

1 High-order dynamic neural correlations reflect naturalistic 2 processing in humans

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

April 5, 2019

Abstract

Our thoughts arise from coordinated activity patterns across our brain. We examined high-order dynamic correlations in functional neuroimaging data collected as human participants listened to different auditory stimuli varying in cognitive richness, along with an additional resting state condition. Our approach combines a kernel-based method for estimating the dynamic functional correlations that are similar (within task) across participants, along with a dimensionality reduction approach that enables us to efficiently compute high-order correlations in the data. We trained classifiers to decode the precise time, relative to the start of the stimulus, when a given neural pattern was recorded. We trained these classifiers using the neural activity timeseries, first-order dynamic correlations, and higher-order correlations (up to tenth-order correlations), and asked which types of features led to the highest decoding accuracy. We found that second-order correlations consistently yielded the highest decoding accuracy in all of the listening conditions of the experiment, whereas first-order correlations yielded the highest decoding accuracy at rest.

18 Introduction

19 The dynamics of the observable universe are meaningful in three respects. First, the behaviors of the *atomic*
20 *units* that exhibit those dynamics are highly interrelated. The actions of one unit typically have implications
21 for one or more other units. In other words, there is non-trivial *correlational structure* defining how different
22 units interact with and relate to each other. Second, that correlational structure is *hierarchical* in the sense
23 that it exists on many spatiotemporal scales. The way one group of units interacts may relate to how another
24 group of units interact, and the interactions between those groups may exhibit some rich structure. Third,
25 the structure at each level of this correlational hierarchy changes from moment to moment, reflecting the
26 “behavior” of the full system.

27 These three properties (rich correlations, hierarchical organization, and dynamics) are major hallmarks
28 of many complex systems. For example, within a single cell, the cellular components interact at many
29 spatiotemporal scales, and those interactions change according to what that single cell is doing. Within a

30 single human brain, the individual neurons interact within each brain structure, and the structures interact
31 to form complex networks. The interactions at each scale vary according to the functions our brains are
32 carrying out. And within social groups, interactions at different scales (e.g., between individuals, family
33 units, communities, etc.) vary over time according to changing goals and external constraints.

34 Although many systems exhibit rich dynamic correlations at many scales, a major challenge to studying
35 such patterns is that typically neither the correlations nor the hierarchical organizations of those correlations
36 may be directly observed. Rather, these fundamental properties must be inferred indirectly by examining
37 the observable parts of the system— e.g., the behaviors of the individual atomic units of that system. In
38 the *Methods* section, we propose a series of mathematical operations that may be used to recover dynamic
39 correlations at a range of scales (i.e., orders of interaction). In the *Results* section, we demonstrate how our
40 approach may be applied to multi-dimensional timeseries data: a synthetic dataset where the underlying
41 dynamic correlations are known (we use this dataset to validate our approach), and a neuroimaging dataset
42 comprising data collected as participants listened to a story (Simony et al., 2016). In different experimental
43 conditions in the neuroimaging study, participants listened to altered versions of the story that varied in
44 cognitive richness: the intact story (fully engaging), a scrambled version of the story where the paragraphs
45 were presented in a randomized order (moderately engaging), a second scrambled condition where the
46 words were presented in a random order (minimally engaging), and a “rest” condition where the participants
47 did not listen to any version of the story (control condition). We use the neuroimaging dataset to examine
48 how higher-order structure in brain data varies as a function of the cognitive richness of the stimulus.

49 Methods

50 There are two basic steps to our approach (Fig. 2). In the first step, we take a number-of-timepoints (T) by
51 number-of-features (F) *matrix of observations* (\mathbf{X}) and we return a T by $\frac{F^2-F}{2}$ *matrix of dynamic correlations* (\mathbf{Y}).
52 Here \mathbf{Y}_0 describes, at each moment, how all of the features (columns of \mathbf{X}) are inferred to be interacting.
53 (Since the interactions are assumed to be non-recurrent and symmetric, only the upper triangle of the full
54 correlation matrix is computed.) In the second step, we project \mathbf{Y}_0 onto an F -dimensional space, resulting in
55 a new T by F matrix \mathbf{Y}_1 . Note that \mathbf{Y}_1 contains information about the correlation dynamics present in \mathbf{X} , but
56 represented in a compressed number of dimensions. By repeatedly applying these two steps in sequence,
57 we can examine and explore higher order dynamic correlations in \mathbf{X} .

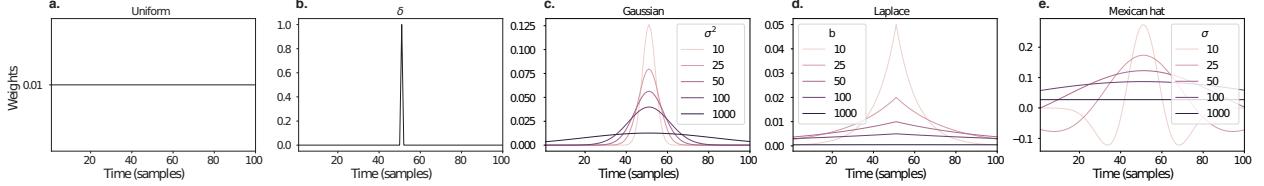


Figure 1: **Examples of time-varying weights.** Each panel displays per-timepoint weights at $t = 50$, evaluated for 100 timepoints (1, ..., 100). **a. Uniform weights.** The weights are timepoint-invariant; observations at all timepoints are weighted equally, and do not change as a function of t . This is a special case of weight function that reduces dynamic correlations to static correlations. **b. Dirac delta function.** Only the observation at timepoint t is given weight (of 1), and weights for observations at all other timepoints are set to 0. **c. Gaussian weights.** Each observation's weights fall off in time according to a Gaussian probability density function centered on $\mu = t$. Weights derived using several different example variance parameters (σ^2) are displayed. **d. Laplace weights.** Each observation's weights fall off in time according to a Laplace probability density function centered on $\mu = t$. Weights derived using several different example scale parameters (b) are displayed. **e. Mexican hat (Ricker wavelet) weights.** Each observation's weights fall off in time according to a Ricker wavelet centered on 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.

58 Dynamic correlations

Given a matrix of observations, we can compute the (static) correlations between any pair of observations, \mathbf{X}_i and \mathbf{X}_j using:

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

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

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

59 We can generalize this formula to compute time-varying correlations by incorporating a *weight function*
60 that takes a time t as input, and returns how much the observed data every timepoint (including t) contribute
61 to the correlations at time t (Fig. 1).

Given a weight function $w(t)$ for timepoint t , evaluated at timepoints in the interval $[1, \dots, T]$, we can

extend the static correlation formula in Equation 2 to reflect an *instantaneous correlation* at timepoint t :

$$\text{timecorr}(\mathbf{X}_i, \mathbf{X}_j, t) = \frac{\sum_{t=1}^T (\mathbf{X}_i(t) - \tilde{\mathbf{X}}_i(t))(\mathbf{X}_j(t) - \tilde{\mathbf{X}}_j(t))}{\sqrt{\sum_{t=1}^T \tilde{\sigma}_{\mathbf{X}_i}^2(t) \tilde{\sigma}_{\mathbf{X}_j}^2(t)}}, \text{ where} \quad (4)$$

$$\tilde{\mathbf{X}}_k(t) = \sum_{i=1}^T w(t, i) \mathbf{X}_k(i), \quad (5)$$

$$\tilde{\sigma}_{\mathbf{X}_k}^2(t) = \sum_{i=1}^T (\mathbf{X}_k(i) - \tilde{\mathbf{X}}_k(t))^2, \quad (6)$$

- and $w(t, i)$ is shorthand for $w(t)$ evaluated at timepoint i . Equation 5 may be used to estimate the instantaneous correlations between every pair of observations, at each timepoint (i.e., \mathbf{Y}).

64 Inter-subject dynamic correlations

Equation 5 provides a means of taking a single observation matrix, \mathbf{X} and estimating the dynamic correlations from moment to moment, \mathbf{Y} . 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 $\{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_P\}$ reflect the T by F observation matrices 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:

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

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 $\mathbf{Y}_i(t)$ denotes the correlation matrix (Eqn. 2) between each column of \mathbf{X}_i and each column of the average observations from all *other* participants, $\bar{\mathbf{X}}_{\setminus i}$:

$$\bar{\mathbf{X}}_{\setminus i} = R \left(\frac{1}{P-1} \sum_{i \in \setminus i} Z(\mathbf{X}_i) \right), \quad (10)$$

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

67 Higher-order correlations

Given a timeseries of dynamic correlations (e.g., obtained using Eqn. 5), higher-order correlations reflect the dynamic correlations between columns of \mathbf{Y} . Given unlimited computing resources, one could use repeated applications of Equation 5 to estimate these higher-order correlations (i.e., substituting in the previous output, \mathbf{Y} , for the input, \mathbf{X} in the equation). However, because each output \mathbf{Y} has $O(F^2)$ columns relative to F columns in the input \mathbf{X} , the output of Equation 5 grows with the square of the number of repeated 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, this approach quickly becomes intractable.

To make progress in computing \mathbf{Y}_{n+1} , we can approximate \mathbf{Y}_n by computing an $O(F)$ -dimensional embedding 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 us to maintain $O(n)$ scaling with respect to n , rather than exponential scaling via the direct approach.

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

81 Dimensionality reduction-based approaches to computing $\hat{\mathbf{Y}}_n$

Commonly used dimensionality reduction algorithms include Principal Components Analysis (PCA; Pearson, 1901), Probabilistic PCA (PPCA; Tipping & Bishop, 1999), Exploratory Factor Analysis (EFA; Spearman, 1904), Independent Components Analysis (ICA; Jutten & Herault, 1991; Comon et al., 1991), t -Stochastic Neighbor Embedding (t -SNE; van der Maaten & Hinton, 2008), Uniform Manifold Approximation and Projection (UMAP; McInnes & Healy, 2018), non-negative matrix factorization (NMF; Lee & Seung, 1999), Topographic Factor Analysis (TFA) Manning et al. (2014), Hierarchical Topographic Factor analysis (HTFA) Manning et al. (2018), Topographic Latent Source Analysis (TLSA) Gershman et al. (2011), Dictionary learning (J. B. Mairal et al., 2009; J. Mairal et al., 2009), deep autoencoders (Hinton & Salakhutdinov,

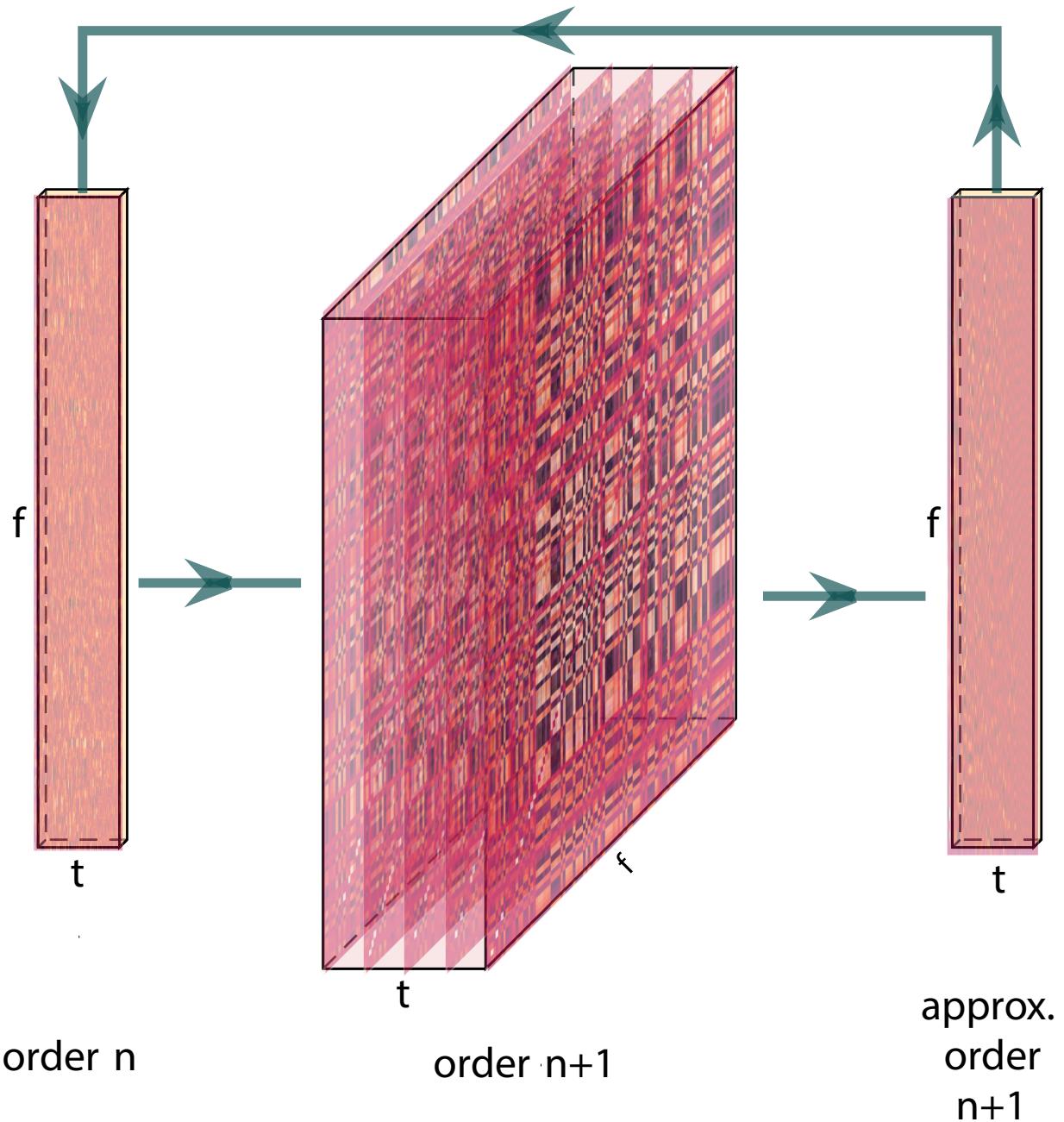


Figure 2: **Computing higher order correlations.** Correlations are computed then approximated to the same size as original data. This process is repeated to computer

90 2006), among others. While complete characterizations of each of these algorithms is beyond the scope of
91 the present manuscript, the general intuition driving these approaches is to compute the $\hat{\mathbf{Y}}$ with i columns
92 that is closest to the original \mathbf{Y} with j columns, and where (typically) $i \ll j$. The different approaches place
93 different constraints on what properties $\hat{\mathbf{Y}}$ must satisfy and which aspects of the data are compared (and
94 how) to characterize the match between $\hat{\mathbf{Y}}$ and \mathbf{Y} .

95 Applying dimensionality reduction algorithms to \mathbf{Y} yields a $\hat{\mathbf{Y}}$ whose columns reflect weighted combi-
96 nations (or nonlinear transformations) of the original columns of \mathbf{Y} . This has two main consequences. First,
97 with each repeated dimensionality reduction, the resulting $\hat{\mathbf{Y}}_n$ has lower and lower fidelity (with respect to
98 what the “true” \mathbf{Y}_n might have looked like without using dimensionality reduction to maintain scalability).
99 In other words, computing $\hat{\mathbf{Y}}_n$ is a lossy operation. Second, whereas the columns of \mathbf{Y}_n may be mapped
100 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
101 reweightings and/or nonlinear transformations.

102 Graph theory-based approaches to computing $\hat{\mathbf{Y}}_n$

103 Graph theoretic measures take as input a matrix of interactions (e.g., using the above notation, an $F \times F$
104 correlation matrix or binarized correlation matrix reconstituted from a single timepoint’s row of \mathbf{Y}) and
105 return as output a set of F measures describing how each node (feature) sits within that interactions matrix
106 with respect to the rest of the population. Common measures include betweenness centrality (the proportion
107 of shortest paths between each pair of nodes in the population that involves the given node in question; e.g.,
108 Newman, 2005; Opsahl et al., 2010; Barthélemy, 2004; Geisberger et al., 2008; Freeman, 1977); diversity and
109 dissimilarity (characterizations of how differently connected a given node is from others in the population;
110 e.g., Rao, 1982; Lin, 2009; Ricotta & Szeidl, 2006); Eigenvector centrality and pagerank centrality (measures of
111 how influential a given node is within the broader network; e.g., Newman, 2008; Bonacich, 2007; Lohmann
112 et al., 2010; Halu et al., 2013); transfer entropy and flow coefficients (a measure of how much information is
113 flowing from a given node to other nodes in the network; e.g., Honey et al., 2007; Schreiber, 2000); k -coreness
114 centrality (a measure of the connectivity of a node within its local sub-graph; e.g., Alvarez-Hamelin et al.,
115 2005; Christakis & Fowler, 2010); within-module degree (a measure of how many connections a node has to
116 its close neighbors in the network; e.g., Rubinov & Sporns, 2010); participation coefficient (a measure of the
117 diversity of a node’s connections to different sub-graphs in the network; e.g., Rubinov & Sporns, 2010); and
118 sub-graph centrality (a measure of a node’s participation in all of the network’s sub-graphs; e.g., Estrada &
119 Rodríguez-Velázquez, 2005).

120 As an alternative to the above dimensionality reduction approach to embedding \mathbf{Y}_n in a lower-dimensional

space, but still allowing for scalable explorations of higher-order structure in the data, we also explore using the above graph theoretic measures as a means of obtaining $\hat{\mathbf{Y}}_n$. In particular: for a given graph theoretic 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 corresponding graph-theoretic properties of each column. Whereas the dimensionality reduction approach to computing $\hat{\mathbf{Y}}_n$ is lossy, the graph-theory approach is lossless. However, whereas the dimensionality reduction approach maintains ties (direct or indirect) to the original activity patterns reflected in \mathbf{Y}_{n-1} , the graph-theory approach does not. Instead, the graph-theory characterizes the nature and timecourse of each feature's *participation* in the network.

129 Evaluation metrics

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

139 Generating synthetic data

140 Ramping dataset and block dataset. Add details (Fig. 3)

141 Recovery of ground truth parameters from synthetic data

To explore recovery of a constant correlation, we generated synthetic data sampled from a constant covariance matrix. To do this, we created a random correlation matrix K of size s , and added a small amount of Gaussian noise to create a timeseries of for s subjects. $\sim N(\mu, \sigma^2)$

142 We applied timecorr with a given kernel (Fig. 1) to synthetic data, then correlate each recovered correlation
143 matrix with the ground truth. Explore how recovery varies with the kernel, kernel parameters, and specific
144 structure of the data (e.g. slow changes as in the ramping dataset, versus rapid changes as in the block
145 dataset).

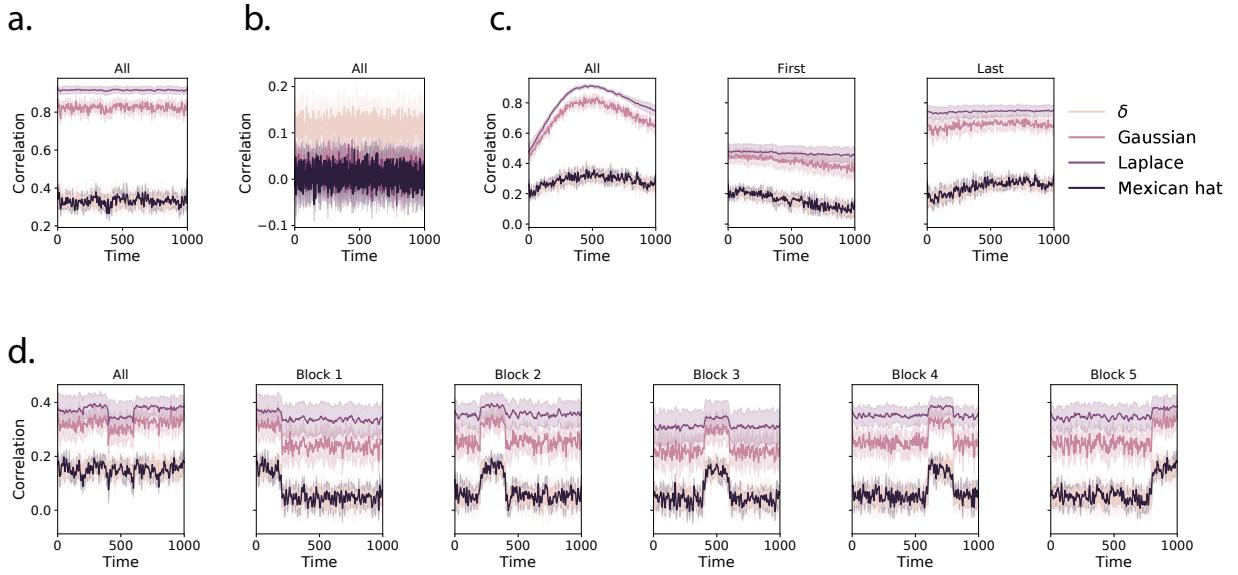


Figure 3: **Dynamic correlation recovery with synthetic data.** **a.** Recovery using a constant dataset. Using constant dataset, how well can we recover using different kernels. **b.** Recovery using random dataset. **c.** Ramping recovery. **d.** Block recovery

146 Timepoint decoding

147 To explore how higher-order structure varies with stimulus structure and complexity, we used a previous
 148 neuroimaging dataset Simony et al. (2016) in which participants listened to an audio recording of a story;
 149 36 participants listen to an intact version of the story, 17 participants listen to time-scrambled recordings of
 150 the same story where paragraphs were scrambled, 36 participants listen to word-scrambled version and 36
 151 participants lay in rest condition.

152 Prior work has shown participants share similar neural responses to richly structured stimuli when
 153 compared to stimuli with less structure. To assess whether the moment-by-moment higher order correlations
 154 were reliably preserved across participants, we used inter-subject functional connectivity (ISFC) to isolate
 155 the time-varying correlational structure (functional connectivity patterns that were specifically driven by
 156 the story participants listened to. Following the analyses conducted by (HTFA) Manning et al. (2018), we
 157 first applied *hierarchical topographic factor analysis* (HTFA) to the fMRI datasets to obtain a time series of
 158 700 node activities for every participant. We then computed the dynamic weighted ISFC using a gaussian
 159 kernel with a width of 5. We then approximated these dynamic correlation using PCA and computed the
 160 dynamic weighted ISFC on the approximations. We repeated this process up to 10th order approximated
 161 correlations.

162 To assess decoding accuracy, we randomly divided participants for each stimulus into training and

163 testing groups. For the zeroth order, we computed the mean factor activity for each group. For all subsequent
164 orders up to the tenth order, we computed the mean approximated dynamic ISFC of factor activity for each
165 group. To assess how additional higher-order correlations contribute to decoding accuracy, for each order
166 we included a weighted-mixture (described below) of the activity patterns of all previous orders. For each
167 group of participants in turn, we compared these activity patterns (using Pearson correlations) to estimate
168 the story times each pattern corresponded to. Specifically, we asked, for each timepoint: what are the
169 correlations between the first group's and second group's activity patterns at each order. We note that the
170 decoding test we used is a conservative in which we count a timepoint label as incorrect if it is not an exact
171 match.

172 For each order we obtained the weighted-mixture of the correlation matrices for the current order and
173 all previous orders using mixing parameter ϕ , where $0 < \phi < 1$ reflects a weighted mixture of order
174 based decoding Fig. ?? Panel C.). We calculated ϕ , by subdividing the training group and using the quasi-
175 Newton method of Broyden, Fletcher, Goldfarb, and Shanno (BFGS) for optimization. We repeated this
176 cross-validation process 100 times.

177 Results

178 Synthetic data

179 Figure: overall timecourse of recovery, also recovery near event boundaries.

180 Neuroimaging dataset (Simony et al., 2016)

181 For our decoding analysis, we used HTFA-derived node activities Manning et al. (2018) from fMRI data
182 collected as participants listened to an audio recording of a story (intact condition; 36 participants), lis-
183 tened to time scrambled recordings of the same story (17 participants in the paragraph-scrambled condition
184 listened to the paragraphs in a randomized order and 36 in the word-scrambled condition listened to
185 the words in a randomized order), or lay resting with their eyes open in the scanner (rest condition; 36
186 participants). We sought to demonstrate how higher-order correlations may be used to examine dynamic
187 interactions of brain patterns in (real) multi-subject fMRI datasets. This story listening dataset was col-
188 lected as part of a separate study, where the full imaging parameters, image preprocessing methods, and
189 experimental details may be found (Simony et al., 2016). The dataset is available at <http://arks.princeton.edu/ark:/88435/dsp015d86p269k>. Bars of each color display cross-validated decoding performance for
190 decoders trained using different sets of neural features: whole-brain patterns of voxel activities

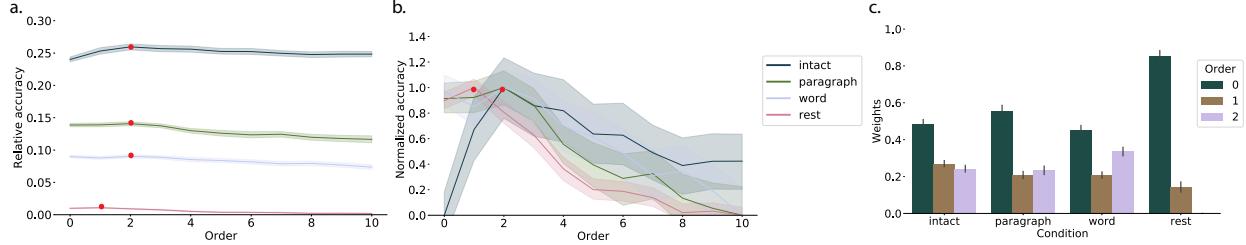


Figure 4: Decoding by order. **a. Relative decoding accuracy by order.** Ribbons of each color display cross-validated decoding performance for each condition (intact, paragraph, word, and rest). Decoders were trained using increasingly more higher-order information and this ribbons are displayed relative to chance at 0. The red dots indicates maximum decoding accuracy for each condition. **b. Normalized decoding accuracy by order.** We normalized the decoding accuracy by order to better visualize the order with the maximum decoding accuracy for each condition. **c. Optimized weights.** Bar heights indicate the optimized mixing paramete ϕ of each contributing order up to and including the order with the maximum decoding accuracy for each contributing order For the order with maximum decoding accuracy by condition, we show barplots of the optimized weights ϕ for each contributing order.

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

198 Separately for each experimental condition, we divided participants into two groups. For the zeroth
 199 order, we computed the mean factor activity for each group. For all subsequent orders up to the tenth
 200 order, we computed the mean approximated dynamic ISFC of factor activity for each group (see *Materials*
 201 and *methods*), and combined in a weighted mixutre with all previous orders (i.e. cross-validation for the
 202 second order contained a weighted-mixture of zeroth, first, and second order c. **Optimized weights.**) For
 203 each order, we correlated the group 1 activity patterns with group 2 activity patterns. We then subdivided
 204 the group 1 to obtain an optimal weighting parameter for each order's correlation matrix using the same
 205 cross validation method. We used the optimal weighting parameters to obtain a weighted-mixture (see
 206 *Materials and methods*) of each order's correlation matrix. Using these correlations, we labeled the group 1
 207 timepoints using the group 2 timepoints with which they were most highly correlated; we then computed
 208 the proportion of correctly labeled group 1 timepoints. (We also performed the symmetric analysis whereby
 209 we labeled the group 2 timepoints using the group 1 timepoints as a template.) We repeated this procedure
 210 100 times (randomly re-assigning participants to the two groups each time) to obtain a distribution of
 211 decoding accuracies for each experimental condition. (There were 272 timepoints for paragraph condition,
 212 300 timepoints for intact and word conditions, and 400 timepoints for rest condition, so chance performance

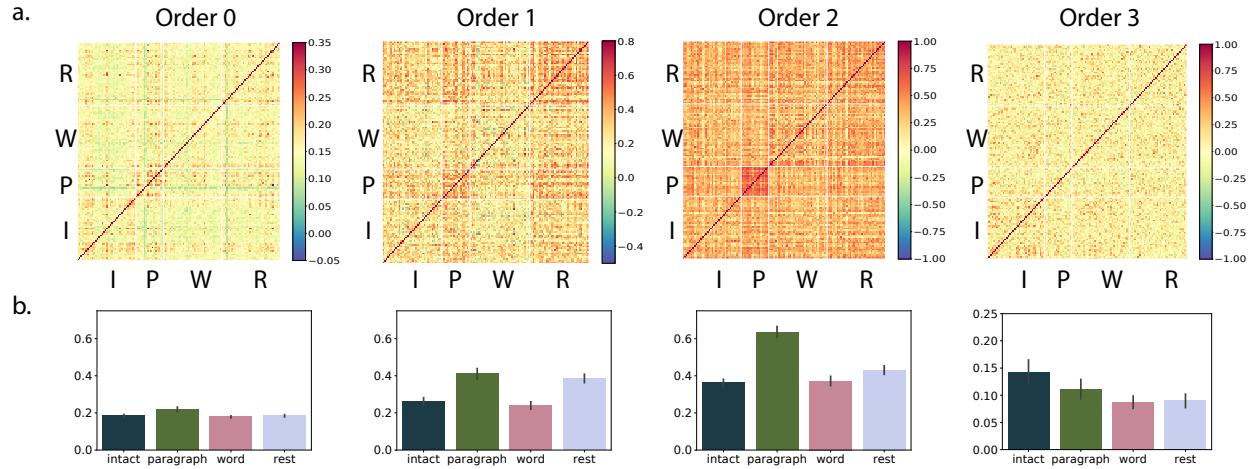


Figure 5: Inter-subject correlations by order. **a. Inter-subject correlation heatmaps by order.** Heatmaps for each order of how correlated each subject is with every other subject grouped by condition (intact (I), paragraph (P), word (W), and rest (R)). **b. Mean correlation for each condition by order.** Bar heights indicate inter-subject correlations by condition for each order.

213 on this decoding test is was $\frac{1}{272}$, $\frac{1}{300}$, and $\frac{1}{400}$ respectively.

214 Discussion

- 215 Decoding accuracy best for level 2 data for all but rest condition. Could be that the brain is 2nd order
216 or that fMRI can only reliable give 2nd order.
- 217 multiple timescale representations (a la Hasson group) implies first-order network interactions.
218 higher-order interactions imply generalizations between interacting representations (e.g. mirrored
219 schema, a la Norman/Baldassano/Hasson). possibly cite NTB 2013 science review, using as evidence
220 that this is where the field is going (voxels → patterns (L0) → interactions (L1) → higher order patterns
221 (L2+)).
- 222 related approaches: sliding window, phase-based correlations, within-ROI spatial correlations at each
223 timepoint, granger causality, other explit models (e.g. virtual brain).
- 224 other applications: molecular interactions (protein folding?), diagnosis (e.g. psychiatric disorders
225 as network flow problems- grattan work?), social network dynamics (e.g. financial markets, social
226 media interactions)

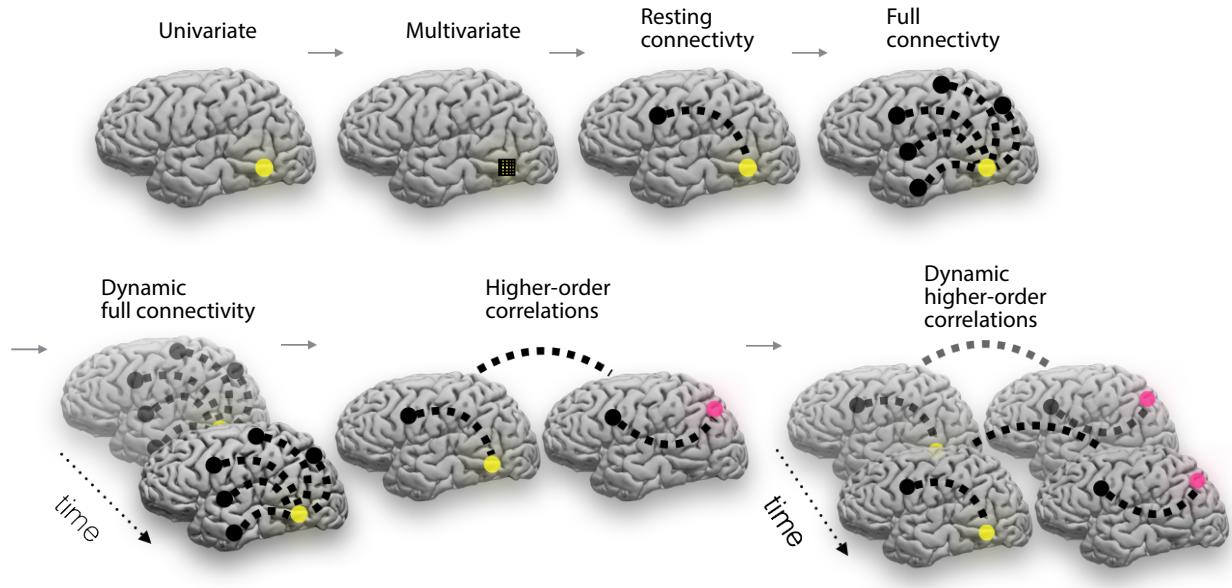


Figure 6: **Direction of the field.**

227 Concluding remarks

228 the universe is complicated and we need scalable approaches to studying how the pieces are interacting to
229 make sense of it. one small step for mankind, and so on.

230 Acknowledgements

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

237 Author contributions

238 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.

239 **References**

- 240 Alvarez-Hamelin, I., Dall'Asta, L., Barrat, A., & Vespignani, A. (2005). *k*-corr decomposition: a tool for the
241 visualization of large scale networks. *arXiv*, cs/0504107v2.
- 242 Barthélemy, M. (2004). Betweenness centrality in large complex networks. *European Physical Journal B*, 38,
243 163–168.
- 244 Bonacich, P. (2007). Some unique properties of eigenvector centrality. *Social Networks*, 29(4), 555–564.
- 245 Christakis, N. A., & Fowler, J. H. (2010). Social network sensors for early detection of contagious outbreaks.
246 *PLoS One*, 5(9), e12948.
- 247 Comon, P., Jutten, C., & Herault, J. (1991). Blind separation of sources, part II: Problems statement. *Signal
248 Processing*, 24(1), 11 - 20.
- 249 Estrada, E., & Rodríguez-Velázquez, J. A. (2005). Subgraph centrality in complex networks. *Physical Review
250 E*, 71(5), 056103.
- 251 Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry*, 40(1), 35–41.
- 252 Geisberger, R., Sanders, P., & Schultes, D. (2008). Better approximation of betweenness centrality. *Proceedings
253 of the meeting on Algorithm Engineering and Experiments*, 90–100.
- 254 Gershman, S., Blei, D., Pereira, F., & Norman, K. (2011). A topographic latent source model for fMRI data.
255 *NeuroImage*, 57, 89–100.
- 256 Halu, A., Mondragón, R. J., Panzarasa, P., & Bianconi, G. (2013). Multiplex PageRank. *PLoS One*, 8(10),
257 e78293.
- 258 Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks.
259 *Science*, 313(5786), 504–507.
- 260 Honey, C. J., Kötter, R., Breakspear, M., & Sporns, O. (2007). Network structure of cerebral cortex shapes
261 functional connectivity on multiple time scales. *Proceedings of the National Academy of Science USA*, 104(24),
262 10240–10245.
- 263 Jutten, C., & Herault, J. (1991). Blind separation of sources, part I: An adaptive algorithm based on
264 neuromimetic architecture. *Signal Processing*, 24(1), 1–10.
- 265 Lee, D. D., & Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature*,
266 401, 788–791.

- 267 Lin, J. (2009). Divergence measures based on the Shannon entropy. *IEEE Transactions on Information Theory*,
268 37(1), 145–151.
- 269 Lohmann, G., Margulies, D. S., Horstmann, A., Pleger, B., Lepsien, J., Goldhahn, D., ... Turner, R. (2010).
270 Eigenvector centrality mapping for analyzing connectivity patterns in fMRI data of the human brain.
271 *PLoS One*, 5(4), e10232.
- 272 Mairal, J., Ponce, J., Sapiro, G., Zisserman, A., & Bach, F. R. (2009). Supervised dictionary learning. *Advances
273 in Neural Information Processing Systems*, 1033–1040.
- 274 Mairal, J. B., Bach, F., Ponce, J., & Sapiro, G. (2009). Online dictionary learning for sparse coding. *Proceedings
275 of the 26th annual international conference on machine learning*, 689–696.
- 276 Manning, J. R., Ranganath, R., Norman, K. A., & Blei, D. M. (2014). Topographic factor analysis: a Bayesian
277 model for inferring brain networks from neural data. *PLoS One*, 9(5), e94914.
- 278 Manning, J. R., Zhu, X., Willke, T. L., Ranganath, R., Stachenfeld, K., Hasson, U., ... Norman, K. A. (2018).
279 A probabilistic approach to discovering dynamic full-brain functional connectivity patterns. *NeuroImage*,
280 180, 243–252.
- 281 McInnes, L., & Healy, J. (2018). UMAP: Uniform manifold approximation and projection for dimension
282 reduction. *arXiv*, 1802(03426).
- 283 Newman, M. E. J. (2005). A measure of betweenness centrality based on random walks. *Social Networks*, 27,
284 39–54.
- 285 Newman, M. E. J. (2008). The mathematics of networks. *The New Palgrave Encyclopedia of Economics*, 2, 1–12.
- 286 Opsahl, T., Agneessens, F., & Skvoretz, J. (2010). Node centrality in weighted networks: generalizing degree
287 and shortest paths. *Social Networks*, 32, 245–251.
- 288 Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. *The London, Edinburgh,
289 and Dublin Philosophical Magazine and Journal of Science*, 2, 559–572.
- 290 Rao, C. R. (1982). Diversity and dissimilarity coefficients: a unified approach. *Theoretical Population Biology*,
291 21(1), 24–43.
- 292 Ricotta, C., & Szeidl, L. (2006). Towards a unifying approach to diversity measures: Bridging the gap
293 between the Shannon entropy and Rao's quadratic index. *Theoretical Population Biology*, 70(3), 237–243.

- 294 Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpreta-
295 tions. *NeuroImage*, 52, 1059–1069.
- 296 Schreiber, T. (2000). Measuring information transfer. *Physical Review Letters*, 85(2), 461–464.
- 297 Simony, E., Honey, C. J., Chen, J., & Hasson, U. (2016). Uncovering stimulus-locked network dynamics
298 during narrative comprehension. *Nature Communications*, 7(12141), 1–13.
- 299 Spearman, C. (1904). General intelligence, objectively determined and measured. *American Journal of
300 Psychology*, 15, 201–292.
- 301 Tipping, M. E., & Bishop, C. M. (1999). Probabilistic principal component analysis. *Journal of Royal Statistical
302 Society, Series B*, 61(3), 611–622.
- 303 van der Maaten, L. J. P., & Hinton, G. E. (2008). Visualizing high-dimensional data using t-SNE. *Journal of
304 Machine Learning Research*, 9, 2579–2605.
- 305 Zar, J. H. (2010). *Biostatistical analysis*. Prentice-Hall/Pearson.