When and Where does our Working Memory Take Place?

A multi-level sub-graph analysis of brain functional connectivities.

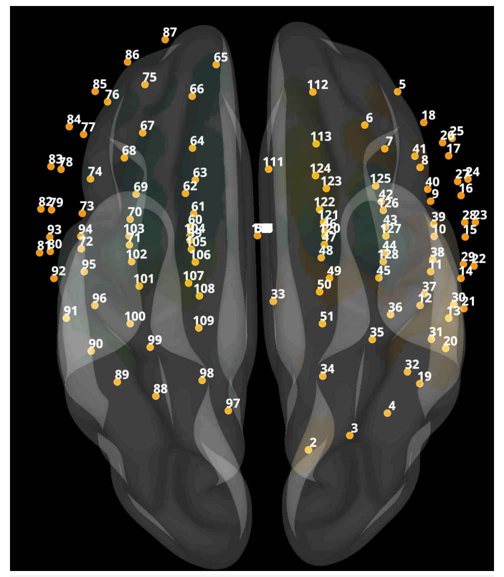
Maoran Xu, Leo Duan

Department of Statistics, University of Florida

EEG: Measuring our Working Memory

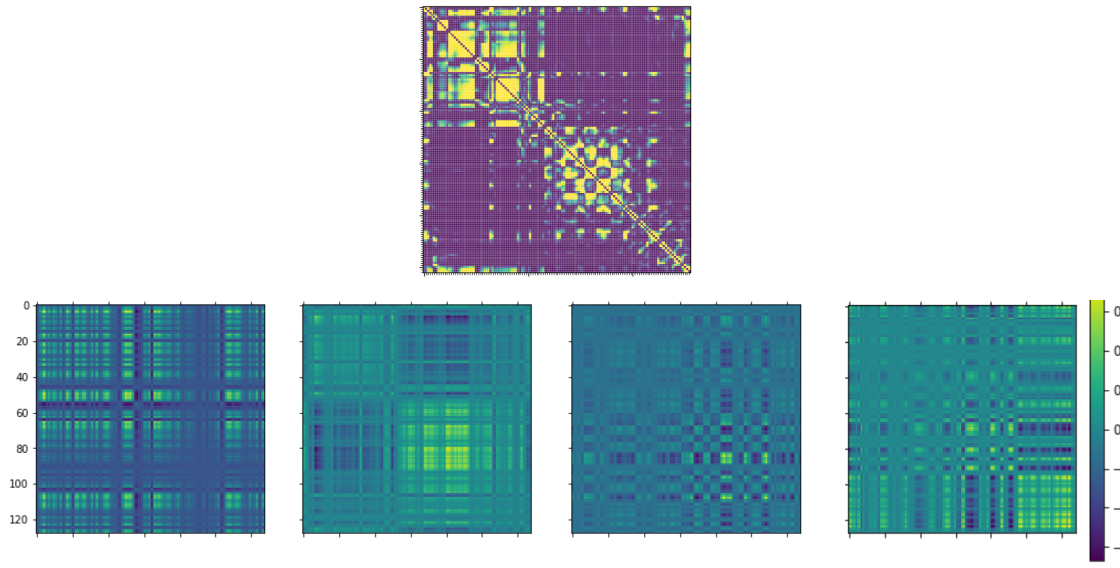
Dataset

- The data was collected in a study of human brain functional connectivities in working memory.
- 20 healthy subjects were asked to do the Sternberg verbal working memory task with three levels of memory load (load 2, 4 and 6). The subjects were wearing EEG head caps with 128 channels of electrodes.



Motivation

- We extract the correlation between each pair of the channels. We study the adjacency matrices of measurement over time period.
- The brain network is large. However, the brain connectivity displays coarse-grain structures that may relate to latent low-dimensional estimation.
- During different period of memory, and loads, our brains display different connectivity.



Treed-SVD: Partition the Brian Network

Framework:

- For a given adjacency matrix A, we want to find out the Treed-SVD decomposition as

$$A = P\Lambda P^{\intercal} + \mathcal{E},$$

where P is an orthonormal matrix and every column of P displays a level partition. We call such matrix space a Tree-Spanning Stiefel Space.

$$\mathcal{V}_{\mathcal{T}}^{d\times n} = \mathcal{T}_d \cap \mathcal{V}^{d\times n}.$$

- The structure is generated level-by-level:
  1. Compute the eigenvector of A corresponding to the largest eigenvalue, and partition by the sign as $\mathcal{T}_{l,s}$.
  2. Maximize the conditional likelihood

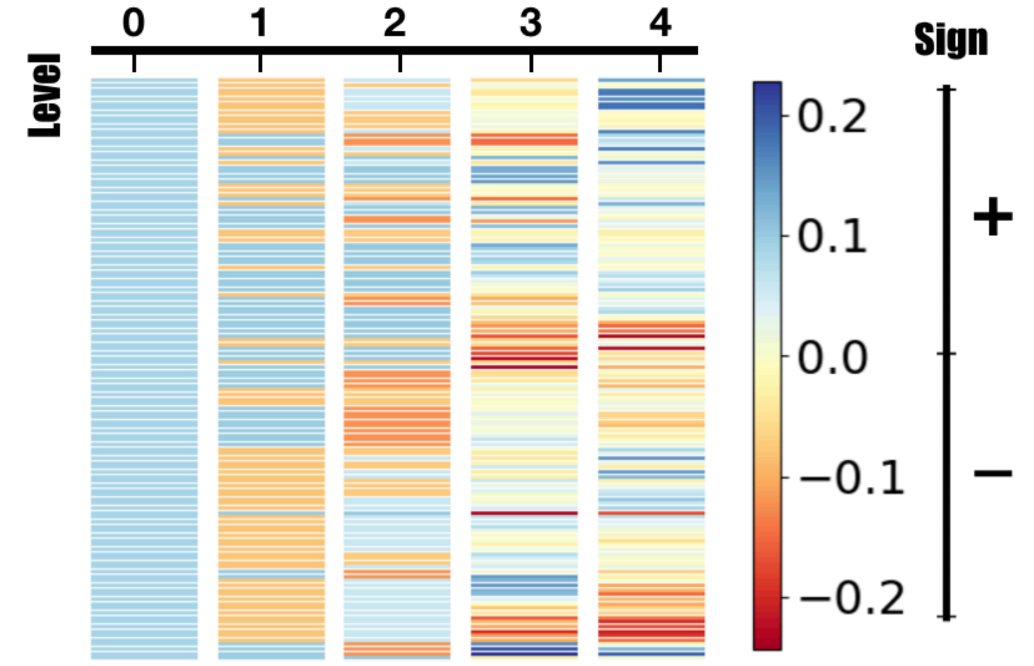
$$L(P_{k+1}, \lambda_{k+1} | \tilde{A}_{1:k}, A) = \exp\{-||(A - \tilde{A}_{1:k}) - \lambda_{k+1} P_{k+1} P_{k+1}^{\intercal}||_F\}.$$

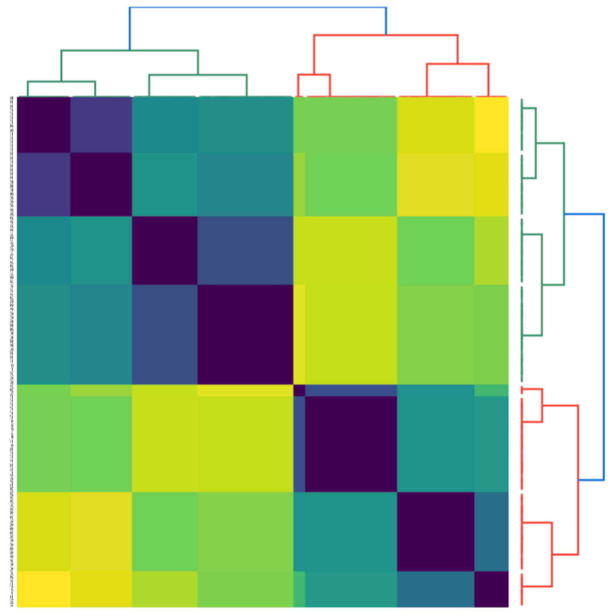
by solving the equivalent quadratic programming:

$$\max_{P_{k+1}} P_{k+1}^{\intercal}\left(A - \tilde{A}_{1:k}\right)P_{k+1}$$

$$\text{subject to } P \in \mathcal{V}_{\mathcal{T}}^{(k+1)\times n},$$

and set $\lambda_{k+1} = P_{k+1}^{T} A P_{k+1}$.

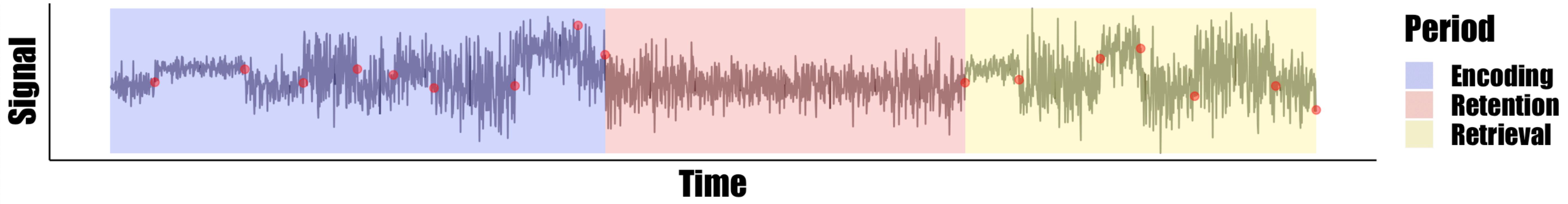




A Case Study: Change Point Detection

We model the change point detection task as a combination of Hidden Markov Model and clustering problem. Suppose there are K hidden states, where the center for the state is $\mu^{(k)}$. A_t belongs to state z_t. Denote the transition probability from k to l by $p_{k,l}$.
The likelihood is therefore

$$L(A_t, z_t, \mu^{(z_t)}) = \prod_{t=1}^{T} p_{z_{t-1}, z_t} \exp\{-\|A_t - \mu^{(z_t)}\|\},$$



The Hidden Markov Clustering

1. Conditioned on $\{z_t\}_t$, update $\{\mu^{(k)}\}_k$ using Treed-SVD algorithm.
2. Conditioned on $\{\mu^{(k)}\}_k$, update the latent assignment one-at-a-time, from the categorical distribution with

$$pr(z_t = k \mid z_{t-1}, z_{t+1}) = \frac{p_{z_{t-1},k} p_{k,z_{t+1}} \exp\{-\|A_t - \mu^{(k)}\|\}}{\sum_{l=1}^{K} p_{z_{t-1},l} p_{l,z_{t+1}} \exp\{-\|A_t - \mu^{(l)}\|\}}$$

3. Conditioned on $\{z_t\}_t$, update the transition probabilities by

$$(p_{k,1}, \dots, p_{k,K}) \sim \text{Dir}\{\alpha + \sum_{t=1}^{T} \mathbf{1}(z_{t-1} = k, z_t = 1), \dots, \alpha + \sum_{t=1}^{T} \mathbf{1}(z_{t-1} = k, z_t = K)\}$$

Reference:

[1] Durante, D., & Dunson, D. B. (2014). Nonparametric Bayes dynamic modelling of relational data. Biometrika, 101(4), 883-898.
[2] Fornito, A., Zalesky, A., & Breakspear, M. (2013). Graph analysis of the human connectome: promise, progress, and pitfalls. Neuroimage, 80, 426-444.
[3] Yu, Y., Wang, T., & Samworth, R. J. (2014). A useful variant of the Davis – Kahan theorem for statisticians. Biometrika, 102(2), 315-323.
[4] Fuh, C. D. (2004). Asymptotic operating characteristics of an optimal change point detection in hidden Markov models. The Annals of Statistics, 32(5), 2305-2339.

Download Poster

My Homepage