The high-order brain dynamics (HOBD) model

Jeremy R. Manning and Qiang Liu
Dartmouth College

April 8, 2016

1 Introduction

When we record from brains, the resulting data can appear incredibly complex. But brain data are far from random: rather, the dynamic patterns of activity our brains exhibit are highly structured. Presumably this mirrors the complex but highly structured nature of our internal thoughts and experiences. For example, neuroscientists have begun to use functional magnetic resonance imaging (fMRI) to obtain high resolution 3D snapshots of brain activity approximately once per second during cognitive experiments. By treating each brain image as a feature vector, machine learning algorithms trained on a subset of the images may be used to decode cognitive information (e.g. which part of a movie someone is watching) from the held-out images (Norman et al., 2006). These approaches typically treat each brain image in isolation, and attempt to identify patterns of activity associated with each of several candidate brain states.

Recently, it has become clear that important cognitive information is contained in higher-order brain patterns, such as the correlational structure of the data. Functional connectivity analyses entail computing the correlations between the time series of activations each pair of brain regions exhibits. One can then examine how (or whether) the correlational structure of these activity patterns (across a given set of brain regions) varies according to the cognitive task an experimental participant is performing (Turk-Browne, 2013).

Despite a recent surge of interest in functional connectivity analyses of brain data, there are two fundamental limitations to analyzing brain data in this way:

1. **Correlation matrices are not scalable.** Examining pairwise correlations in brain data produces a correlation matrix with $O(n^2)$ entries (where n is the number of brain regions). When n is large (e.g. the number of voxels in an fMRI volume), the full correlation matrix can become unwieldy to compute with. Further, if one wishes to examine higher-order patterns (e.g. how correlations between correlations change over time), the storage requirements of the resulting patterns increase exponentially.

2. **Correlation matrices are not well suited to studying dynamic activity.** Computing correlations requires a time series. Therefore, studying how correlations change over time requires using sliding time windows and examining correlations within each window. In addition to adding another free parameter to the analysis (window duration), the sliding window approach provides only a poor approximation of the moment-by-moment patterns at the heart of these representations.

Here we propose the *high-order brain dynamics* (HOBD) model. We designed the HOBD model to address the above limitations of standard correlation-based analyses. First, HOBD provides an efficient (scalable) means of describing and computing with high-order patterns of brain activity. Within the model space, activity-based patterns requires the same amount of memory as correlation-based patterns, which require the same amount of memory as higher-order patterns. Second, HOBD provides moment-by-moment estimates of (lower-order and higher-order) brain patterns. For example, HOBD incorporates representations of interactions between brain regions (analogous to functional connectivity) at each moment during an experiment. In turn, this allows one to examine interactions between brain regions (or higher-order patterns such as interactions between interactions between brain regions) at each moment during an experiment (rather than relying on sliding window-based analyses).

2 Model description

HOBD is a hierarchical factor analysis model: each brain image in a dataset is represented in the model as a weighted sum of K *factors*. A formal description of the model may be found in Algorithm 1 (generative process) and Figure 1 (graphical model). Here we also provide a high-level description of the model along with intuitions for its various components.

Drawing inspiration from another factor analysis model of brain data, Topographic Factor Analysis (Manning et al., 2014), each factor in HOBD is constrained in shape according to a pre-specified spatial function. In the current formulation, each factor is a Gaussian Radial Basis Function (RBF) parameterized by a center parameter, c_k and a width parameter, w_k. In multi-person datasets, the center and width parameters of each factor are constrained within the model to be similar across people.

To obtain each factor image, that factor's RBF is evaluated at the location of each voxel. Given the factor images (which are held fixed for all time), each brain image is described as a weighted sum of the factor images (i.e., a vector of K weights). Since K is typically much less than the number of voxels in the images (V), the weight matrix provides a low-dimensional embedding of the original data that is efficient to compute with (Manning et al., 2014).

The "trick" in HOBD that allows for efficient representations of higher-order brain patterns comes from how the per-timepoint factor weights are defined. Specifically, each factor is associated with a set of N *feature vectors* (corresponding to activity patterns at N "levels"). These N feature vectors change over time, such that each of T timepoints are each associated with N feature vectors ($\beta^{1\ldots N}_{(1\ldots T)ks}$). At the lowest level ($n = 0$), the feature vectors are mapped directly to the factor weights $\omega_{(1\ldots T)ks}$ via a subject-specific mapping vector, π^0_s:

$$\omega_{tks} = \pi^0_s \left(\beta^0_{tks}\right)'. \tag{1}$$

As n increases, each level's feature vectors are determined using a similar mapping procedure applied to the next-highest level's feature vectors:

$$\beta^n_{tks} \sim \mathcal{N}\left(\phi^n_s \beta^{n+1}_{tks}, \kappa_{\beta_s}\right), \tag{2}$$

where ϕ^n_s is a subject-specific mapping matrix that is learned for each feature level.

Each level of feature vector corresponds to brain activity at a different order of analysis. For example, we may compute the degree to which different factors (i and j) relate to each other (within the feature space) at a given timepoint:

$$\gamma^n_{i,j}(t) = \eta\left(\beta^n_{tjs}\left(\beta^n_{tis}\right)'\right), \tag{3}$$

where the function $\eta(\cdot)$ maps its real-valued argument onto the interval $[-1, 1]$:

$$\eta(x) = \frac{2}{1 + e^{-mx}} - 1, \tag{4}$$

where $m \in \mathcal{R}^+$ is a constant (hyperparameter) that determines how steeply the off-zero values go towards -1 or 1.

When these mapping functions are evaluated at $n = 0$ (i.e., $\gamma^0_{i,j}$), this yields the analog of the functional connectivity (correlation) between i and j. Similarly, $\gamma^1_{p,q}$ is analogous to the correlation between the *correlations* p and q (which in turn each correspond to the correlations between some pair of factors at the lowest-level representation of the data). The products between feature vectors at higher values of n reflect the analog of correlations between the corresponding feature vectors at the next-highest level $(n - 1)$.

Algorithm 1: Generative process. Here F is the number of dimensions (features) in the feature vectors, and $\mathbf{I}^x$ is the x-dimensional identity matrix. The function $\mathtt{RBF}_s(c, w)$ returns a V-dimensional vector corresponding to an RBF evaluated at the location of each voxel in subject s's imaged brain volume. All other non-standard symbols are defined in the text or in the algorithm itself.

```
for k = 1 to K do
    Draw global center c̄_k ~ N(μ_c, κ_c);
    Draw global log width w̄_k ~ N(μ_w, σ_w^2);
    Draw global features for first timepoint β̄^N_{1k} ~ N(μ_β, κ_β);
end
for s = 1 to S do
    for k = 1 to K do
        Draw subject center c_{ks} ~ N(c̄_k, κ_c);
        Draw subject log width w_{ks} ~ N(w̄_k, σ_w^2);
        Draw subject features β^N_{1ks} ~ N(β̄^N_{1k}, κ_β);
        for n = (N − 1) to 0 do
            Draw subject mapping matrix φ^n_s ~ W(I^F, F);
            Draw subject features β^n_{1ks} ~ N(β^{n+1}_{1ks} φ^n_s, κ_β);
            for t = 2 to (T − 1) do
                Draw subject features β^n_{tks} ~ N(ρ(β^{n+1}_{tks} φ^n_s) + (1 − ρ)(β^n_{(t−1)ks}), κ_β);
            end
        end
        Draw subject mapping vector π^0_s ~ Dir(α_π);
        Draw subject noise parameter σ^2_{ys} ~ Gam(a_{σ_y}, b_{σ_y});
        for t = 1 to T do
            Draw brain image y_{ts} ~ N(Σ_k (π^0_s (β^0_{tks})' RBF_s(c_{ks}, exp(w_{ks}))), σ^2_{ys} I^V);
        end
    end
end
```

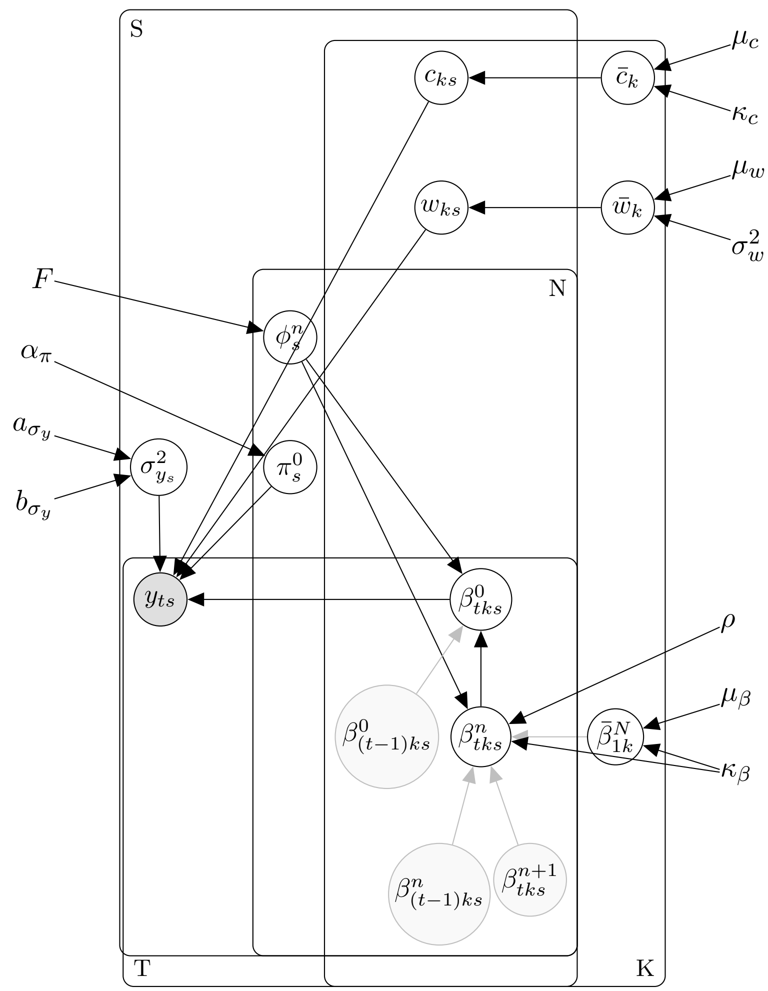

Figure 1: **Graphical model.** See Algorithm 1 for additional details. The gray latent variable and arrow denote (unobserved) higher order relations in the model (i.e. self references across levels within the same plate). Arrows to the gray latent variables are omitted, since they are redundant with arrows to the self-referenced node they point to.

References

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

Norman, K. A., Polyn, S. M., Detre, G. J., and Haxby, J. V. (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences*, 10(9):424–430.

Turk-Browne, N. B. (2013). Functional interactions as big data in the human brain. *Science*, 342:580–584.