

Multi-scale Graph Principal Component Analysis of Brain Connectomes

David Dunson

Department of Statistical Science
Duke University

12/4/2020

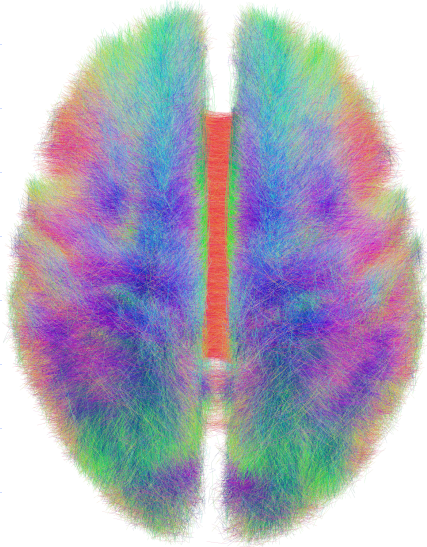
Duke
UNIVERSITY



Background: Connectomes/Connectivity

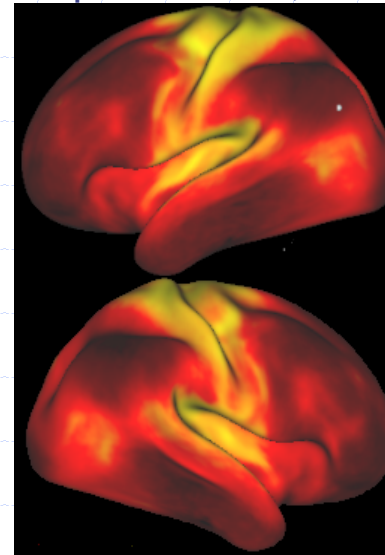
➤ Structural Connectivity

- A pattern of anatomical links, **dMRI**



➤ Functional Connectivity

- Statistical Dependencies, **fMRI, EEG, MEG**



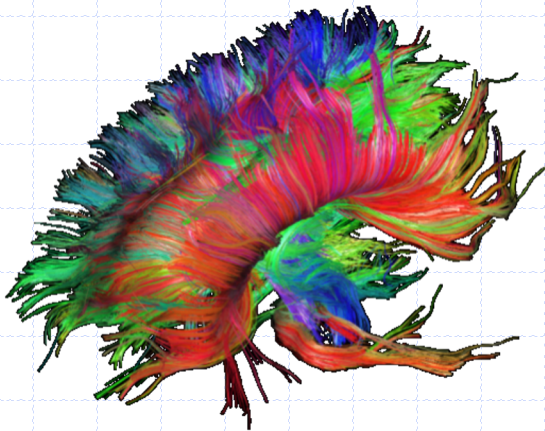
The Human Connectome Project

- HCP focuses on elucidating the neural pathways that underlie brain function and behavior.

The Heavily Connected Brain

Peter Stern, “**Connection, connection, connection...**”,

Science, Nov. 1 2013: Vol. 342 no. 6158 P.577



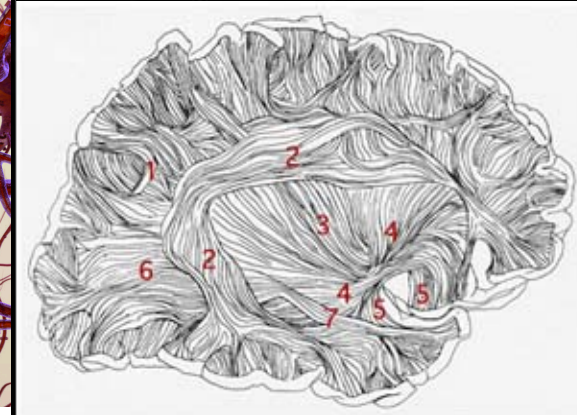
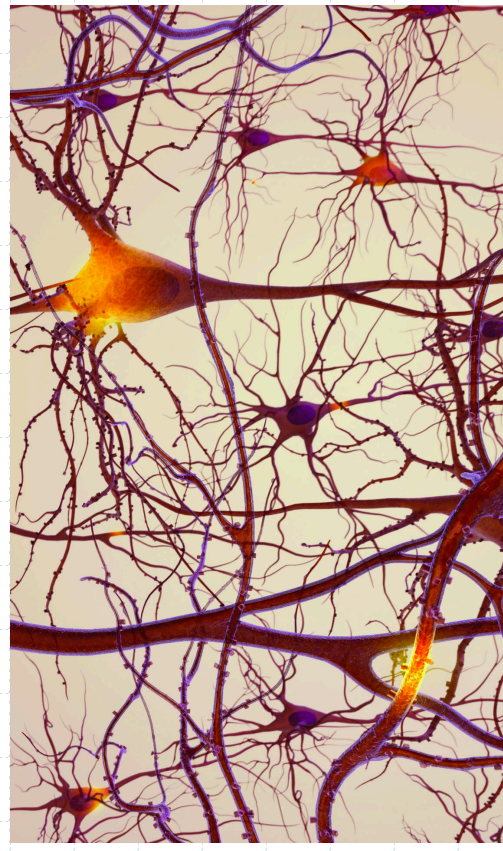
- High quality brain images: functional MRI (fMRI), diffusion MRI, structural MRI, Magnetoencephalography (MEG) and electroencephalography (EEG)
- Rich demographic and behavioral data: cognition, perception, and personality measurements.

- Diffusion MRI is routinely collected in many/most brain studies

- UK Biobank
- Adolescent Brain Cognitive Development Study (ABCD)
- ...

Diffusion Imaging Acquisition

- Axons have $\sim\mu\text{m}$ diameters
- Axons group together in bundles that traverse the white matter in brain
- We can not image individual axons, but we can **indirectly** image bundles with **diffusion MRI** technique

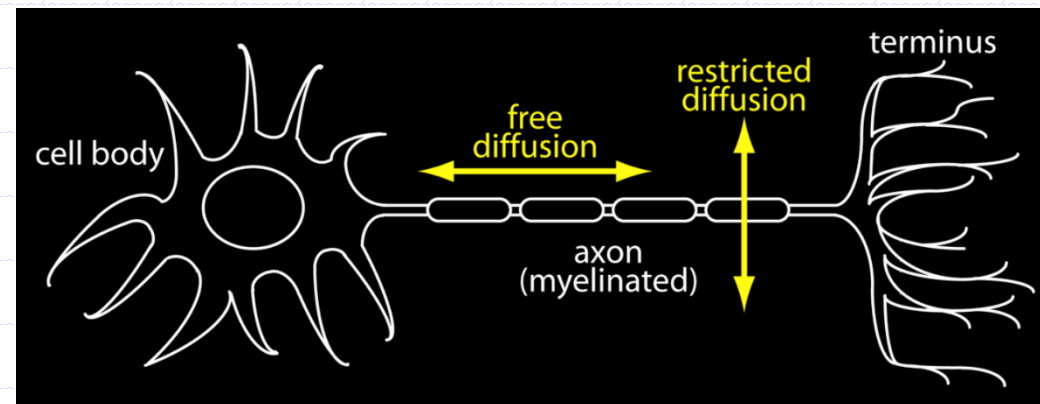
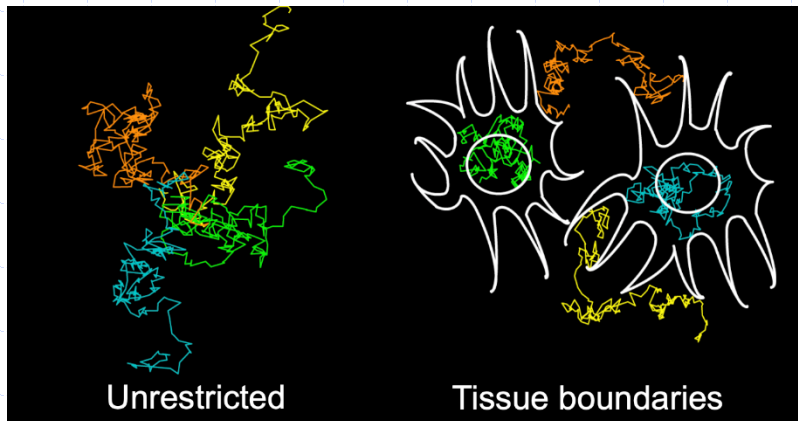
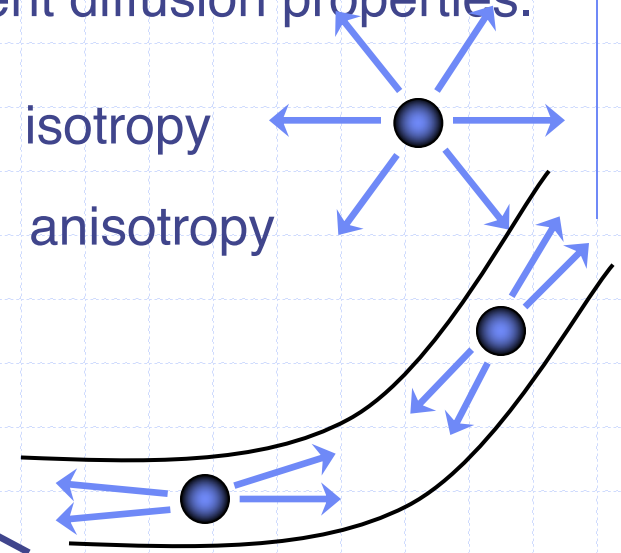
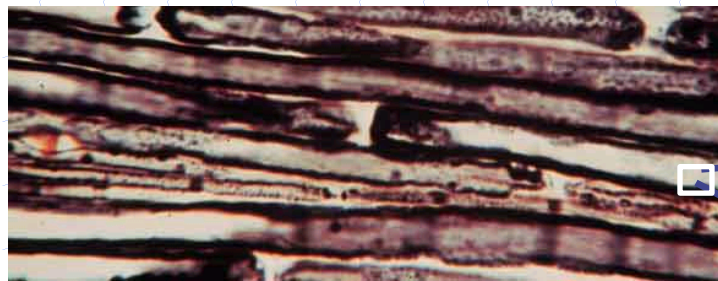


(From UMD website)

Diffusion in Brain Tissue

➤ Water molecules in different tissues have different diffusion properties.

- **Gray matter:** Diffusion is unrestricted ↔ isotropy
- **White matter:** Diffusion is restricted ↔ anisotropy



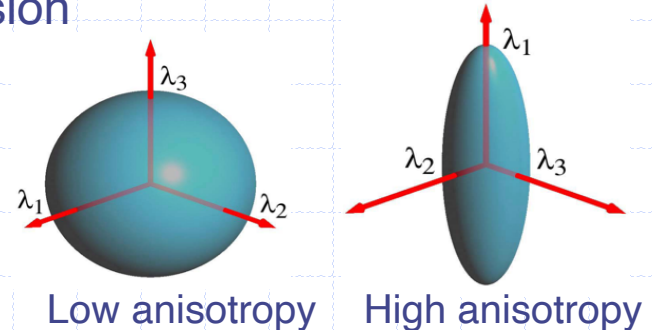
Reconstruction of Local WM Configuration

➤ At each voxel, we want to infer:

- The **orientation** and the **magnitude** of the diffusion

(1) Diffusion tensor image (DTI)

$$D = \begin{pmatrix} d_{1,1} & d_{2,1} & d_{3,1} \\ d_{2,1} & d_{2,2} & d_{3,2} \\ d_{3,1} & d_{3,2} & d_{3,3} \end{pmatrix}$$



(2) High angular resolution diffusion imaging (HARDI)

- Orientation distribution function (ODF) [Tuch et al. 04]
- Fiber ODF [Descoteaux et al. 09]
- ...

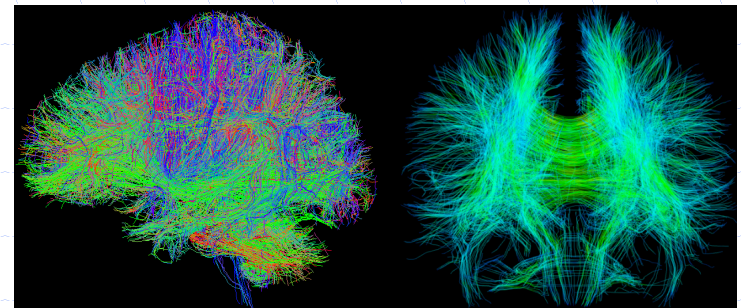


➤ **Fiber reconstruction** using stochastic differential equation:

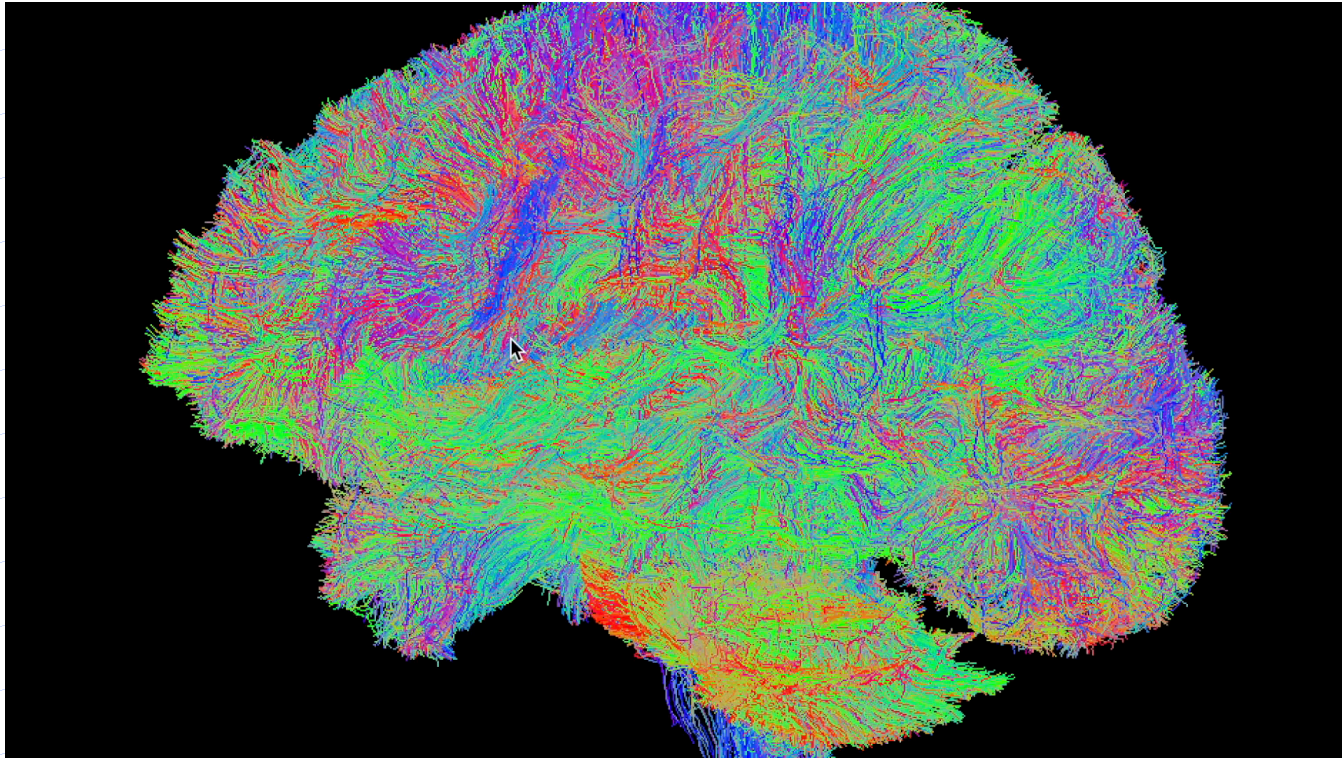
$$\frac{d\mathbf{v}(t)}{dt} = \mathbf{e}(\mathbf{v}), \quad t \geq 0 \quad \text{with} \quad \mathbf{v}(0) = \mathbf{v}_0,$$

$$\hat{\mathbf{e}}(\mathbf{v}) = \mathbf{e}(\mathbf{v}) + \epsilon(\mathbf{v}),$$

$\mathbf{v}(t)$ is the reconstructed fiber track ➡



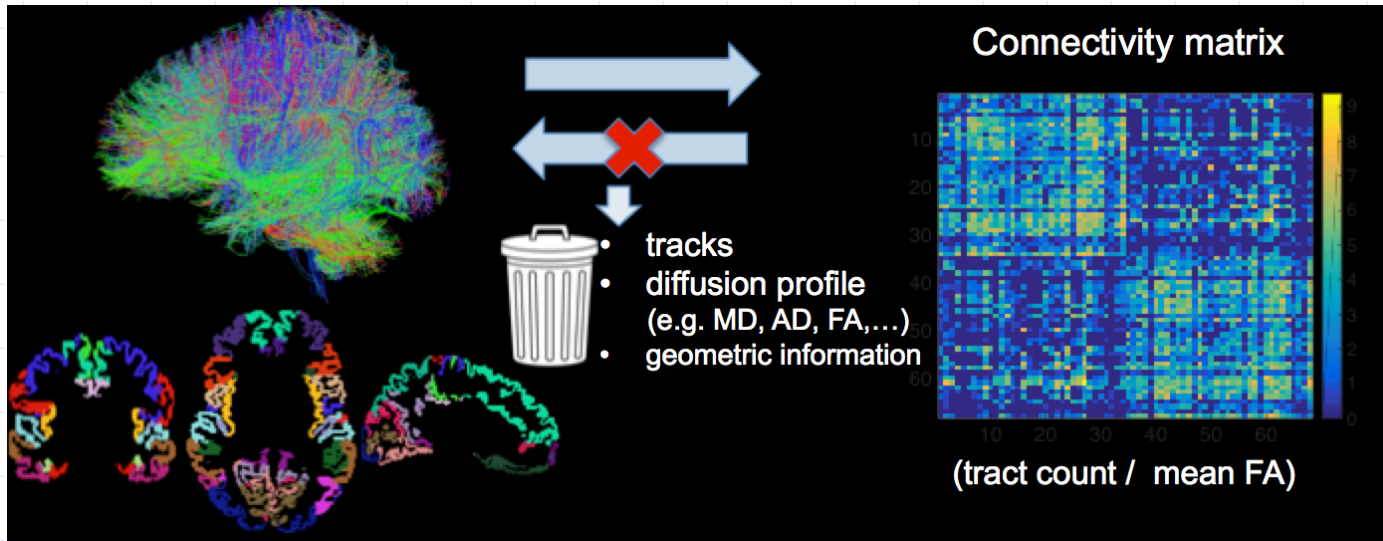
From Connectivity to Knowledge



- 1 HCP Subject
- $\sim 10^6$ curves
- ~ 3 Gbs

- Beautiful picture (video) inference?
- **Any systematic variation (with traits) in normal/disease subjects?**

Traditional Connectome Mapping



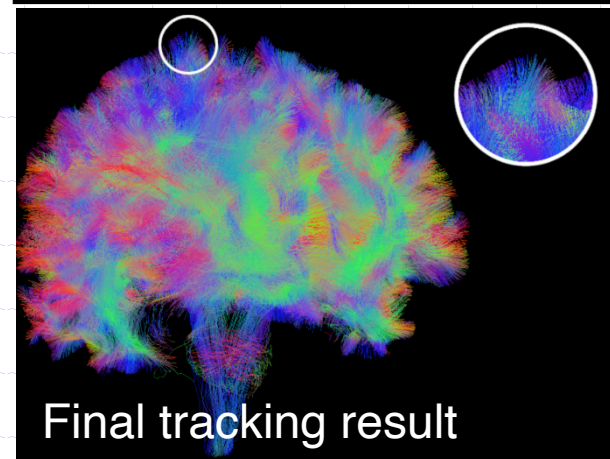
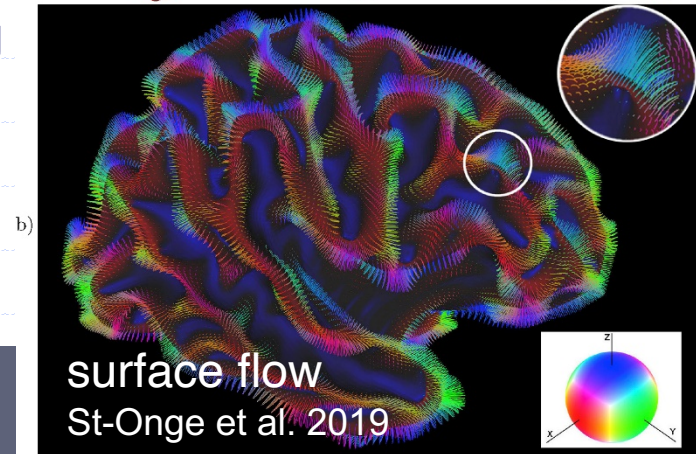
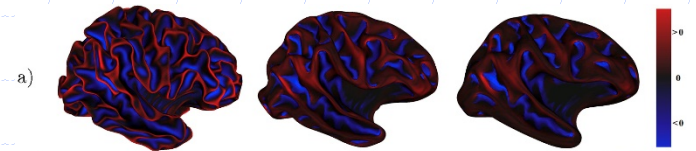
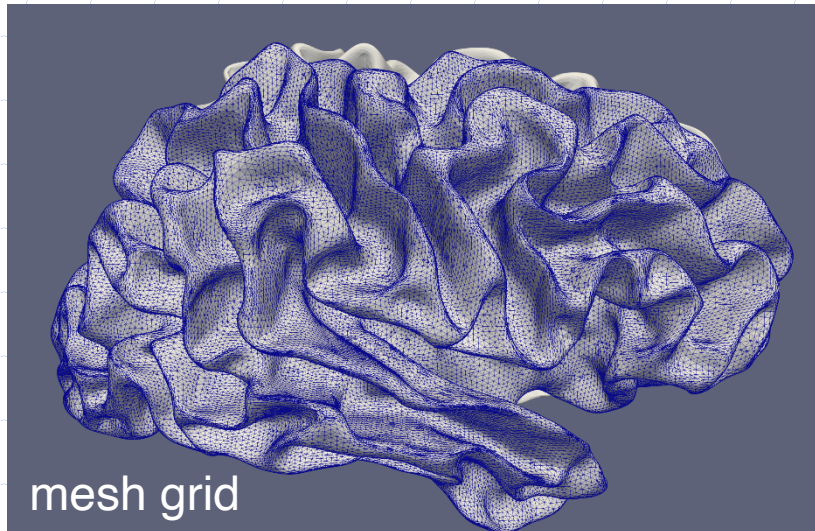
- Traditional ROIs are volume based – lack of flexibility for changing connectome resolution
- Traditional seeding for tractography is in the volume space – producing gyral biases or bias caused by large fiber bundles
- We propose to instead utilize the **white surface** to construct connectome

Surface-based Connectome Mapping

➤ Seeding on the white surface & use “surface flow” to go to white matter for fiber tracking

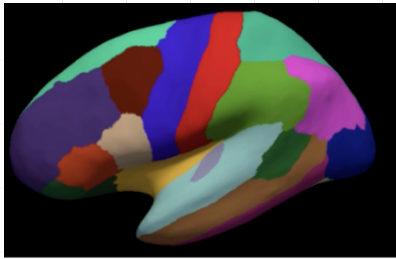
➤ All constructed streamlines are connecting white surface

➤ We put a uniform mesh grid on the white surface

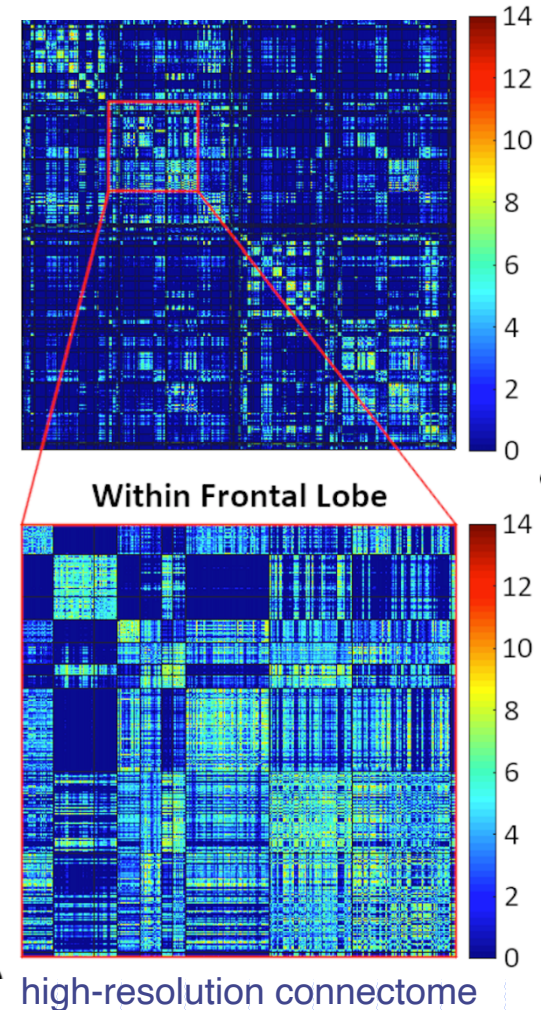
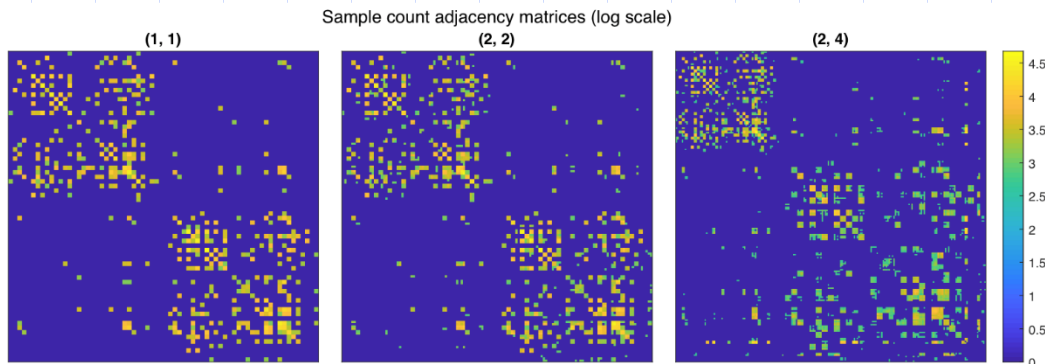


Surface-based Connectome Mapping

- Connections between mesh triangles form a high-resolution connectome
- Given any parcellation of the brain surface, e.g., Desikan, we can easily obtain a low-resolution connectivity matrix



- We can easily manipulate the parcellation to get HIGH or LOW resolution matrices



Multi-scale Graph Principal Component Analysis

Data Description

➤ Dataset: Human Connectome Project (HCP)

The HCP dataset contains:



- **Image data:** 1065 subjects with diffusion MRI and structural MRI. All are preprocessed with our PSC pipeline.
- **Traits:** Rich demographic and behavioral traits, including cognition, motion, personality measurements substance use and so on.

We extracted 175 different trait measures for each subject

Example Traits:

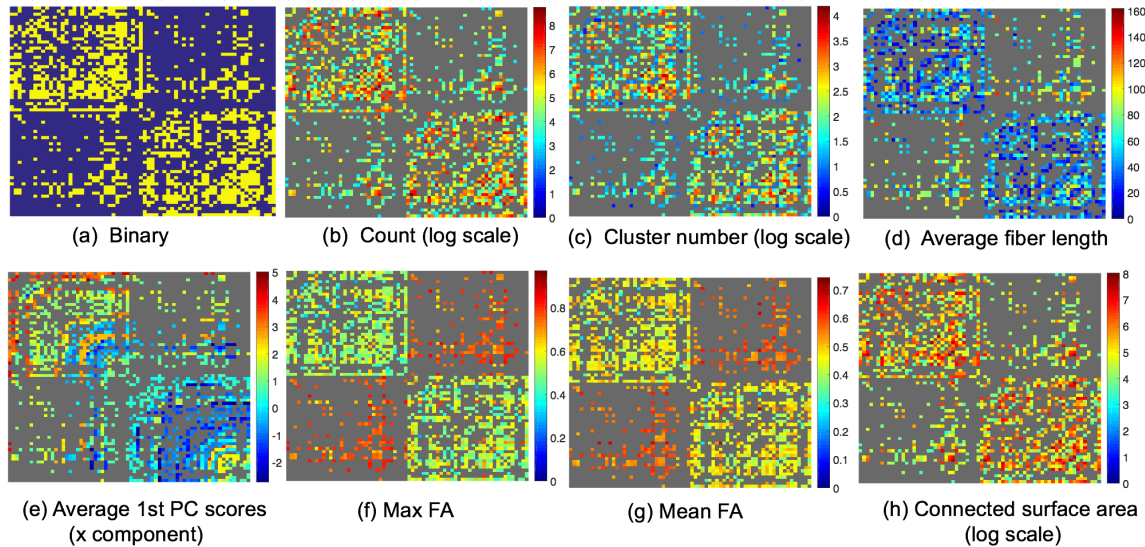
Cognition: *NIH Toolbox Oral Reading Recognition Test, Penn Word Memory Test,...*

Substance use: *Drinks per day in heaviest 12-month period, Max drinks in a single day in past 12 months,...*

Sensory: *Odor Identification, Regional Taste Intensity, ...*

Tensor Representation

- For each subject, if we stack their different weighted networks together, we obtain a 3-way tensor with dimensionality of $v \times v \times m$



- Similarly, if we stack n subjects data together, we get a 4-way tensor with dimensionality of $v \times v \times m \times n$
- Each tensor is semi-symmetric because of the symmetry of connection.

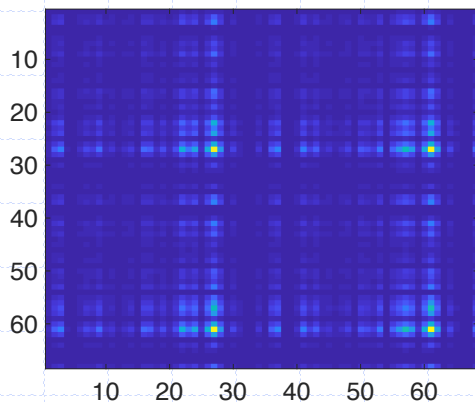
Within Scale Decompositions

- Use a semi-symmetric CP decomposition within each scale:

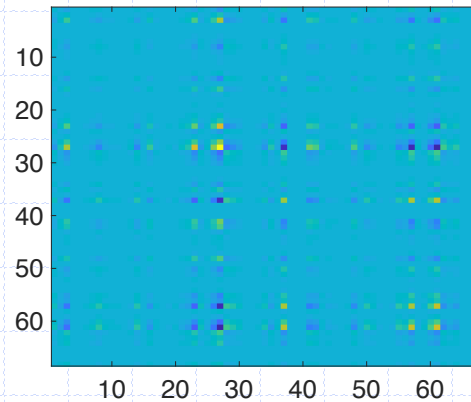
$$\mathcal{X} \approx \sum_{k=1}^K d_k \mathbf{v}_k \circ \mathbf{v}_k \circ \mathbf{u}_k,$$

$\mathcal{X} \in \mathcal{R}^{v \times v \times n}$ - v # of nodes, n subjects

- $\mathbf{v}_k \in \mathcal{R}^v$ is called network mode
- $\mathbf{u}_k \in \mathcal{R}^n$ is called subject mode
- Enforcing orthogonality for \mathbf{u}_k s



$v_1 \circ v_1$



$v_2 \circ v_2$

...

Multi Scale Tensor Principal Component Analysis

- Link single-scale models through common subject modes:

$$\mathcal{X}^{(j)} \approx \sum_{k=1}^K d_k^{(j)} \mathbf{v}_k^{(j)} \circ \mathbf{v}_k^{(j)} \circ \mathbf{u}_k, \quad j = 1, \dots, R.$$

- Anchoring fine-scale to coarse-scale data greatly reduces the effects of noise.
- Model is no more restrictive than the single-scale alternative (if K is allowed to increase).
- Need a sensible objective to optimize.

Optimization I

- Idea: minimize the squared error within each scale:

$$\begin{aligned} & \underset{d_h, \mathbf{v}_h, \mathbf{u}_h}{\text{minimize}} \left\| \mathcal{X} - \sum_{h=1}^K d_h \mathbf{v}_h \circ \mathbf{v}_h \circ \mathbf{u}_h \right\|_2 & \text{(i)} \\ & \text{subject to } \|\mathbf{u}_h\|_2 = 1, \|\mathbf{v}_h\|_2 = 1, \mathbf{v}_h^T \mathbf{v}_{h'} = \delta_{h,h'} \end{aligned}$$

- Express this as a series of rank 1 maximization problems with the n-mode product. If $\mathcal{X} \in \mathcal{R}^{I_1 \times I_2 \times \dots \times I_M}$ and $A \in \mathcal{R}^{J_n \times I_n}$ then

$$(\mathcal{X} \times_n A)_{i_1, \dots, i_{n-1}, j, i_{n+1}, \dots, i_M} = \sum_{i_n} x_{i_1, i_2, \dots, i_M} a_{j, i_n}$$

- (i) is equivalent to

$$\begin{aligned} & \underset{\mathbf{v}_h, \mathbf{u}_h}{\text{maximize}} \mathcal{X} \times_1 \mathbf{P}_{h-1} \mathbf{v}_h \times_2 \mathbf{P}_{h-1} \mathbf{v}_h \times_3 \mathbf{u}_h & \text{(ii)} \\ & \text{subject to } \|\mathbf{u}_h\|_2 = 1, \|\mathbf{v}_h\|_2 = 1 \end{aligned}$$

- \mathbf{P} is the projection onto the orthogonal complement of $[\mathbf{v}_1, \dots, \mathbf{v}_k]$
- Looks complicated, but can be written entirely in terms of dot products.

Optimization II

- Our rank 1 multi-scale problem is the sum of the squares of the single-scale problems:

$$\begin{aligned} & \underset{\mathbf{u}_h, \mathbf{v}_h^{(1)}, \dots, \mathbf{v}_h^{(R)}}{\text{maximize}} \sum_{j=1}^R \left(\mathcal{X}^{(j)} \times_1 \mathbf{P}_{h-1}^{(j)} \mathbf{v}_h^{(j)} \times_2 \mathbf{P}_{h-1} \mathbf{v}_h^{(j)} \times_3 \mathbf{u}_h \right)^2 \\ & \text{subject to } \|\mathbf{u}_h\|_2 = 1, \|\mathbf{v}_h^{(1)}\|_2 = 1, \dots, \|\mathbf{v}_h^{(R)}\|_2 = 1. \end{aligned}$$

- Use block coordinate ascent, iteratively updating

$$\begin{aligned} \hat{\mathbf{u}}_h | \mathbf{v}_h^{(1)}, \dots, \mathbf{v}_h^{(R)} &= E_{\max} \left(\sum_{j=1}^R \left(\mathcal{X}^{(j)} \times_1 \mathbf{P}_{h-1}^{(j)} \mathbf{v}_h^{(j)} \times_2 \mathbf{P}_{h-1} \mathbf{v}_h^{(j)} \right) \right. \\ & \quad \left. \left(\mathcal{X}^{(j)} \times_1 \mathbf{P}_{h-1}^{(j)} \mathbf{v}_h^{(j)} \times_2 \mathbf{P}_{h-1} \mathbf{v}_h^{(j)} \right)^T \right) \\ \hat{\mathbf{v}}_h^{(j)} | \mathbf{u}_h &= E_{\max} \left(\mathbf{P}_{h-1}^{(j)} (\mathcal{X}^{(j)} \times_3 \mathbf{u}_h) \mathbf{P}_{h-1}^{(j)} \right), \end{aligned}$$

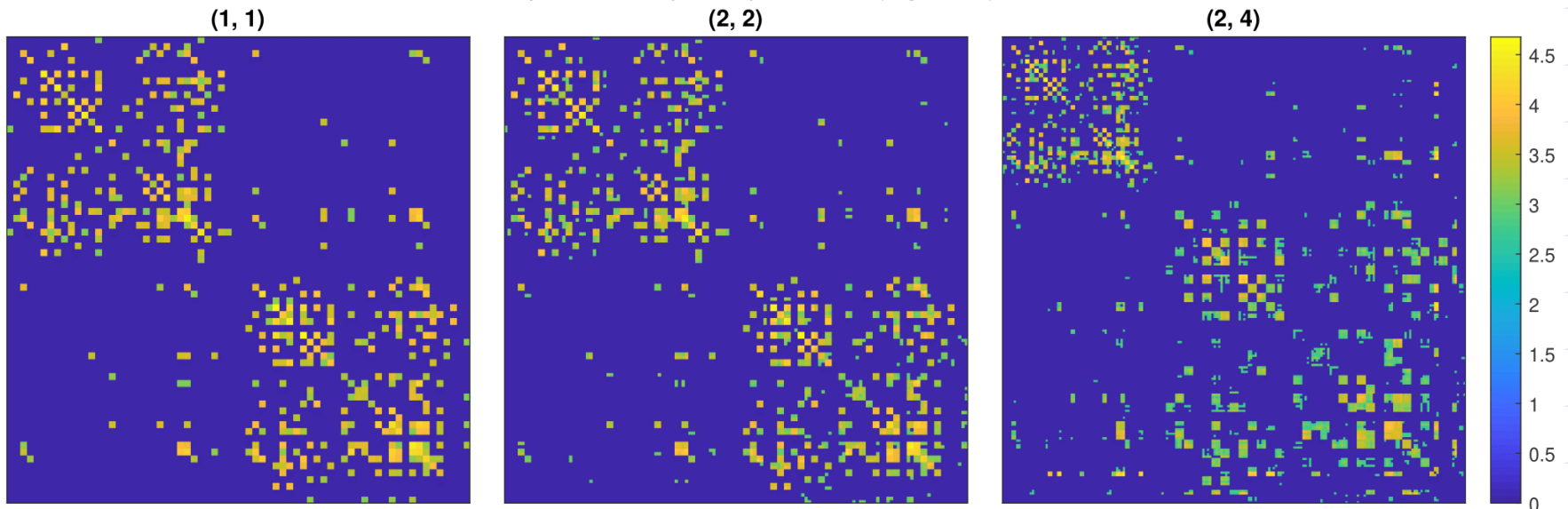
where $E_{\max}(\mathbf{A})$ is the eigenvector of \mathbf{A} with the largest eigenvalue.

- Multi-scale modelling combines spectral information across scales!

Applications to HCP

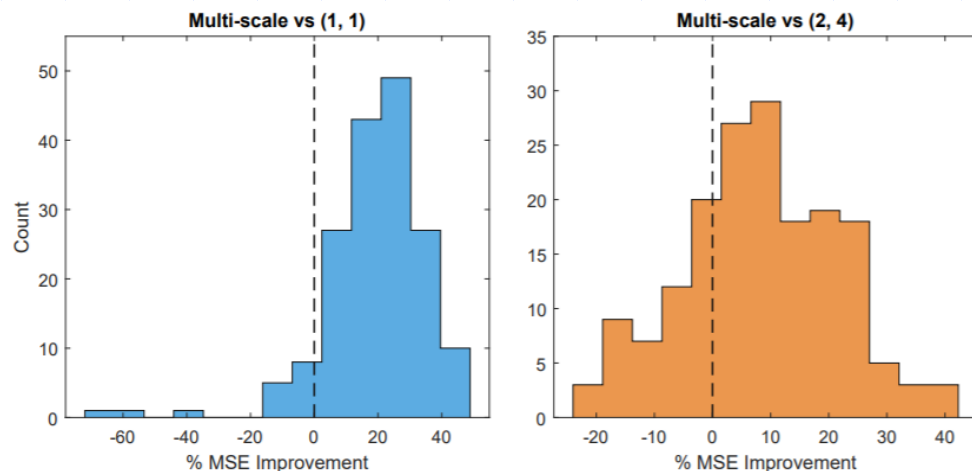
- Tested our model on 118 individuals.
- Notation: (l, r) is the parcellation created by splitting each Desikan region in the left (resp. right) hemisphere into l (resp. r) regions.
- Consistently saw the greatest gains with $\{(1,1), (2,4)\}$.

Sample count adjacency matrices (log scale)

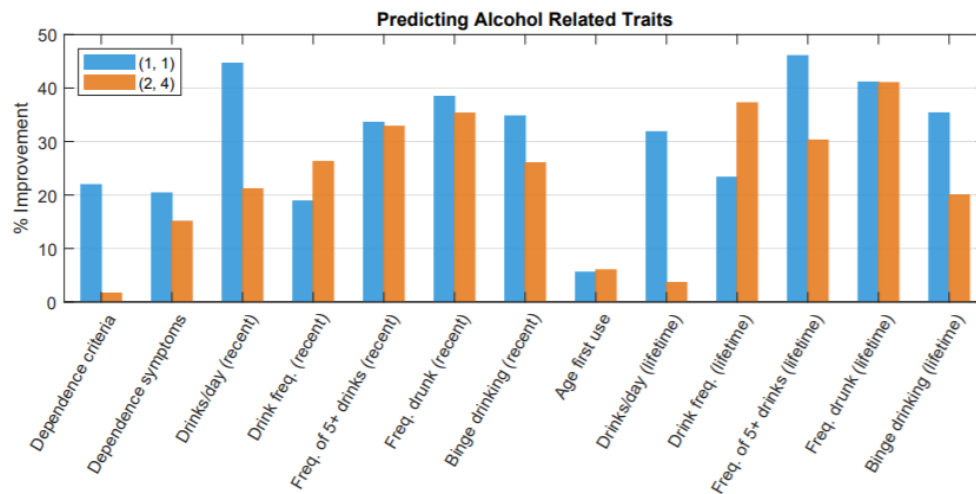


Improved Trait Predictions

- Used $K=70$ latent factors as inputs to ridge regression.
- Trained on 70% of data, computed MSE for trait predictions on other 30%.



(a) Histograms of MSE improvements. The left plot has one omitted outlier around -200 .



(b) Highlighting improvements in predicting traits specifically related to alcohol use.

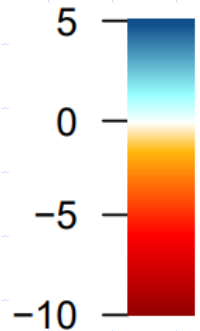
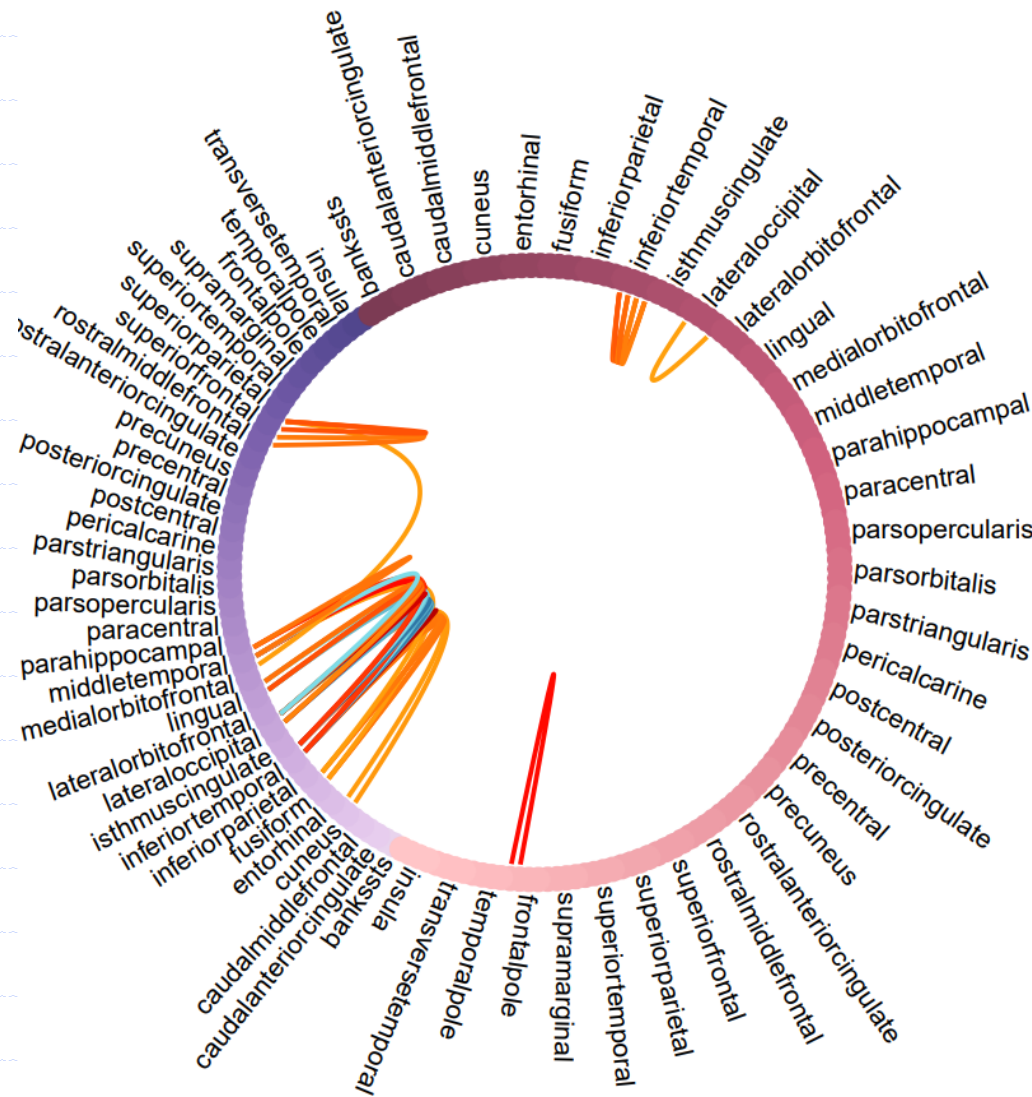
Inference on Group Differences

- Interested in understanding how the connectome changes with traits.
- Two stage process:
 1. Find a unit direction \mathbf{w} such that the projection of the latent factors onto \mathbf{w} are maximally correlated with the traits.
 2. Map back onto networks:

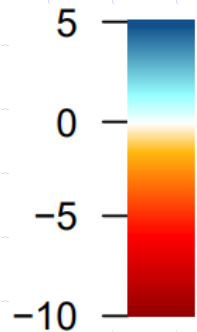
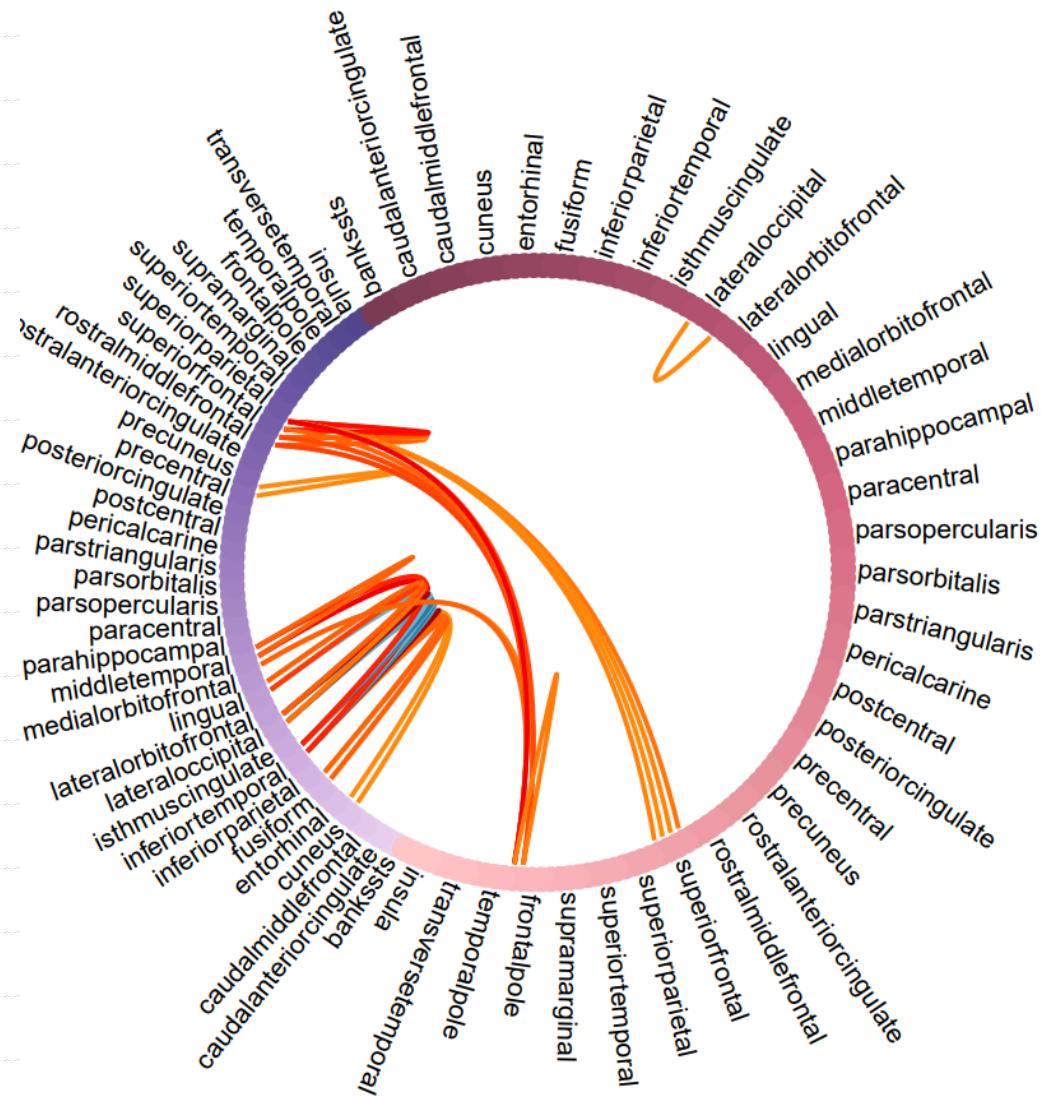
$$\Delta_{\mathbf{X}}(s) = s \sum_{k=1}^K d_k \mathbf{w}(k) v_k \circ v_k, \text{ for } s \in [-1, 1],$$

- Tested on the HCP trait “worst lifetime binge drinking” with $\{(1,1), (2,4)\}$.
- Next slides show 100 largest changes when moving from low to high binge drinking ($K=10$).

Effects of Binge Drinking (Single-Scale)



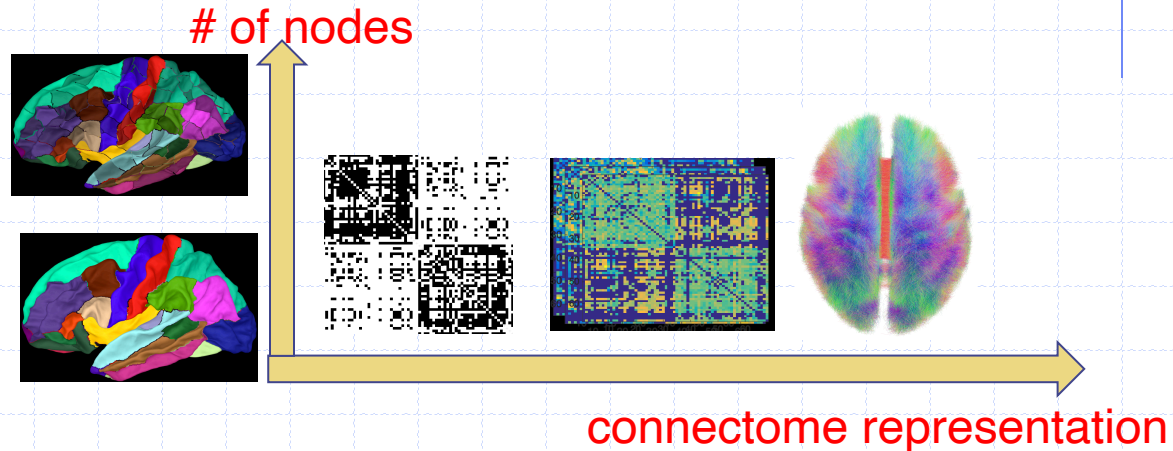
Effects of Binge Drinking (Multi-Scale)



Summary

➤ We developed a **multiscale + multiresolution** population-based structural connectome analysis framework

- Reproducible
- Preserves the geometry and diffusion information



➤ Novel statistical methods for new connectome data analysis:

- To understand the normal connectome variation in healthy subjects
- To relate connectome to covariates of interest and traits
- To predict the risk of neuropsychiatric disorders

Acknowledgements

Collaborators:

Zhengwu Zhang (*UNC STOR*)

Steven Winter (*Duke Statistical Science*)

Funding Support:

National Institute of Health grant 1R01MH118927-01