Geometry-based Brain Structural Connectome Analysis

Zhengwu Zhang

July 20, 2018

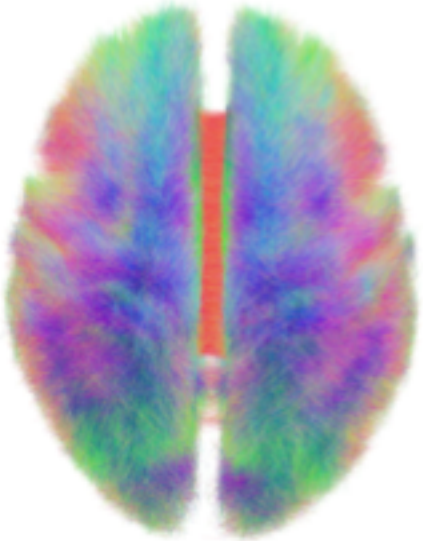
For CBMS Conference
Elastic Functional and Shape Data Analysis

Department of Biostatistics and Computational Biology

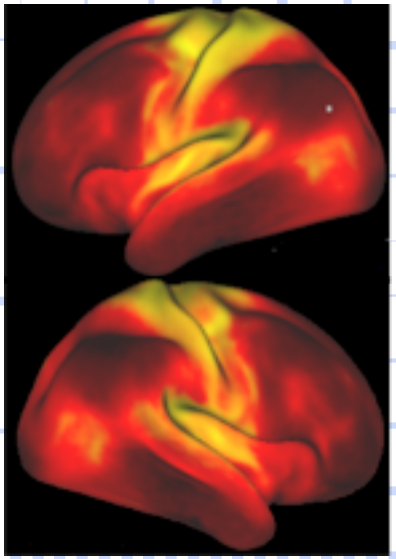
UNIVERSITY of ROCHESTER

Different Brain Connectomes

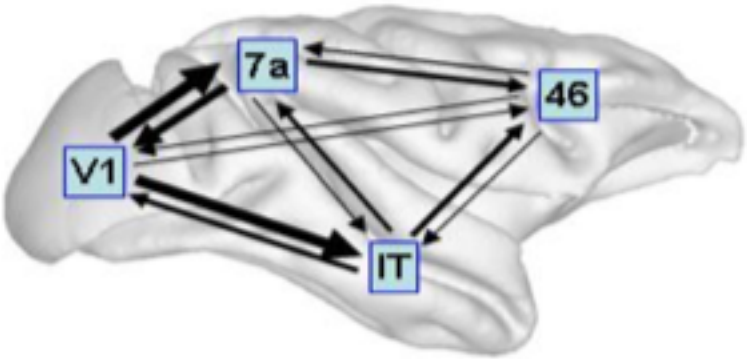
- **Structural Connectivity**
  - A pattern of anatomical links, dMRI



- Functional Connectivity
  - Statistical Dependencies, fMRI, EEG, MEG



- Effective Connectivity
  - Causal interactions, fMRI, EEG, MEG



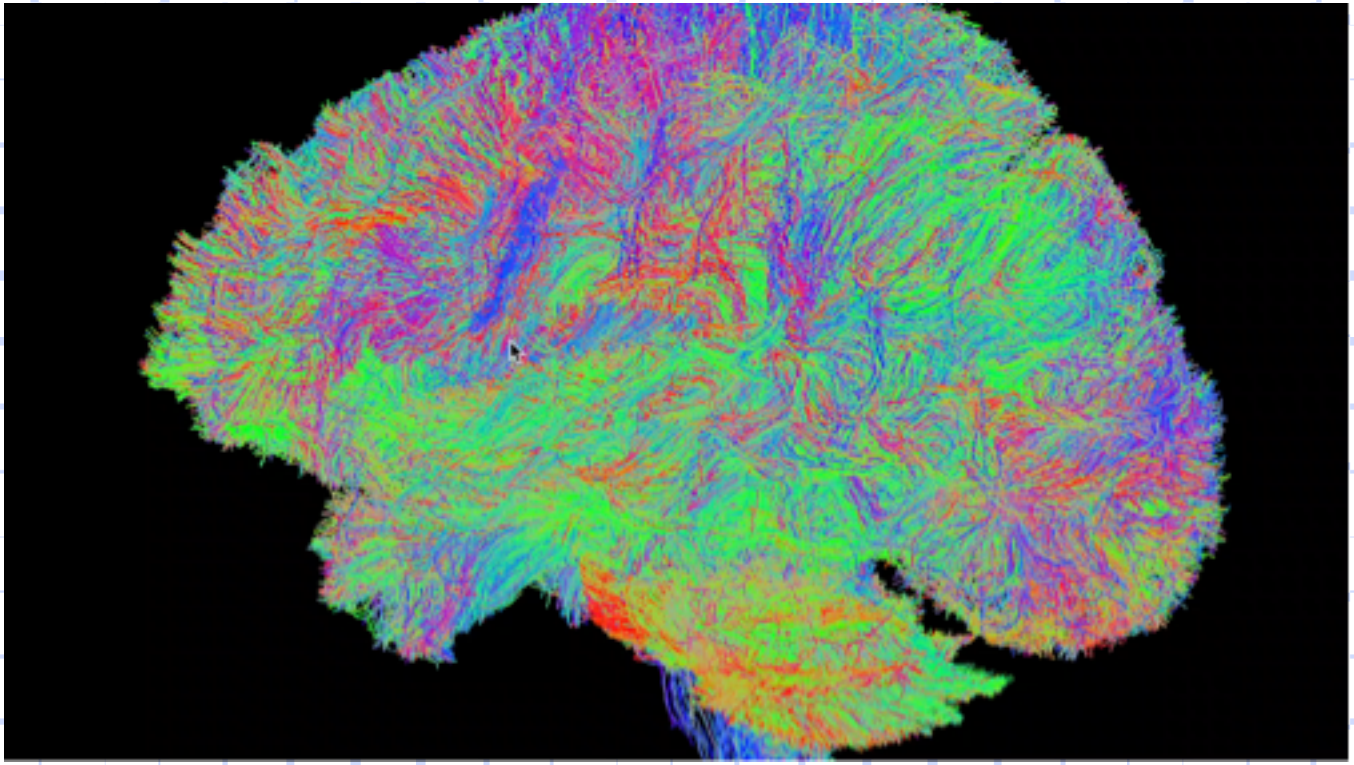
The Human Connectome Project

- The HCP is to elucidate 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, substance use and personality measurements.

- Diffusion MRI now is routinely collected in all brain studies
  - UK Biobank
  - The Adolescent Brain Cognitive Development (ABCD) Study
  - …



- 1 HCP Subject
- ~ 10^6 curves
- ~ 3 Gbs

- **Research problems:** reconstruction, representation and statistical analysis

➢ Reconstruction: Reliably and accurately recover white matter tracts

➢ Representation: Represent in forms that can facilitate population-based analysis

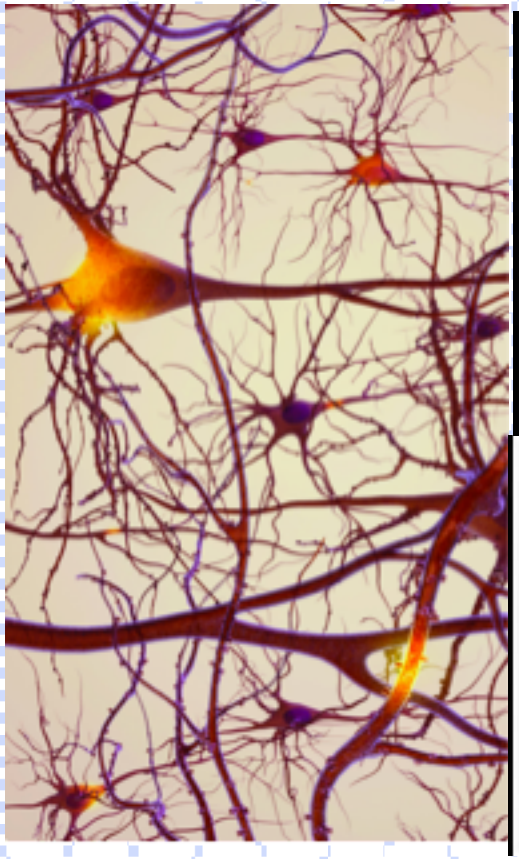
➢ Statistical analysis: Find systematic variation (with traits) in normal/disease subjects

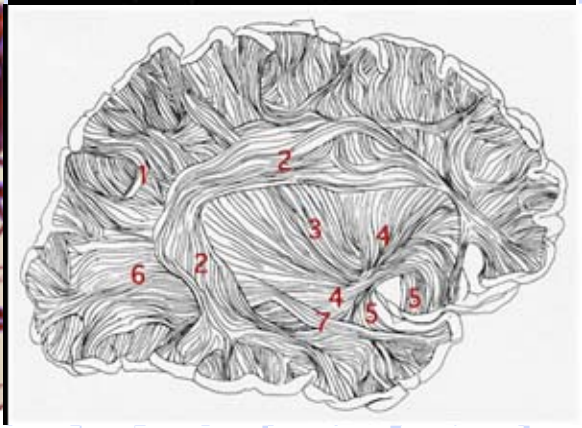
Outline

- Introduction to diffusion MRI
- Construction of geometric connectomes
- Geometric representations of connectomes
- Statistical analysis of connectomes
- Software demonstration

Diffusion Imaging

- Axons have ~μm diameters
- Axons group together in bundles that traverse the white matter in brain
- We can not image individual axon, but we can image bundles with diffusion MRI technique

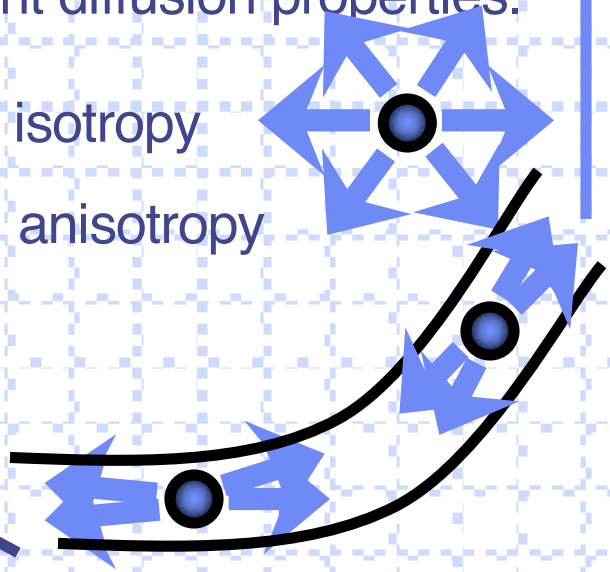




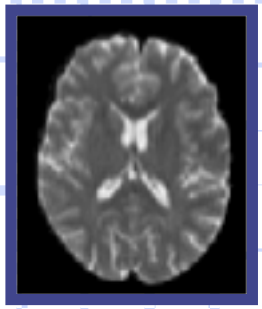
(From UMD website)

Diffusion in Brain Tissue

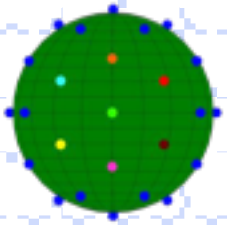
- Water molecules in different tissues have different diffusion properties.
  - Gray matter: Diffusion is unrestricted ⇔ isotropy
  - White matter: Diffusion is restricted ⇔ anisotropy

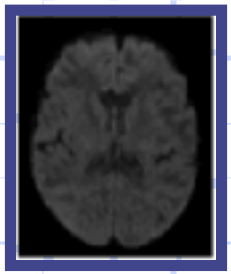


- Diffusion MRI measures the water diffusion movement inside brain

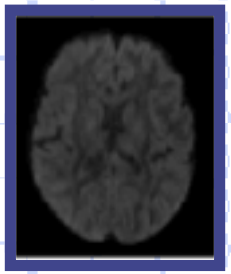


No diffusion encoding

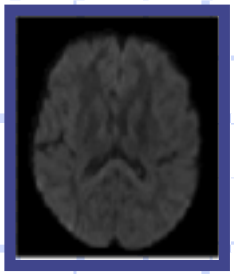




Diffusion in direction g_1



g_2



g_3

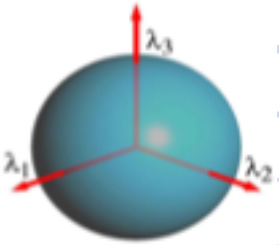
...

Reconstruction of White Matter Bundles

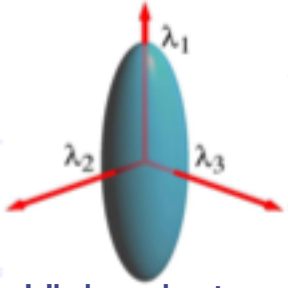
- At each voxel, we want to infer:
  - The orientation and the magnitude of the diffusion

(1) Diffusion tensor image (DTI)

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



Low anisotropy



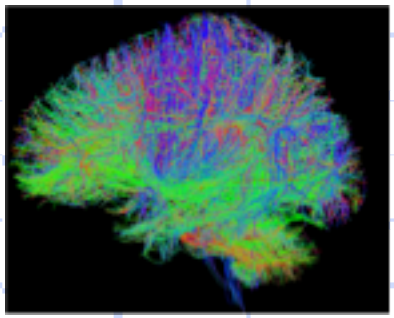
High anisotropy

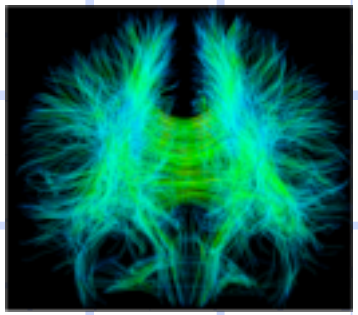
(2) High angular resolution diffusion imaging (HARDI)

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

fODF

- Fiber reconstruction: use local diffusion info to recover fibers





Outline

- Introduction to diffusion MRI
- Construction of geometric connectomes
- Geometric representations of connectomes
- Statistical analysis of connectomes
- Software demonstration

Current Tractography Approach

- Most of the existing (Tractography) methods are based on an ordinary differential equation to grow fiber β from a seed point:

$$\frac{d\beta(t)}{dt} = \hat{\mathbf{e}}(\beta(t)), \beta(0) = \mathbf{v}_0, t \geq 0$$

where $\hat{\mathbf{e}}(\beta(t))$ represents the estimated local fiber orientation.

- There are many techniques / algorithms to improve estimation $\hat{\mathbf{e}}(\beta(t))$
  - Mixture of tensors [*Wong et al. 2016*]
  - Fiber ODF [*Descoteaux et al. 2009*]
  - Incorporate spatial information [*Raoa et al. 2016*]
  - Sparsity [*Daducci et al. 2014*]

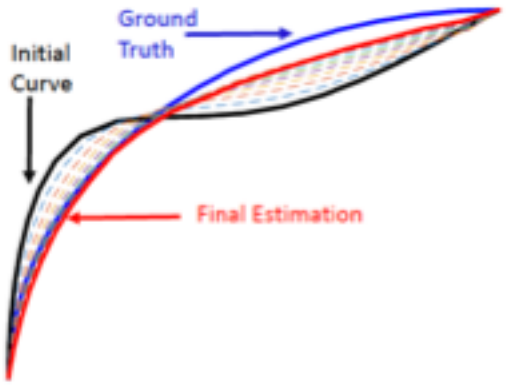
- Here, we proposed two novel procedures to improve the fiber curve construction process

  (1) A Bayesian active contour approach

  (2) A multiscale approach

Method 1: Active Contour Tractography

- **Main idea**: prior + data to reduce false positives fibers
- Geometric prior (shape) is learned from atlas data (e.g., *Yeh et al. 2018*)
- Bayesian active contour methods to recover long fiber curves



- Given two fixed points, we seek parameterized curve $\hat{\beta}(t)$ connecting the two fixed points that minimizes

$$\hat{\beta} = \mathit{argmin}_{\beta \in \mathcal{B}} E_{total}(\beta) \ ,$$

where

$$E_{total}(\beta) = \lambda_1 E_{data}(\beta) + \lambda_2 E_{smooth}(\beta) + \lambda_3 E_{prior}(\beta).$$

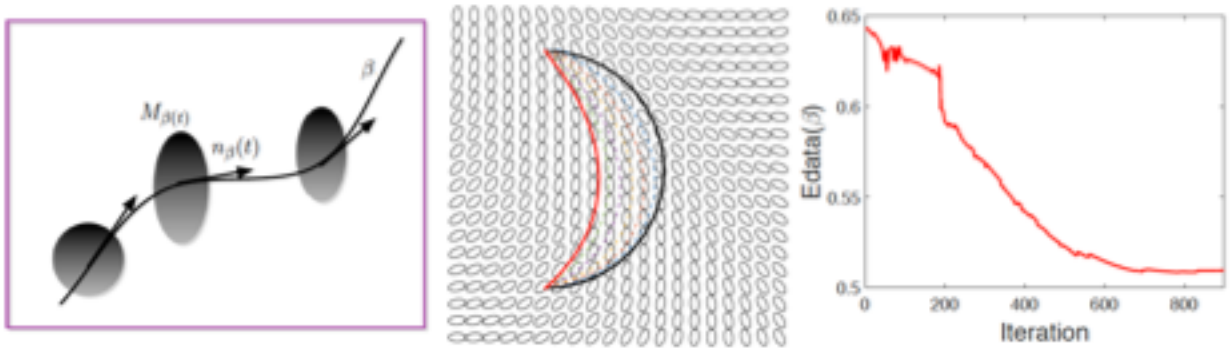
(1) Data-Likelihood term: $E_{data}[\beta] = \int_0^1 n_\beta(t)^T M_{\beta(t)}^{-1} n_\beta(t) \ dt$, where $n_\beta(t) = \frac{\dot{\beta}(t)}{|\dot{\beta}(t)|}$

(2) Smoothness: $E_{smooth}(\beta) = \int_0^1 |\dot{\beta}(t)| dt$

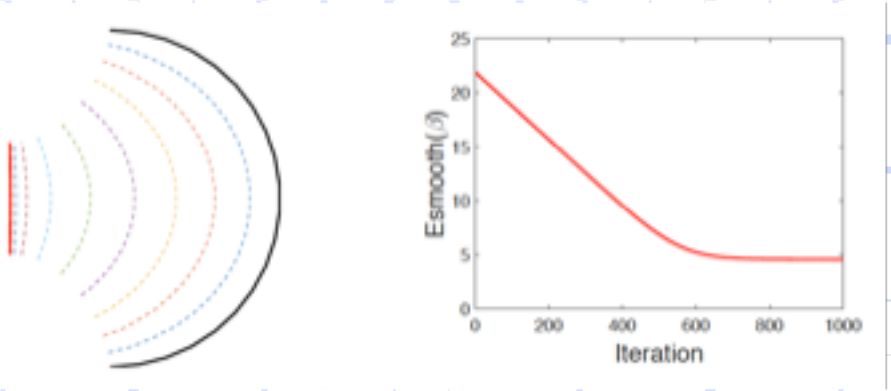
(3) Shape prior: normal distribution in shape space (*square-root velocity function*)

Some Simulation Results

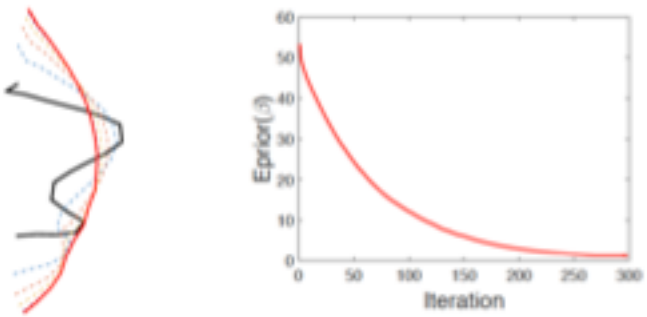
- Data term:



- Smoothness:



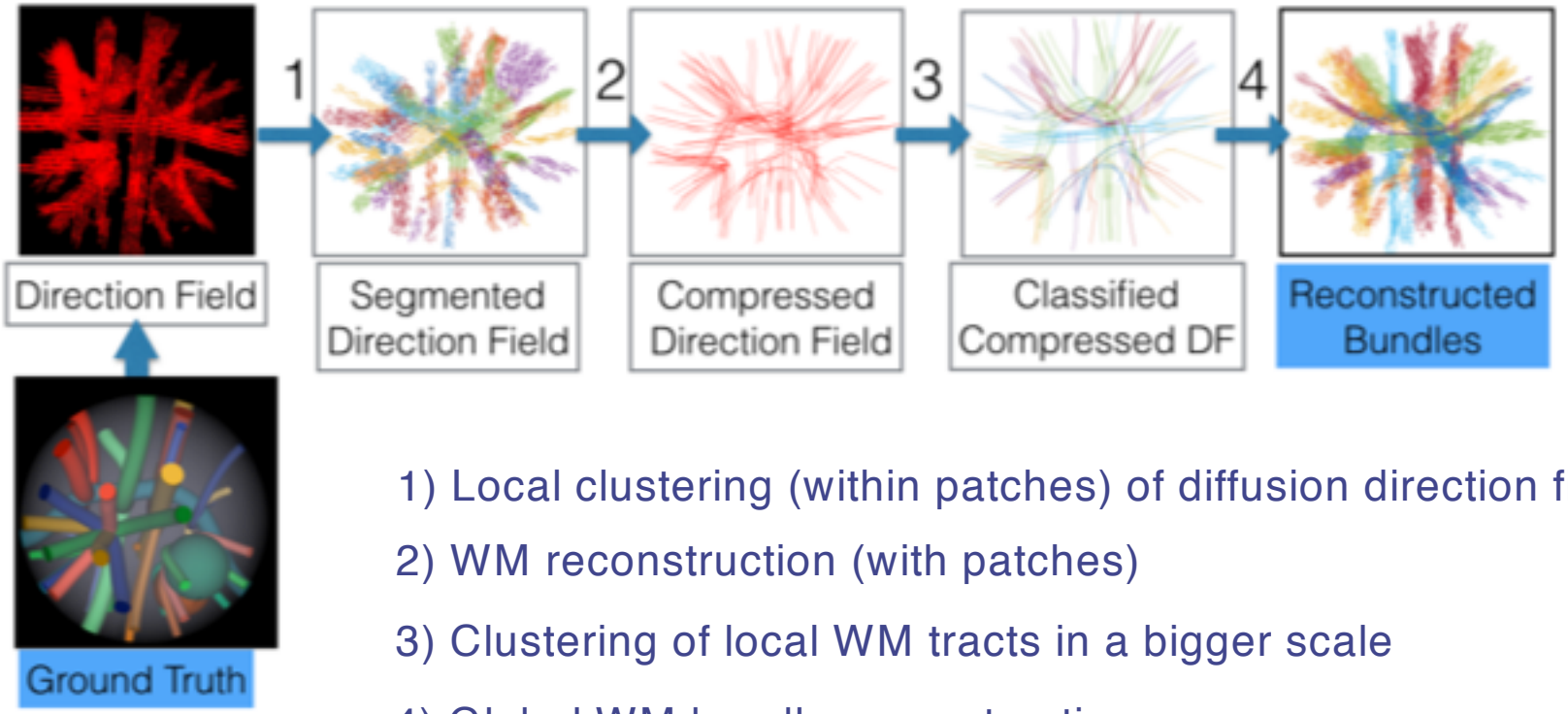
- Shape term:



- Black: initial curve
- Red: final curve

Method 2: A Multiscale Approach

➢ **Main idea**: (1) local white matter (WM) configurations are much easier to handle comparing with the global one;
(2) multiscale approach is used to bridge local and global WM config.



1) Local clustering (within patches) of diffusion direction field

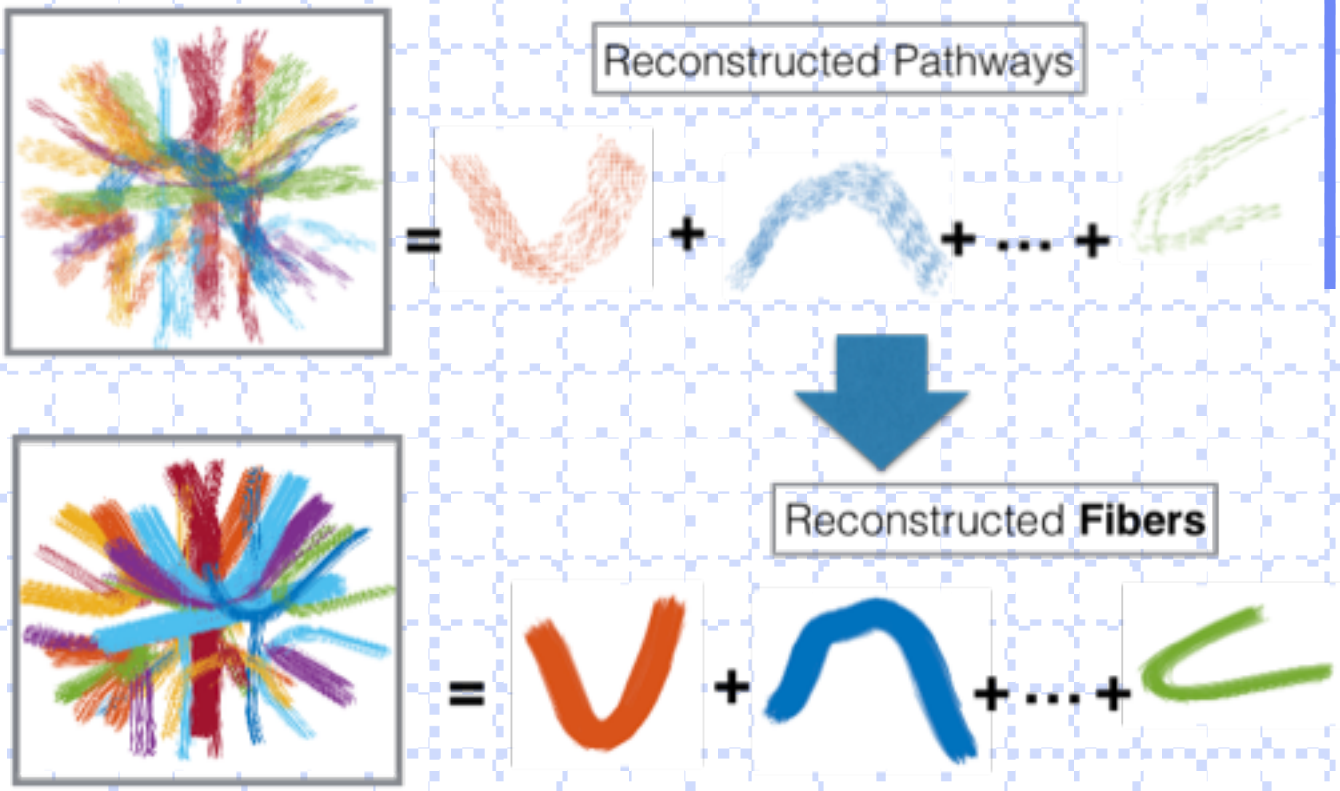
2) WM reconstruction (with patches)

3) Clustering of local WM tracts in a bigger scale

4) Global WM bundle reconstruction

Some Simulation Results

- Fiber Reconstruction



- Evaluation:

	#t	VC(%)	IC(%)	NC(%)	VC+IC(%)	$\frac{VC}{VC+IC}$(%)	VB	IB
MSMT-iFOD2	8511	35.9	22.7	41.4	58.6	61.3	27	86
MSMT-GT	5756	15.9	7.0	77.1	22.9	69.4	**27**	50
DMDT	8124*	**99.6**	**0**	**0.4**	**99.6**	**100**	**27**	**0**

MSMT: Multi-shell multi-tissue global tracking method (Christiaens et al. 2015)

DMDT: Our method - deep multiscale diffusion tracking

Future Directions

- A few key factors affecting tractography:
  - Image resolution
  - Local WM fiber configuration estimation (e.g., fODF estimation)
  - Prior ground truth knowledge
  - Supervised / semi-supervised methods to incorporate priors

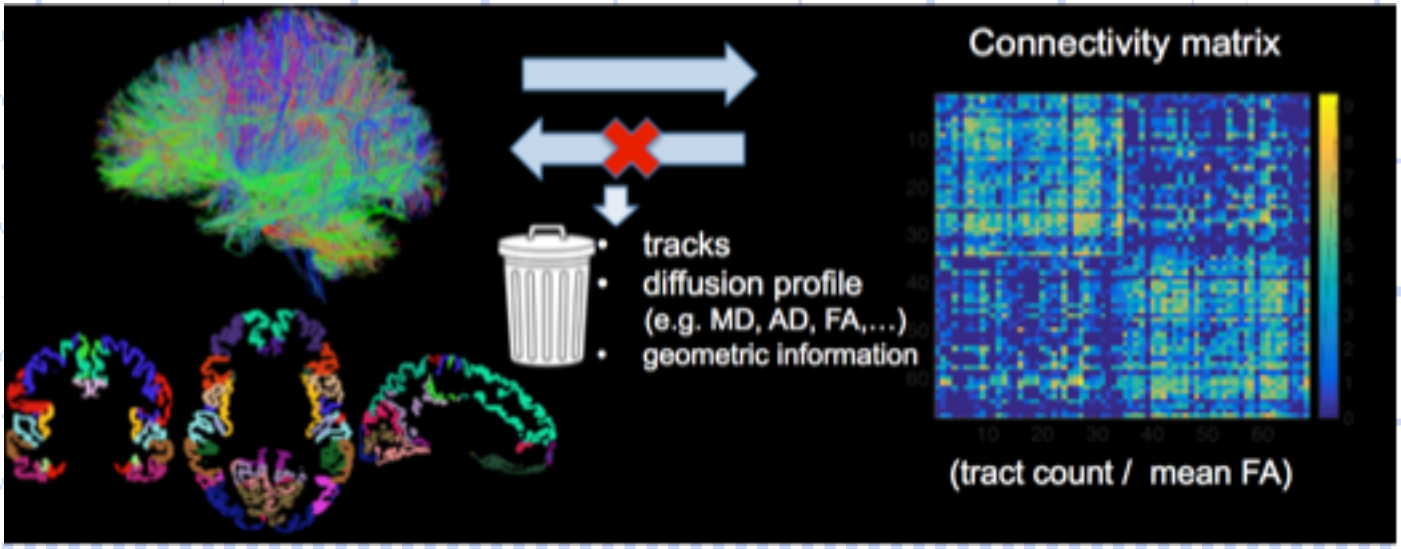
- Our working directions:
  - Collaborating with radiologists (e.g. Dr. Allen Song at Duke) to obtain high resolution dMRI data
  - Better methods to estimate fODF / diffusion tensors by borrowing geometric information inside the brain (e.g., spatial location, brain tissue type)
  - Better prior geometric knowledge from experts / animal data
  - Novel tractography methods that can incorporate priors + data

Outline

- Introduction to diffusion MRI
- Construction of geometric connectomes
- Geometric representations of connectomes
- Statistical analysis of connectomes
- Software demonstration

Diffusion MRI to Connectome

- Traditional pipelines reduce the rich information into a matrix

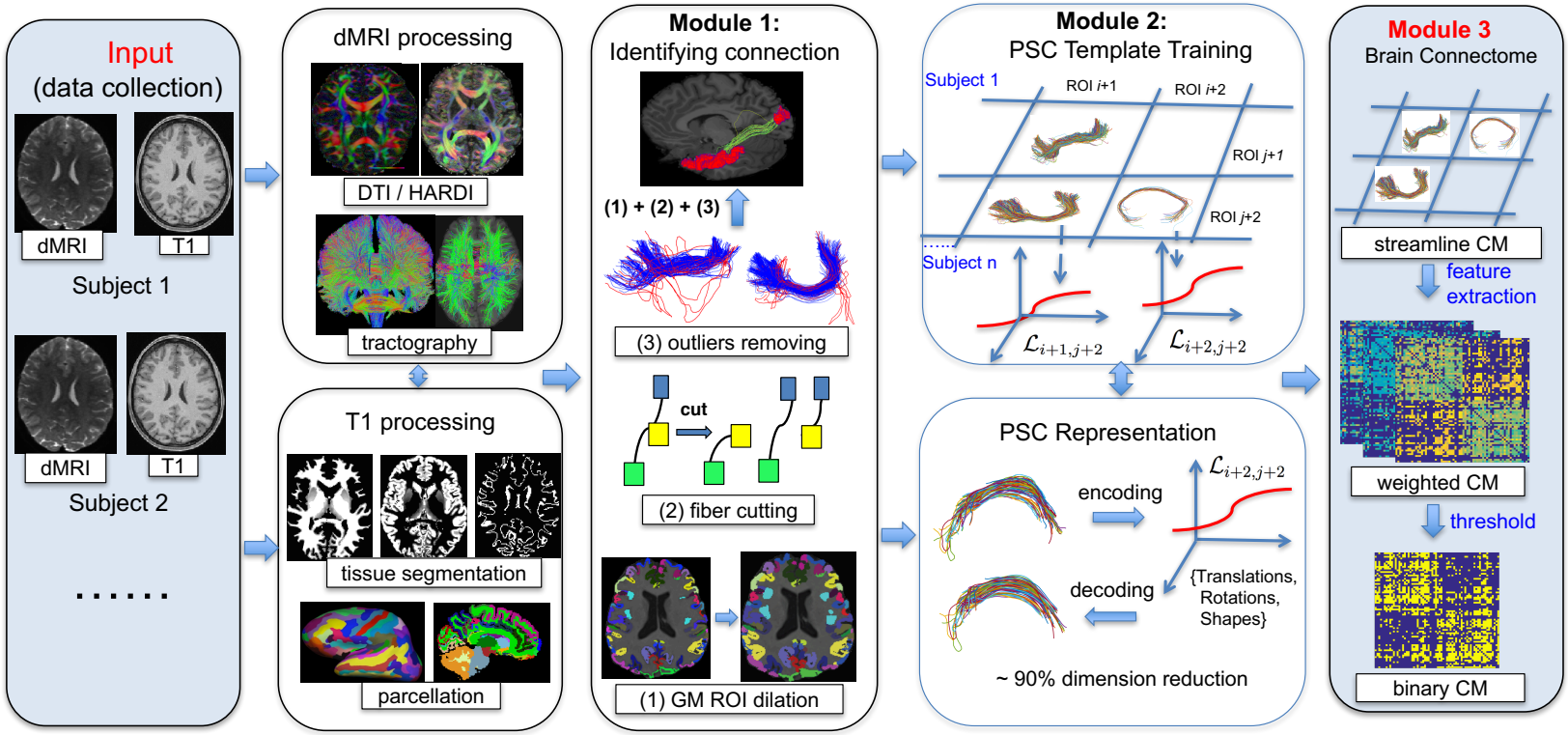


- Information loss
- How to define meaningful values in the connectivity matrix
- Not reproducible because of the noise in the data

A New Connectome Mapping Framework

Structural Connectome Mapping

- We developed a new population-based structural connectomes (PSC) mapping framework



(1) Provides multiscale representations

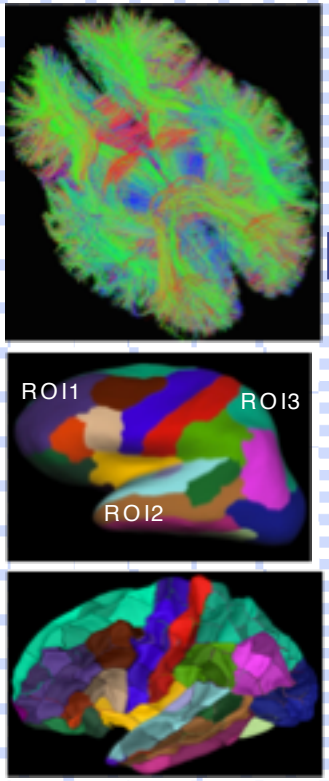
(2) Preserves more information

(3) Improves robustness and reproducibility

Multi-Scale Connectome Representation

- Connectome analysis at different levels

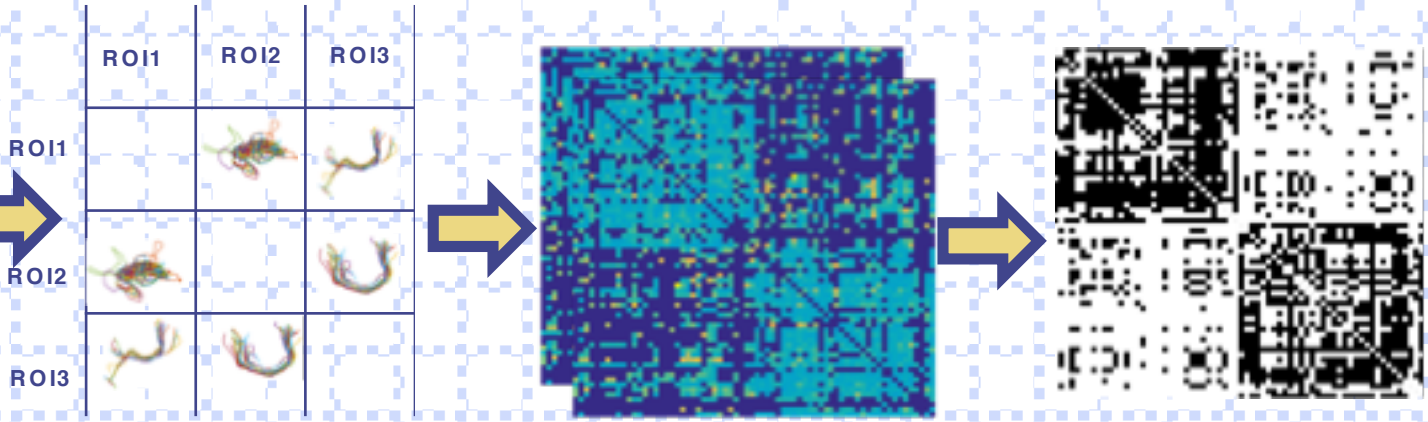
Tractography
+
Parcellation



finer resolution

Complex → Simple

Streamline level | Weighted network level | Binary network level



Streamline Connectivity Cell Matrix (SCCM)

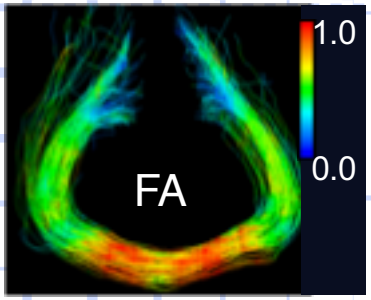
Scalar Matrices

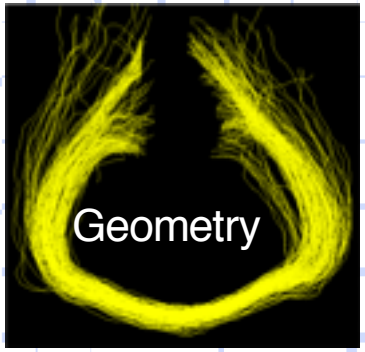
Binary Matrix

Z. Zhang, M. Descoteaux, A. Srivastava, D. Dunson, H. Zhu, et al.
Mapping Population-based Structural Connectomes, *Neuroimaging*

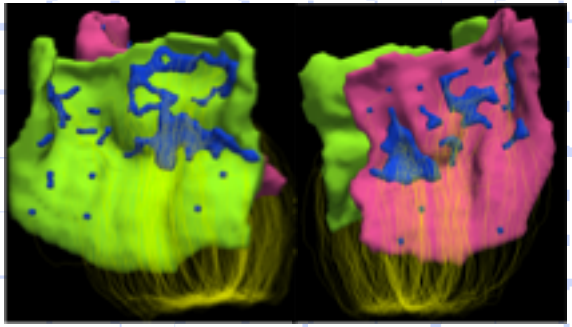
New Features for Connectome Analysis

- PSC extracts different features reflecting different aspects about one connection



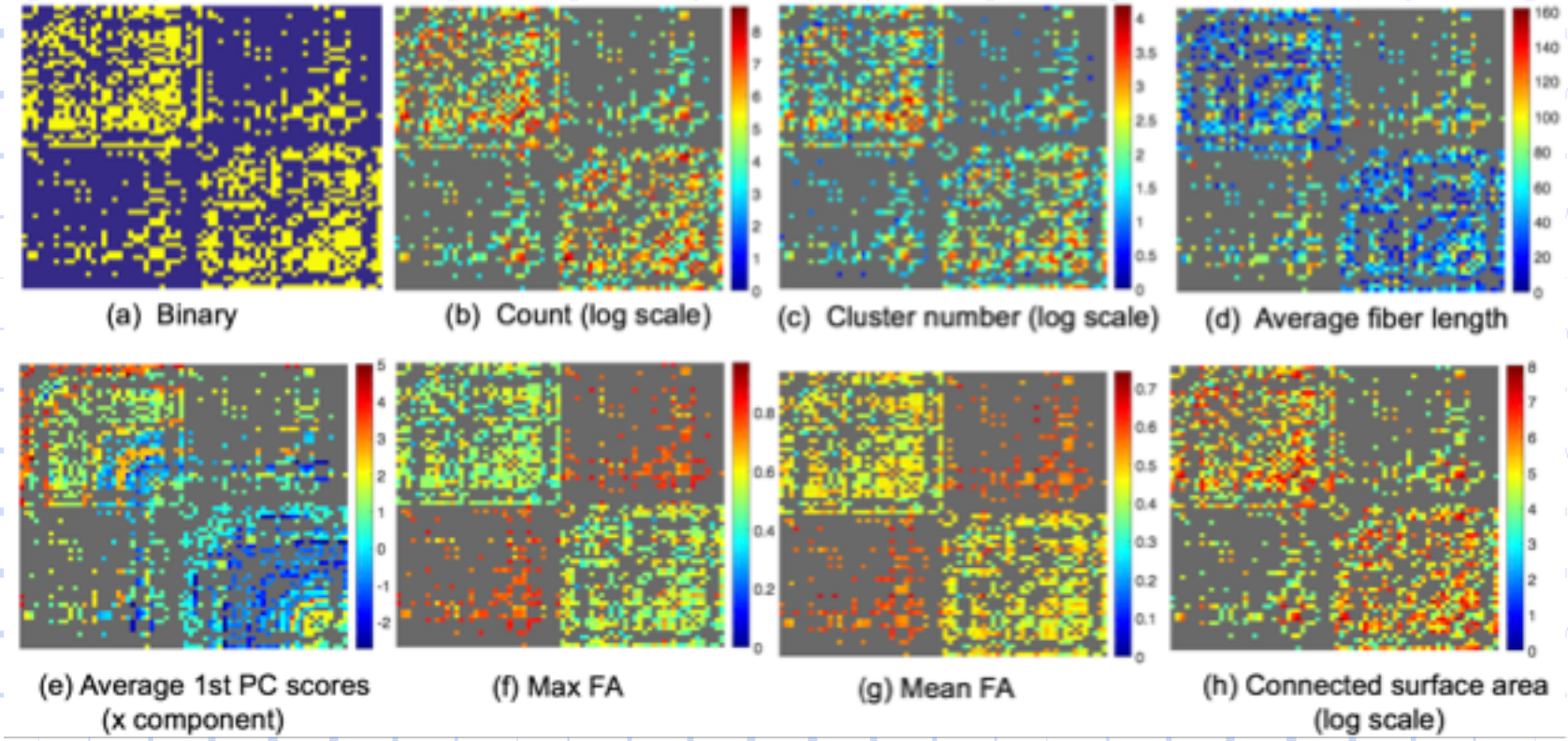


- Diffusion-related features
  - ✓ DTI metrics, such as Fractional Anisotropy (FA), Mean Diffusivity (MD), et al.
  - ✓ ODF metrics, such as Generalized Fractional Anisotropy (GFA), Apparent Fiber Density (AFD), et al.
- Geometry-related features
  - ✓ Average fiber length
  - ✓ # of clusters
  - ✓ Average deviation from the mean fiber
  - ✓ Topological features – Persistent homology
- Endpoint-related features
  - ✓ Fiber count
  - ✓ Connected surface area
  - ✓ Weighted connected surface area



Connected surface area

Examples of Weighted Networks



Test-Retest Dataset to Improve Reproducibility

- Sherbrooke Test-Retest Dataset (clinical-like acquisition):
  - 11 subjects, and 3 scans per subject – with 1 month interval
  - 1.5 Tesla, 2 *mm* isotropic resolution, single shell, 64 diffusion weighting directions
- Human Connectome Project (HCP) Test-Retest Dataset:
  - 44 subjects, and 2 scans per subject
  - 3 Tesla, 1.25 *mm* isotropic resolution, 3 shells, 270 diffusion weighting directions
- Quantitative evaluation of the reproducibility
  - Distance-based intraclass correlation coefficient (dICC)

$$\mathrm{dICC} = \frac{\bar{d}_{bs}^2 - \bar{d}_{ws}^2}{\bar{d}_{bs}^2}$$

$\bar{d}_{bs}^2$ -- average distance between subjects

$\bar{d}_{ws}^2$ -- average distance within subjects (multiple scans)

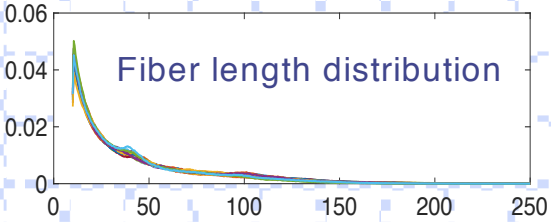
  - Distance is obtained based on L2 norm before network adjacency matrices

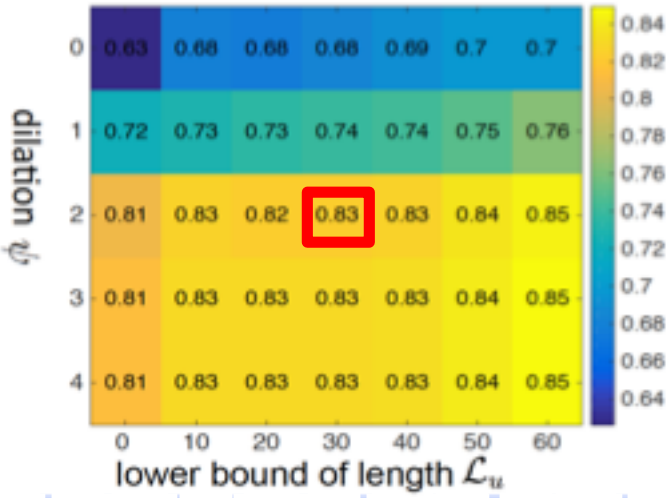
PSC Parameter Optimization

- DICC helps to select the optimal parameters

(1) Dilation parameter ψ

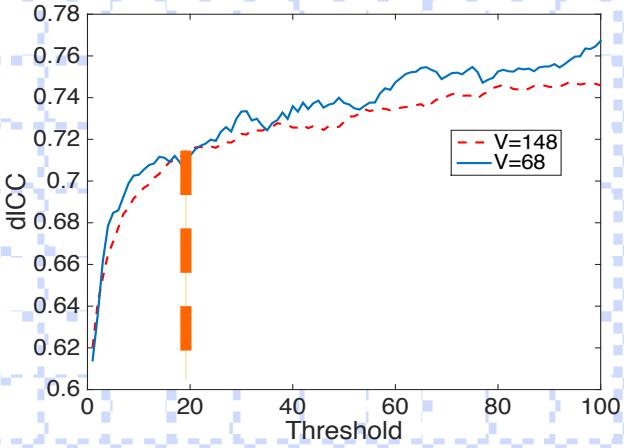
(2) Remove short fibers?





- Dilation 2 mm
- Lower bound of fiber length: 20 mm

(3) Threshold to get a binary network



Threshold vs dICC

- Threshold between 10 and 20

Comparison with Traditional Framework

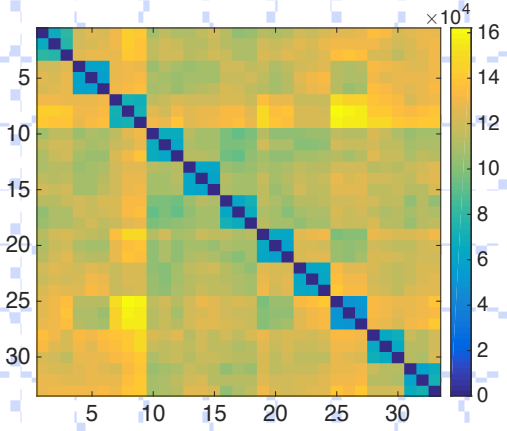
- PSC V.S. traditional pipeline (MIGRAINE [Roncal et al., 2013]) using **count feature**

PSC

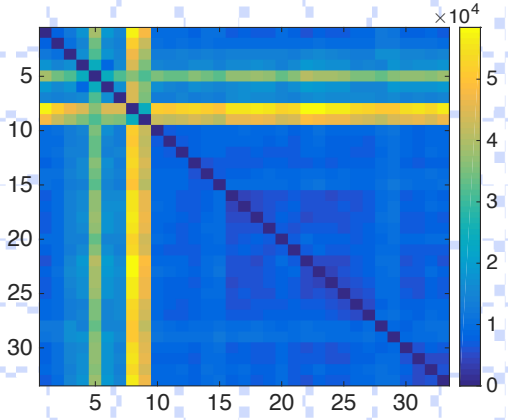
Traditional method

- Pairwise distance

(Sherbrooke Test-Retest Dataset)



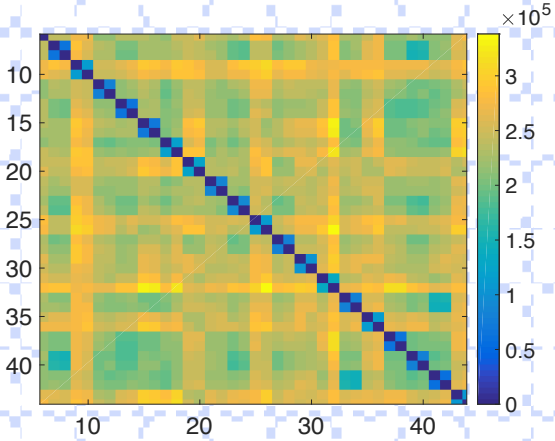
dICC = 0.79



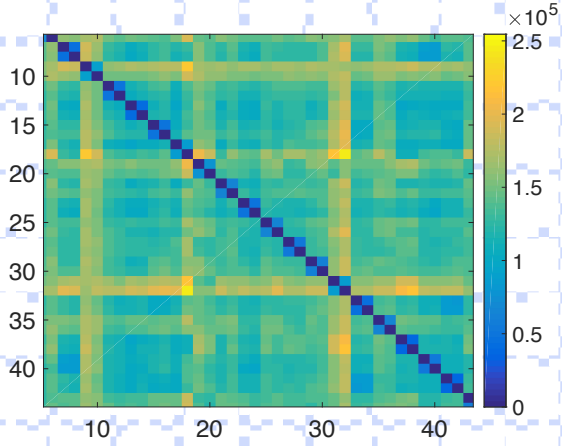
dICC = 0.40

- Pairwise distance

(HCP Test-Retest Dataset)



dICC = 0.87



dICC = 0.82

Outline

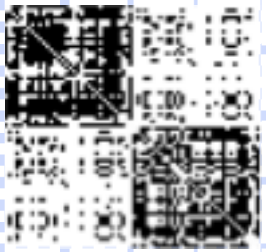
- Introduction to diffusion MRI
- Construction of geometric connectomes
- Geometric representations of connectomes
- Statistical analysis of connectomes
- Software demonstration

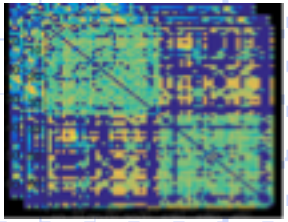
Structural Connectome Statistical Analysis

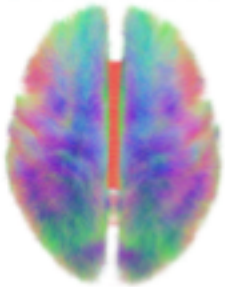
Simple

Data

Complex

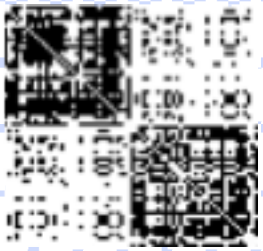


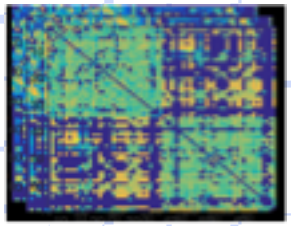


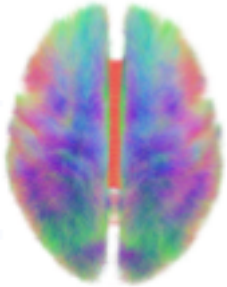


Ongoing Studies

Simple







Complex

- Common and Individual Structure of Multiple Networks

  With Lu Wang and David Dunson

- Heritability of structural and functional connectomes

  With Ben Risk and Hongtu Zhu

- **Tensor network factorizations: Relationships between brain structural connectomes and traits**

  With Genevera Allen and David Dunson

- **Discovering brain subgraphs related to human traits**

  With Lu Wang and David Dunson

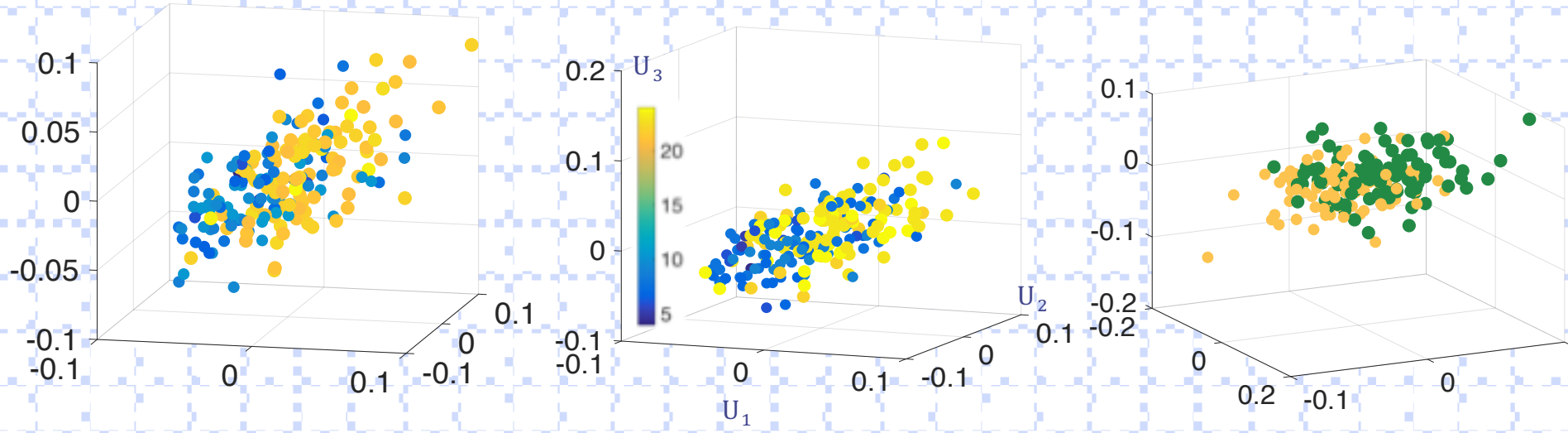
- **Statistical models of fiber curves connecting brain regions**

  With Maxime Descoteaux and David Dunson

- Parcellation of brain cortical surface based on fiber geometry

  With David Dunson

Tensor network factorizations: Relationships between brain structural connectomes and traits



Dataset 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.

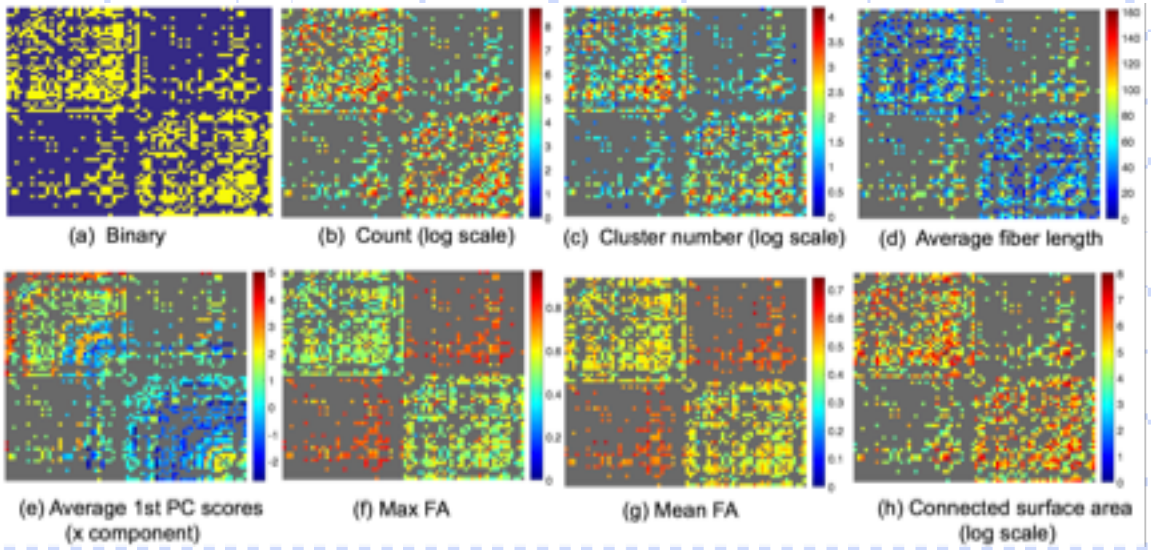
We extracted 175 different trait measures for each subject

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 dimension $v \times v \times m \times n$

- Each tensor is semi-symmetric because of the symmetry of connection.

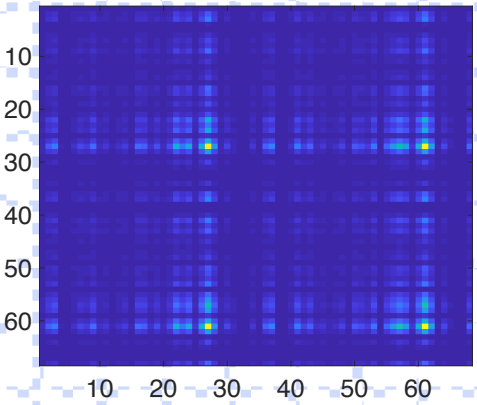
Tensor Principle Component Analysis

- Semi-symmetric tensor decomposition for three way-tensor (or higher):

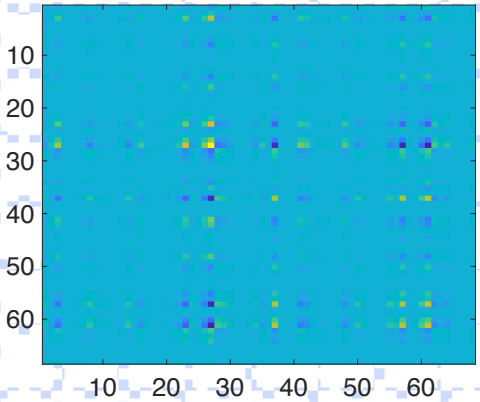
$$\boldsymbol{\mathcal{X}} \approx \sum_{k=1}^{K} d_k \, \mathbf{v}_k \circ \mathbf{v}_k \circ \mathbf{u}_k,$$

$\mathcal{X} \in \mathcal{R}^{v \times v \times n}$ - ***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



$v_1 \circ v_1$



$v_2 \circ v_2$

...

Tensor Principle Component Analysis

- We solve the decomposition through the following optimization:

$$\underset{d_k,\mathbf{v}_k,\mathbf{u}_k}{\text{minimize}} \| \boldsymbol{\mathcal{X}} - \sum_{k=1}^{K} d_k \mathbf{v}_k \circ \mathbf{v}_k \circ \mathbf{u}_k \|_2^2$$

$$\text{subject to } \mathbf{u}_k^T \mathbf{u}_k = 1, \mathbf{v}_k^T \mathbf{v}_k = 1, \mathbf{v}_k^T \mathbf{v}_j = 0 \; \forall j < k.$$

- It is non-convex but is instead bi-convex in **v** and **u**
- We utilize a block coordinate descent method
- Because of the additional orthogonality constraint, we use a greedy one-at-a-time strategy that sequentially solves a rank-one problem

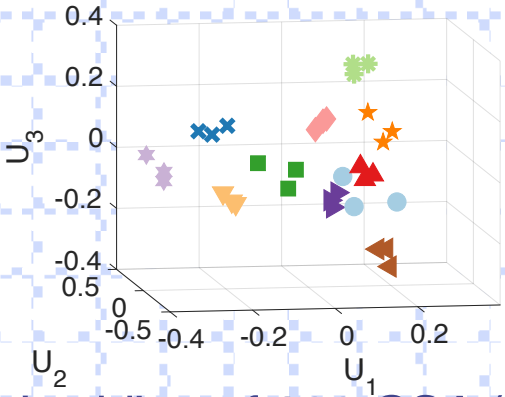
$$\underset{\mathbf{u}_k,\mathbf{v}_k}{\text{maximize}} \quad \boldsymbol{\mathcal{X}} \times_1 (\mathbf{P}_{k-1} \mathbf{v}_k) \times_2 (\mathbf{P}_{k-1} \mathbf{v}_k) \times_3 \mathbf{u}_k$$

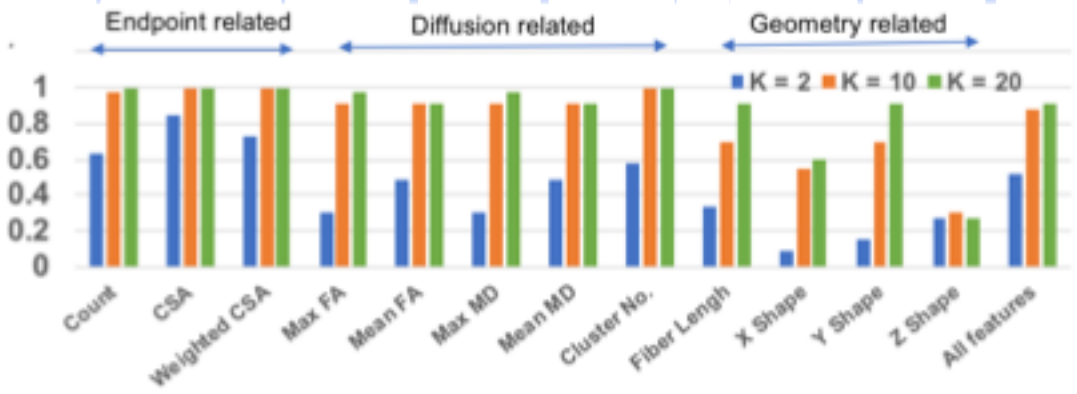
$$\text{subject to} \quad \mathbf{u}_k^T \mathbf{u}_k = 1, \; \mathbf{v}_k^T \mathbf{v}_k = 1,$$

where $\mathbf{P}_{k-1} = \mathbf{I} - \mathbf{V}_{k-1} \mathbf{V}_{k-1}^T$ with $\mathbf{V}_{k-1} = [\mathbf{v}_1, \ldots \mathbf{v}_{k-1}]$

Exploratory Analysis

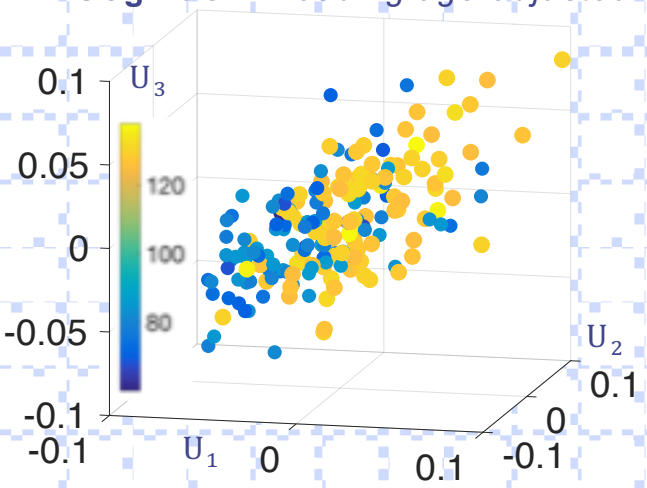
- Embedding and nearest neighbor classification on Sherbrooke test-retest data:



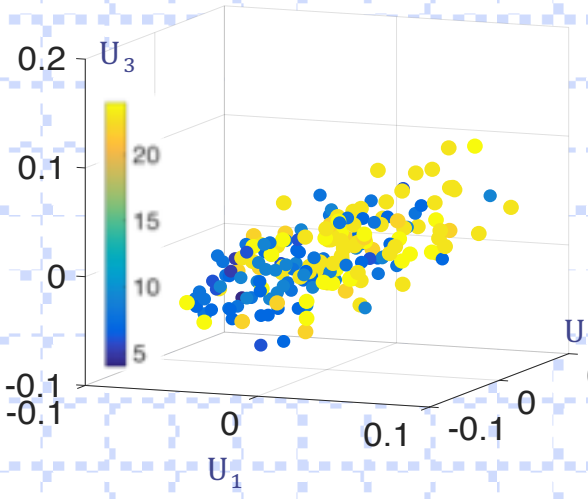


- Embedding of 200 CSA (connected surface area) networks in the HCP dataset (100 subjects with high scores, 100 with low scores):

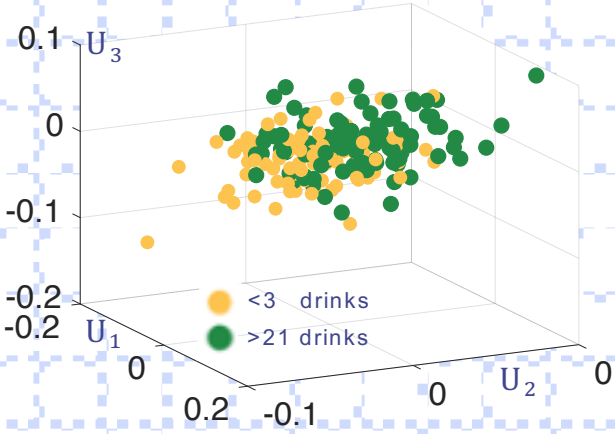
Cognition: *Reading age-adjusted*



Cognition: *Fluid Intelligence*

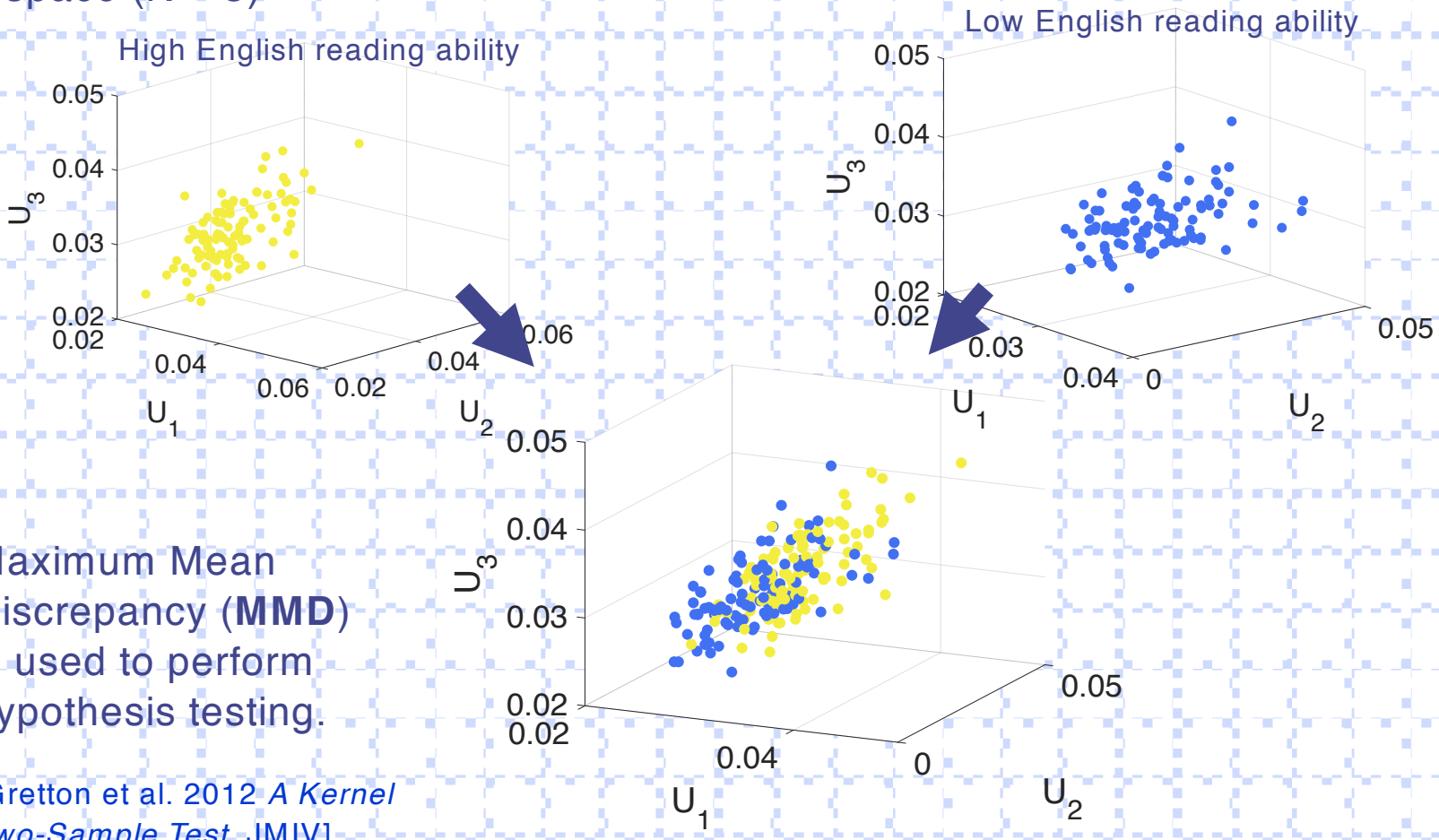


Substance use:
Max drinks in single day



Connectomes vs. Traits

- Hypothesis testing - whether connectomes are associated with traits
  - For each weighted connectivity matrix – embed to a low dimensional vector space (K = 3)

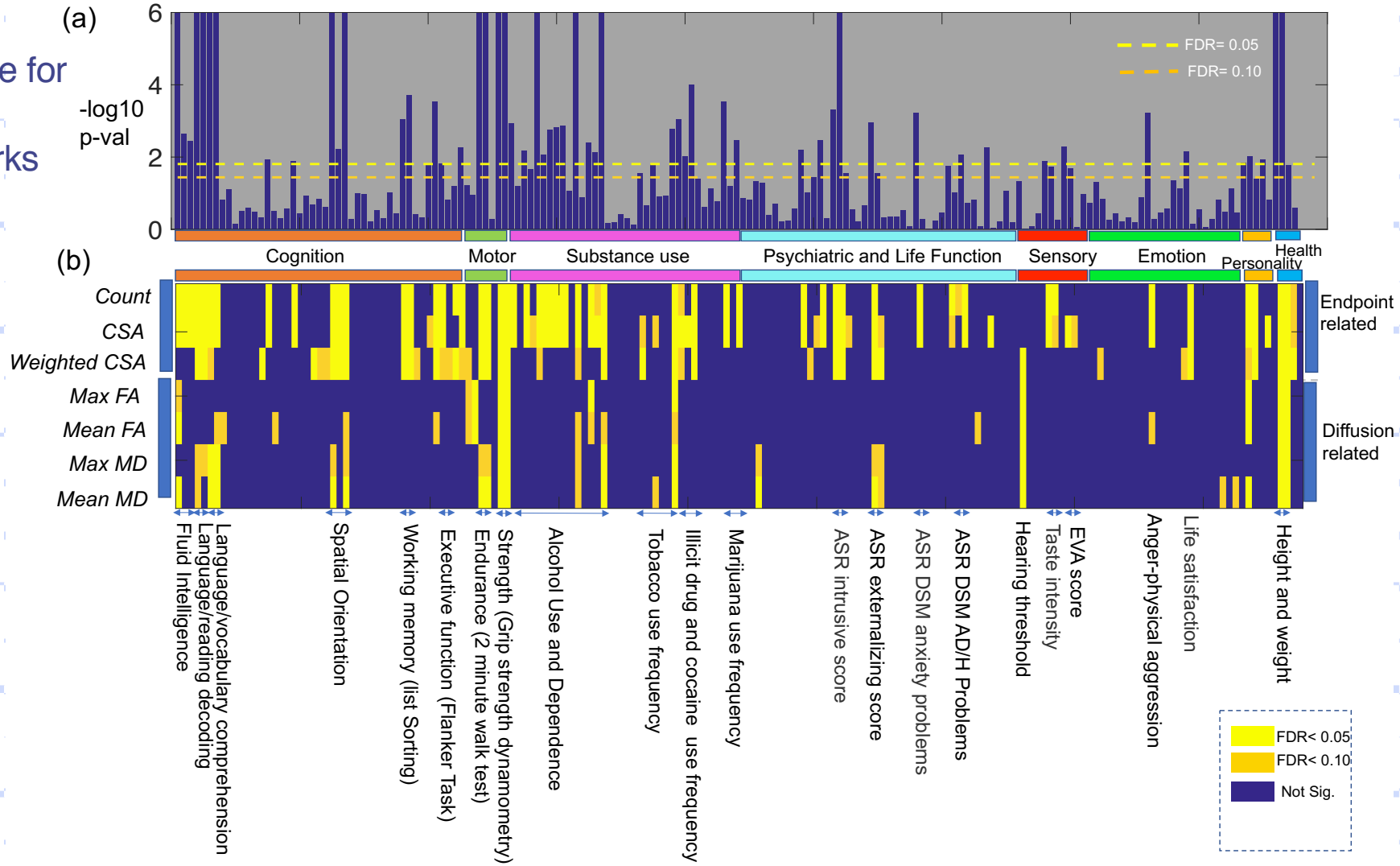


- Maximum Mean Discrepancy (**MMD**) is used to perform hypothesis testing.

[Gretton et al. 2012 *A Kernel Two-Sample Test*, JMIV]

Connectomes vs. Traits

- Hypothesis testing - whether connectomes are associated with traits (K=30)



Connectomes vs. Traits

- Prediction - whether connectomes can predict traits?
  - Baseline model: $\hat{y}_i^b \sim f([\text{age, gender}])$
  - Comparison: $\hat{y}_i \sim f([\text{age, gender, connectomes}])$

Various trait scores

Various machine learning methods

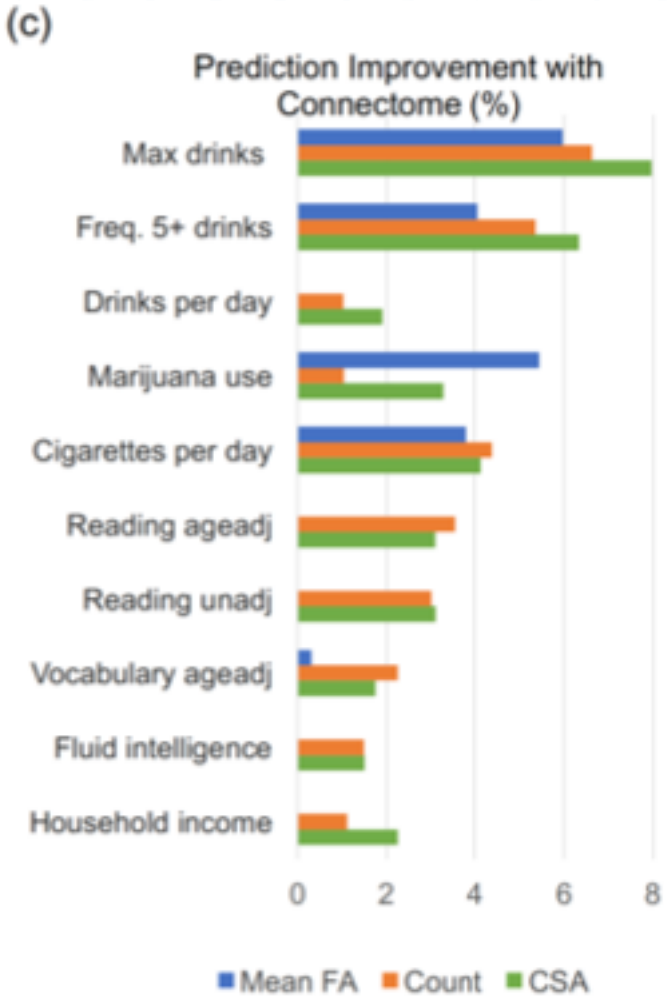
- Evaluate the prediction **improvement** with structural connectomes.
  - The root-mean-square error (RMSE) is used to evaluate the prediction accuracy
  - Prediction improvement ratio:

(RMSE_*baseline* - RMSE_*connectomes)* / RMSE_*baseline*

- Various machine learning methods are used to fit the data, the best model is selected based the validation dataset.
  - 2/3 for training (> 690 subjects), 1/3 for validation (>330), 1/3 for testing (>330)

Connectomes vs. Traits

- Prediction results (top 10 traits that can be predicted better by structural connectomes):



- 5 of them are related to substance use
- 5 of them are related to cognition

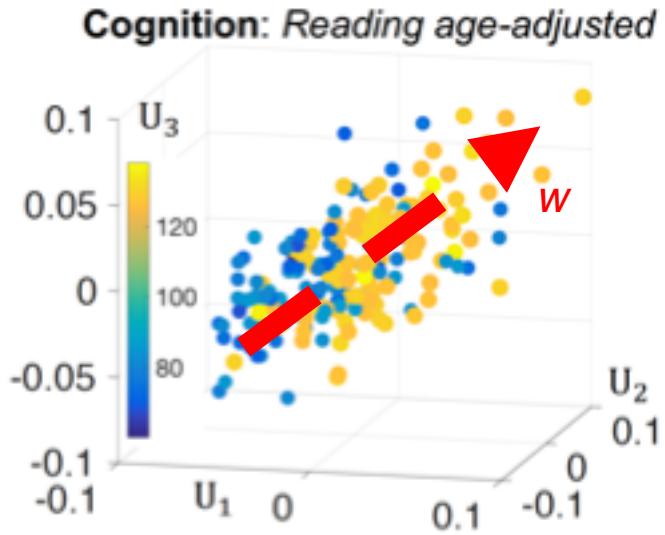
Connectomes vs. Traits

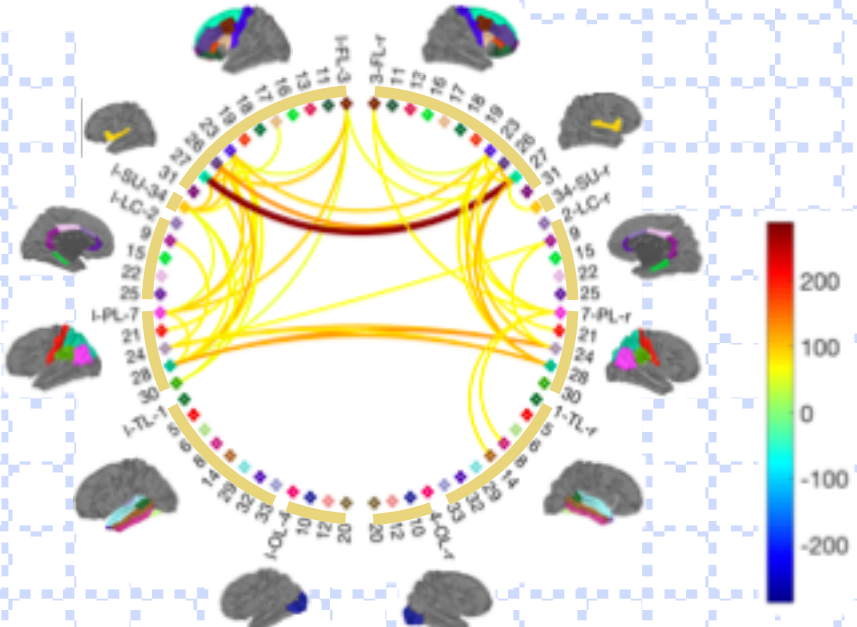
- For a particular weighted network (e.g. CSA), how does the network change with increasing of a trait?
  - Find a unit direction *w* in $\mathbb{R}^K$ such that correlation between the trait scores $\{y_i\}$ and and projection $\{u_{proj}(i) = \mathbf{U}_K(i,:)\mathbf{w}\}$

$$\text{argmax}_{\mathbf{w}\in\mathbb{R}^K}\text{COV}(y, u_{proj}) = \text{argmax}_{\mathbf{w}\in\mathbb{R}^K}\frac{1}{N}\mathbf{w}^T\mathbf{U}_K^T\mathbf{Y} \qquad \text{s.t. } \mathbf{w}^T * \mathbf{w} = 1.$$

  - Bring *w* back to the network representation

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

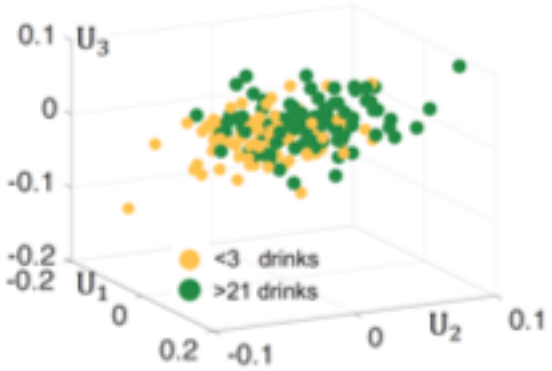


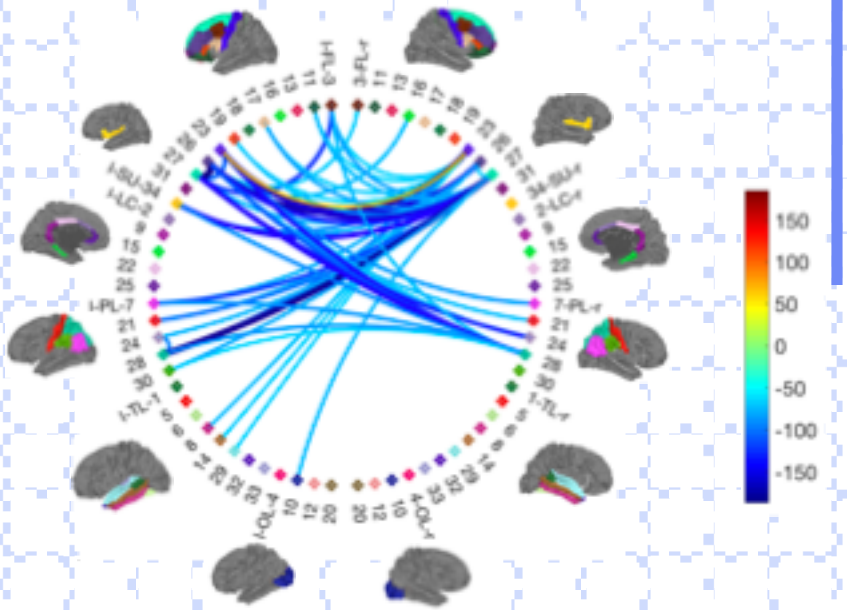


Connectomes vs. Traits

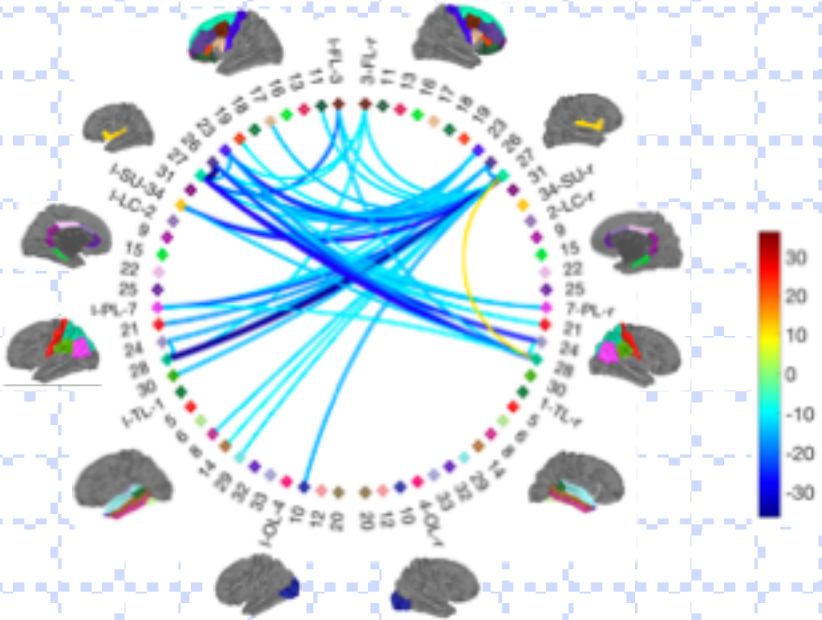
- More results

Max drinks in single day (classification rate = 80.99%)



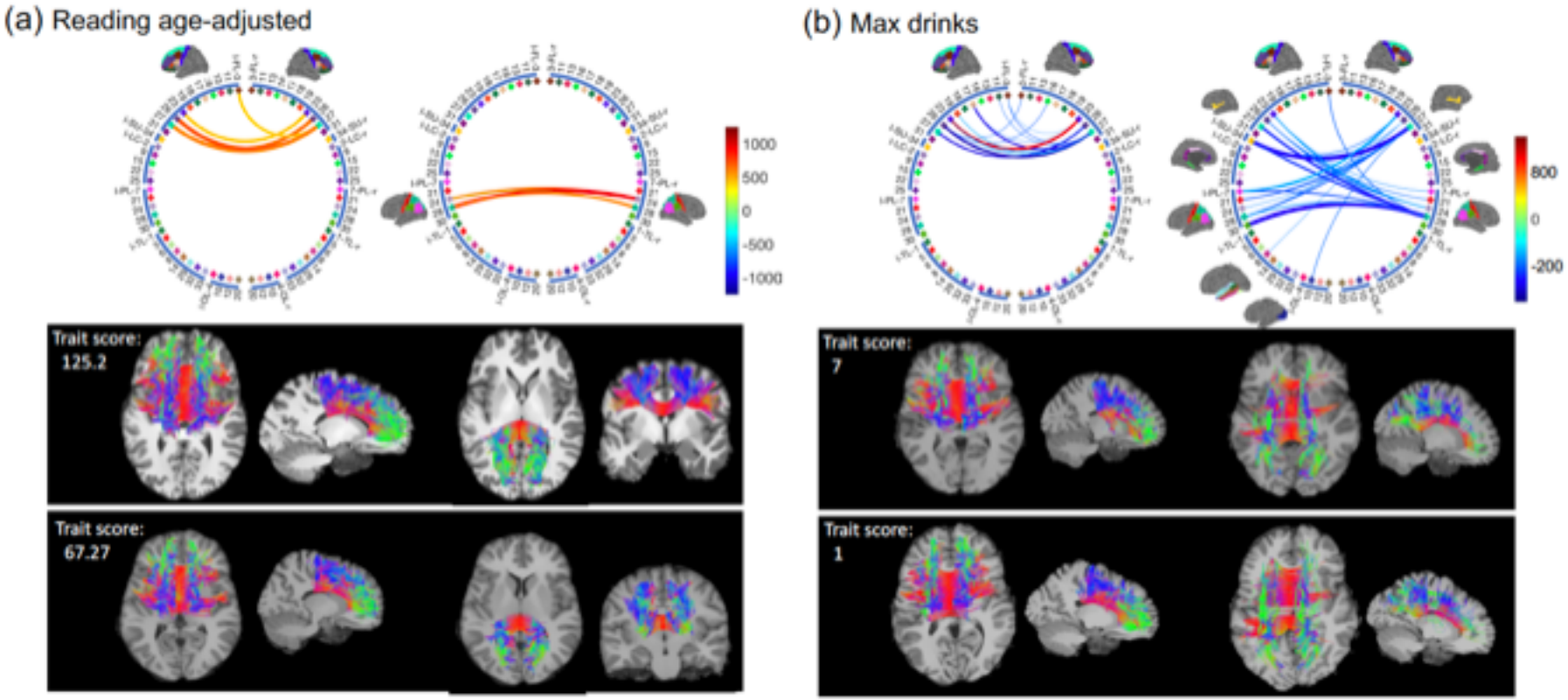


Times used Marijuana (classification rate = 59.68%)

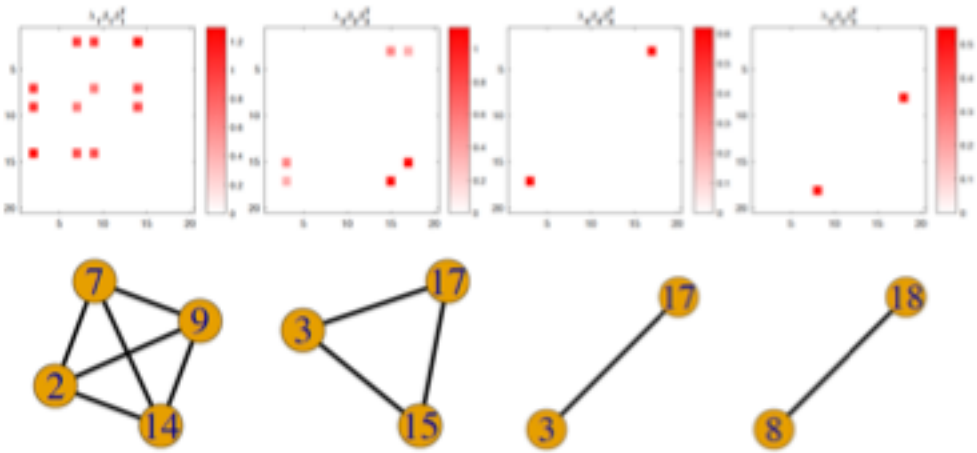


Corresponding WM Tracts

- We display the corresponding tracts using selected subjects



Symmetric Bilinear Regression for Signal Subgraph Estimation



Supervised Subnetwork Identification

- To identify subnetworks that related to traits, a supervised one-step method might work better (than the unsupervised tensor decomposition)

- Individuals over- or under-expressing a subnetwork have higher or lower values of trait y_i on average

- We start with a *Symmetric Bilinear Regression* (SBR):

$$E(y_i \mid X_i) = \alpha + \langle \theta, X_i \rangle ,$$

where $\langle \theta, X \rangle = \text{trace}(\theta^\top X) = \text{vec}(\theta)^\top \text{vec}(X)$

  - $X_i \in \mathcal{R}^{v \times v}$ network for *i*-th subject

  - Large p, small n problem (# of parameters to estimate:1 + v(v-1)/2; e.g., V = 68 → p = 2279, n = 1000)

Optimization Problem

- Suppose θ admits a rank-K CP decomposition with sparsity penalty on $\{\lambda_h \boldsymbol{\beta}_h \boldsymbol{\beta}_h^\top\}_{h=1}^K$, we have $\theta = \sum_{h=1}^{K} \lambda_h \boldsymbol{\beta}_h \boldsymbol{\beta}_h^\top$

$$E(y_i \mid W_i) = \alpha + \left\langle \sum_{h=1}^{K} \lambda_h \boldsymbol{\beta}_h \boldsymbol{\beta}_h^\top, X_i \right\rangle = \alpha + \sum_{h=1}^{K} \lambda_h \boldsymbol{\beta}_h^\top X_i \boldsymbol{\beta}_h$$

- Reduce parameters from 1 + v*(v-1)/2 to 1 + v + Kv
- Maintain flexibility: if set K = v(v-1)/2 and $\{\boldsymbol{\beta}_h\}_{h=1}^K = \{\boldsymbol{e}_u + \boldsymbol{e}_v\}_{u<v}$, the problem becomes unstructured linear model
- Interpretation: nonzero entries in each $\lambda_h \boldsymbol{\beta}_h \boldsymbol{\beta}_h^\top$ identify a clique subgraph

Optimization Problem

- Suppose θ admits a rank-K CP decomposition with sparsity penalty on $\{\lambda_h \boldsymbol{\beta}_h \boldsymbol{\beta}_h^\top\}_{h=1}^K$, we have

$$\theta = \sum_{h=1}^{K} \lambda_h \boldsymbol{\beta}_h \boldsymbol{\beta}_h^\top$$

$$E(y_i \mid W_i) = \alpha + \left\langle \sum_{h=1}^{K} \lambda_h \boldsymbol{\beta}_h \boldsymbol{\beta}_h^\top, X_i \right\rangle = \alpha + \sum_{h=1}^{K} \lambda_h \boldsymbol{\beta}_h^\top X_i \boldsymbol{\beta}_h$$

- Our objective function now becomes:

$$\frac{1}{2n} \sum_{i=1}^{n} \left(y_i - \alpha - \sum_{h=1}^{K} \lambda_h \boldsymbol{\beta}_h^\top X_i \boldsymbol{\beta}_h \right)^2 + \gamma \sum_{h=1}^{K} |\lambda_h| \sum_{u=1}^{R} \sum_{v<u} |\beta_{hu} \beta_{hv}|$$

  - Avoid scaling problems between λ_h and $\boldsymbol{\beta}_h$ compared to simply penalizing $\sum_{h=1}^{K} \|\boldsymbol{\beta}_h\|_1$ → sufficient to identify each matrix $\lambda_h \boldsymbol{\beta}_h \boldsymbol{\beta}_h^\top$
  - Efficient coordinate descent algorithm (Friedman et al. 2010) can be derived having analytic updates & with active set speed up
  - Can choose K as an upper bound & zero out unnecessary components

Simulation

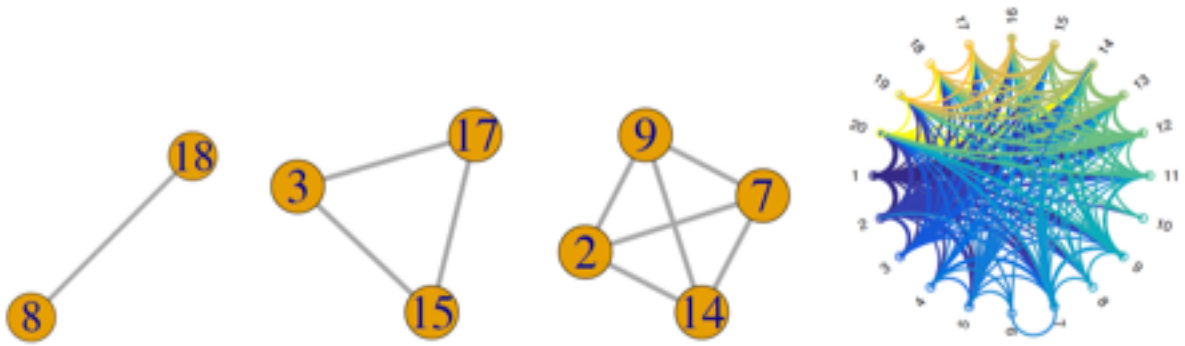
- Considered a variety of data generating processes for

$$(X_i, y_i), i = 1, \ldots, n.$$

- X_i is generated via individual-specific weights on common subnetworks + Gaussian noise

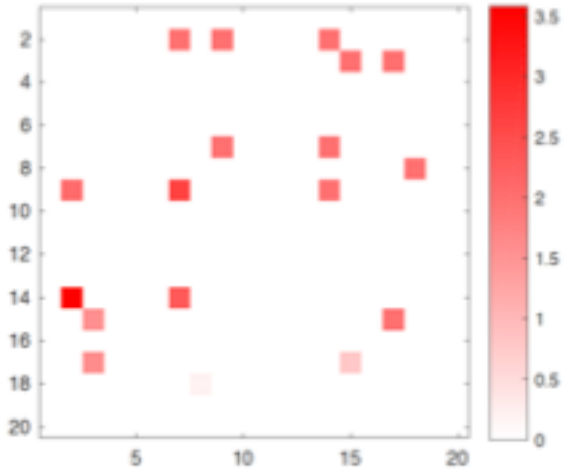
- A subset of these subnetworks are related to the response y_i

- Different signal-to-noise scenarios + compared with LASSO Tensor PCA

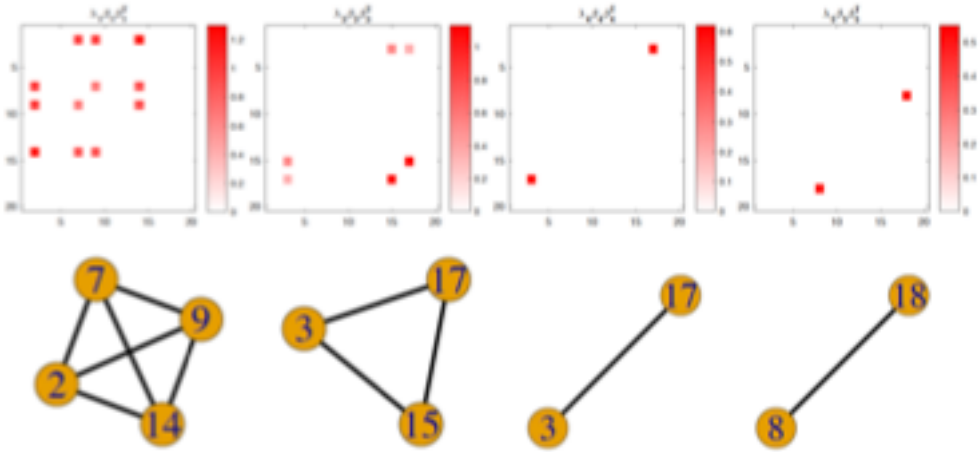


Low Noise

Coefficients of LASSO



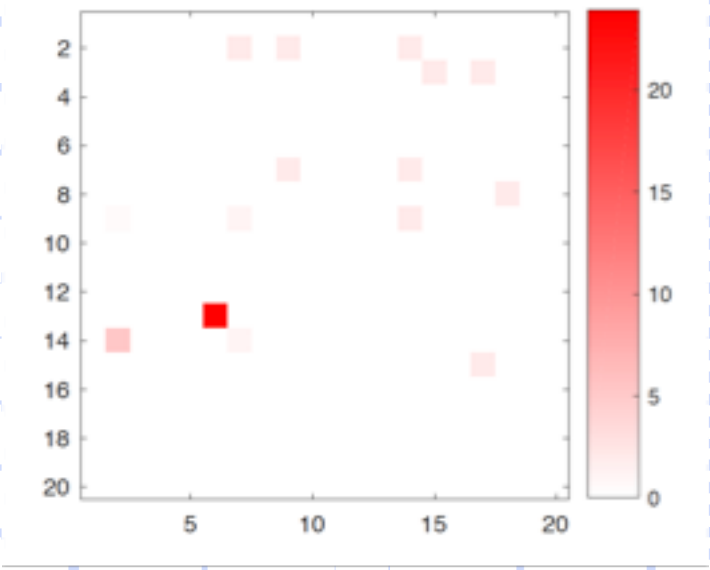
Coefficients and selected subgraphs of SBL



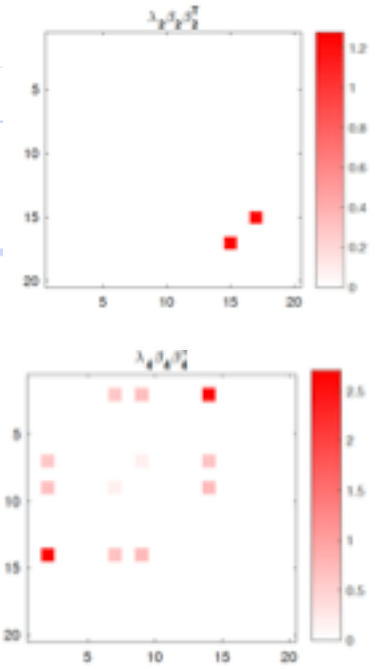
	MSE	TPR	FPR
lasso	10.98±4.40	0.837±0.138	**0.002±0.005**
TN-PCA	**10.04±4.66**	0.449±0.499	0.449±0.499
SBL	10.08±4.51	**0.848±0.169**	0.005±0.007

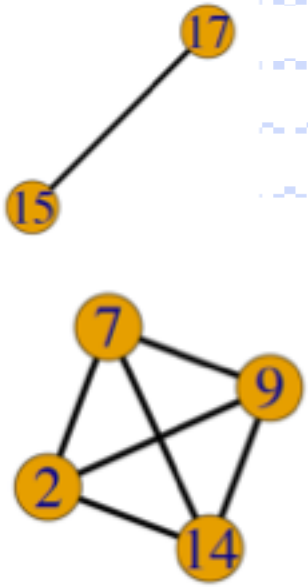
High Noise

Coefficients of LASSO



Coefficients and selected subgraphs of SBL



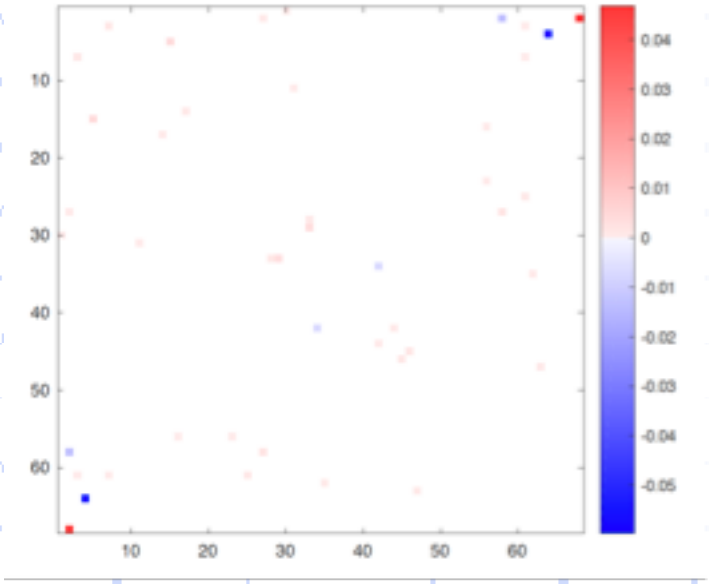


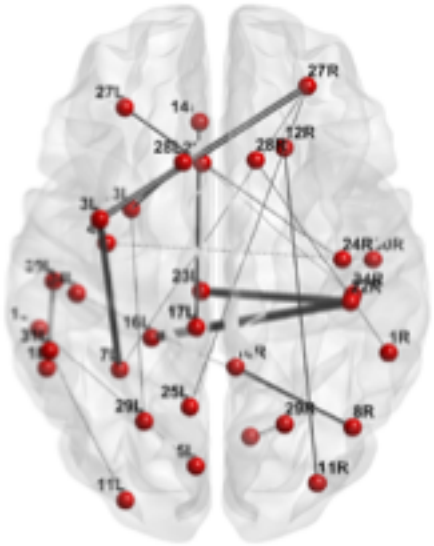
	MSE	TPR	FPR
lasso	448.3±195.3	0.445±0.141	**0.025±0.037**
TN-PCA	624.0±287.8	0.060±0.239	0.060±0.238
SBL	**393.7±159.2**	**0.539±0.210**	0.029±0.038

Real Data Analysis

- Age-adjusted picture vocabulary (PV) score from 1065 HCP subjects
  - presented with an audio recording of a word and 4 images
  - select the picture that most closely matches the word
- Weighted brain network of between counts among 68 regions were used; 565 subject for training and 500 for testing.

- Estimated coefficients from LASSO

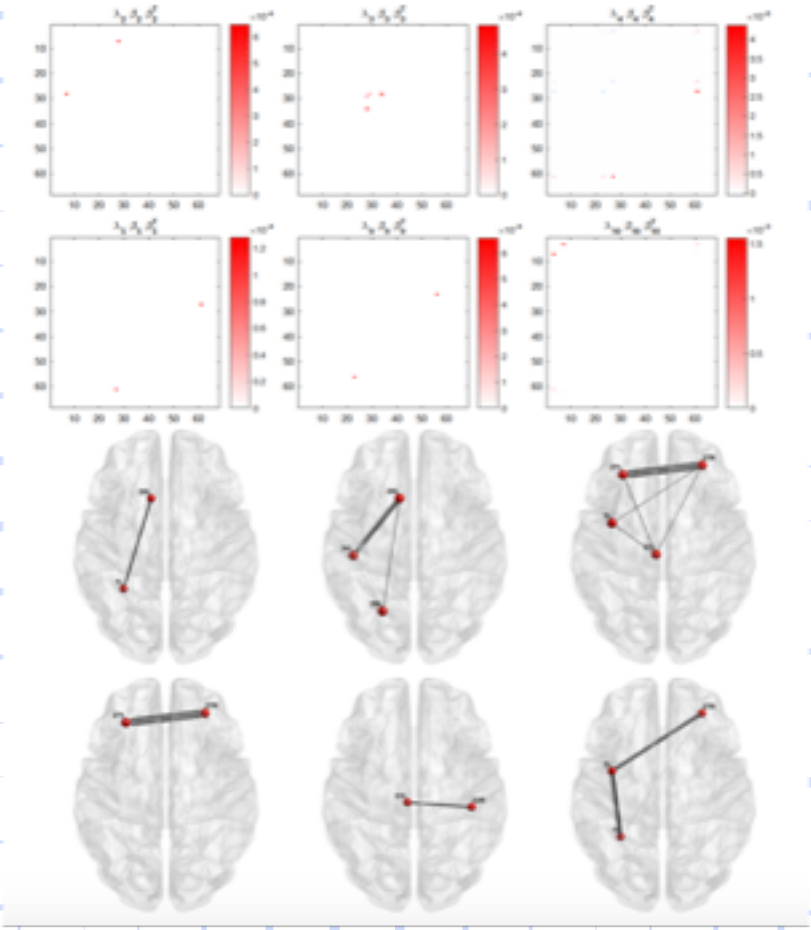




Real Data Analysis

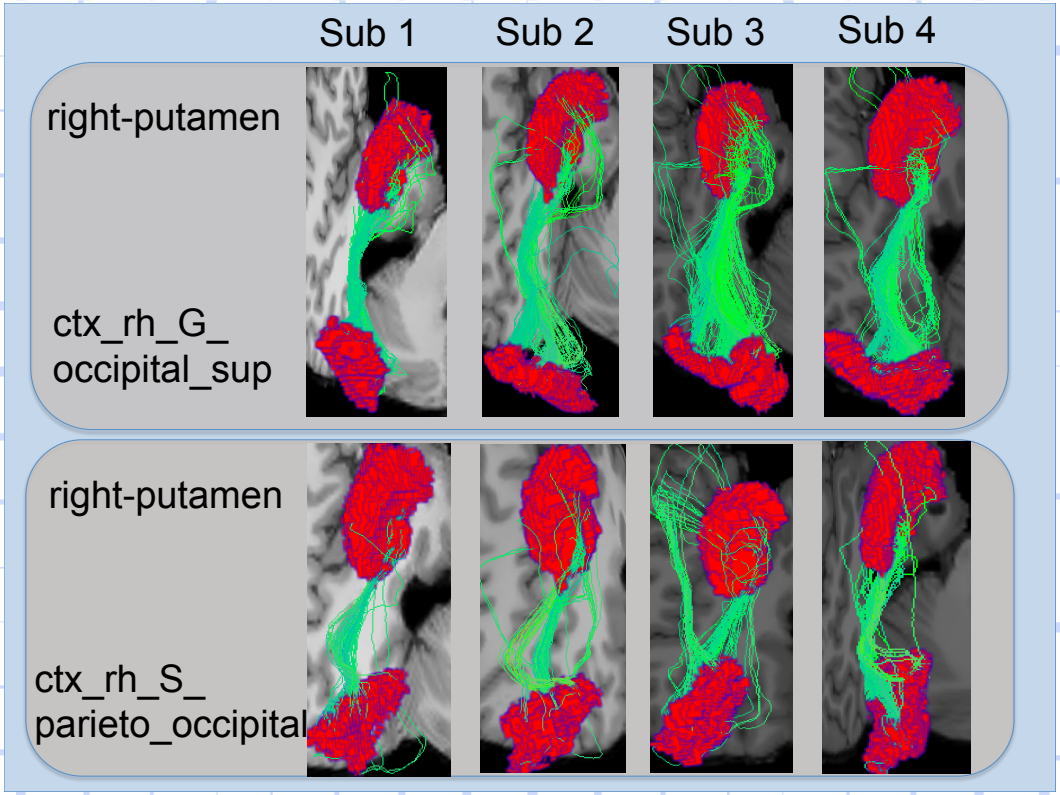
- Results from SBL

6 nonempty coefficient components out of $\{\lambda_h \boldsymbol{\beta}_h \boldsymbol{\beta}_h^\top\}_{h=1}^{10}$



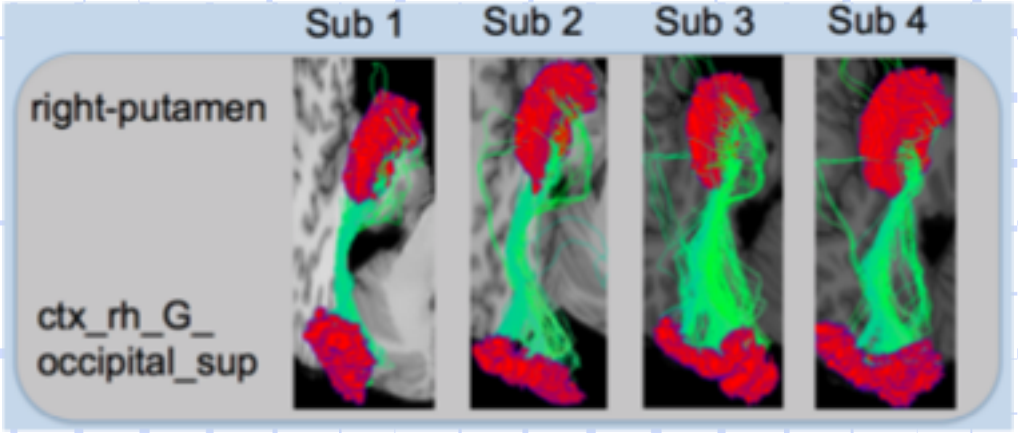
27L, 27R (left and right superior frontal gyrus), 7L (left inferior parietal gyrus) and 29L (left superior temporal gyrus) are among activated regions when shifting from listening to meaningless pseudo sentences to listening to meaningful sentences (Saur et al., 2008; Dronkers, 2011).

Statistical models of fiber curves connecting brain regions



Ongoing Studies

- Fiber curves connecting two brain regions contain rich information
  - Functional data
  - Clear clustering pattern
  - Heterogeneity
  - Big data



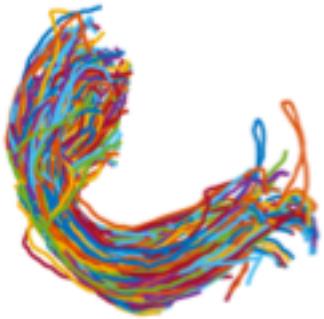
- We are interested to:
  - utilize the geometric information to model the brain connectome

- However, the challenges are:
  - The complexity of the data form: $y_i : [0,1] \to \mathbb{R}^3$
  - Big data issue: hundreds ~ thousands of fibers connecting two regions
  - Miss alignment issue: different subjects have different coordinate system

Z. Zhang, M. Descoteaux, D. Dunson
Nonparametric Bayes Models of Fiber Curves Connecting Brain Regions,
Revision at JASA 2017+

Variation Decomposition

- To more efficiently represent fibers in connection (r_a, r_b), we perform a variation decomposition w.r.t. a template fiber:
  - Shapes
  - Rotations
  - Translations
  - scaling
  - re-parameterizations



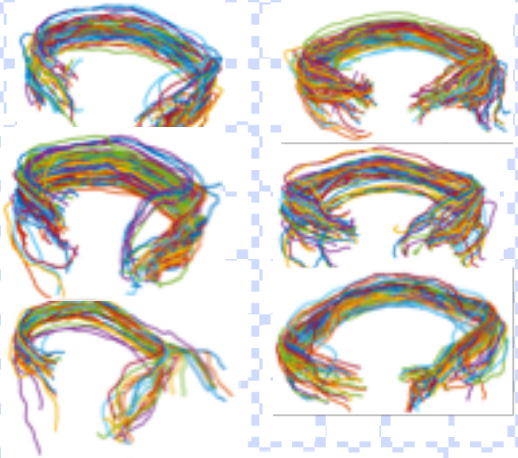
Elastic shape analysis [Srivastava et al. 2011]

Shapes:

Rotations: $\in SO(3)$

Translations: $\in \mathbb{R}^3$

- How to represent the shape part?



Connection (r_a, r_b) of different subjects

Alignment

Learn a low dimensional structure

Mean Basis functions

$$L_{(r_a,r_b)} = \{y_l, \{\phi_l, l = 1, ..., T\}\}$$

(shared by all subjects)

Variation Decomposition

- Any shape of streamline in (r_a, r_b) can be represented as:

$$g(s) \approx y_\mu(s) + \sum_{l=1}^{T} x_l \phi_l(s)$$

Coefficient

Basis function

- A streamline is represented by components: shape + translation + rotation

$$y_i := \{c_i^{(1)}, c_i^{(2)}, c_i^{(3)}\}$$

Shape: $c_i^{(1)} \in \mathbb{R}^T$ Translation: $c_i^{(2)} \in \mathbb{R}^3$ Rotation: $c_i^{(3)} \in SO(3)$

- Recovery of a streamline:

$$y_i = (c_i^{(3)})^T * \left(y_\mu + \sum_{l=1}^{T} c_i^{(1)}(l)\phi_l \right) + c_i^{(2)}$$

Variation Decomposition

- Any shape of streamline in (r_a, r_b) can be represented as:

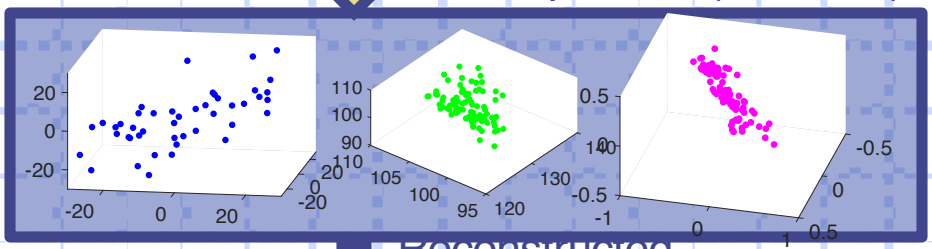
$$g(s) \approx y_\mu(s) + \sum_{l=1}^{T} x_l \phi_l(s)$$

Coefficient

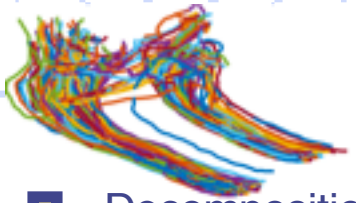
Basis function

- A streamline is represented by components: shape + translation + rotation

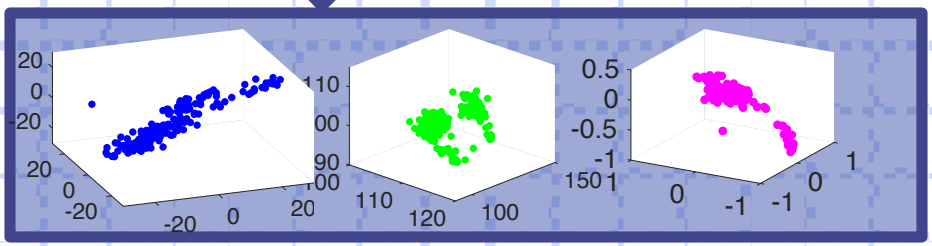
Decomposition (3 + 3 + 3)



Reconstructed



Decomposition (3 + 3 + 3)



Reconstructed

Model for One Individual

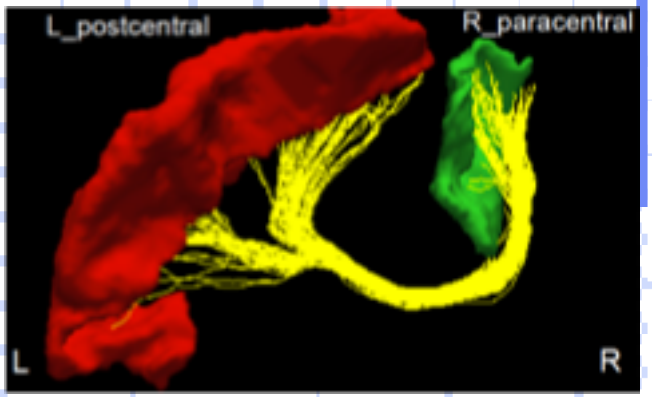
- We try to model fiber curves from a single subject in connection (r_a, r_b)
- Each component $c^{(m)}$ of a fiber has a Euclidean or manifold support
- We use a product kernel mixture model to characterize the m components of fibers in a connection

$$f(y_i) = \int_{\Theta} \prod_{m=1}^{M} \mathcal{K}_m(c_i^{(m)}; \theta^{(m)}) dP(\theta), \quad \theta = \{\theta^{(1)}, ..., \theta^{(M)}\}$$

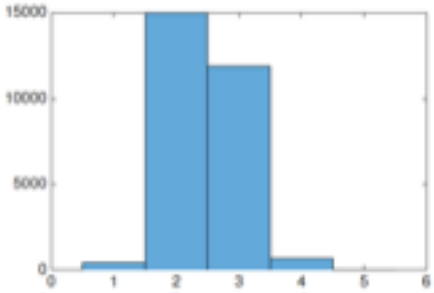
  - $c^{(m)}$ has a support of $\mathcal{Y}_m$
  - $\mathcal{K}_m(\cdot; \theta^{(m)})$ is a parametric probability measure on $\{\mathcal{Y}_m, \mathcal{B}(\mathcal{Y}_m)\}$
  - P is a parametric probability measure over $\{\Theta_m, \mathcal{B}(\Theta_m)\}$
  - A nonparametric approach realized by choosing P as a random probability measure and assigning an appropriate prior

Experimental Results

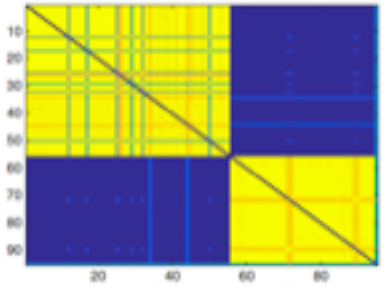
- Consider the connection between *right paracentral lobule (r_pl)* and *left postcentral gyrus (l_pg)* in HCP one subject (with 95 fibers)
- We use the defined mixture models to cluster fibers based on **each** component / **all** components together



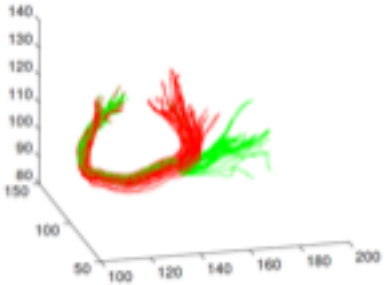
Posterior on K:



Pairwise prob. Matrix:



Final voting result:



All components:

Comparison with manual clustering results:

RI – rand index

ARI – adjust rand index

	(*r_pl*, *l_pg*)			
	Shape	Trans.	Rot.	All
RI	0.9265	0.6656	0.8694	1.0
ARI	0.8533	0.3393	0.7391	1.0

Model for a Set of Individuals

- We model fiber curves from a set of subjects in connection (r_a, r_b)
- Our goal is to:

  (1). model connections across different subjects
  (2). cluster subjects and cluster fibers within each subject

- Miss alignment between subjects is a challenge

- We apply a nested Dirichlet Process (NDP) ([Rodriguez et. al. 2008]) to model $\{y_{j\cdot}\}$

$$f_j(y_{ij}) = \int \prod_{m=1}^{M} \mathcal{K}_m(c_{ij}^{(m)}; \theta^m) dG_j(\theta), \theta = \{\theta^1, ..., \theta^M\},$$

where $G_j \sim \mathbf{NDP}(\alpha, \beta, P_0)$

  - NDP allows clustering fibers within each subject, and also produces clusters between subjects

- Posterior MCMC sampling can be easily developed

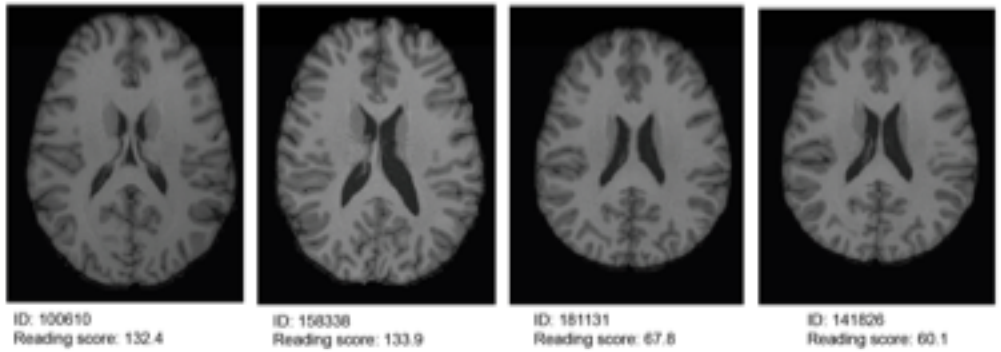
Geometry of Fiber Curves

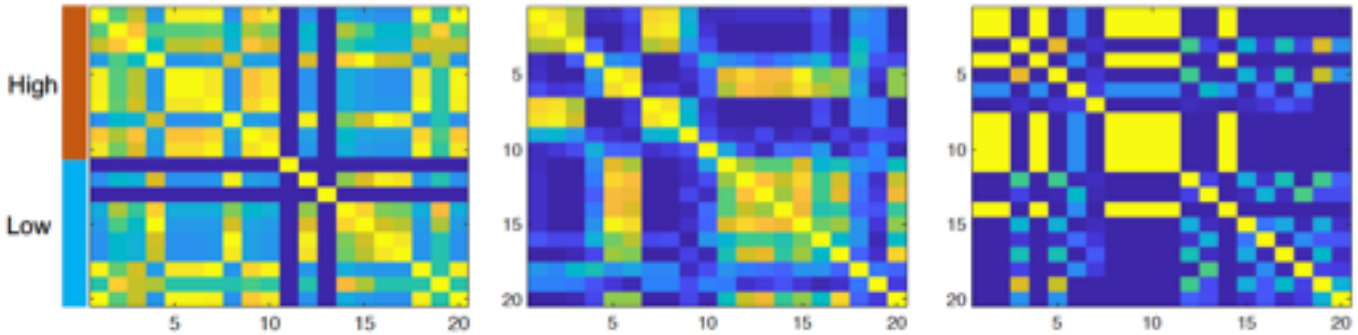
- Discriminative analysis: applied to the test-retest

- RI – Rand Index, calculates the ratio of agreement between the inferred and ground truth
- ARI – Adjusted Rand Index, the corrected-for-chance version of the Rand index

Geometry v.s. Traits

- Discriminative analysis: applied to the test-retest
- Can geometry infer cognitive difference? Seems yes…





(a) Connection (11,47), RI = 0.62, best RI with count

(b) Connection (24,58), RI = 0.64, best RI with shape

(c) Connection (11,47), RI = 0.55, based on shape

- Pairwise probability of clustering 20 HCP subjects with high and low reading scores

Outline

- Introduction to diffusion MRI
- Construction of geometric connectomes
- Geometric representations of connectomes
- Statistical analysis of connectomes
- Software demonstration

Software Development

- Preprocessing and connectome reconstruction are complicated processes
- Skills and knowledge from different disciplines are required - steep learning curves for beginners
- Our goal is to build a user-friendly and extendable software (platform) for people who do not want to know details of the preprocessing and reconstruction
  - Usage: input data folder + one command to run the PSC

```
nextflow run main.nf --subject 1848/ -with-singularity SCIL_Singularity.img
```

  - Utilize two open source software: Singularity + Nextflow

Singularity containers

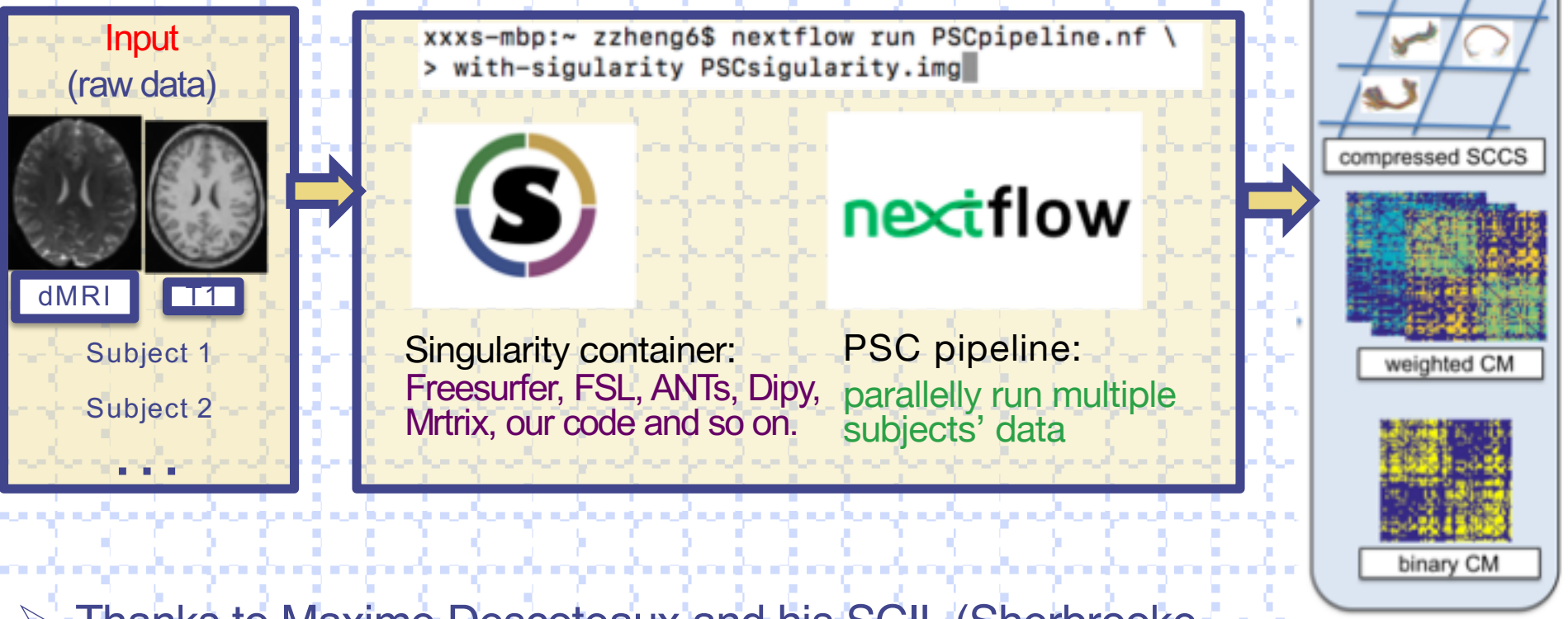
Pipeline

Software Development

- Singularity containers can pack the entire scientific workflows, software, libraries, and data
  - Enclosed all necessary software, no need additional software installation or tedious version control
  - OS independent
  - Easy to install
- Nextflow enables scalable and reproducible scientific workflows using software containers (e.g., singularity).
  - Compatible with Singularity container
  - Simplifies the implementation and the deployment of complex parallel and reactive workflows on clouds and clusters
  - Especially useful when there are many small steps in the workflow + some steps can be run parallelly

Software Development

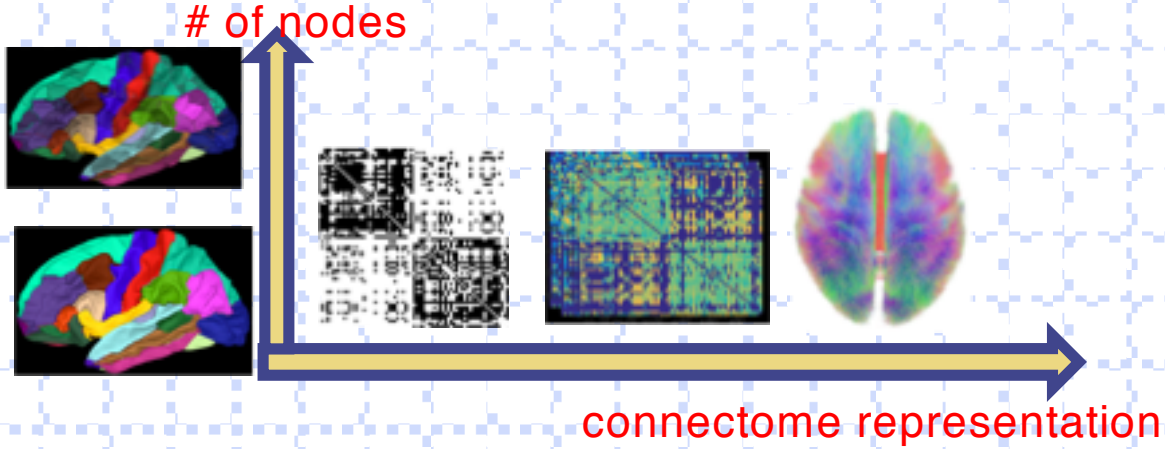
- Current version:



- Thanks to Maxime Descoteaux and his SCIL (Sherbrooke Connectivity Imaging Lab)
- Will be released soon in GitHub.

Summary & Discussion

- We are trying to incorporate more geometry elements in structural brain connectome analysis
- We have developed a **robust** structural connectome extraction framework
  - Reproducible
  - Invertible
  - Preserves the geometry and diffusion information



- New statistical methods for various connectome data:
  - To understand the normal connectome variation in healthy subjects
  - To relate connectome to covariates of interest and traits
- A lot of more interesting work can be done…

Thank You