

1 High-level cognition is supported by at least second order
2 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

³Amazon.com, Seattle, WA

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

4 August 23, 2019

5 **Abstract**

6 Our thoughts arise from coordinated patterns of interactions between brain structures that change
7 with our ongoing experiences. High-order dynamic correlations in brain 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 supported by high-order dynamic correlations in brain activity
10 patterns. We developed an approach to estimating high-order dynamic correlations in timeseries data,
11 and we applied the approach to neuroimaging data collected as human participants either listened to a
12 ten-minute story or a temporally scrambled version of the story, or underwent a resting state scan. We
13 trained across-participants pattern classifiers to decode (in held-out data) when in the session each activity
14 snapshot was collected. We found that classifiers trained to decode from high-order dynamic correlations
15 yielded better performance on data collected as participants listened to the (unscrambled) story. By
16 contrast, classifiers trained to decode data from scrambled versions of the story or during the resting
17 state scan yielded the best performance when they were trained using first-order dynamic correlations
18 or raw activity patterns. We suggest that as our thoughts become more complex, they are supported by
19 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., 2018;
26 Tong & Pratte, 2012). Training decoding models on different types of neural features can also help to elucidate
27 which specific aspects of neural activity patterns are informative about cognition– and, by extension, which
28 types of neural activity patterns might comprise the neural code. For example, prior work has used region
29 of interest analyses to estimate the anatomical locations of specific neural representations (e.g., Etzel et al.,
30 2009), or to compare the relative contributions to the neural code of multivariate activity patterns versus

31 patterns of dynamic correlations between neural activity patterns (e.g., Fong et al., 2019; Manning et al.,
32 2018). An emerging theme in this literature is that cognition is mediated by complex dynamic interactions
33 between brain structures (Bassett et al., 2006; Demertzi et al., 2019; Sporns & Honey, 2006; Turk-Browne,
34 2013).

35 Studies of the neural code to date have primarily focused on univariate or multivariate neural pat-
36 terns (for review see , NormEtal06) or (more recently) on patterns of dynamic first-order correlations (i.e.,
37 interactions between pairs of brain structures; Fong et al., 2019; Manning et al., 2018). We wondered what
38 the future of this line of work might hold. For example, is the neural code mediated by higher-order
39 interactions between brain structures? Second-order correlations reflect *homologous* patterns of correlation.
40 In other words, if the changing patterns of correlations between two regions, *A* and *B*, are similar to those
41 between two other regions, *C* and *D*, this would be reflected in the second-order correlations between (*A*–*B*)
42 and (*C*–*D*). In this way, second-order correlations identify similarities and differences between subgraphs
43 of the brain’s connectome. Analogously, third-order correlations reflect homologies between second-order
44 correlations– i.e., homologous patterns of homologous interactions between brain regions. More generally,
45 higher-order correlations reflect homologies between patterns of lower-order correlations. We can then ask:
46 which “orders” of interaction are most reflective of high-level cognitive processes?

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

58 To gain insights into the above aspects of the neural code, we developed a computational framework
59 for estimating dynamic high-order correlations in timeseries data. This framework provides an important
60 advance, in that it enables us to examine patterns in higher-order correlations that are computationally
61 intractable to estimate via conventional methods. Given a multivariate timeseries, our framework provides
62 timepoint-by-timepoint estimates of the first-order correlations, second-order correlations, and so on (up to
63 tenth-order correlations in this manuscript). Our approach combines a kernel-based method for computing

64 dynamic correlations in timeseries data with a dimensionality reduction step that projects the resulting dy-
65 namic correlations into a low-dimensional space. We explored two dimensionality reduction approaches:
66 principle components analysis (PCA; Pearson, 1901), which preserves an approximately invertable transfor-
67 mation back to the original data; and a second non-invertible algorithm that explored patterns in eigenvector
68 centrality (Landau, 1895). This latter approach characterizes correlations between each feature dimension's
69 relative *position* in the network in favor of the specific activity histories of different features.

70 We validated our approach using synthetic data where the underlying correlations were known. We
71 then applied our framework to a neuroimaging dataset collected as 125 participants listened to either an
72 audio recording of a ten-minute story or a temporally scrambled version of the story, or underwent a resting
73 state scan (Simony et al., 2016). We used a subset of the data to train across-participant classifiers to decode
74 listening times using a blend of neural features (comprising neural activity patterns, as well as different
75 orders of correlations between those patterns that were inferred using our computational framework).
76 We found that both the PCA-based and eigenvector centrality-based approaches yielded neural patterns
77 that could be used to decode accurately. Both approaches also yielded the best decoding accuracy for
78 data collected during (intact) story listening when high-order (PCA: second-order; eigenvector centrality:
79 fourth-order) dynamic correlation patterns were included as features. When we trained classifiers on the
80 scrambled stories or resting state data, only lower-order dynamic patterns were informative to the decoders.
81 Taken together, our results indicate that high-level cognition is supported by high-order dynamic patterns
82 of communication between brain structures.

83 Methods

84 Our general approach to comprises four general steps. First, we derive a kernel-based approach to comput-
85 ing dynamic pairwise correlations in a T (timepoints) by K (features) multivariate timeseries, \mathbf{X}_0 . This yields
86 a T by $O(K^2)$ matrix of dynamic correlations, \mathbf{Y}_1 , where each row comprises the upper triangle of the correla-
87 tion matrix at a single timepoint, reshaped into a row vector. Second, we apply a dimensionality reduction
88 step to project the matrix of dynamic correlations back onto a K -dimensional space. This yields a T by K
89 matrix, \mathbf{X}_1 , that reflects an approximation of the dynamic correlations reflected in the original data. Third,
90 we use repeated applications of the kernel-based dynamic correlation step to \mathbf{X}_n and the dimensionality
91 reduction step to the resulting \mathbf{Y}_{n+1} to estimate high-order dynamnic correlations. Each application of these
92 steps to a T by K time series \mathbf{X}_n yields a T by K matrix, \mathbf{X}_{n+1} , that reflects the dynamic correlations between
93 the columns of \mathbf{X}_n . In this way, we refer to n as the *order* of the timeseries, where \mathbf{X}_0 (order 0) denotes the
94 original data and \mathbf{X}_n denotes n^{th} -order dynamic correlations between the columns of \mathbf{X}_0 . Finally, we use

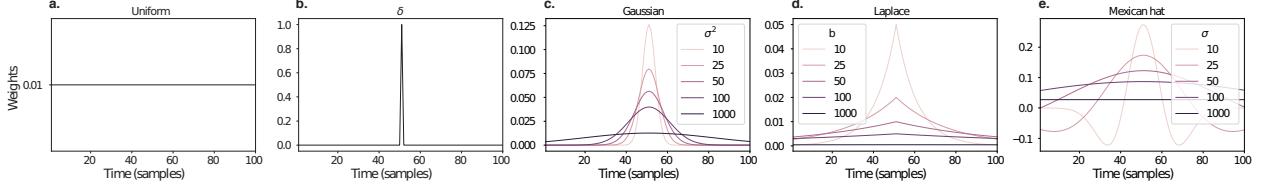


Figure 1: Examples of kernel functions. Each panel displays per-timepoint weights at $t = 50$, evaluated for 100 timepoints ($\tau \in [1, \dots, 100]$). **a. Uniform kernel.** The weights are timepoint-invariant; observations at all timepoints are weighted equally, and do not change as a function of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

95 a cross-validation-based decoding approach to evaluate how well information contained in a given order
 96 (or weighted mixture of orders) may be used to decode relevant cognitive states. If including a given \mathbf{X}_n in
 97 the feature set yields higher classification accuracy on held-out data, we interpret this as evidence that the
 98 given cognitive states are reflected in patterns of n^{th} -order correlations.

99 Kernel-based approach for computing dynamic correlations

Given a matrix of observations, 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):

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

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

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

100 We can generalize this formula to compute time-varying correlations by incorporating a *kernel function* that
 101 takes a time t as input, and returns how much the observed data at each timepoint $\tau \in [-\infty, \infty]$ contributes
 102 to the estimated instantaneous correlation at time t (Fig. 1).

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

static correlation formula in Equation 2 to estimate the *instantaneous correlation* at timepoint t :

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

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

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

- 103 Here $\text{timecorr}(\mathbf{X}(\cdot, i), \mathbf{X}(\cdot, j), \kappa_t)$ reflects the correlation at time t between columns i and j of \mathbf{X} , estimated using
104 the kernel κ_t .

105 **Inter-subject dynamic correlations**

Equation 5 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 et al., 2016) to compute the degree of stimulus-driven correlations reflected in the multi-participant dataset at a given timepoint t using:

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

where M extracts and vectorizes the diagonal and upper triangle 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 $\mathbf{Y}_n^p(t)$ denotes the correlation matrix (Eqn. 2) between each column of \mathbf{X}_n^p and each column of the average \mathbf{X}_n from all *other* participants, $\bar{\mathbf{X}}_n^{\setminus p}$:

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

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

108 Low-dimensional representations of dynamic correlations

109 Given a T by $O(K^2)$ matrix of dynamic correlations, \mathbf{Y}_n , we propose two general approaches to computing
110 a T by K low-dimensional representation of these correlations, \mathbf{X}_n . The first approach uses dimensionality
111 reduction algorithms to project \mathbf{Y}_n onto a K -dimensional space. The second approach uses graph-theoretic
112 measures to characterize the relative positions of each feature ($k \in [1, \dots, K]$) in the network defined by the
113 correlation matrix at each timepoint. **JRM STOPPED HERE**

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

115 Commonly used dimensionality reduction algorithms include Principal Components Analysis (PCA; Pearson,
116 1901), Probabilistic PCA (PPCA; Tipping & Bishop, 1999), Exploratory Factor Analysis (EFA; Spearman,
117 1904), Independent Components Analysis (ICA; Comon et al., 1991; Jutten & Herault, 1991), t -Stochastic
118 Neighbor Embedding (t -SNE; van der Maaten & Hinton, 2008), Uniform Manifold Approximation and
119 Projection (UMAP; McInnes & Healy, 2018), non-negative matrix factorization (NMF; Lee & Seung,
120 1999), Topographic Factor Analysis (TFA) Manning et al. (2014), Hierarchical Topographic Factor analysis
121 (HTFA) Manning et al. (2018), Topographic Latent Source Analysis (TLSA) Gershman et al. (2011), Dictionary
122 learning (J. Mairal et al., 2009; J. B. Mairal et al., 2009), deep autoencoders (Hinton & Salakhutdinov,
123 2006), among others. While complete characterizations of each of these algorithms is beyond the scope of
124 the present manuscript, the general intuition driving these approaches is to compute the $\hat{\mathbf{Y}}$ with i columns
125 that is closest to the original \mathbf{Y} with j columns, and where (typically) $i \ll j$. The different approaches place
126 different constraints on what properties $\hat{\mathbf{Y}}$ must satisfy and which aspects of the data are compared (and
127 how) to characterize the match between $\hat{\mathbf{Y}}$ and \mathbf{Y} .

128 Applying dimensionality reduction algorithms to \mathbf{Y} yields a $\hat{\mathbf{Y}}$ whose columns reflect weighted combi-
129 nations (or nonlinear transformations) of the original columns of \mathbf{Y} . This has two main consequences. First,
130 with each repeated dimensionality reduction, the resulting $\hat{\mathbf{Y}}_n$ has lower and lower fidelity (with respect to

131 what the “true” \mathbf{Y}_n might have looked like without using dimensionality reduction to maintain scalability).
132 In other words, computing $\hat{\mathbf{Y}}_n$ is a lossy operation. Second, whereas the columns of \mathbf{Y}_n may be mapped
133 directly onto pairs of columns of \mathbf{Y}_{n-1} , that mapping either becomes less cleanly defined in $\hat{\mathbf{Y}}_n$ due to the
134 reweightings and/or nonlinear transformations.

135 **Graph theoretic approaches to computing \mathbf{X}_n**

136 Graph theoretic measures take as input a matrix of interactions (e.g., using the above notation, an $F \times F$
137 correlation matrix or binarized correlation matrix reconstituted from a single timepoint’s row of \mathbf{Y}) and
138 return as output a set of F measures describing how each node (feature) sits within that interactions matrix
139 with respect to the rest of the population. Common measures include betweenness centrality (the proportion
140 of shortest paths between each pair of nodes in the population that involves the given node in question; e.g.,
141 Barthélémy, 2004; Freeman, 1977; Geisberger et al., 2008; Newman, 2005; Opsahl et al., 2010); diversity and
142 dissimilarity (characterizations of how differently connected a given node is from others in the population;
143 e.g., Lin, 2009; Rao, 1982; Ricotta & Szeidl, 2006); Eigenvector centrality and pagerank centrality (measures of
144 how influential a given node is within the broader network; e.g., Bonacich, 2007; Halu et al., 2013; Lohmann
145 et al., 2010; Newman, 2008); transfer entropy and flow coefficients (a measure of how much information is
146 flowing from a given node to other nodes in the network; e.g., Honey et al., 2007; Schreiber, 2000); k -coreness
147 centrality (a measure of the connectivity of a node within its local sub-graph; e.g., Alvarez-Hamelin et al.,
148 2005; Christakis & Fowler, 2010); within-module degree (a measure of how many connections a node has to
149 its close neighbors in the network; e.g., Rubinov & Sporns, 2010); participation coefficient (a measure of the
150 diversity of a node’s connections to different sub-graphs in the network; e.g., Rubinov & Sporns, 2010); and
151 sub-graph centrality (a measure of a node’s participation in all of the network’s sub-graphs; e.g., Estrada &
152 Rodríguez-Velázquez, 2005).

153 As an alternative to the above dimensionality reduction approach to embedding \mathbf{Y}_n in a lower-dimensional
154 space, but still allowing for scalable explorations of higher-order structure in the data, we also explore using
155 the above graph theoretic measures as a means of obtaining $\hat{\mathbf{Y}}_n$. In particular: for a given graph theoretic
156 measure, $\eta : \mathcal{R}^{F \times F} \rightarrow \mathcal{R}^F$, we can use η to transform each row of \mathbf{Y}_n in a way that characterizes the cor-
157 responding graph-theoretic properties of each column. Whereas the dimensionality reduction approach
158 to computing $\hat{\mathbf{Y}}_n$ is lossy, the graph-theory approach is lossless. However, whereas the dimensionality
159 reduction approach maintains ties (direct or indirect) to the original activity patterns reflected in \mathbf{Y}_{n-1} , the
160 graph-theory approach does not. Instead, the graph-theory characterizes the nature and timecourse of each
161 feature’s *participation* in the network.

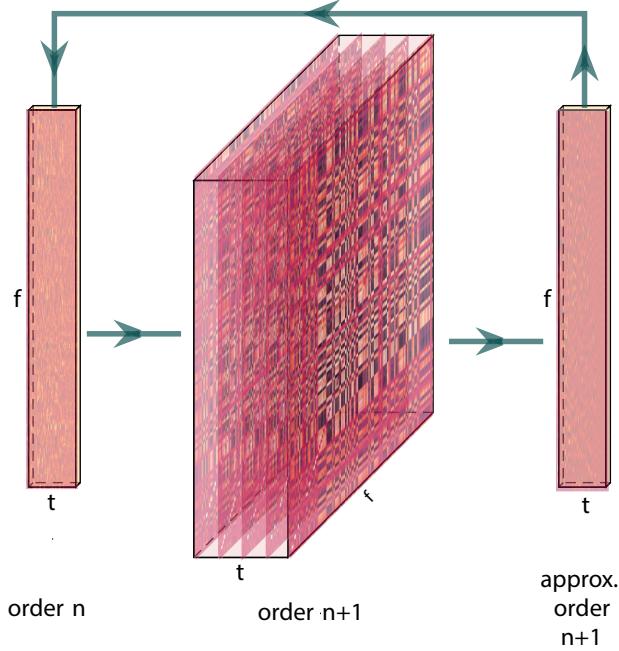


Figure 2: Computing higher order correlations. Dynamic correlations are computed then approximated to the same size as original data using dimensionality reduction. This process is repeated, and can be used to compute up to any arbitrary order with computations scaling linearly (as opposed to exponentially) with order.

162 Higher-order correlations

163 Given a timeseries of dynamic correlations (e.g., obtained using Eqn. 5), higher-order correlations reflect the
 164 dynamic correlations between columns of \mathbf{Y} . Given unlimited computing resources, one could use repeated
 165 applications of Equation 5 to estimate these higher-order correlations (i.e., substituting in the previous
 166 output, \mathbf{Y} , for the input, \mathbf{X} in the equation). However, because each output \mathbf{Y} has $O(F^2)$ columns relative
 167 to F columns in the input \mathbf{X} , the output of Equation 5 grows with the square of the number of repeated
 168 applications (total cost of computing n^{th} order correlations is $O(F^{2n})$ for $n \in \mathcal{J}, n > 0$). When F or n is large,
 169 this approach quickly becomes intractable.

170 To make progress in computing \mathbf{Y}_{n+1} , we can approximate \mathbf{Y}_n by computing an $O(F)$ -dimensional em-
 171 bedding of \mathbf{Y}_n , termed $\hat{\mathbf{Y}}_n$, and then we can apply Equation 5 to $\hat{\mathbf{Y}}_n$ rather than directly to \mathbf{Y}_n . This enables
 172 us to maintain $O(n)$ scaling with respect to n , rather than exponential scaling via the direct approach.

173 There are many possible methods for computing $\hat{\mathbf{Y}}_n$ from \mathbf{Y}_n , including traditional dimensionality
 174 reduction approaches and graph theory based approaches as described next. In the *Discussion* section we
 175 elaborate on other potential approaches.

176 **Evaluation metrics**

177 We evaluate our approach to extracting dynamic correlations and higher-order correlations using several
178 metrics detailed next. First, we generated synthetic data using known time-varying correlations, and then
179 we evaluated the fidelity with which Equation 5 could recover those correlations (for synthetic datasets
180 with different properties, and using different kernels to define the weights; Fig. 1). We then turned to a
181 series of analyses on a (real) neuroimaging dataset where the ground truth correlations were *not* known.
182 We evaluated whether the recovered correlations could be used to accurately label held-out neuroimaging
183 data with the time at which it was collected. We used this latter evaluations (using timepoint decoding)
184 as a proxy for gauging how much explanatory power the recovered correlations held with respect to the
185 observed data.

186 **Generating synthetic data**

187 To explore recovery of a constant covariance (Fig. 3, a.), we generated synthetic data sampled from a constant
188 covariance matrix. To do this, we created one random covariance matrix, K , with 50 features, and for each
189 of the 300 timepoints we sampled from a Gaussian distribution centered on K . Similarly, we generated
190 synthetic data sampled from a random covariance matrix (Fig. 3, b.) by creating a new random covariance
191 matrix $K(t)$, for each of the 300 timepoints and sampled from a Gaussian distribution centered on $K(t)$.

To generate synthetic data from a dynamically changing covariance matrix (Fig. 3, c.), we generated two
random covariance matrices, K_1 and K_2 . We then computed a weighted average covariance matrix for each
of the 300 timepoint, $K(t)$, by taking the linearly spaced weights (w) of the two random matrices,

$$K(t) = w(t) * K_1 + (1 - w(t)) * K_2, \quad (11)$$

$$(12)$$

192 and for each of the 300 timepoints sampled from a Gaussian distribution centered on $K(t)$.

193 Lastly, for the synthetic data containing block structure (Fig. 3, d.), we followed the same process of
194 creating synthetic data sampled from a constant covariance matrix (see above) but sampled from a new
195 random covariance matrix after 60 consecutive timepoints. We then pieced the blocks together to create a
196 synthetic dataset with 300 total timepoints but drawn from 5 separate covariance matrices.

197 **Recovery of ground truth parameters from synthetic data**

198 We applied timecorr, using delta and gaussian (width = 10) kernels Fig. 1) to each of these synthetic datasets,
199 then correlated each recovered correlation matrix with the ground truth. We repeated this process 10 times
200 and explored how recovery varies with the kernel and the specific structure of the data. For the ramping
201 synthetic dataset (Fig. 3, c.) and for the block synthetic dataset (Fig. 3, d.) we made further comparisons
202 of the timecorr recovered correlation matrices. We compared the ramping recovered correlation matrices to
203 only the first random covariance matrix K_1 (First, Fig. 3, c.) and to only the last random covariance matrix
204 K_2 (Last, Fig. 3, c.) from Equation 12. We also compared the block recovered correlation matrices in to the
205 block specific covariance matrix (Block 1-5, Fig. 3, d.).

206 **Timepoint decoding**

207 To explore how higher-order structure varies with stimulus structure and complexity, we used a previous
208 neuroimaging dataset Simony et al. (2016) in which participants listened to an audio recording of a story;
209 36 participants listen to an intact version of the story, 17 participants listen to time-scrambled recordings of
210 the same story where paragraphs were scrambled, 36 participants listen to word-scrambled version and 36
211 participants lay in rest condition.

212 Prior work has shown participants share similar neural responses to richly structured stimuli when
213 compared to stimuli with less structure. To assess whether the moment-by-moment higher order correlations
214 were reliably preserved across participants, we used inter-subject functional connectivity (ISFC) to isolate
215 the time-varying correlational structure (functional connectivity patterns that were specifically driven by
216 the story participants listened to. Following the analyses conducted by (HTFA) Manning et al. (2018), we
217 first applied *hierarchical topographic factor analysis* (HTFA) to the fMRI datasets to obtain a time series of
218 700 node activities for every participant. We then computed the dynamic weighted ISFC using a range of
219 kernels and widths. Specifically, we used Gaussian, Laplace, and mexican hat kernels, as well as widths of
220 5, 10, 20, and 50. We then approximated these dynamic correlation using two reduction measures, PCA and
221 eigenvector centrality, and computed the dynamic weighted ISFC on the approximations. We repeated this
222 process up to 10th order approximated correlations.

223 To assess decoding accuracy, we randomly divided participants for each stimulus into training and
224 testing groups. For the zeroth order, we computed the mean factor activity for each group. For all subsequent
225 orders up to the tenth order, we computed the mean approximated dynamic ISFC of factor activity for each
226 group. To assess how additional higher-order correlations contribute to decoding accuracy, for each order
227 we included a weighted-mixutre (described below) of the activity patterns of all previous orders. For each

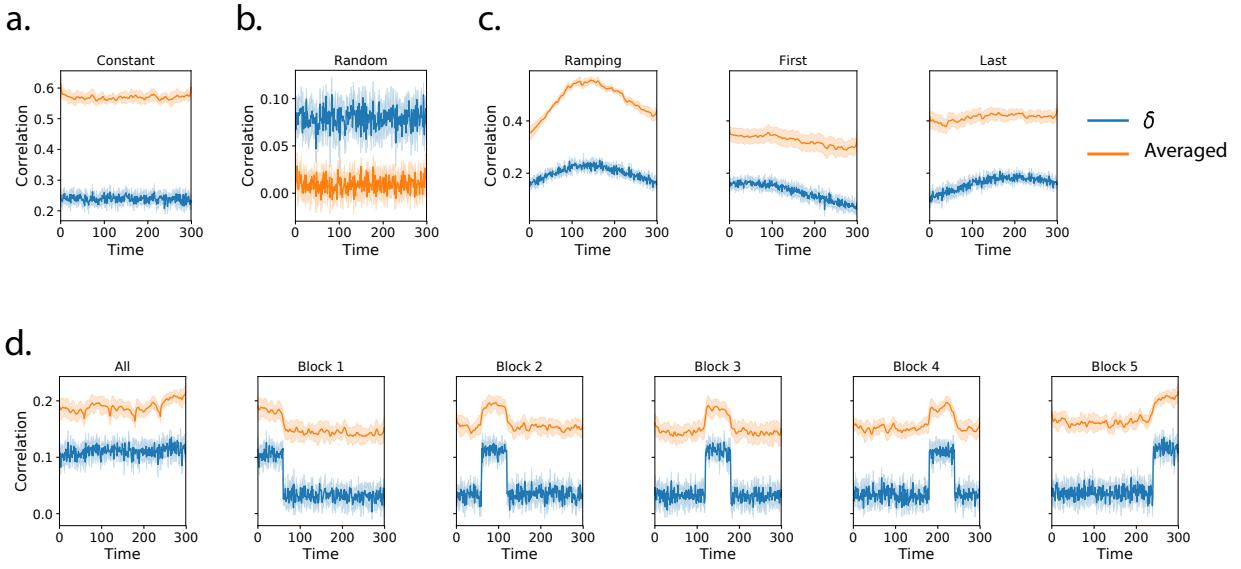


Figure 3: Dynamic correlation recovery with synthetic data. Using synthetic data containing different underlying correlational structure, we test how well we can recover dynamic correlation matrices using different kernels when compared to ground truth. We compare the results using a delta kernel with averaged results from several kernels (Gaussian, Laplace, and mexican hat) and several widths (5, 10, 20, and 50). We plot recovery using of datasets containing the following underlying structure: **a. Constant.** **b. Random.** **c. Ramping.** **d. Block.**

228 group of participants in turn, we compared these activity patterns (using Pearson correlations) to estimate
 229 the story times each pattern corresponded to. Specifically, we asked, for each timepoint: what are the
 230 correlations between the first group's and second group's activity patterns at each order. We note that the
 231 decoding test we used is a conservative in which we count a timepoint label as incorrect if it is not an exact
 232 match.

233 For each order we obtained the weighted-mixture of the correlation matrices for the current order and
 234 all previous orders using mixing parameter, ϕ , where $0 < \phi < 1$ reflects a weighted mixture of order based
 235 decoding Fig. 4 Panel c.). We calculated ϕ , by subdividing the training group and using the quasi-Newton
 236 method of Broyden, Fletcher, Goldfarb, and Shanno (BFGS (Nocedal & Wright, 2006)) for optimization. We
 237 repeated this cross-validation process 10 times for each parameter set.

238 Results

239 Synthetic data

240 To assess the performance of dynamic correlation recovery using timecorr, we varied width the kernel and
 241 the specific structure of the data. We applied timecorr, using delta and gaussian kernels Fig. 1) to each of

242 the following synthetic datasets: constant, random, ramping, and block. We then correlated each recovered
243 correlation matrix with the ground truth.

244 For the constant synthetic dataset, a gaussian kernel (width=10) outperformed the delta kernel (Fig. 3,
245 a.). This is in contrast with the random synthetic dataset, for which the delta kernel best captures the rapidly
246 changing structure (Fig. 3, b.). For the ramping synthetic dataset, the slow changing strucutre within the
247 data is best captured by the gaussian kernel and the best recovery occurs in the middle (Ramping, Fig. 3,
248 c.). In addition to comparing the timecorr recovered correlation matrices to the ground truth, we further
249 compared the ramping recovered correlation matrices to only the first random covariance matrix K_1 (First,
250 Fig. 3, c.) and to only the last random covariance matrix K_2 (Last, Fig. 3, c.), both of which perform best at
251 the beginning and end respectively.

252 Similary for the block sythetic dataset, we compared the timecorr recovered correlation matrices to
253 the ground truth as well as to each block-specific covariance matrix (Block 1-5, Fig. 3, d.). Although the
254 structure is changing by block, the gaussian kernel once again outperforms the delta kernel. Performance
255 does however drop near even boundaries for when using the gaussian kernel.

256 **Neuroimaging dataset (Simony et al., 2016)**

257 For our decoding analysis, we used HTFA-derived node activities Manning et al. (2018) from fMRI data
258 collected as participants listened to an audio recording of a story (intact condition; 36 participants), lis-
259 tened to time scrambled recordings of the same story (17 participants in the paragraph-scrambled condition
260 listened to the paragraphs in a randomized order and 36 in the word-scrambled condition listened to
261 the words in a randomized order), or lay resting with their eyes open in the scanner (rest condition; 36
262 participants). We sought to demonstrate how higher-order correlations may be used to examine dynamic
263 interactions of brain patterns in (real) multi-subject fMRI datasets. This story listening dataset was col-
264 lected as part of a separate study, where the full imaging parameters, image preprocessing methods, and
265 experimental details may be found (Simony et al., 2016). The dataset is available at <http://arks.princeton.edu/ark:/88435/dsp015d86p269k>.

267 We next evaluated if our model of high-order correlations in brain activity can capture cognitively
268 relevant brain patterns. We performed a decoding analysis, using cross validation to estimate (using other
269 participants' data) which parts of the story each weighted-mixture of higher-order brain activity pattern
270 corresponded to (see *Materials and methods*). We note that our primary goal was not to achieve perfect
271 decoding accuracy, but rather to use decoding accuracy as a benchmark for assessing whether different
272 neural features specifically capture cognitively relevant brain patterns.

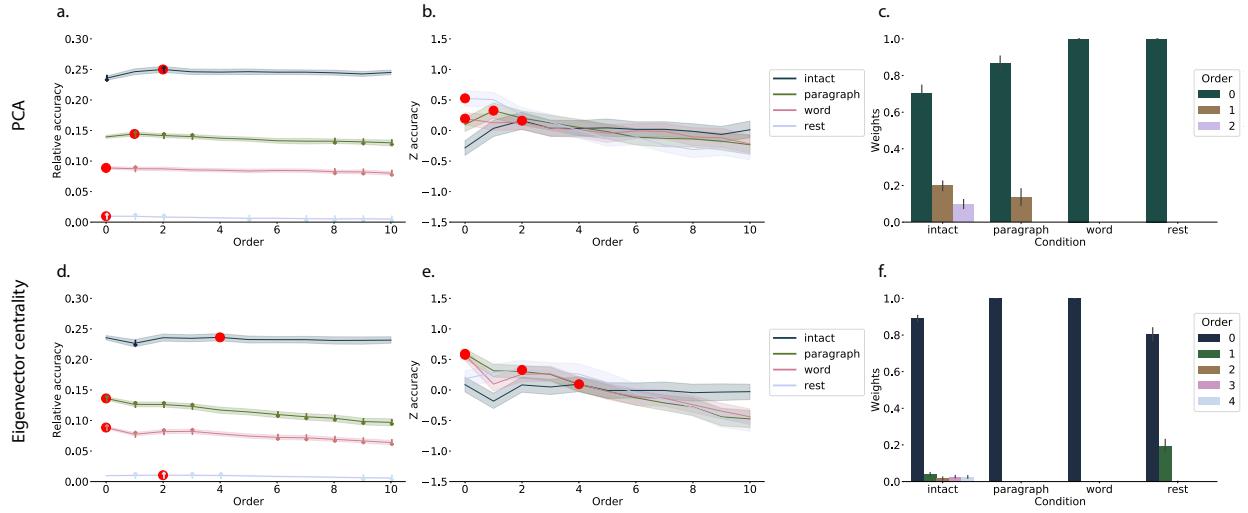


Figure 4: **Decoding by order.** **a.&d. Relative decoding accuracy by order.** Ribbons of each color display cross-validated decoding performance averaged over all parameters for each condition (intact, paragraph, word, and rest) using PCA (a.) or eigenvector centrality (d.) to approximate correlations. Decoders were trained using increasingly more higher-order information and this ribbon are displayed relative to chance at 0. The red dots indicates maximum decoding accuracy for each condition. **b.&e. Z-transformed decoding accuracy by order.** We Z-transformed the decoding accuracy by order to better visualize the order with the maximum decoding accuracy for each condition using PCA (b.) or eigenvector centrality (e.) to approximate correlations. **c.&f. Optimized weights.** Bar heights indicate the optimized mixing parameter ϕ of each contributing order up to and including the order with the maximum decoding accuracy for each contributing order using PCA (c.) or eigenvector centrality (f.) to approximate correlations. For the order with maximum decoding accuracy by condition, we show barplots of the optimized weights ϕ for each contributing order.

Separately for each experimental condition, we divided participants into two groups. For the zeroth order, we computed the mean factor activity for each group. For all subsequent orders up to the tenth order, we computed the mean approximated dynamic ISFC of factor activity for each group (see *Materials and methods*), and combined in a weighted mixture with all previous orders (i.e. cross-validation for the second order contained a weighted-mixture of zeroth, first, and second order (Fig. 4, c.&f.). For each order, we correlated the group 1 activity patterns with group 2 activity patterns. We then subdivided the group 1 to obtain an optimal weighting parameter for each order's correlation matrix using the same cross validation method. We used the optimal weighting parameters to obtain a weighted-mixture (see *Materials and methods*) of each order's correlation matrix. Using these correlations, we labeled the group 1 timepoints using the group 2 timepoints with which they were most highly correlated; we then computed the proportion of correctly labeled group 1 timepoints. (We also performed the symmetric analysis whereby we labeled the group 2 timepoints using the group 1 timepoints as a template.) We repeated this procedure 100 times (randomly re-assigning participants to the two groups each time) to obtain a distribution of decoding accuracies for each experimental condition. There were 272 timepoints for paragraph condition, 300 timepoints for intact and word conditions, and 400 timepoints for rest condition, so chance performance on this decoding test is was $\frac{1}{272}$, $\frac{1}{300}$, and $\frac{1}{400}$ respectively.

We repeated this process for each set of parameters, varying kernel type and width, and averaged over the reduction technique used to approximate the higher-order correlations (PCA Fig. 4, a.-c. and eigenvector centrality Fig. 4, d.-f.). Since there is no ground truth in these analyses, and we did not know which parameters best capture the data, we instead report a robustness search by averaging over the parameters and reporting which results consistently showed up across all parameters.

The two methods used to approximate the higher-order correlations (PCA Fig. 4, a.-c. and eigenvector centrality Fig. 4, d.-f.) capture different facets of the activity patterns. Using PCA, the higher-order correlations are all linked to the original activity patterns, whereas eigenvector centrality breaks the immediate link with specific brain areas and instead characterizes the position of the nodes in the network that are similar over time.

We found for both PCA and eigenvector centrality, during the intact condition in the experiment, classifiers that incorporated higher-order correlations yielded consistently higher accuracy than classifiers trained only on lower-order patterns (Fig. 4, a.&d.). We plot the average correlations for up to the fourth order for the intact condition (Fig. 5) representing the degree of agreement by location pair over time. By contrast, we found that incorporating higher-order (greater than first order) correlations did not further improve decoding accuracy for the other listening conditions or rest condition. This suggests that the cognitive processing that supported the most cognitively rich condition involved higher-order network

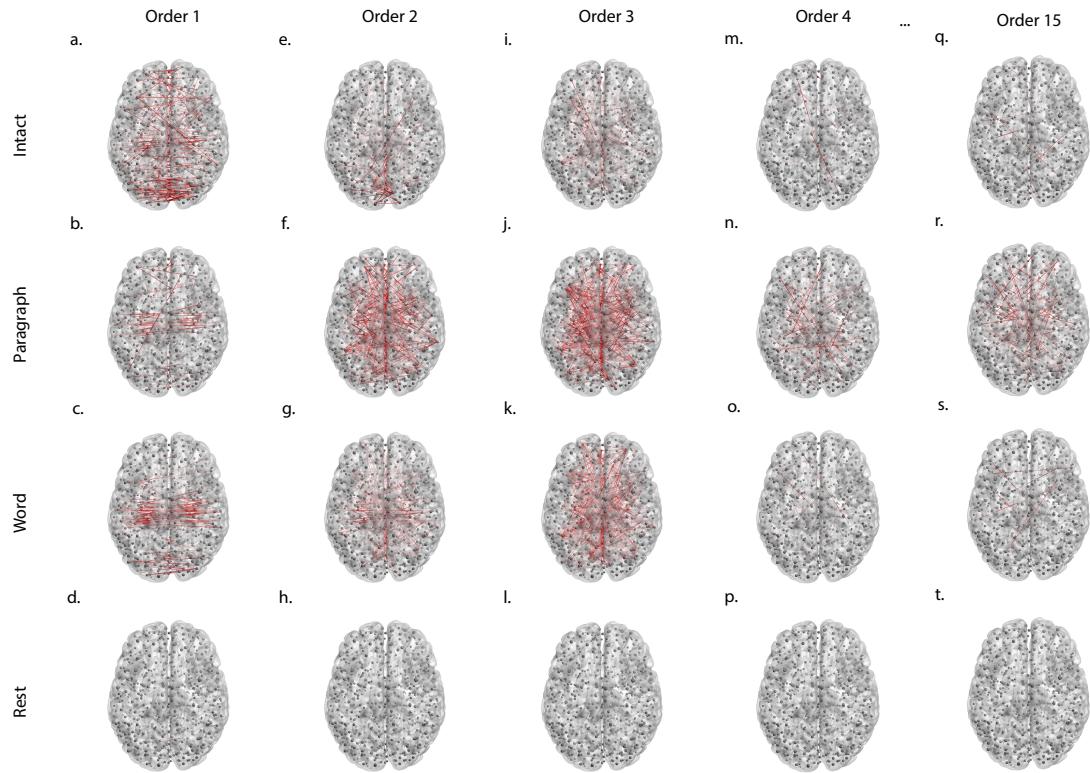


Figure 5: Average correlations by order for the intact listening condition. Using eigenvector centrality to approximate higher-order correlations for the intact, paragraph scrambled, word scrambled, and rest condition. We plot the strongest 50% absolute value mean correlation for **a.-d. first order, e.-h. second order, i.-l. third order, and m.-p. fourth order**, representing the degree of agreement by location pair over time. To demonstrate how this method is computationally scalable, we also approximated **a.-d. fifteenth order** dynamic correlation, which would be possible to compute using conventional methods since it would require more bits to represent the solution than there are molecules in the universe.

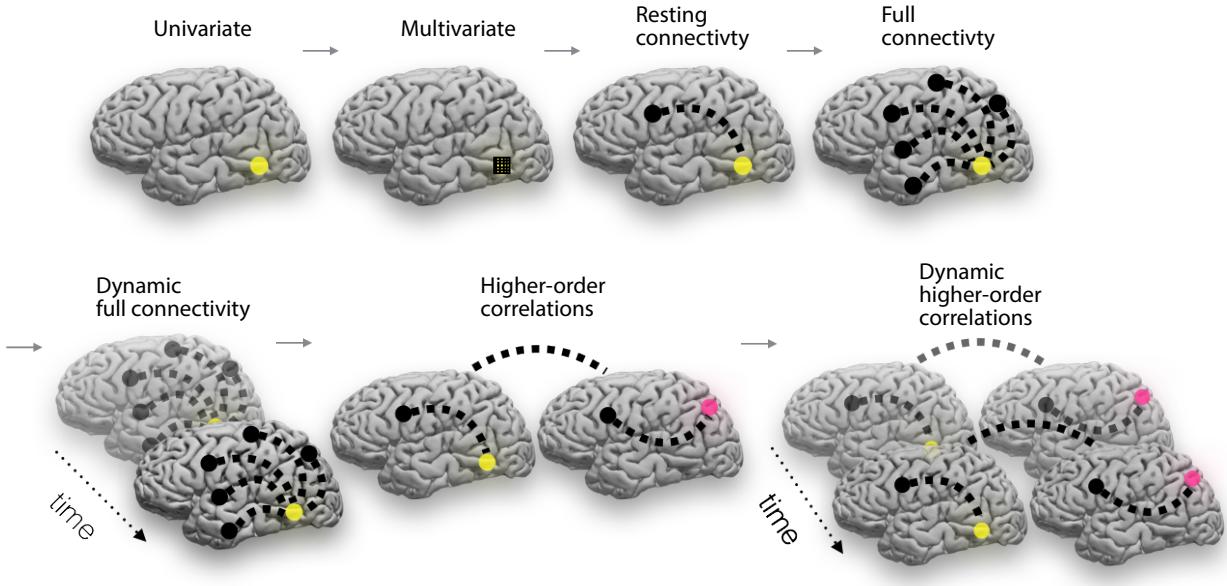


Figure 6: Direction of the field (adapted from (Turk-Browne, 2013)). The evolution of fMRI analyses started with univariate activation, which refers to the average amplitude of BOLD activity evoked by events of an experimental condition. Next, multivariate classifiers are trained on patterns of activation across voxels to decode distributed representations for specific events. The next, resting connectivity, is the temporal correlation of one or more seed regions with the remainder of the brain during rest. Additionally, task-based connectivity examines how these correlations differ by cognitive state. Following this increasing trajectory of increasing complexity, full connectivity considers all pairwise correlations in the brain, most commonly at rest. Next, dynamic full connectivity considers how full connectivity changes over time. Continuing this line of reasoning, we expect higher-order network dynamics might provide even richer insights into the neural basis of cognition.

306 dynamics.

307 Discussion

308 Based on prior work (Demertzis et al., 2019) and following the direction of the field (Turk-Browne, 2013)
 309 we think our thoughts might be encoded in dynamic network patterns, and possibly higher order network
 310 patterns (Fig. 6). We sought to test this hypothesis by developing an approach to inferring high-order
 311 network dynamics from timeseries data.

312 One challenge in studying dynamic interactions is the computational resources required to calculate
 313 higher-order correlations. We developed a computationally tractable model of network dynamics (Fig. 2)
 314 that takes in a feature timeseries and outputs approximated first-order dynamics (i.e., dynamic functional
 315 correlations), second-order dynamics (reflecting homologous networks that dynamically form and disperse),
 316 and higher-order network dynamics (up to tenth-order dynamic correlations).

317 We first validated our model using synthetic data, and explored how recovery varied with different

318 underlying data structures and kernels. We then applied the approach to an fMRI dataset (Simony et al.,
319 2016) in which participants listened to an audio recording of a story, as well as scrambled versions of the
320 same story (where the scrambling was applied at different temporal scales). We trained classifiers to take
321 the output of the model and decode the timepoint in the story (or scrambled story) that the participants
322 were listening to. We found that, during the intact listening condition in the experiment, classifiers that
323 incorporated higher-order correlations yielded consistently higher accuracy than classifiers trained only
324 on lower-order patterns (Fig. 4, a.&d.). By contrast, these higher-order correlations were not necessary
325 to support decoding the other listening conditions and (minimally above chance) during a control rest
326 condition. This suggests that the cognitive processing that supported the most cognitively rich listening
327 conditions involved second-order (or higher) network dynamics.

328 Although we found decoding accuracy was best when incorporating higher-order network dynamics
329 for all but rest condition, it is unclear if this is a product of the brain or the data collection technique. It could
330 be that the brain is second-order or it could be that fMRI can only reliably give second-order interactions.
331 Exploring this method with other data collection technique will be important to disentangle this question.

332 **Concluding remarks**

333 How can we better understand how brain patterns change over time? How can we quantify the potential
334 network dynamics that might be driving these changes? One way to judge the techniques of the future is
335 to look at the trajectory of the fMRI field so far has taken so far (Fig. 2). The field started with univariate
336 activation, measuring the average activity for each voxel. Analyses of multivariate activation followed,
337 looking at spatial patterns of activity over voxels. Next, correlations of activity were explored, first with
338 measures like resting connectivity that take temporal correlation between a seed voxel and all other voxels
339 then with full connectivity that measure all pairwise correlations. Additionally, this path of increasing
340 complexity also moved from static to dynamic measurements. One logical next step in this trajectory would
341 be dynamic higher-order correlations. We have created a method to support these calculations by scalably
342 approximating dynamic higher-order correlations.

343 **Acknowledgements**

344 We acknowledge discussions with Luke Chang, Hany Farid, Paxton Fitzpatrick, Andrew Heusser, Eshin
345 Jolly, Qiang Liu, Matthijs van der Meer, Judith Mildner, Gina Notaro, Stephen Satterthwaite, Emily Whitaker,
346 Weizhen Xie, and Kirsten Ziman. Our work was supported in part by NSF EPSCoR Award Number 1632738

347 to J.R.M. and by a sub-award of DARPA RAM Cooperative Agreement N66001-14-2-4-032 to J.R.M. The
348 content is solely the responsibility of the authors and does not necessarily represent the official views of our
349 supporting organizations.

350 **Author contributions**

351 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.

352 **References**

- 353 Alvarez-Hamelin, I., Dall'Asta, L., Barrat, A., & Vespignani, A. (2005). *k*-corr decomposition: a tool for the
354 visualization of large scale networks. *arXiv*, cs/0504107v2.
- 355 Barthélemy, M. (2004). Betweenness centrality in large complex networks. *European Physical Journal B*, 38,
356 163–168.
- 357 Bassett, D., Meyer-Lindenberg, A., Achard, S., Duke, T., & Bullmore, E. (2006, December). Adaptive
358 reconfiguration of fractal small-world human brain functional networks. *Proceedings of the National
359 Academy of Sciences, USA*, 103(51), 19518-23.
- 360 Bonacich, P. (2007). Some unique properties of eigenvector centrality. *Social Networks*, 29(4), 555–564.
- 361 Bullmore, E., & Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and
362 functional systems. *Nature Reviews Neuroscience*, 10(3), 186–198.
- 363 Christakis, N. A., & Fowler, J. H. (2010). Social network sensors for early detection of contagious outbreaks.
364 *PLoS One*, 5(9), e12948.
- 365 Comon, P., Jutten, C., & Herault, J. (1991). Blind separation of sources, part II: Problems statement. *Signal
366 Processing*, 24(1), 11 - 20.
- 367 Demertzi, A., Tagliazucchi, E., Dehaene, S., Deco, G., Barttfeld, P., Raimondo, F., . . . Sitt, J. D. (2019). Human
368 consciousness is supported by dynamic complex patterns of brain signal coordination. *Science Advances*,
369 5(2), eaat7603.
- 370 Estrada, E., & Rodríguez-Velázquez, J. A. (2005). Subgraph centrality in complex networks. *Physical Review
371 E*, 71(5), 056103.

- 372 Etzel, J. A., Gazzola, V., & Keysers, C. (2009). An introduction to anatomical ROI-based fMRI classification.
373 *Brain Research*, 1282, 114–125.
- 374 Fong, A. H. C., Yoo, K., Rosenberg, M. D., Zhang, S., Li, C.-S. R., Scheinost, D., … Chun, M. M. (2019).
375 Dynamic functional connectivity during task performance and rest predicts individual differences in
376 attention across studies. *NeuroImage*, 188, 14–25.
- 377 Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry*, 40(1), 35–41.
- 378 Geisberger, R., Sanders, P., & Schultes, D. (2008). Better approximation of betweenness centrality. *Proceedings*
379 *of the meeting on Algorithm Engineering and Experiments*, 90–100.
- 380 Gershman, S., Blei, D., Pereira, F., & Norman, K. (2011). A topographic latent source model for fMRI data.
381 *NeuroImage*, 57, 89–100.
- 382 Halu, A., Mondragón, R. J., Panzarasa, P., & Bianconi, G. (2013). Multiplex PageRank. *PLoS One*, 8(10),
383 e78293.
- 384 Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and
385 overlapping representations of faces and objects in ventral temporal cortex. *Science*, 293, 2425–2430.
- 386 Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks.
387 *Science*, 313(5786), 504–507.
- 388 Honey, C. J., Kotter, R., Breakspear, M., & Sporns, O. (2007). Network structure of cerebral cortex shapes
389 functional connectivity on multiple time scales. *Proceedings of the National Academy of Science USA*, 104(24),
390 10240–10245.
- 391 Huth, A. G., de Heer, W. A., Griffiths, T. L., Theunissen, F. E., & Gallant, J. L. (2016). Natural speech reveals
392 the semantic maps that tile human cerebral cortex. *Nature*, 532, 453–458.
- 393 Huth, A. G., Nisimoto, S., Vu, A. T., & Gallant, J. L. (2012). A continuous semantic space describes
394 the representation of thousands of object and action categories across the human brain. *Neuron*, 76(6),
395 1210–1224.
- 396 Jutten, C., & Herault, J. (1991). Blind separation of sources, part I: An adaptive algorithm based on
397 neuromimetic architecture. *Signal Processing*, 24(1), 1–10.
- 398 Kamitani, Y., & Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nature*
399 *Neuroscience*, 8, 679–685.

- 400 Landau, E. (1895). Zur relativen Wertbemessung der Turnierresultate. *Deutsches Wochenschach*, 11, 366–369.
- 401 Lee, D. D., & Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature*,
402 401, 788–791.
- 403 Lin, J. (2009). Divergence measures based on the Shannon entropy. *IEEE Transactions on Information Theory*,
404 37(1), 145–151.
- 405 Lohmann, G., Margulies, D. S., Horstmann, A., Pleger, B., Lepsien, J., Goldhahn, D., ... Turner, R. (2010).
406 Eigenvector centrality mapping for analyzing connectivity patterns in fMRI data of the human brain.
407 *PLoS One*, 5(4), e10232.
- 408 Mairal, J., Ponce, J., Sapiro, G., Zisserman, A., & Bach, F. R. (2009). Supervised dictionary learning. *Advances
409 in Neural Information Processing Systems*, 1033–1040.
- 410 Mairal, J. B., Bach, F., Ponce, J., & Sapiro, G. (2009). Online dictionary learning for sparse coding. *Proceedings
411 of the 26th annual international conference on machine learning*, 689–696.
- 412 Manning, J. R., Ranganath, R., Norman, K. A., & Blei, D. M. (2014). Topographic factor analysis: a Bayesian
413 model for inferring brain networks from neural data. *PLoS One*, 9(5), e94914.
- 414 Manning, J. R., Zhu, X., Willke, T. L., Ranganath, R., Stachenfeld, K., Hasson, U., ... Norman, K. A. (2018).
415 A probabilistic approach to discovering dynamic full-brain functional connectivity patterns. *NeuroImage*,
416 180, 243–252.
- 417 McInnes, L., & Healy, J. (2018). UMAP: Uniform manifold approximation and projection for dimension
418 reduction. *arXiv*, 1802(03426).
- 419 Mitchell, T., Shinkareva, S., Carlson, A., Chang, K., Malave, V., Mason, R., & Just, M. (2008). Predicting
420 human brain activity associated with the meanings of nouns. *Science*, 320(5880), 1191.
- 421 Newman, M. E. J. (2005). A measure of betweenness centrality based on random walks. *Social Networks*, 27,
422 39–54.
- 423 Newman, M. E. J. (2008). The mathematics of networks. *The New Palgrave Encyclopedia of Economics*, 2, 1–12.
- 424 Nishimoto, S., Vu, A., Naselaris, T., Benjamini, Y., Yu, B., & Gallant, J. (2011). Reconstructing visual
425 experience from brain activity evoked by natural movies. *Current Biology*, 21, 1 - 6.
- 426 Nocedal, J., & Wright, S. J. (2006). *Numerical optimization*. New York: Springer.

- 427 Norman, K. A., Newman, E., Detre, G., & Polyn, S. M. (2006). How inhibitory oscillations can train neural
428 networks and punish competitors. *Neural Computation*, 18, 1577–1610.
- 429 Opsahl, T., Agneessens, F., & Skvoretz, J. (2010). Node centrality in weighted networks: generalizing degree
430 and shortest paths. *Social Networks*, 32, 245–251.
- 431 Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. *The London, Edinburgh,
432 and Dublin Philosophical Magazine and Journal of Science*, 2, 559–572.
- 433 Pereira, F., Lou, B., Pritchett, B., Ritter, S., Gershman, S. J., Kanwisher, N., ... Fedorenko, E. (2018). Toward
434 a universal decoder of linguistic meaning from brain activation. *Nature Communications*, 9(963), 1–13.
- 435 Rao, C. R. (1982). Diversity and dissimilarity coefficients: a unified approach. *Theoretical Population Biology*,
436 21(1), 24–43.
- 437 Ricotta, C., & Szeidl, L. (2006). Towards a unifying approach to diversity measures: Bridging the gap
438 between the Shannon entropy and Rao's quadratic index. *Theoretical Population Biology*, 70(3), 237–243.
- 439 Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpreta-
440 tions. *NeuroImage*, 52, 1059–1069.
- 441 Schreiber, T. (2000). Measuring information transfer. *Physical Review Letters*, 85(2), 461–464.
- 442 Simony, E., Honey, C. J., Chen, J., & Hasson, U. (2016). Uncovering stimulus-locked network dynamics
443 during narrative comprehension. *Nature Communications*, 7(12141), 1–13.
- 444 Spearman, C. (1904). General intelligence, objectively determined and measured. *American Journal of
445 Psychology*, 15, 201–292.
- 446 Sporns, O., & Honey, C. J. (2006). Small worlds inside big brains. *Proceedings of the National Academy of
447 Science USA*, 103(51), 19219–19220.
- 448 Tipping, M. E., & Bishop, C. M. (1999). Probabilistic principal component analysis. *Journal of Royal Statistical
449 Society, Series B*, 61(3), 611–622.
- 450 Tong, F., & Pratte, M. S. (2012). Decoding patterns of human brain activity. *Annual Review of Psychology*, 63,
451 483–509.
- 452 Turk-Browne, N. B. (2013). Functional interactions as big data in the human brain. *Science*, 342, 580–584.
- 453 van der Maaten, L. J. P., & Hinton, G. E. (2008). Visualizing high-dimensional data using t-SNE. *Journal of
454 Machine Learning Research*, 9, 2579–2605.

⁴⁵⁵ Zar, J. H. (2010). *Biostatistical analysis*. Prentice-Hall/Pearson.