simulated higher-order correlations and tested how well Equation 4
376 could recover these correlations using the best kernel from the previous synthetic data analyses. We then
377 applied our approach to a functional neuroimaging dataset to test the hypothesis that ongoing cognitive
378 processing is reflected in high-order dynamic correlations. We used an across-participant classification test
379 to estimate whether dynamic correlations of different orders contain information about which timepoint in
380 a story participants were listening to.

381 **Synthetic data: simulating dynamic first-order correlations**

382 We constructed a total of 40 different multivariate timeseries, collectively reflecting a total of 4 qualitatively
383 different patterns of dynamic first-order correlations (i.e., 10 datasets reflecting each type of dynamic pat-
384 tern). Each timeseries comprised 50 features (dimensions) that varied over 300 timepoints. The observations
385 at each timepoint were drawn from a zero-mean multivariate Gaussian distribution with a covariance matrix
386 defined for each timepoint as described below. We drew the observations at each timepoint independently
387 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 ,
388 $p(s_t) = p(s_t | s_{\setminus t})$.

Constant. We generated data with stable underlying correlations to evaluate how Equation 4 characterized correlation “dynamics” when the ground truth correlations were static. We constructed 10 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 :

$$\Sigma_m = \mathbf{C}\mathbf{C}^\top, \text{ where} \quad (11)$$

$$\mathbf{C}(i, j) \sim \mathcal{N}(0, 1), \text{ and where} \quad (12)$$

389 $i, j \in [1, 2, \dots, 50]$. In other words, all of the observations (for each of the 300 timepoints) within each dataset
390 were drawn from a multivariate Gaussian distribution with the same covariance matrix, and the 10 datasets
391 each used a different covariance matrix.

392 **Random.** We generated a second set of 10 synthetic datasets whose observations at each timepoint were
 393 drawn from a Gaussian distribution with a new randomly constructed (using Eqn. 11) covariance matrix.
 394 Because each timepoint’s covariance matrix was drawn independently from the covariance matrices for all
 395 other timepoints, these datasets provided a test of reconstruction accuracy in the absence of any meaningful
 396 underlying temporal structure in the dynamic correlations underlying the data.

Ramping. We generated a third set of 10 synthetic datasets whose underlying correlations changed gradually over time. For each dataset, we constructed two *anchor* covariance matrices using Equation 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

$$\Sigma_t = \left(1 - \frac{t}{299}\right)\Sigma_{\text{start}} + \frac{t}{299}\Sigma_{\text{end}}. \quad (13)$$

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

400 **Event.** We generated a fourth set of 10 synthetic datasets whose underlying correlation matrices exhibited
 401 prolonged intervals of stability, interspersed with abrupt changes. For each dataset, we used Equation 11
 402 to generate 5 random covariance matrices. We constructed a timeseries where each set of 60 consecutive
 403 samples was drawn from a Gaussian with the same covariance matrix. These datasets were intended to
 404 simulate a system that exhibits periods of stability punctuated by occasional abrupt state changes.

405 Synthetic data: simulating dynamic high-order correlations

406 We developed an iterative procedure for constructing timeseries data that exhibits known dynamic high-
 407 order correlations. The procedure builds on our approach to generating dynamic first-order correlations.
 408 Essentially, once we generate a timeseries with known first-order correlations, we can use the known first-
 409 order correlations as a template to generate a new timeseries of second-order correlations. In turn, we can
 410 generate a timeseries of third-order correlations from the second-order correlations, and so on. In general,
 411 we can generate order n correlations given a timeseries of order $n - 1$ correlations, for any $n > 1$. Finally,
 412 given the order n timeseries, we can reverse the preceding process to generate an order $n - 1$ timeseries, an
 413 order $n - 2$ order timeseries, and so on, until we obtain an order 0 timeseries of simulated data that reflects
 414 the chosen high-order dynamics.

The central mathematical operations in our procedure are two functions, $\text{vec}(\cdot)$ and $\text{mat}(\cdot)$. The $\text{vec}(\cdot)$ function takes as input a $K \times K$ symmetric matrix and returns as output a $(\frac{K^2-K}{2} + K)$ -dimensional column vector containing the entries in the upper triangle and diagonal. The $\text{mat}(\cdot)$ function inverts $\text{vec}(\cdot)$ by taking as input a $(\frac{K^2-K}{2} + K)$ -dimensional column vector and returning a $K \times K$ symmetric matrix as output. We can then generate an order n correlation matrix (for one timepoint, t) from an order $n - 1$ template (from the same timepoint) as follows:

$$\Sigma_n(t) = \text{mat}(\text{vec}(\Sigma_{n-1}(t)) \otimes \text{vec}(\Sigma_{n-1}(t))^\top). \quad (14)$$

Given a timeseries of order n correlation matrices, we can draw an order $n - 1$ correlation matrix for each timepoint t using

$$\sigma_{n-1}(t) \sim \mathcal{N}(\mathbf{0}, \Sigma_n(t)) \quad (15)$$

$$\Sigma_{n-1}(t) = \text{mat}(\sigma_{n-1}(t)). \quad (16)$$

415 We can then use repeated applications of Equations 15 and 16 in order to obtain a synthetic dataset.

416 When the template first-order correlations are constructed to exhibit different temporal profiles (e.g.,
417 using the constant, random, ramping, and event procedures described above), the resulting high-order
418 correlations and synthetic data will exhibit the same category of temporal profile.

419 Following our approach to generating synthetic data exhibiting known first-order correlations, we
420 constructed a total of 40 additional multivariate timeseries, collectively reflecting a total of 4 qualitatively
421 different patterns of dynamic correlations (i.e., 10 datasets reflecting each type of dynamic pattern: constant,
422 random, ramping, and event).

423 Functional neuroimaging data collected during story listening

424 We examined an fMRI dataset collected by Simony et al. (2016) that the authors have made publicly available
425 at arks.princeton.edu/ark:/88435/dsp015d86p269k. The dataset comprises neuroimaging data collected as
426 participants listened to an audio recording of a story (intact condition; 36 participants), listened to temporally
427 scrambled recordings of the same story (17 participants in the paragraph-scrambled condition listened to
428 the paragraphs in a randomized order and 36 in the word-scrambled condition listened to the words in a
429 randomized order), or lay resting with their eyes open in the scanner (rest condition; 36 participants). Full
430 neuroimaging details may be found in the original paper for which the data were collected (Simony et al.,
431 2016).

432 **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 described by 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.

442 Temporal decoding

443 We sought to identify neural patterns that reflected participants’ ongoing cognitive processing of incoming
444 stimulus information. As reviewed by Simony et al. (2016), one way of homing in on these stimulus-driven
445 neural patterns is to compare activity patterns across individuals (e.g., using ISFC analyses). In particular,
446 neural patterns will be similar across individuals to the extent that the neural patterns under consideration
447 are stimulus-driven, and to the extent that the corresponding cognitive representations are reflected in
448 similar spatial patterns across people (also see Simony & Chang, 2020). Following this logic, we used an
449 across-participant temporal decoding test developed by Manning et al. (2018) to assess the degree to which
450 different neural patterns reflected ongoing stimulus-driven cognitive processing across people (Fig. 10). The
451 approach entails using a subset of the data to train a classifier to decode stimulus timepoints (i.e., moments
452 in the story participants listened to) from neural patterns. We use decoding (forward inference) accuracy
453 on held-out data, from held-out participants, as a proxy for the extent to which the inputted neural patterns
454 reflected stimulus-driven cognitive processing in a similar way across individuals.

455 Forward inference and decoding accuracy

456 We used an across-participant correlation-based classifier to decode which stimulus timepoint matched
457 each timepoint’s neural pattern (Fig. 10). We first divided the participants into two groups: a template group,
458 $\mathcal{G}_{\text{template}}$ (i.e., training data), and a to-be-decoded group, $\mathcal{G}_{\text{decode}}$ (i.e., test data). We used Equation 7 to
459 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
460 $\bar{\mathbf{C}}_{\text{template}}$ and $\bar{\mathbf{C}}_{\text{decode}}$ to form a number-of-timepoints by number-of-timepoints decoding matrix, Λ . In this
461 way, the rows of Λ reflected timepoints from the template group, while the columns reflected timepoints

462 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
463 $\bar{\mathbf{C}}_{\text{template}}$ with which it was most highly correlated. We then repeated this decoding procedure, but using
464 $\mathcal{G}_{\text{decode}}$ as the template group and $\mathcal{G}_{\text{template}}$ as the to-be-decoded group. Given the true timepoint labels (for
465 each group), we defined the *decoding accuracy* as the average proportion of correctly decoded timepoints,
466 across both groups. We defined the *relative decoding accuracy* as the difference between the decoding accuracy
467 and chance accuracy (i.e., $\frac{1}{T}$).

468 **Feature weighting and testing**

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

472 One challenge to fairly evaluating high-order correlations is that if the kernel used in Equation 4 is
473 wider than a single timepoint, each repeated application of the equation will result in further temporal
474 blur. Because our primary assessment metric is temporal decoding accuracy, this unfairly biases against
475 detecting meaningful signal in higher-order correlations (relative to lower-order correlations). We attempted
476 to mitigate temporal blur in estimating each \mathbf{X}_n by using a Dirac δ function kernel (which places all of its
477 mass over a single timepoint; Fig. 9b, 10a) to compute each lower-order correlation ($\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_{n-1}$). We
478 then used a new (potentially wider, as described below) kernel to compute \mathbf{X}_n from \mathbf{X}_{n-1} . In this way,
479 temporal blurring was applied only in the last step of computing \mathbf{X}_n . We note that, because each \mathbf{X}_n is a
480 low-dimensional representation of the corresponding \mathbf{Y}_n , the higher-order correlations we estimated reflect
481 true correlations in the data with lower-fidelity than estimates of lower-order correlations. Therefore, even
482 after correcting for temporal blurring, our approach is still biased against finding meaningful signal in
483 higher-order correlations.

484 After computing each $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_{n-1}$ for each participant, we divided participants into two equally sized
485 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 $\mathcal{G}_{\text{train}_1}$
486 and $\mathcal{G}_{\text{train}_2}$. We then computed Λ (temporal correlation) matrices for each type of neural feature, using $\mathcal{G}_{\text{train}_1}$
487 and $\mathcal{G}_{\text{train}_2}$. This resulted in $n + 1$ Λ matrices (one for the original timeseries of neural activations, and one
488 for each of n orders of dynamic correlations). Our objective was to find a set of weights for each of these
489 Λ matrices such that the weighted average of the $n + 1$ matrices yielded the highest decoding accuracy.
490 We used quasi-Newton gradient ascent (Nocedal & Wright, 2006), using decoding accuracy (for $\mathcal{G}_{\text{train}_1}$ and
491 $\mathcal{G}_{\text{train}_2}$) as the objective function to be maximized, to find an optimal set of training data-derived weights,
492 $\phi_{0,1,\dots,n}$, where $\sum_{i=0}^n \phi_i = 1$ and where $\phi_i \geq 0 \forall i \in [0, 1, \dots, n]$.

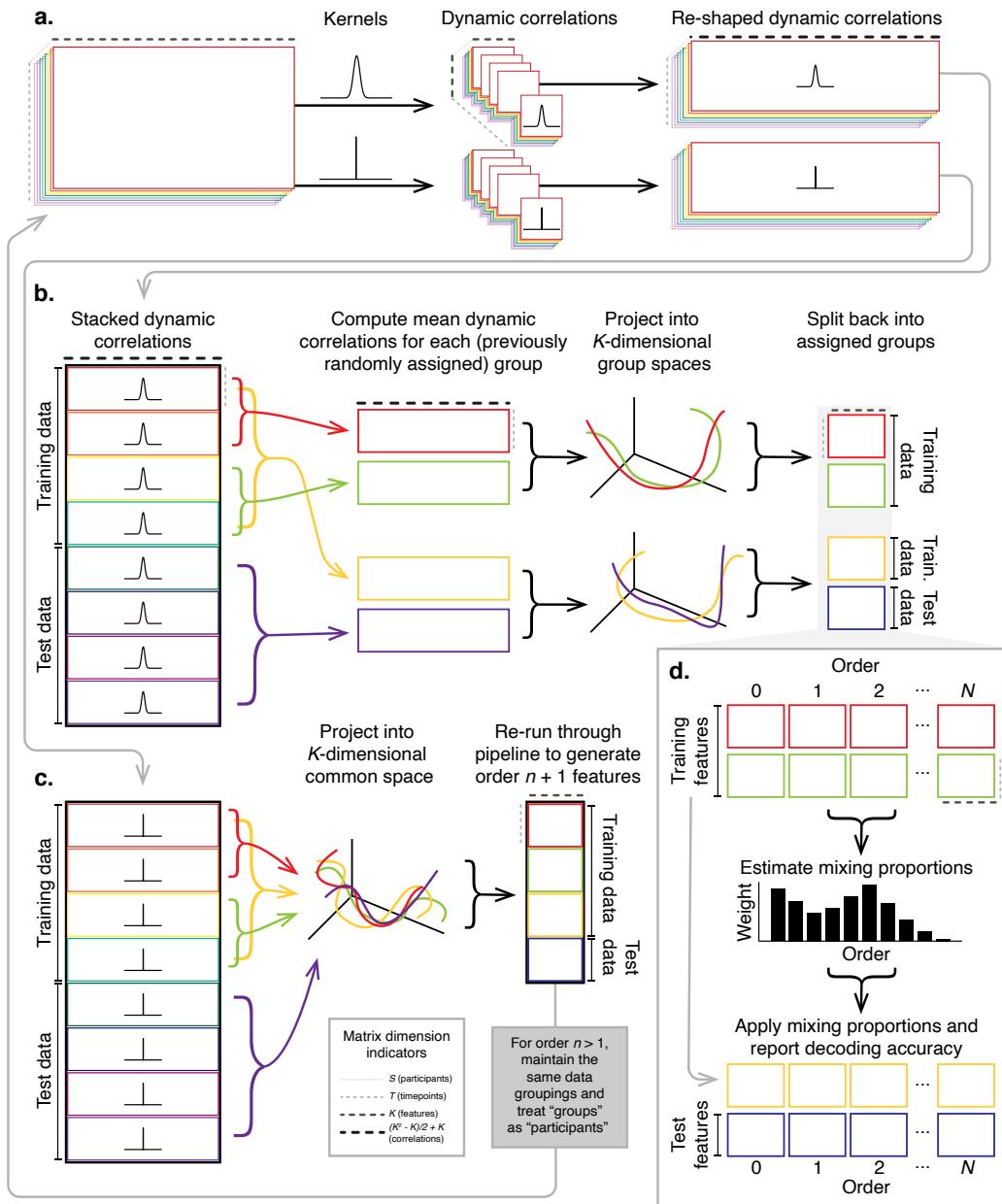


Figure 10: Decoding analysis pipeline. **a. Computing dynamic correlations from timeseries data.** Given a timeseries of observations as a $T \times K$ matrix (or a set of S such matrices), we use Equation 4 to compute each participant's DISFC (relative to other participants in the training or test sub-group, 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.

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

497 We used a permutation-based procedure to form stable estimates of decoding accuracy for each set of
498 neural features. In particular, we computed the decoding accuracy for each of 10 random group assignments
499 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
500 features.

501 Identifying robust decoding results

502 The temporal decoding procedure we use to estimate which neural features support ongoing cognitive
503 processing is governed by several parameters. In particular, Equation 4 requires defining a kernel function,
504 which can take on different shapes and widths. For a fixed set of neural features, each of these parameters
505 can yield different decoding accuracies. Further, the best decoding accuracy for a given timepoint may be
506 reliably achieved by one set of parameters, whereas the best decoding accuracy for another timepoint might
507 be reliably achieved by a different set of parameters, and the best decoding accuracy across *all* timepoints
508 might be reliably achieved by still another different set of parameters. Rather than attempting to maximize
509 decoding accuracy, we sought to discover the trends in the data that were robust to classifier parameters
510 choices. Specifically, we sought to characterize how decoding accuracy varied (under different experimental
511 conditions) as a function of which neural features were considered.

512 To identify decoding results that were robust to specific classifier parameter choices, we repeated our
513 decoding analyses after substituting into Equation 4 each of a variety of kernel shapes and widths. We
514 examined Gaussian (Fig. 9c), Laplace (Fig. 9d), and Mexican Hat (Fig. 9e) kernels, each with widths of 5, 10,
515 20, and 50 samples. We then report the average decoding accuracies across all of these parameter choices.
516 This enabled us to (partially) factor out performance characteristics that were parameter-dependent, within
517 the set of parameters we examined.

518 Reverse inference

519 The dynamic patterns we examined comprise high-dimensional correlation patterns at each timepoint. To
520 help interpret the resulting patterns in the context of other studies, we created summary maps by computing
521 the across-timepoint average pairwise correlations at each order of analysis (first order, second order, etc.).
522 We selected the 10 strongest (absolute value) correlations at each order. Each correlation is between the

523 dynamic activity patterns (or patterns of dynamic high-order correlations) measured at two RBF nodes
524 (see *Hierarchical Topographic Factor Analysis*). Therefore, the 10 strongest correlations involved up to 20 RBF
525 nodes. Each RBF defines a spatial function whose activations range from 0 to 1. We constructed a map
526 of RBF components that denoted the endpoints of the 10 strongest correlations (we set each RBF to have a
527 maximum value of 1). We then carried out a meta analysis using Neurosynth (Rubin et al., 2017) to identify
528 the 10 terms most commonly associated with the given map. This resulted in a set of 10 terms associated
529 with the average dynamic correlation patterns at each order.

530 Acknowledgements

531 We acknowledge discussions with Luke Chang, Vassiki Chauhan, Hany Farid, Paxton Fitzpatrick, Andrew
532 Heusser, Eshin Jolly, Aaron Lee, Qiang Liu, Matthijs van der Meer, Judith Mildner, Gina Notaro, Stephen
533 Satterthwaite, Emily Whitaker, Weizhen Xie, and Kirsten Ziman. Our work was supported in part by NSF
534 EPSCoR Award Number 1632738 to J.R.M. and by a sub-award of DARPA RAM Cooperative Agreement
535 N66001-14-2-4-032 to J.R.M. The content is solely the responsibility of the authors and does not necessarily
536 represent the official views of our supporting organizations.

537 Author contributions

538 Concept: J.R.M. Implementation: T.H.C., L.L.W.O., and J.R.M. Analyses: L.L.W.O. and J.R.M. Writing:
539 L.L.W.O. and J.R.M.

540 References

- 541 Allen, E. A., Damaraju, E., Plis, S. M., Erhardt, E. B., Eichele, T., & Calhoun, V. D. (2012). Tracking
542 whole-brain connectivity dynamics in the resting state. *Cerebral Cortex*, 24(3), 663–676.
- 543 Alvarez-Hamelin, I., Dall'Asta, L., Barrat, A., & Vespignani, A. (2005). k -corr decomposition: a tool for the
544 visualization of large scale networks. *arXiv*, cs/0504107v2.
- 545 Baldassano, C., Chen, J., Zadbood, A., Pillow, J. W., Hasson, U., & Norman, K. A. (2017). Discovering event
546 structure in continuous narrative perception and memory. *Neuron*, 95(3), 709–721.
- 547 Barthélémy, M. (2004). Betweenness centrality in large complex networks. *European Physical Journal B*, 38,
548 163–168.

- 549 Bassett, D., Meyer-Lindenberg, A., Achard, S., Duke, T., & Bullmore, E. (2006, December). Adaptive
550 reconfiguration of fractal small-world human brain functional networks. *Proceedings of the National
551 Academy of Sciences, USA*, 103(51), 19518–23.
- 552 Betzel, R. F., Byrge, L., Esfahlani, F. Z., & Kennedy, D. P. (2019). Temporal fluctuations in the brain's modular
553 architecture during movie-watching. *bioRxiv*, doi.org/10.1101/750919.
- 554 Bonacich, P. (2007). Some unique properties of eigenvector centrality. *Social Networks*, 29(4), 555–564.
- 555 Bullmore, E., & Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and
556 functional systems. *Nature Reviews Neuroscience*, 10(3), 186–198.
- 557 Capota, M., Turek, J., Chen, P.-H., Zhu, X., Manning, J. R., Sundaram, N., . . . Shin, Y. S. (2017). *Brain imaging
558 analysis kit*. Retrieved from <https://doi.org/10.5281/zenodo.59780>
- 559 Chang, C., & Glover, G. H. (2010). Time-frequency dynamics of resting-state brain connectivity measured
560 with fMRI. *NeuroImage*, 50, 81–98.
- 561 Chien, H.-Y. S., & Honey, C. J. (2019). Constructing and forgetting temporal context in the human cerebral
562 cortex. *bioRxiv*, doi.org/10.1101/761593.
- 563 Christakis, N. A., & Fowler, J. H. (2010). Social network sensors for early detection of contagious outbreaks.
564 *PLoS One*, 5(9), e12948.
- 565 Combrisson, E., Vallat, R., O'Reilly, C., Jas, M., Pascarella, A., l Saive, A., . . . Jerbi, K. (2019). Visbrain: a
566 multi-purpose GPU-accelerated open-source suite for multimodal brain data visualization. *Frontiers in
567 Neuroinformatics*, 13(14), 1–14.
- 568 Comon, P., Jutten, C., & Herault, J. (1991). Blind separation of sources, part II: Problems statement. *Signal
569 Processing*, 24(1), 11 - 20.
- 570 Demertzi, A., Tagliazucchi, E., Dehaene, S., Deco, G., Barttfeld, P., Raimondo, F., . . . Sitt, J. D. (2019). Human
571 consciousness is supported by dynamic complex patterns of brain signal coordination. *Science Advances*,
572 5(2), eaat7603.
- 573 Estrada, E., & Rodríguez-Velázquez, J. A. (2005). Subgraph centrality in complex networks. *Physical Review
574 E*, 71(5), 056103.
- 575 Etzel, J. A., Gazzola, V., & Keysers, C. (2009). An introduction to anatomical ROI-based fMRI classification.
576 *Brain Research*, 1281, 114–125.

- 577 Fallon, J., Ward, P., Parkes, L., Oldham, S., Arnatkevičiūtė, A., Fornito, A., & Fulcher, B. D. (2019).
578 Timescales of spontaneous activity fluctuations relate to structural connectivity in the brain. *bioRxiv*,
579 doi.org/10.1101/655050.
- 580 Fong, A. H. C., Yoo, K., Rosenberg, M. D., Zhang, S., Li, C.-S. R., Scheinost, D., ... Chun, M. M. (2019).
581 Dynamic functional connectivity during task performance and rest predicts individual differences in
582 attention across studies. *NeuroImage*, 188, 14–25.
- 583 Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry*, 40(1), 35–41.
- 584 Geisberger, R., Sanders, P., & Schultes, D. (2008). Better approximation of betweenness centrality. *Proceedings*
585 *of the meeting on Algorithm Engineering and Experiments*, 90–100.
- 586 Gershman, S., Blei, D., Pereira, F., & Norman, K. (2011). A topographic latent source model for fMRI data.
587 *NeuroImage*, 57, 89–100.
- 588 Gonzalez-Castillo, J., Caballero-Gaudes, C., Topolski, N., Handwerker, D. A., Pereira, F., & Bandettini, P. A.
589 (2019). Imaging the spontaneous flow of thought: distinct periods of cognition contribute to dynamic
590 functional connectivity during rest. *NeuroImage*, 202(116129).
- 591 Halu, A., Mondragón, R. J., Panzarasa, P., & Bianconi, G. (2013). Multiplex PageRank. *PLoS One*, 8(10),
592 e78293.
- 593 Hasson, U., Chen, J., & Honey, C. J. (2015). Hierarchical process memory: memory as an integral component
594 of information processing. *Trends in Cognitive Science*, 19(6), 304–315.
- 595 Hasson, U., Yang, E., Vallines, I., Heeger, D. J., & Rubin, N. (2008). A hierarchy of temporal receptive windows
596 in human cortex. *Journal of Neuroscience*, 28(10), 2539–2550. doi: 10.1523/JNEUROSCI.5487-07.2008
- 597 Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and
598 overlapping representations of faces and objects in ventral temporal cortex. *Science*, 293, 2425–2430.
- 599 Honey, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks.
600 *Science*, 313(5786), 504–507.
- 601 Honey, C. J., Kötter, R., Breakspear, M., & Sporns, O. (2007). Network structure of cerebral cortex shapes
602 functional connectivity on multiple time scales. *Proceedings of the National Academy of Science USA*, 104(24),
603 10240–10245.

- 604 Honey, C. J., Thesen, T., Donner, T. H., Silbert, L. J., Carlson, C. E., Devinsky, O., ... Hasson, U. (2012). Slow
605 cortical dynamics and the accumulation of information over long timescales. *Neuron*, 76, 423-434. doi:
606 10.1016/j.neuron.2012.08.011
- 607 Huth, A. G., de Heer, W. A., Griffiths, T. L., Theunissen, F. E., & Gallant, J. L. (2016). Natural speech reveals
608 the semantic maps that tile human cerebral cortex. *Nature*, 532, 453–458.
- 609 Huth, A. G., Nisimoto, S., Vu, A. T., & Gallant, J. L. (2012). A continuous semantic space describes
610 the representation of thousands of object and action categories across the human brain. *Neuron*, 76(6),
611 1210-1224.
- 612 Jutten, C., & Herault, J. (1991). Blind separation of sources, part I: An adaptive algorithm based on
613 neuromimetic architecture. *Signal Processing*, 24(1), 1–10.
- 614 Kamitani, Y., & Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nature*
615 *Neuroscience*, 8, 679–685.
- 616 Landau, E. (1895). Zur relativen Wertbemessung der Turnierresultate. *Deutsches Wochenschach*, 11, 366–369.
- 617 Lee, C. S., Aly, M., & Baldassano, C. (2020). Anticipation of temporally structured events in the brain.
618 *bioRxiv*, 10.1101/2020.10.14.338145.
- 619 Lee, D. D., & Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature*,
620 401, 788–791.
- 621 Lerner, Y., Honey, C. J., Katkov, M., & Hasson, U. (2014). Temporal scaling of neural responses to compressed
622 and dilated natural speech. *Journal of Neurophysiology*, 111, 2433–2444.
- 623 Lerner, Y., Honey, C. J., Silbert, L. J., & Hasson, U. (2011). Topographic mapping of a hierarchy of
624 temporal receptive windows using a narrated story. *Journal of Neuroscience*, 31(8), 2906-2915. doi: 10.1523/
625 JNEUROSCI.3684-10.2011
- 626 Liégeois, R., Li, J., Kong, R., Orban, C., De Ville, D. V., Ge, T., ... Yeo, B. T. T. (2019). Resting brain
627 dynamics at different timescales capture distinct aspects of human behavior. *Nature Communications*,
628 10(2317), doi.org/10.1038/s41467-019-10317-7.
- 629 Lin, J. (2009). Divergence measures based on the Shannon entropy. *IEEE Transactions on Information Theory*,
630 37(1), 145–151.

- 631 Lohmann, G., Margulies, D. S., Horstmann, A., Pleger, B., Lepsius, J., Goldhahn, D., ... Turner, R. (2010).
632 Eigenvector centrality mapping for analyzing connectivity patterns in fMRI data of the human brain.
633 *PLoS One*, 5(4), e10232.
- 634 Lurie, D., Kessler, D., Bassett, D., Betzel, R., Breakspear, M., Keilholz, S., ... Calhoun, V. (2018). On the
635 nature of time-varying functional connectivity in resting fMRI. *PsyArXiv*, doi.org/10.31234/osf.io/xtzre.
- 636 Mack, M. L., Preston, A. R., & Love, B. C. (2017). Medial prefrontal cortex compresses concept representations
637 through learning. *bioRxiv*, doi.org/10.1101/178145.
- 638 Mairal, J., Ponce, J., Sapiro, G., Zisserman, A., & Bach, F. R. (2009). Supervised dictionary learning. *Advances*
639 *in Neural Information Processing Systems*, 1033–1040.
- 640 Mairal, J. B., Bach, F., Ponce, J., & Sapiro, G. (2009). Online dictionary learning for sparse coding. *Proceedings*
641 *of the 26th annual international conference on machine learning*, 689–696.
- 642 Manning, J. R., Ranganath, R., Norman, K. A., & Blei, D. M. (2014). Topographic factor analysis: a Bayesian
643 model for inferring brain networks from neural data. *PLoS One*, 9(5), e94914.
- 644 Manning, J. R., Zhu, X., Willke, T. L., Ranganath, R., Stachenfeld, K., Hasson, U., ... Norman, K. A. (2018).
645 A probabilistic approach to discovering dynamic full-brain functional connectivity patterns. *NeuroImage*,
646 180, 243–252.
- 647 McInnes, L., Healy, J., & Melville, J. (2018). UMAP: Uniform manifold approximation and projection for
648 dimension reduction. *arXiv*, 1802(03426).
- 649 McIntosh, A. R., & Jirsa, V. K. (2019). The hidden repertoire of brain dynamics and dysfunction. *Network*
650 *Neuroscience*, doi.org/10.1162/netn_a_00107.
- 651 Mitchell, T., Shinkareva, S., Carlson, A., Chang, K., Malave, V., Mason, R., & Just, M. (2008). Predicting
652 human brain activity associated with the meanings of nouns. *Science*, 320(5880), 1191.
- 653 Newman, M. E. J. (2005). A measure of betweenness centrality based on random walks. *Social Networks*, 27,
654 39–54.
- 655 Newman, M. E. J. (2008). The mathematics of networks. *The New Palgrave Encyclopedia of Economics*, 2, 1–12.
- 656 Nishimoto, S., Vu, A., Naselaris, T., Benjamini, Y., Yu, B., & Gallant, J. (2011). Reconstructing visual
657 experience from brain activity evoked by natural movies. *Current Biology*, 21, 1 - 6.
- 658 Nocedal, J., & Wright, S. J. (2006). *Numerical optimization*. New York: Springer.

- 659 Norman, K. A., Newman, E., Detre, G., & Polyn, S. M. (2006). How inhibitory oscillations can train neural
660 networks and punish competitors. *Neural Computation*, 18, 1577–1610.
- 661 Opsahl, T., Agneessens, F., & Skvoretz, J. (2010). Node centrality in weighted networks: generalizing degree
662 and shortest paths. *Social Networks*, 32, 245–251.
- 663 Park, H.-J., Friston, K. J., Pae, C., Park, B., & Razi, A. (2018). Dynamic effective connectivity in resting state
664 fMRI. *NeuroImage*, 180, 594–608.
- 665 Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. *The London, Edinburgh,
666 and Dublin Philosophical Magazine and Journal of Science*, 2, 559–572.
- 667 Pereira, F., Lou, B., Pritchett, B., Ritter, S., Gershman, S. J., Kanwisher, N., … Fedorenko, E. (2018). Toward
668 a universal decoder of linguistic meaning from brain activation. *Nature Communications*, 9(963), 1–13.
- 669 Preti, M. G., Bolton, T. A. W., & Van De Ville, D. (2017). The dynamic functional connectome: state-of-the-art
670 and perspectives. *NeuroImage*, 160, 41–54.
- 671 Rao, C. R. (1982). Diversity and dissimilarity coefficients: a unified approach. *Theoretical Population Biology*,
672 21(1), 24–43.
- 673 Reimann, M. W., Nolte, M., Scolamiero, M., Turner, K., Perin, R., Chindemi, G., … Markram, H. (2017).
674 Cliques of neurons bound into cavities provide a missing link between structure and function. *Frontiers
675 in Computational Neuroscience*, 11(48), 1–16.
- 676 Ricotta, C., & Szeidl, L. (2006). Towards a unifying approach to diversity measures: Bridging the gap
677 between the Shannon entropy and Rao's quadratic index. *Theoretical Population Biology*, 70(3), 237–243.
- 678 Roy, D. S., Park, Y.-G., Ogawa, S. K., Cho, J. H., Choi, H., Kamensky, L., … Tonegawa, S. (2019). Brain-
679 wide mapping of contextual fear memory engram ensembles supports the dispersed engram complex
680 hypothesis. *bioRxiv*, doi.org/10.1101/668483.
- 681 Rubin, T. N., Kyoejo, O., Gorgolewski, K. J., Jones, M. N., Poldrack, R. A., & Yarkoni, T. (2017). Decoding
682 brain activity using a large-scale probabilistic functional-anatomical atlas of human cognition. *PLoS
683 Computational Biology*, 13(10), e1005649.
- 684 Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpreta-
685 tions. *NeuroImage*, 52, 1059–1069.
- 686 Schreiber, T. (2000). Measuring information transfer. *Physical Review Letters*, 85(2), 461–464.

- 687 Shappell, H., Caffo, B. S., Pekar, J. J., & Lindquist, M. A. (2019). Improved state change estimation in
688 dynamic functional connectivity using hidden semi-Markov models. *NeuroImage*, 191, 243–257.
- 689 Simony, E., & Chang, C. (2020). Analysis of stimulus-induced brain dynamics during naturalistic paradigms.
690 *NeuroImage*, 216, 116461.
- 691 Simony, E., Honey, C. J., Chen, J., & Hasson, U. (2016). Dynamic reconfiguration of the default mode
692 network during narrative comprehension. *Nature Communications*, 7(12141), 1–13.
- 693 Sizemore, A. E., Giusti, C., Kahn, A., Vettel, J. M., Betzel, R. F., & Bassett, D. S. (2018). Cliques and cavities
694 in the human connectome. *Journal of Computational Neuroscience*, 44(1), 115–145.
- 695 Solomon, S. H., Medaglia, J. D., & Thompson-Schill, S. L. (2019). Implementing a concept network model.
696 *Behavior Research Methods*, 51, 1717–1736.
- 697 Spearman, C. (1904). General intelligence, objectively determined and measured. *American Journal of
698 Psychology*, 15, 201–292.
- 699 Sporns, O., & Honey, C. J. (2006). Small worlds inside big brains. *Proceedings of the National Academy of
700 Science USA*, 103(51), 19219–19220.
- 701 Thompson, W. H., Richter, C. G., Plavén-Sigray, P., & Fransson, P. (2018). Simulations to benchmark
702 time-varying connectivity methods for fMRI. *PLoS Computational Biology*, 14(5), e1006196.
- 703 Tipping, M. E., & Bishop, C. M. (1999). Probabilistic principal component analysis. *Journal of Royal Statistical
704 Society, Series B*, 61(3), 611–622.
- 705 Toker, D., & Sommer, F. T. (2019). Information integration in large brain networks. *PLoS Computational
706 Biology*, 15(2), e1006807.
- 707 Tong, F., & Pratte, M. S. (2012). Decoding patterns of human brain activity. *Annual Review of Psychology*, 63,
708 483–509.
- 709 Turk-Browne, N. B. (2013). Functional interactions as big data in the human brain. *Science*, 342, 580–584.
- 710 van der Maaten, L. J. P., & Hinton, G. E. (2008). Visualizing high-dimensional data using t-SNE. *Journal of
711 Machine Learning Research*, 9, 2579–2605.
- 712 Vidaurre, D., Abeysuriya, R., Becker, R., Quinn, A. J., Alfaro-Almagro, F., Smith, S. M., & Woolrich, M. W.
713 (2018). Discovering dynamic brain networks from big data in rest and task. *NeuroImage*, 180, 646–656.

- 714 Zar, J. H. (2010). *Biostatistical analysis*. Prentice-Hall/Pearson.
- 715 Zheng, M., Allard, A., Hagmann, P., & Serrano, M. A. (2019). Geometric renormalization unravels self-
716 similarity of the multiscale human connectome. *arXiv*, 1904.11793.
- 717 Zou, Y., Donner, R. V., Marwan, N., Donges, J. F., & Kurths, J. (2019). Complex network approaches to
718 nonlinear time series analysis. *Physics Reports*, 787, 1–97.