2017 SIAM International Conference on DATA MINING

April 27 – 29, 2017

The Westin Galleria Houston
Houston, Texas, USA

Unified and Contrasting Graphical Lasso for Brain Network Discovery

Xinyue Liu*, Xiangnan Kong*, Ann B. Ragin#

*Worcester Polytechnic Institute

#Northwestern University

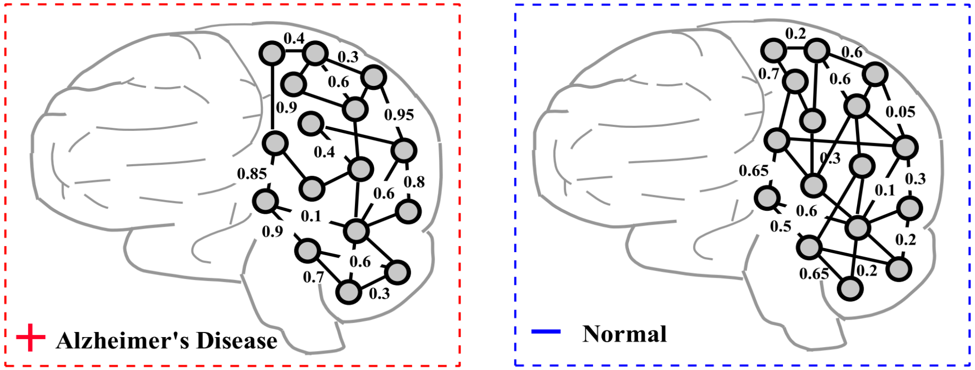
Network Representation of Brain

- ❑ Each node = a brain region
- ❑ Each edge = a functional connection

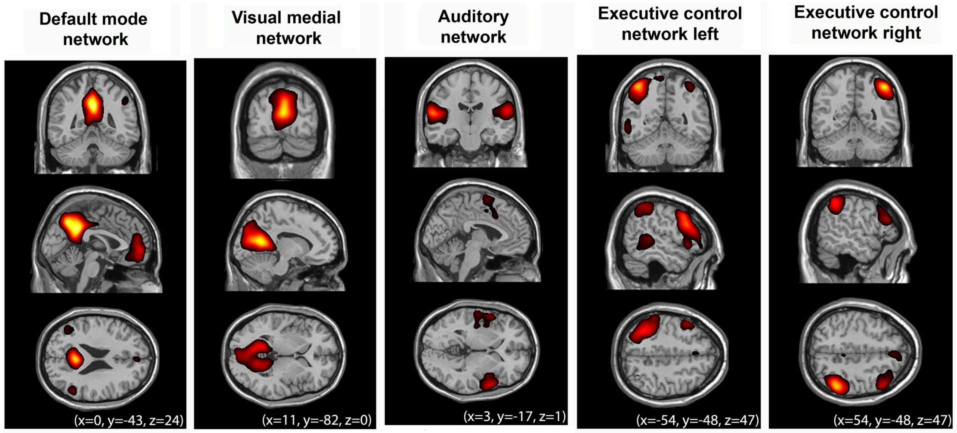
Use of Brain Network

Not Given!
Should be derived from neuroimaging data

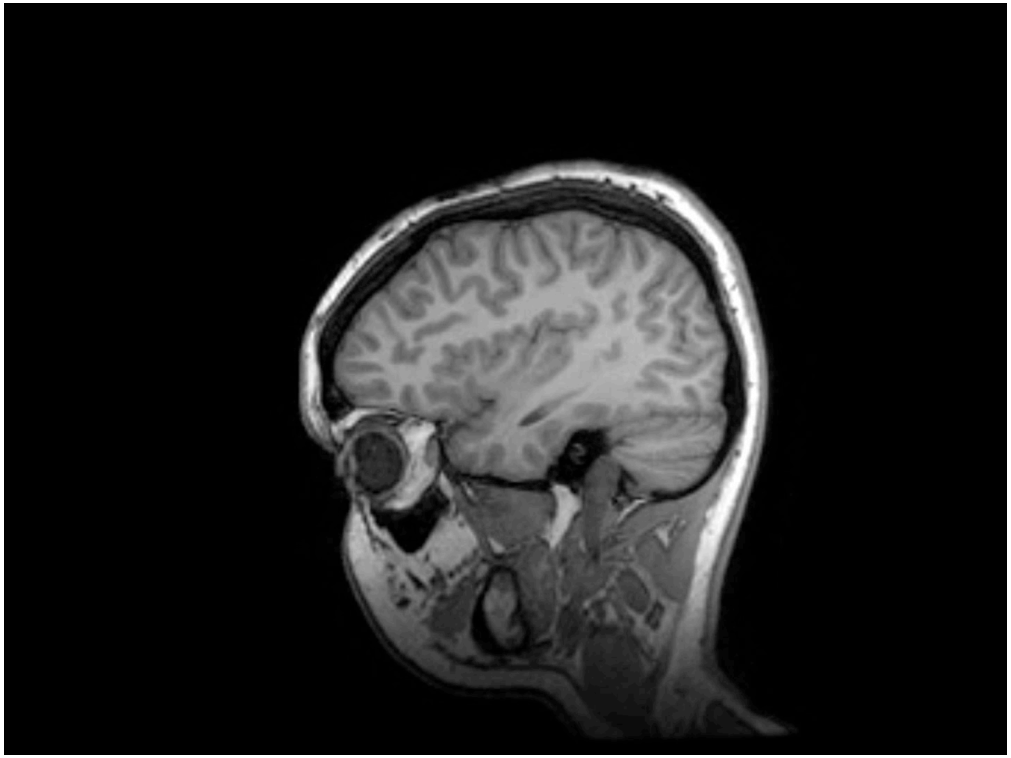
Disease Diagnosis



Functional Study

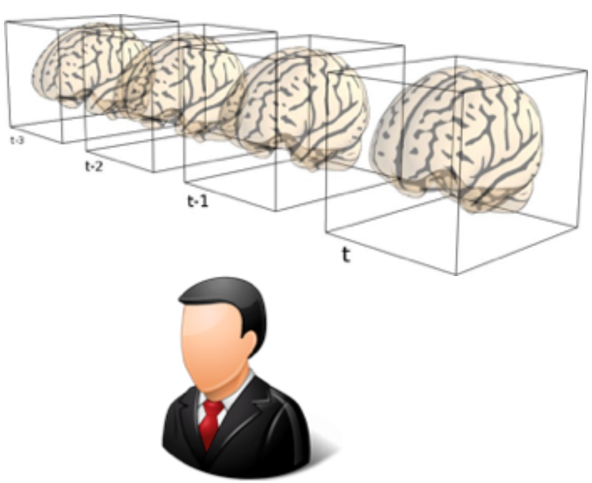


The fMRI data



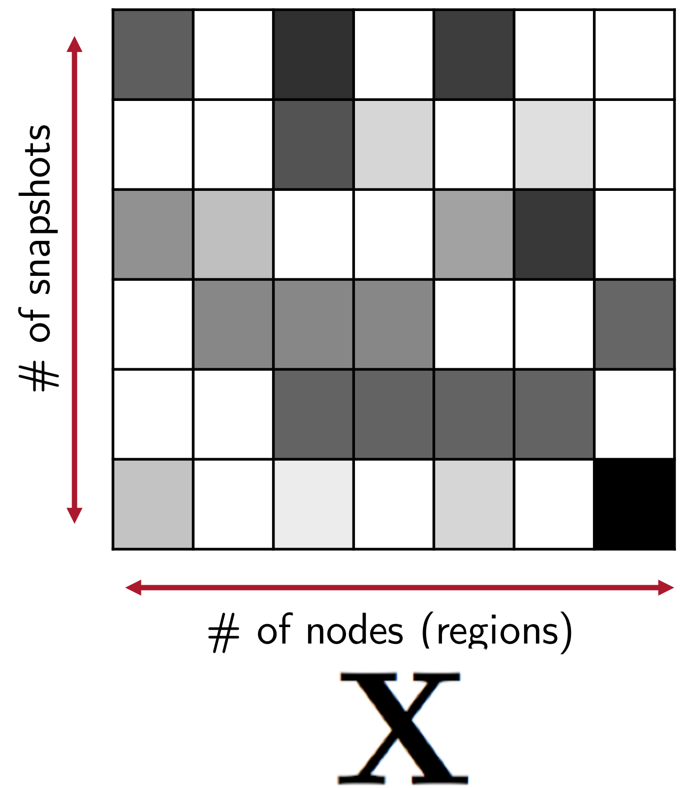
The fMRI Data

Each row refers to a snapshot of the brain!

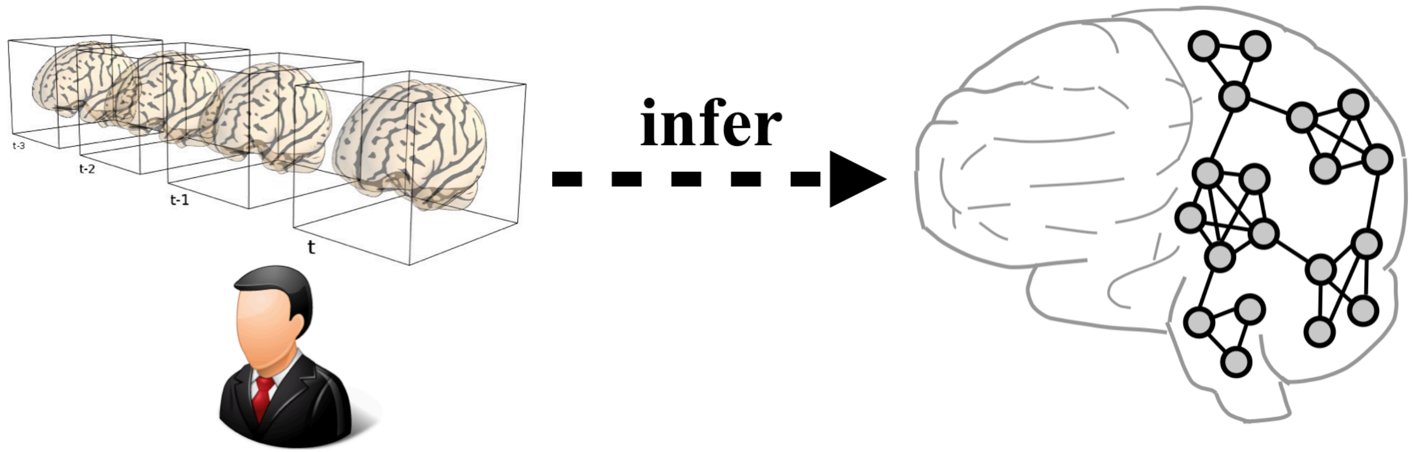


The Data Matrix
$\boldsymbol{X} \in \boldsymbol{R}^{m \times n}$

=



Brain Network Discovery

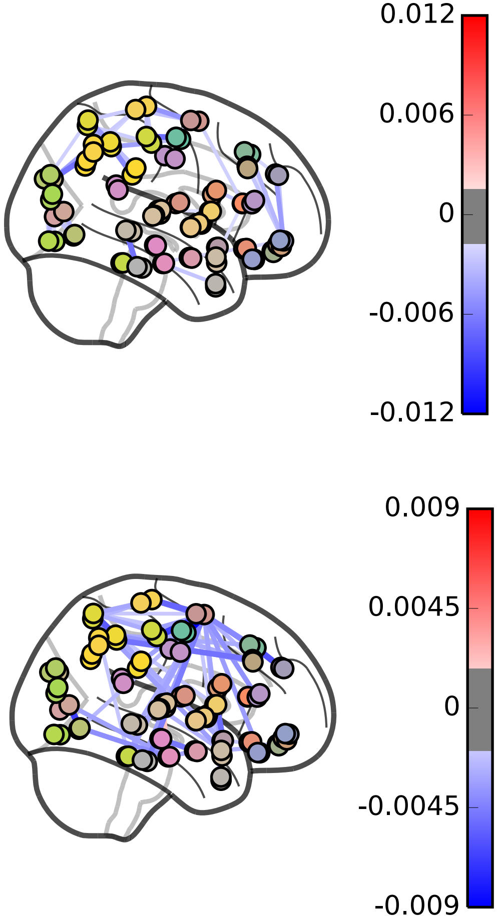


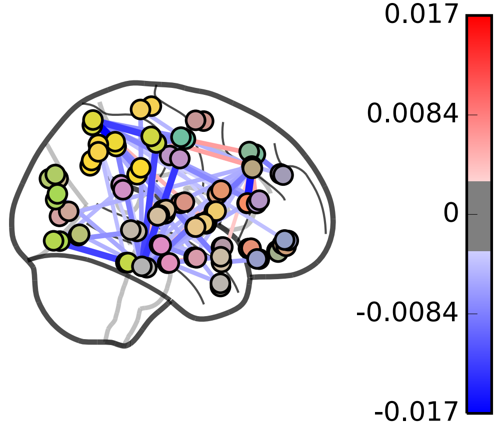
Inferring a network for a single subject
(treating multiple subjects as a single one
by concatenating the data)

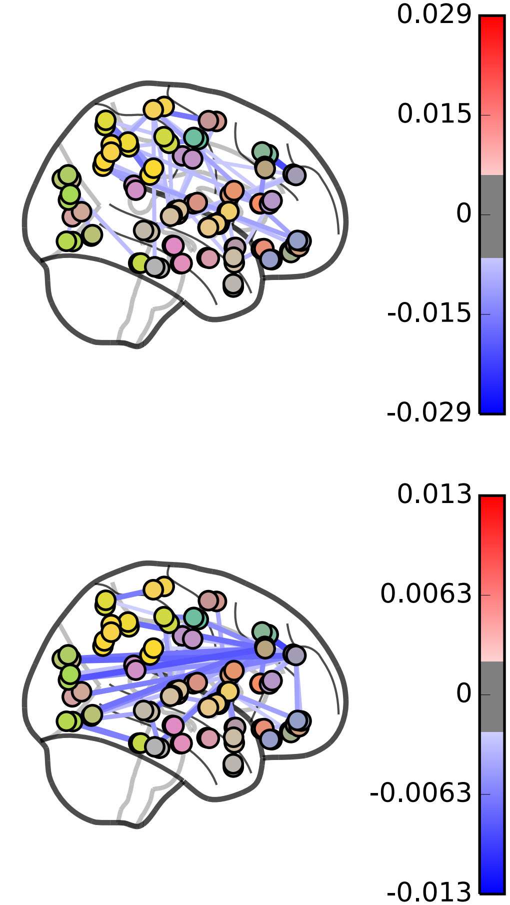
Usually use Graphical Lasso

Problem of Individual Network

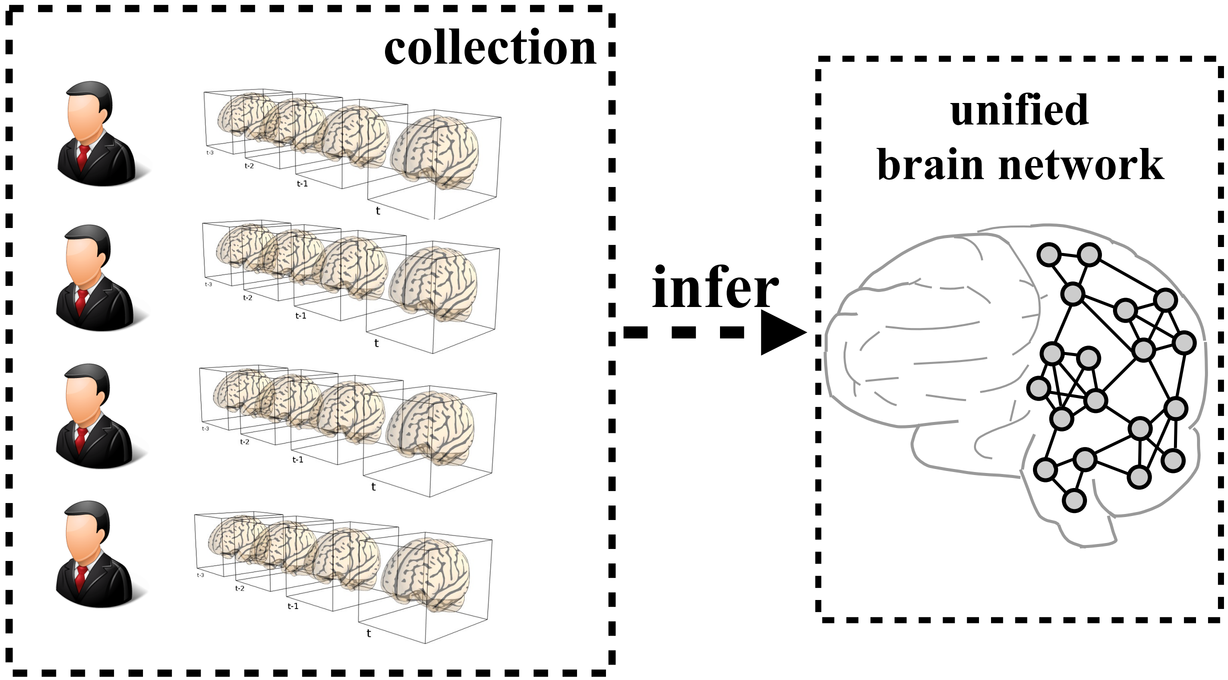
Networks derived by GLasso
for 5 healthy subjects





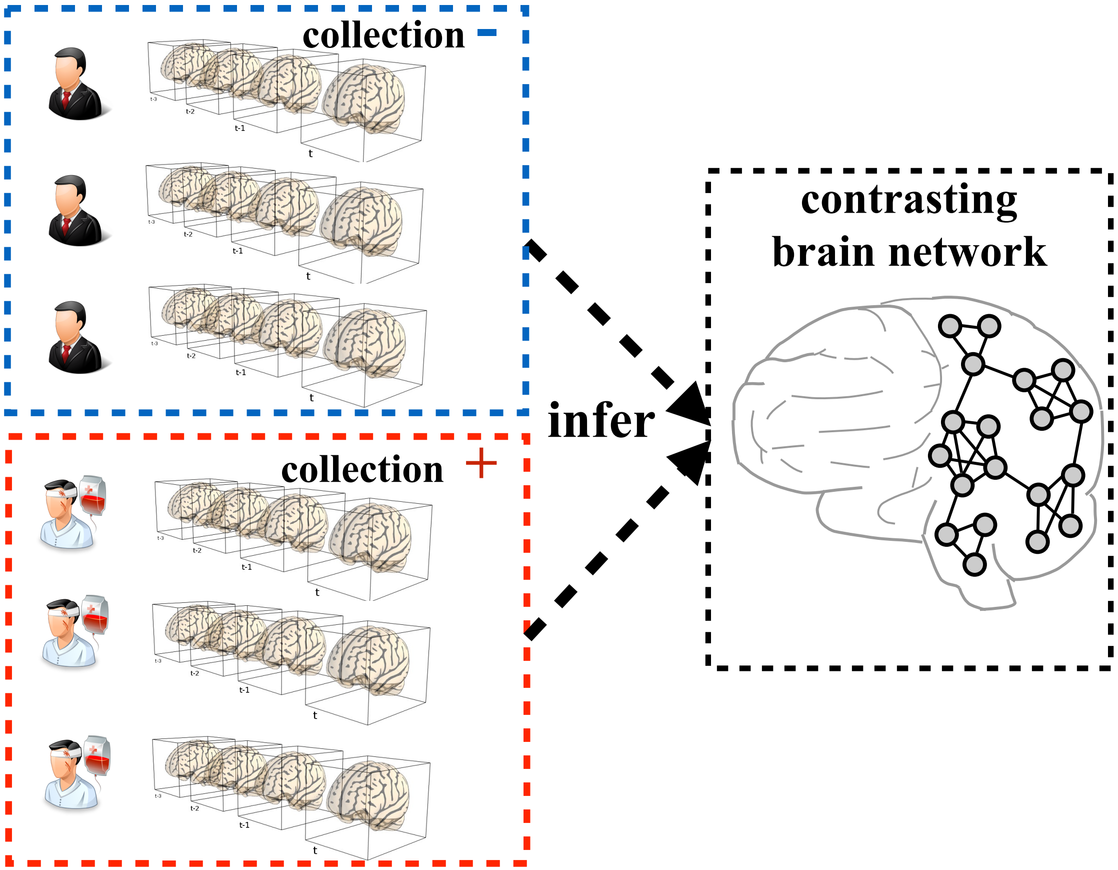


P1: Unified Brain Network Discovery



Finding the most <u>representative</u> network for a collection of subjects

P2: Contrasting Brain Network Discovery



Finding the most discriminative network on two collections of subjects

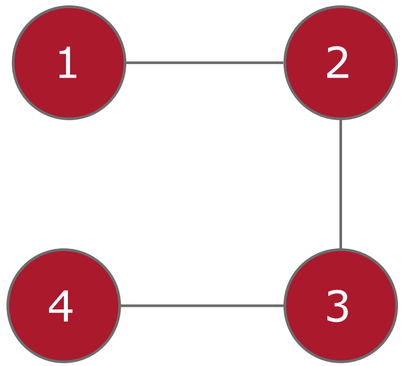
P2: Contrasting Brain Network Discovery

- A common practice:
  - infer a network for each collection
  - compare two networks

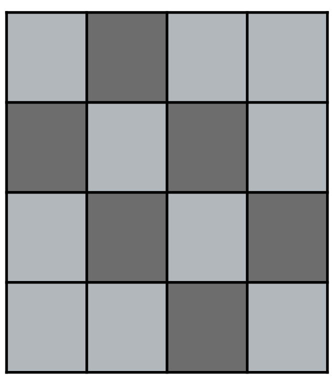
- Infer two networks: time-consuming (>40k nodes)

- Extract the differences: difficult (noises and unreliability)

Preliminary: Inverse Covariance

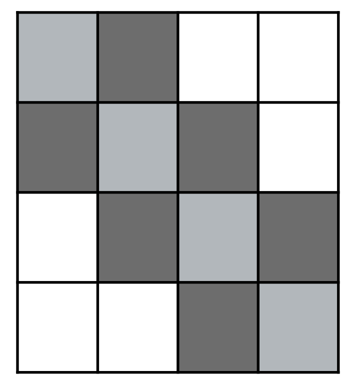


Covariance Matrix



Indirect Connections

Inverse Covariance Matrix



Direct Connections

Preliminary: Compute the Inverse

$$\boxed{\mathbf{S} = \frac{1}{n}\mathbf{X}^T\mathbf{X}}$$

empirical covariance matrix

Find the precision (inverse covariance) matrix

$$\cancel{\boldsymbol{\Theta} \leftarrow \mathbf{S}^{-1}}$$

Very likely to be singular

Preliminary: Graphical Lasso

$$\underset{\boxed{\Theta \succ 0}}{\text{minimize}} \; \underbrace{-\log \det \mathbf{\Theta} + \mathrm{tr}(\mathbf{S}\mathbf{\Theta})}_{\text{Negative Log Likelihood}} + \underbrace{\lambda ||\mathbf{\Theta}||_1}_{\text{L1-norm Regularization}}$$

positive definite constraint

Find a sparse positive definite matrix $\mathbf{\Theta}$ which has high likelihood to be the precision matrix of $\boldsymbol{S}$

P1: Unified Graphical Lasso

Input: $\{\mathbf{S}_1, \ldots, \mathbf{S}_p\}$

$$\underset{\boldsymbol{\Theta} \succ 0}{\text{minimize}} \ \frac{1}{p} \sum_{i=1}^{p} \big(-\log\det \boldsymbol{\Theta} + \mathrm{tr}(\mathbf{S}_i \boldsymbol{\Theta}) \big)$$

average negative log likelihood

$$+ \frac{\alpha}{p} \sum_{i=1}^{p} ||\boldsymbol{\Theta} - \boxed{\hat{\boldsymbol{\Theta}}_i}||_F^2$$

Inferred network for subject *i*

homogeneous regularization

$$+ \lambda ||\boldsymbol{\Theta}||_1$$

sparseness regularization

$$\hat{\boldsymbol{\Theta}}_i = \arg\min_{\boldsymbol{\Theta} \succ 0} -\log\det \boldsymbol{\Theta} + \mathrm{tr}(\mathbf{S}_i \boldsymbol{\Theta}) + \lambda ||\boldsymbol{\Theta}||_1$$

P2: Contrasting Graphical Lasso

Input:

Collection A $\{\mathbf{S}_1^{(A)}, \ldots, \mathbf{S}_p^{(A)}\}$

Collection B $\{\mathbf{S}_1^{(B)}, \ldots, \mathbf{S}_q^{(B)}\}$

The negative log likelihood of $\boldsymbol{\Theta}$ to be the network of collection A

$$L(\boldsymbol{\Theta}, \mathbf{S}_1^{(A)}, \ldots, \mathbf{S}_p^{(A)}) = \frac{1}{p}\sum_{i=1}^{p}\left(\boxed{-\log\det\boldsymbol{\Theta}} + \mathrm{tr}(\mathbf{S}_i^{(A)}\boldsymbol{\Theta})\right)$$

$$L(\boldsymbol{\Theta}, \mathbf{S}_1^{B}, \ldots, \mathbf{S}_q^{(B)}) = \frac{1}{q}\sum_{i=1}^{q}\left(\boxed{-\log\det\boldsymbol{\Theta}} + \mathrm{tr}(\mathbf{S}_i^{(B)}\boldsymbol{\Theta})\right)$$

The negative log likelihood of $\boldsymbol{\Theta}$ to be the network of collection B

P2:
Contrasting Graphical Lasso

$$\underset{\boldsymbol{\Theta} \succ 0}{\text{minimize}} \ \frac{1}{p} \sum_{i=1}^{p} \operatorname{tr}(\mathbf{S}_i^{(A)} \boldsymbol{\Theta}) - \frac{1}{q} \sum_{j=1}^{q} \operatorname{tr}(\mathbf{S}_j^{(B)} \boldsymbol{\Theta}) + \lambda ||\boldsymbol{\Theta}||_1$$

Find a sparse positive definite matrix $\boldsymbol{\Theta}$ that:

1. Has high likelihood to be the network of collection B
2. Has low likelihood to be the network of collection A

Solution

$$\underset{\boldsymbol{\Theta} \succ 0}{\text{minimize}} \quad \underbrace{g(\boldsymbol{\Theta})}_{\substack{\text{differentiable} \\ \text{smooth function}}} + \overbrace{\lambda ||\boldsymbol{\Theta}||_1}^{\substack{\text{non-differentiable} \\ \text{Regularization function}}}$$

Solved by Spectral Projected Gradient Descent

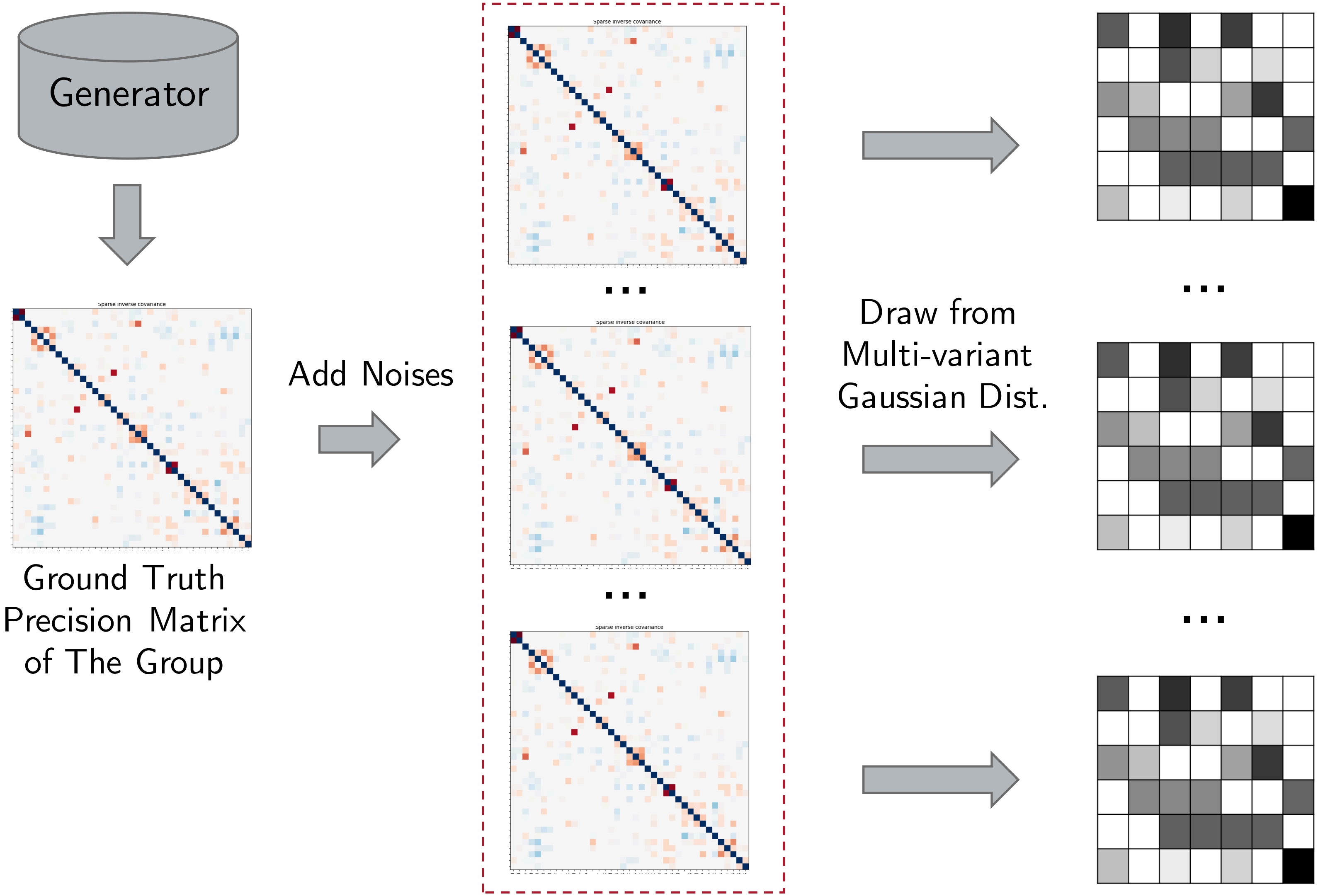
Unified Graphical Lasso

$$\boxed{g(\boldsymbol{\Theta}) = L(\boldsymbol{\Theta}, \mathbf{S}_1, \ldots, \mathbf{S}_p) + \alpha \sum_{i=1}^{p} ||\boldsymbol{\Theta} - \hat{\boldsymbol{\Theta}}_i||_F^2}$$

Contrasting Graphical Lasso

$$\boxed{g(\boldsymbol{\Theta}) = \frac{1}{p} \sum_{i=1}^{p} \operatorname{tr}(\mathbf{S}_i^{(A)} \boldsymbol{\Theta}) - \frac{1}{q} \sum_{j=1}^{q} \operatorname{tr}(\mathbf{S}_j^{(B)} \boldsymbol{\Theta})}$$

Synthetic Data



Synthetic Data: UGLasso

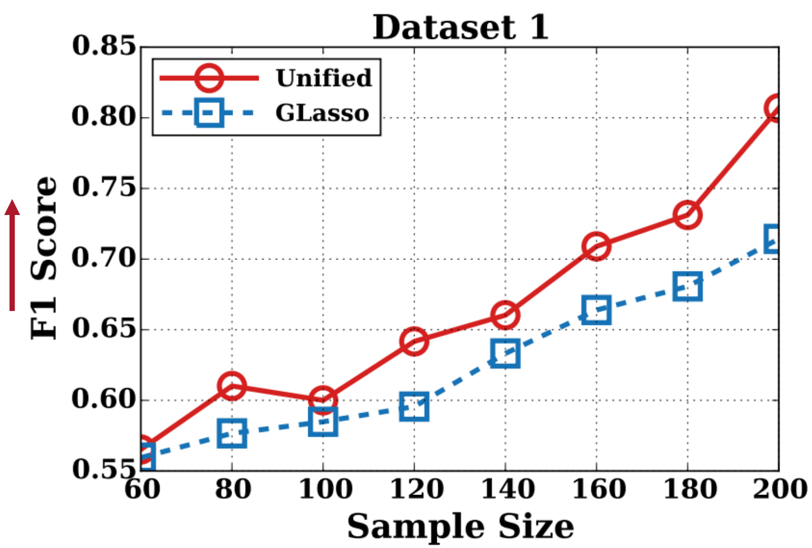
- Test the performance of Unified Graphical Lasso.
  - Data Set 1: Weak Noise
  - Data Set 2: Moderate Noise
  - Data Set 3: Strong Noise

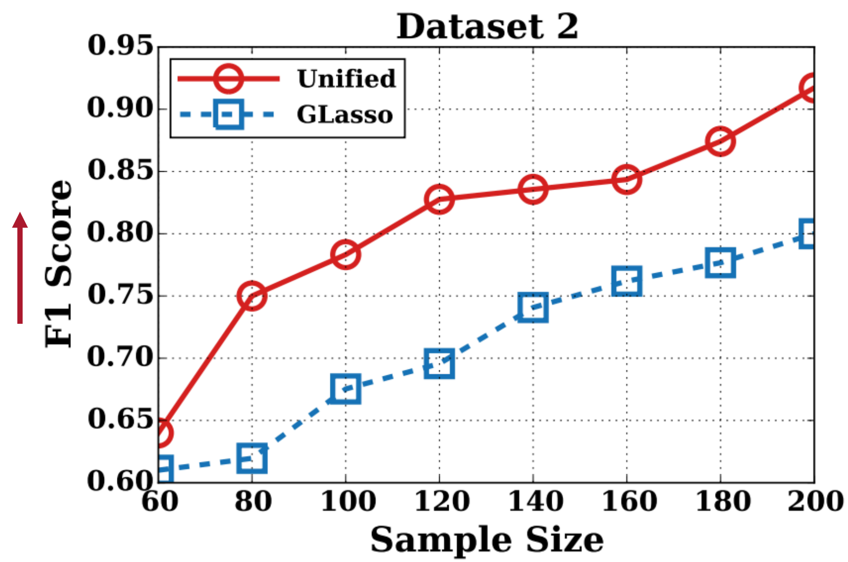
- Sample size (# of snapshot) = 60~200 (with step size 20).

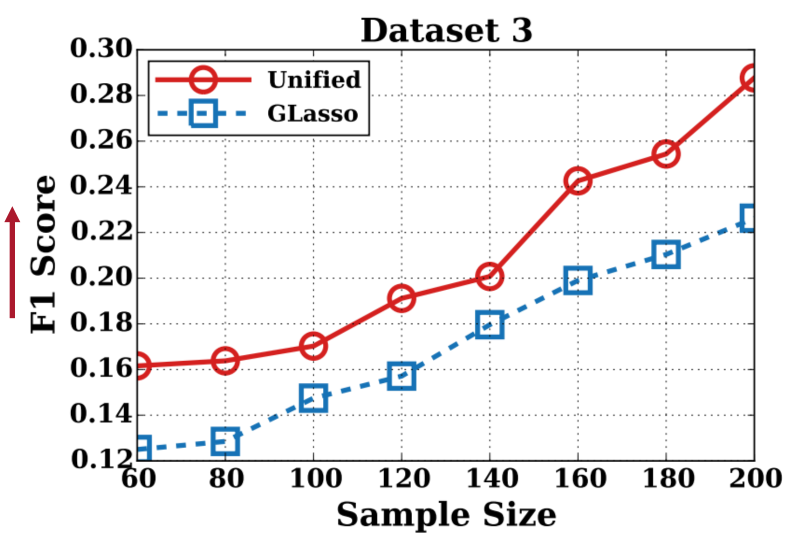
- Evaluation: F1 score of the discovered edges.

- Methods: UGLasso and GLasso.

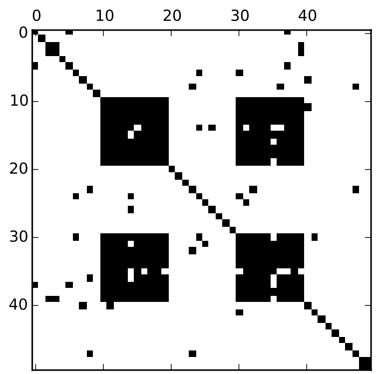
Synthetic Data: UGLasso



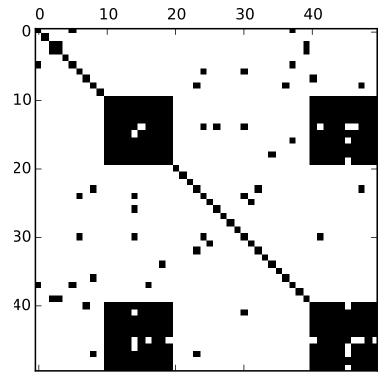




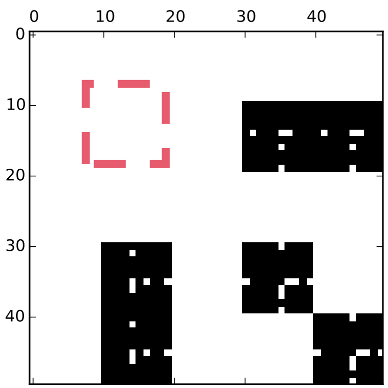
Synthetic Data: CGLasso



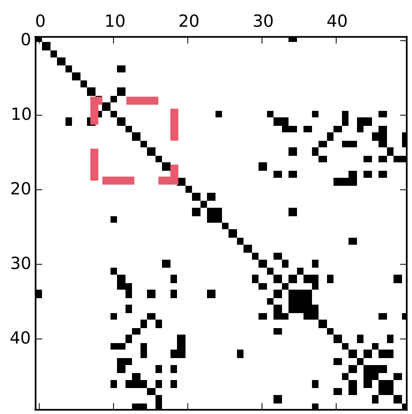
Ground Truth A



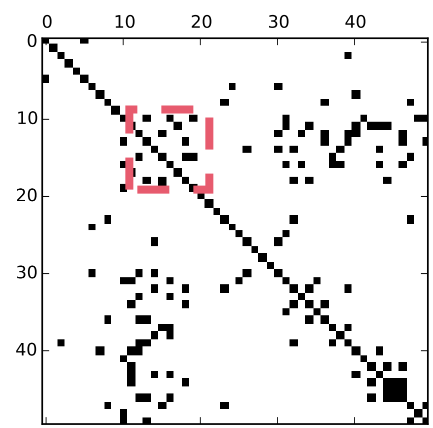
Ground Truth B



Ground Truth A - B

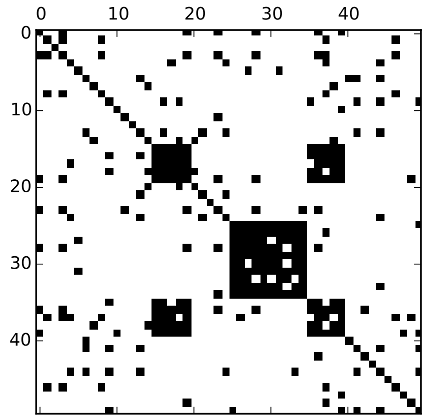


Contrasting GLasso

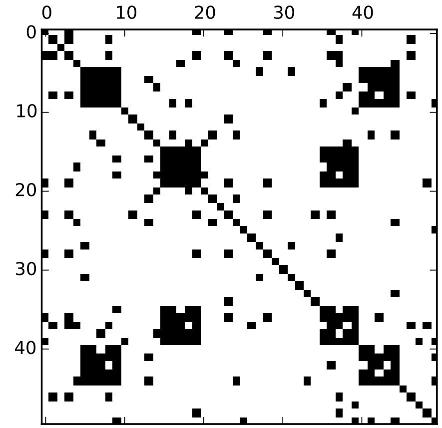


Glasso A - B

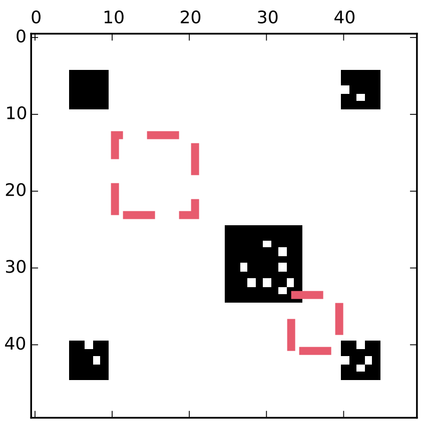
Synthetic Data: CGLasso



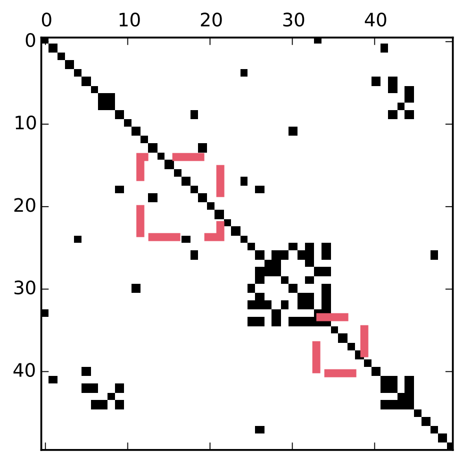
Ground Truth A



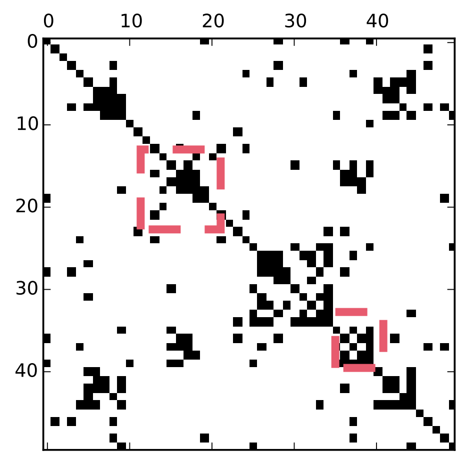
Ground Truth B



Ground Truth A - B



Contrasting GLasso



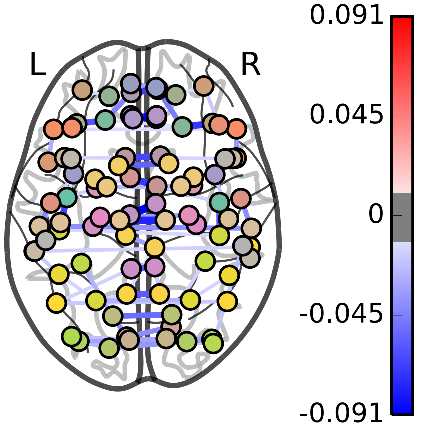
Glasso A - B

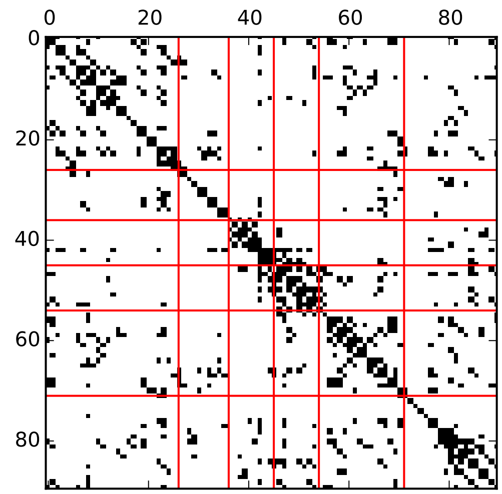
Real fMRI Data

- ADNI (Alzheimer's Disease)
  - Alzheimer's Disease (AD) and Mild Cognitive Impairment (MCI) as positive subjects
  - Normal Controls (NC) as negative subjects
  - Applied Automated Anatomical Labeling (AAL) to extract sequences of the 116 anatomical volumes of interest
- HIV (Human Immunodeficiency Virus Infection)
  - HIV patients as positive subjects
  - Normal Controls as negative subjects
- All results are derived using $\lambda = 2.0$ and $\alpha = 0.5$

ADNI (Unified)

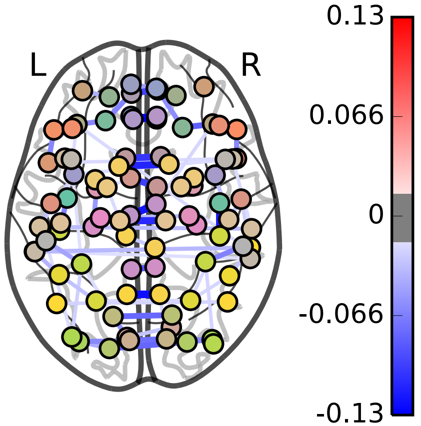
Positive

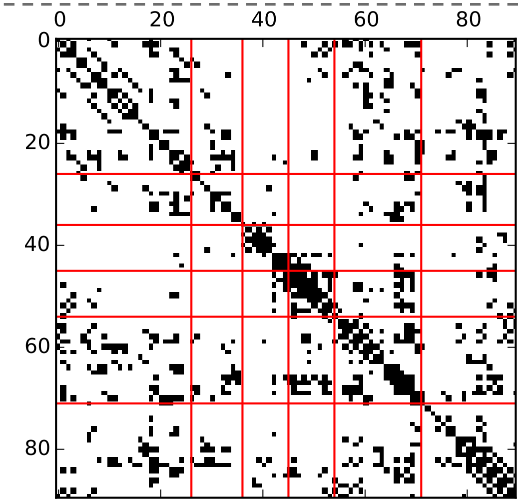




Weaker interconnections

Negative

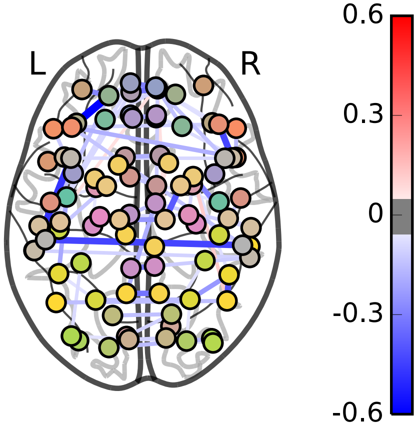


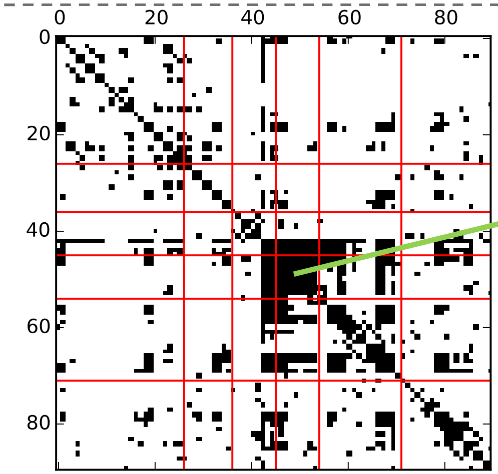


Stronger interconnections

HIV (Unified)

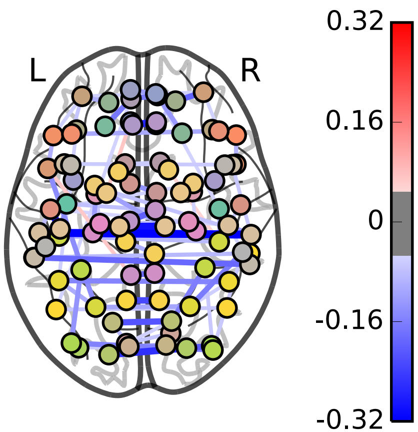
Positive

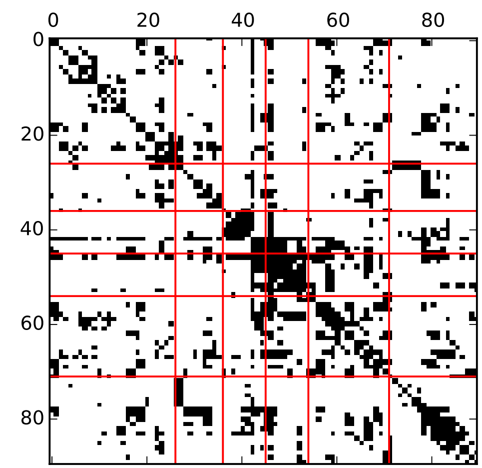




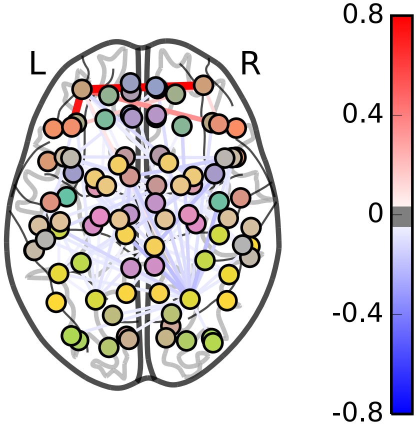
increased connectivity inside occipital lobe

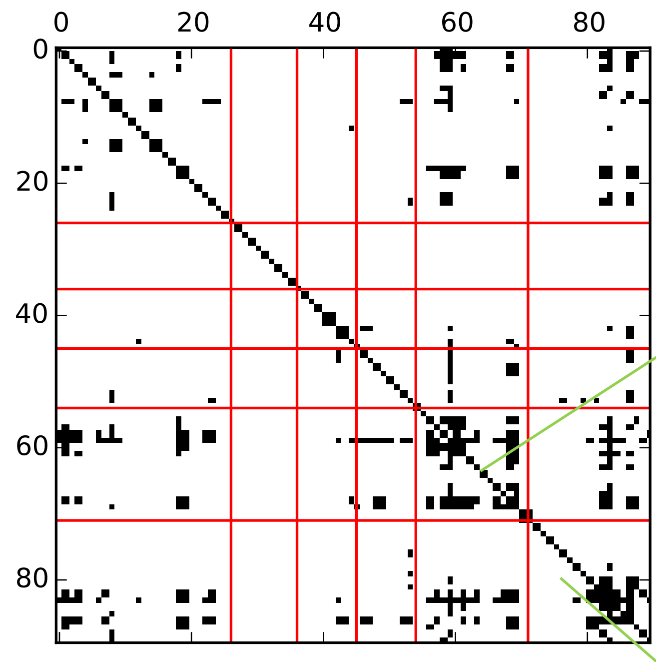
Negative





ADNI (Contrasting)

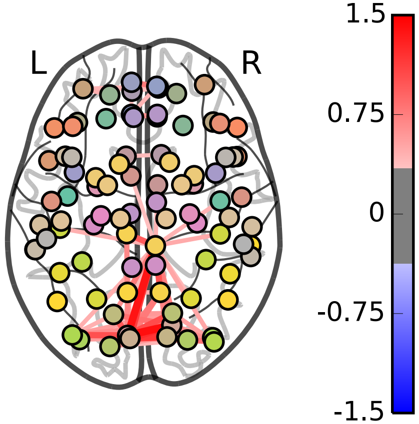


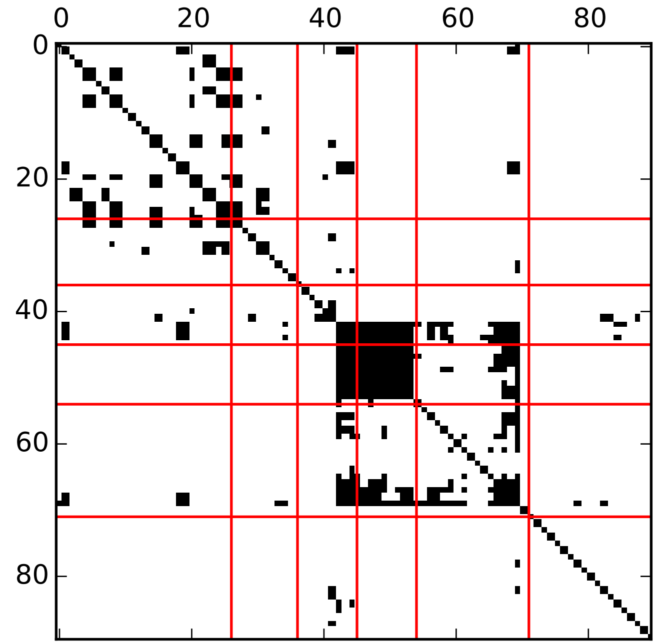


parietal lobe

temporal lobe

HIV (Contrasting)





Summary

- Problem Studied:
  - Unified Brain Network Discovery
  - Contrasting Brain Network Discovery

- Proposed Methods:
  - UGLasso and CGLasso
  - Both solved by SPGD

- Conclusion:
  - UGLasso finds better representative network for a group of subjects.
  - CGLasso extracts the discriminative patterns between two collections of subjects more accurately.

Q&A

Xinyue Liu, Xiangnan Kong, Ann B. Ragin
xliu4@wpi.edu, xkong@wpi.edu, ann-ragin@northwestern.edu