High-order dynamic neural correlations reflect naturalistic processing in humans

Lucy L. W. Owen[1], Thomas H. Chang[1,2], and Jeremy R. Manning[1,†]

[1]Department of Psychological and Brain Sciences,
Dartmouth College, Hanover, NH
[3]Amazon.com, Seattle, WA
[†]Address correspondence to jeremy.r.manning@dartmouth.edu

May 10, 2019

Abstract

Our thoughts arise from coordinated activity patterns across our brain. We examined high-order dynamic correlations in functional neuroimaging data collect 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.

Introduction

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

These three properties (rich correlations, hierarchical organization, and dynamics) are major hallmarks of many complex systems. For example, within a single cell, the cellular components interact at many spatiotemporal scales, and those interactions change according to what that single cell is doing. Within a single human brain, the individual neurons interact within each brain structure, and the structures interact to form complex networks. The interactions at each scale vary according to the functions our brains are carrying out. And within social groups, interactions at different scales (e.g., between individuals, family units, communities, etc.) vary over time according to changing goals and external constraints.

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

Methods

There are two basic steps to our approach (Fig. 2). In the first step, we take a number-of-timepoints (T) by number-of-features (F) *matrix of observations* ($\mathbf{X}$) and we return a T by $\frac{F^2-F}{2}$ *matrix of dynamic correlations* ($\mathbf{Y}$). Here $\mathbf{Y_0}$ describes, at each moment, how all of the features (columns of $\mathbf{X}$) are inferred to be interacting. (Since the interactions are assumed to be non-recurrent and symmetric, only the upper triangle of the full correlation matrix is computed.) In the second step, we project $\mathbf{Y_0}$ onto an F-dimensional space, resulting in 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 represented in a compressed number of dimensions. By repeatedly applying these two steps in sequence, we can examine and explore higher order dynamic correlations in $\mathbf{X}$.

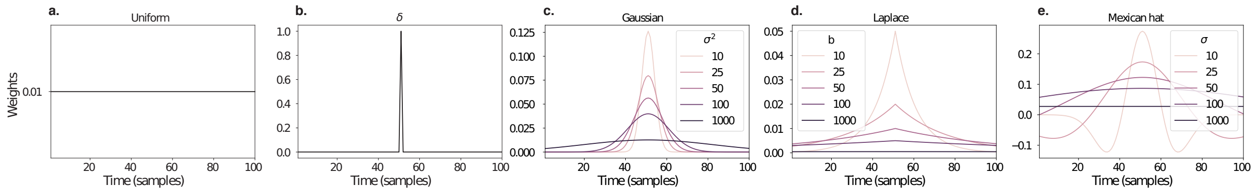

Figure 1: **Examples of time-varying weights. Each panel displays per-timepoint weights at** $t = 50$**, evaluated for 100 timepoints (**$1, ..., 100$**). a. Uniform weights.** The weights are timepoint-invariant; observations at all timepoints are weighted equally, and do not change as a function of 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.

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:

$$\mathrm{corr}(\mathbf{X}_i, \mathbf{X}_j) = \frac{\sum_{t=1}^{T} \left(\mathbf{X}_i(t) - \bar{\mathbf{X}}_i\right)\left(\mathbf{X}_j(t) - \bar{\mathbf{X}}_j\right)}{\sqrt{\sum_{t=1}^{T} \sigma^2_{\mathbf{X}_i} \sigma^2_{\mathbf{X}_j}}}, \text{ where} \tag{1}$$

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

$$\sigma^2_{\mathbf{X}_k} = \sum_{t=1}^{T} \left(\mathbf{X}_k - \bar{\mathbf{X}}_k\right)^2 \tag{3}$$

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

Given a weight function $w(t)$ for timepoint t, evaluated at timepoints in the interval $[1, ..., T]$, we can extend the static correlation formula in Equation 2 to reflect an *instantaneous correlation* at timepoint t:

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

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

$$\widetilde{\sigma}^2_{\mathbf{X}_k}(t) = \sum_{i=1}^{T} \left(\mathbf{X}_k(i) - \widetilde{\mathbf{X}}_k(t)\right)^2, \tag{6}$$

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., $\mathbf{Y}$).

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 $\mathbf{X}_1$, $\mathbf{X}_2$, ..., $\mathbf{X}_P$ 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:

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

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

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

R is the inverse of Z:

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

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, $\bar{\mathbf{X}}_{\setminus i}$:

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

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).

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 $\mathcal{O}(F^2)$ 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 $\mathcal{O}(F^{2n})$ 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 $\mathcal{O}(F)$-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 $\mathcal{O}(n)$ 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.

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,

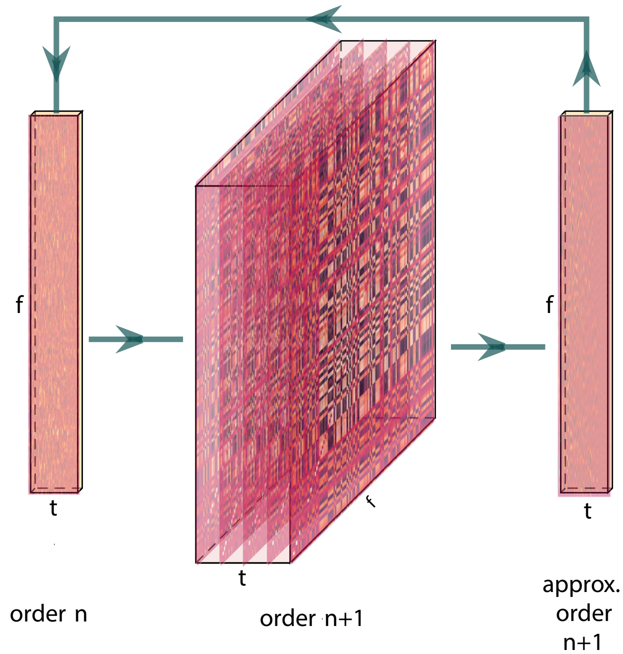

Figure 2: **Computing higher order correlations.** Dynamic correlations are computed then approximated to the same size as original data using dimensionality reduction. This process is repeated, and can be used to compute up to any arbitrary order with compuations scaling linearly (as opposed to exponentially)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

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

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

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

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

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 tranform 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.

Evaluation metrics

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

Generating synthetic data

To explore recovery of a constant covariance (Fig. 3, A.), we generated syntheic data sampled from a constant covariance matrix. To do this, we created one random covariance matrix, K, with 50 features, and for each of the 300 timepoints we sampled from a Gaussian distribution centered on K. Similarly, we generated synthetic data sampled from a random covariance matrix (Fig. 3, B.) by creating a new random covariance 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,

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

$$\tag{12}$$

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

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

Recovery of ground truth parameters from synthetic data

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

Timepoint decoding

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

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

To assess decoding accuracy, we randomly divided participants for each stimulus into training and testing 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. To assess how additional higher-order correlations contribute to decoding accuracy, for each order we included a weighted-mixutre (described below) of the activity patterns of all previous orders. For each group of participants in turn, we compared these activity patterns (using Pearson correlations) to estimate the story times each pattern corresponded to. Specifically, we asked, for each timepoint: what are the correlations between the first group's and second group's activity patterns at each order. We note that the decoding test we used is a conservative in which we count a timepoint label as incorrect if it is not an exact match.

For each order we obtained the weighted-mixture of the correlation matrices for the current order and all previous orders using mixing parameter , ϕ, where $0 < \phi < 1$ reflects a weighted mixture of order based decoding Fig. 4 Panel C.). We calculated ϕ, by subdividing the training group and using the quasi-Newton method of Broyden, Fletcher, Goldfarb, and Shanno (BFGS) for optimization. We repeated this cross-validation process 100 times.

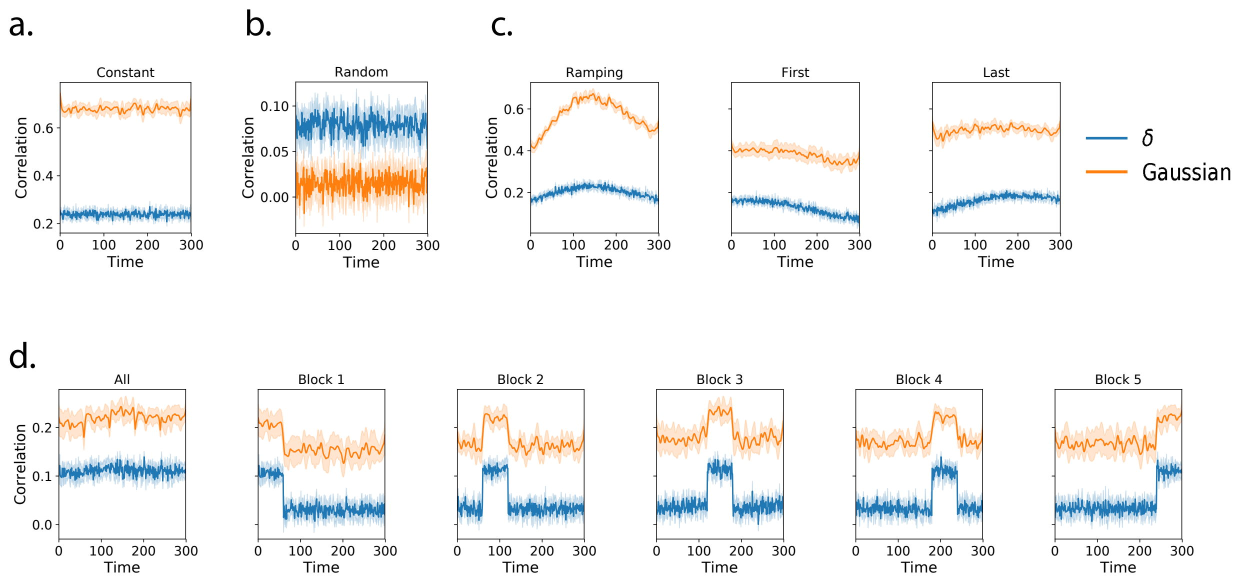

Figure 3: **Dynamic correlation recovery with synthetic data.** Using synthetic data containing different underlying correlational structure, we test how well we can recover dynamic correlation matrices using two kernels (delta and gaussian, width = 10) when compared to groud truth. We plot recovery using of datasets containing the following underlying structure: **a. Constant. b. Random. c. Ramping. d. Block.**

Results

Synthetic data

To assess the performance of dynamic correlation recovery using timecorr, we varied width the kernel and the specific structure of the data. We applied timecorr, using delta and gaussian kernels Fig. 1) to each of the following synthetic datasets: constant, random, ramping, and block. We then correlated each recovered correlation matrix with the ground truth.

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

Similary for the block sythetic dataset, we compared the timecorr recovered correlation matrices to the ground truth as well as to each block-specific covariance matrix (Block 1-5, Fig. 3, D.). Although the

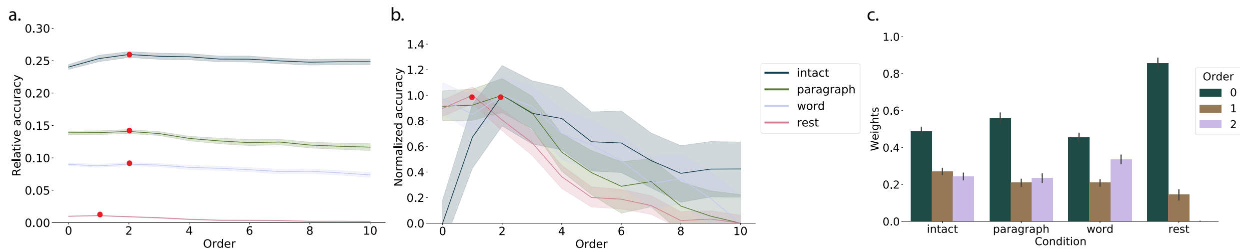

Figure 4: **Decoding by order. a. Relative decoding accuracy by order.** Ribbons of each color display cross-validated decoding performance for each condition (intact, paragraph, word, and rest. Decoders were trained using increasingly more higher-order information and this ribbons are displayed relative 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.

structure is changing by block, the gaussian kernel once again outperforms the delta kernel. Performance does however drop near even boundaries for when using the gaussian kernel.

Neuroimaging dataset (Simony et al., 2016)

For our decoding analysis, we used HTFA-derived node activities Manning et al. (2018) from fMRI data collected as participants listened to an audio recording of a story (intact condition; 36 participants), listened to time scrambled recordings of the same story (17 participants in the paragraph-scrambled condition listened to the paragraphs in a randomized order and 36 in the word-scrambled condition listened to the words in a randomized order), or lay resting with their eyes open in the scanner (rest condition; 36 participants). We sought to demonstrate how higher-order correlations may be used to examine dynamic interactions of brain patterns in (real) multi-subject fMRI datasets. This story listening dataset was collected as part of a separate study, where the full imaging parameters, image preprocessing methods, and 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 decoders trained using different sets of neural features: whole-brain patterns of voxel activities

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

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 mixutre 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.). 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 found that, during each of the listening conditions in the experiment, classifiers that incorporated second-order correlations yielded consistently higher accuracy than classifiers trained only on lower-order patterns (Fig. 4, A.), and incorporating higher-order correlations did not further improve decoding accuracy. By contrast, these higher-order correlations were not necessary to support (minimally above chance) decoding during a control rest condition. This suggests that the cognitive processing that supported the listening conditions involved second-order (or higher) network dynamics.

Discussion

Although dynamic interactions between brain structures underlie our thoughts, the ways in which we can study these patterns are limited. One challenge in studying dynamic interactions is the computational resources required to calculate higher-order correlations. We developed a computationally tractable model of network dynamics (Fig. 2) that takes in a feature timeseries and outputs approximated first-order dynamics (i.e., dynamic functional correlations), second-order dynamics (reflecting homologous networks that dynamically form and disperse), and higher-order network dynamics (up to tenth-order dynamic correlations).

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

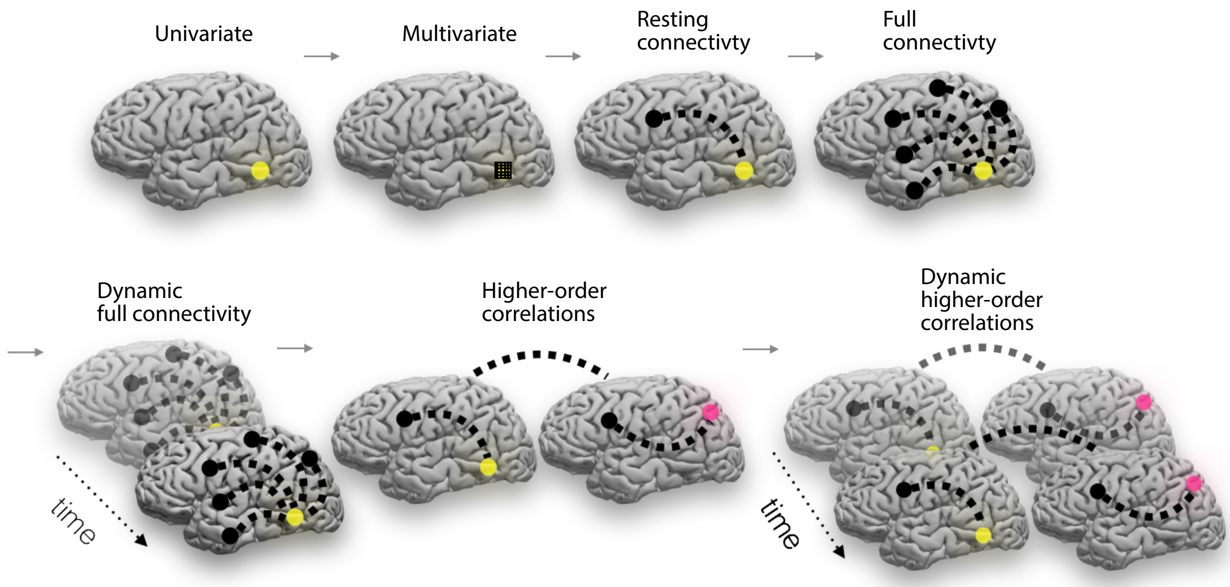

Figure 5: **Direction of the field.**

underlying data structures and kernels. We then applied the approach to an fMRI dataset (Simony et al., 2016) in which participants listened to an audio recording of a story, as well as scrambled versions of the same story (where the scrambling was applied at different temporal scales). We trained classifiers to take the output of the model and decode the timepoint in the story (or scrambled story) that the participants were listening to. We found that, during each of the listening conditions in the experiment, classifiers that incorporated up to second-order correlations yielded consistently higher accuracy than classifiers trained only on lower-order patterns (Fig. 4, A.), and incorporating more higher-order correlations did not further improve decoding accuracy. By contrast, these higher-order correlations were not necessary to support (minimally above chance) decoding during a control rest condition. This suggests that the cognitive processing that supported the listening conditions involved second-order (or higher) network dynamics.

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

Concluding remarks

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

Acknowledgements

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

Author contributions

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.

References

Alvarez-Hamelin, I., Dall'Asta, L., Barrat, A., & Vespignani, A. (2005). *k*-corr decomposition: a tool for the visualiztion of large scale networks. *arXiv*, cs/0504107v2.

Barthélemy, M. (2004). Betweenness centrality in large complex networks. *European Physical Journal B*, *38*, 163–168.

Bonacich, P. (2007). Some unique properties of eigenvector centrality. *Social Networks*, *29*(4), 555–564.

Christakis, N. A., & Fowler, J. H. (2010). Social network sensors for early detection of contagious outbreaks. *PLoS One*, *5*(9), e12948.

Comon, P., Jutten, C., & Herault, J. (1991). Blind separation of sources, part II: Problems statement. *Signal Processing*, *24*(1), 11 - 20.

Estrada, E., & Rodríguez-Velázquez, J. A. (2005). Subraph centrality in complex networks. *Physical Review E*, *71*(5), 056103.

Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry*, *40*(1), 35–41.

Geisberger, R., Sanders, P., & Schultes, D. (2008). Better approximation of betweenness centrality. *Proceedings of the meeting on Algorithm Engineering and Experiments*, 90–100.

Gershman, S., Blei, D., Pereira, F., & Norman, K. (2011). A topographic latent source model for fMRI data. *NeuroImage*, *57*, 89–100.

Halu, A., Mondragón, R. J., Panzarasa, P., & Bianconi, G. (2013). Multiplex PageRank. *PLoS One*, *8*(10), e78293.

Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science*, *313*(5786), 504–507.

Honey, C. J., Kötter, R., Breakspear, M., & Sporns, O. (2007). Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *Proceedings of the National Academy of Science USA*, *104*(24), 10240–10245.

Jutten, C., & Herault, J. (1991). Blind separation of sources, part I: An adaptive algorithm based on neuromimetic architecture. *Signal Processing*, *24*(1), 1–10.

Lee, D. D., & Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature*, *401*, 788–791.

Lin, J. (2009). Divergence measures based on the Shannon entropy. *IEEE Transactions on Information Theory*, *37*(1), 145–151.

Lohmann, G., Margulies, D. S., Horstmann, A., Pleger, B., Lepsien, J., Goldhahn, D., . . . Turner, R. (2010). Eigenvector centrality mapping for analyzing connectivity patterns in fMRI data of the human brain. *PLoS One*, *5*(4), e10232.

Mairal, J., Ponce, J., Sapiro, G., Zisserman, A., & Bach, F. R. (2009). Supervised dictionary learning. *Advances in Neural Information Processing Systems*, 1033–1040.

Mairal, J. B., Bach, F., Ponce, J., & Sapiro, G. (2009). Online dictionary learning for sparse coding. *Proceedings of the 26th annual international conference on machine learning*, 689–696.

Manning, J. R., Ranganath, R., Norman, K. A., & Blei, D. M. (2014). Topographic factor analysis: a Bayesian model for inferring brain networks from neural data. *PLoS One*, *9*(5), e94914.

Manning, J. R., Zhu, X., Willke, T. L., Ranganath, R., Stachenfeld, K., Hasson, U., . . . Norman, K. A. (2018). A probabilistic approach to discovering dynamic full-brain functional connectivity patterns. *NeuroImage*, *180*, 243–252.

McInnes, L., & Healy, J. (2018). UMAP: Uniform manifold approximation and projection for dimension reduction. *arXiv*, *1802*(03426).

Newman, M. E. J. (2005). A measure of betweenness centrality based on random walks. *Social Networks*, *27*, 39–54.

Newman, M. E. J. (2008). The mathematics of networks. *The New Palgrave Encyclopedia of Economics*, *2*, 1–12.

Opsahl, T., Agneessens, F., & Skvoretz, J. (2010). Node centrality in weighted networks: generalizing degree and shortest paths. *Social Networks*, *32*, 245–251.

Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *2*, 559-572.

Rao, C. R. (1982). Diversity and dissimilarity coeficients: a unified approach. *Theoretical Population Biology*, *21*(1), 24–43.

Ricotta, C., & Szeidl, L. (2006). Towards a unifying approach to diversity measures: Bridging the gap between the Shannon entropy and Rao's quadratic index. *Theoretical Population Biology*, *70*(3), 237–243.

Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. *NeuroImage*, *52*, 1059–1069.

Schreiber, T. (2000). Measuring information transfer. *Physical Review Letters*, *85*(2), 461–464.

Simony, E., Honey, C. J., Chen, J., & Hasson, U. (2016). Uncovering stimulus-locked network dynamics during narrative comprehension. *Nature Communications*, *7*(12141), 1–13.

Spearman, C. (1904). General intelligence, objectively determined and measured. *Americal Journal of Psychology*, *15*, 201–292.

Tipping, M. E., & Bishop, C. M. (1999). Probabilistic principal component analysis. *Journal of Royal Statistical Society, Series B*, *61*(3), 611–622.

van der Maaten, L. J. P., & Hinton, G. E. (2008). Visualizing high-dimensional data using t-SNE. *Journal of Machine Learning Research*, *9*, 2579-2605.

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