Latent Variable Graphical Model Selection using Harmonic Analysis: Applications to the Human Connectome Project (HCP)

WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

*Won Hwa Kim[1], *Hyunwoo J. Kim[1], Nagesh Adluru[3], Vikas Singh[2]

[1]Dept. of Computer Sciences, UW-Madison
[2]Dept. of Biostatistics and Medical Informatics
[3]Waisman Center
*First authorship shared

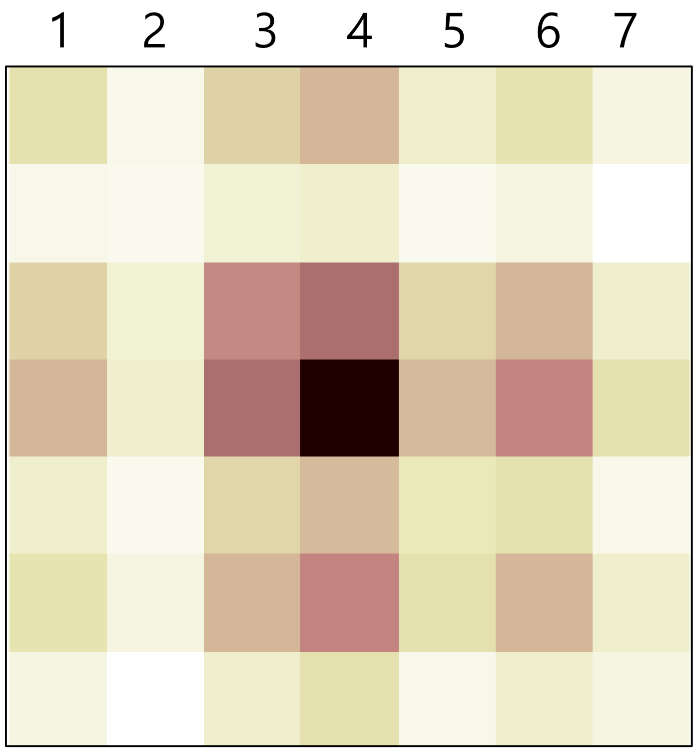
Graphical Model Selection

Estimate conditional independence represented by undirected graphical models

Observation

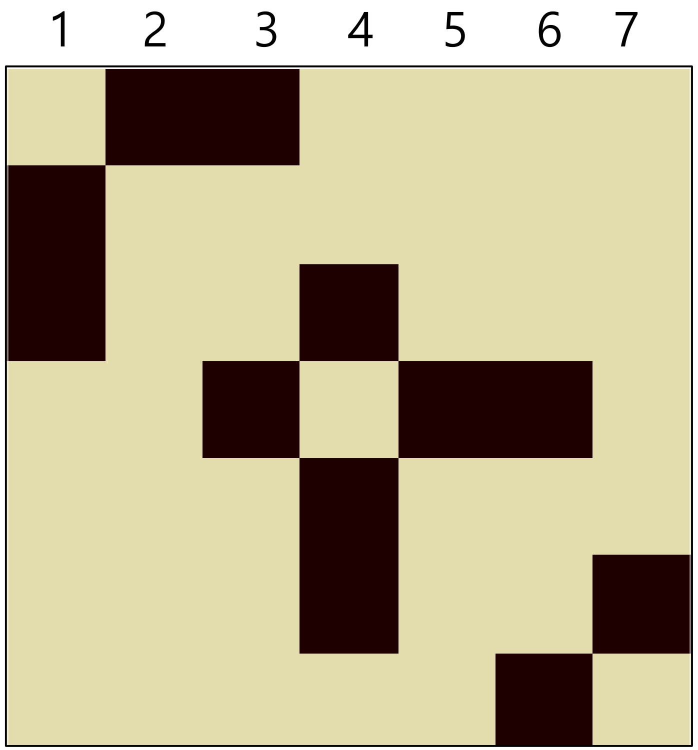
	f1	f2	…	f7
subj 1	0.1	0.1		1
subj 2	1.1	1.1		2
subj 3	2.2	2.2	…	1
subj 4	1.5	1.5		2.2
subj 5	3.6	3.6		1.1
…		…		
…		…		

Covariance Matrix



$$\Sigma = \sum_{\ell=1}^{n} \lambda_\ell V_\ell V_\ell^T$$

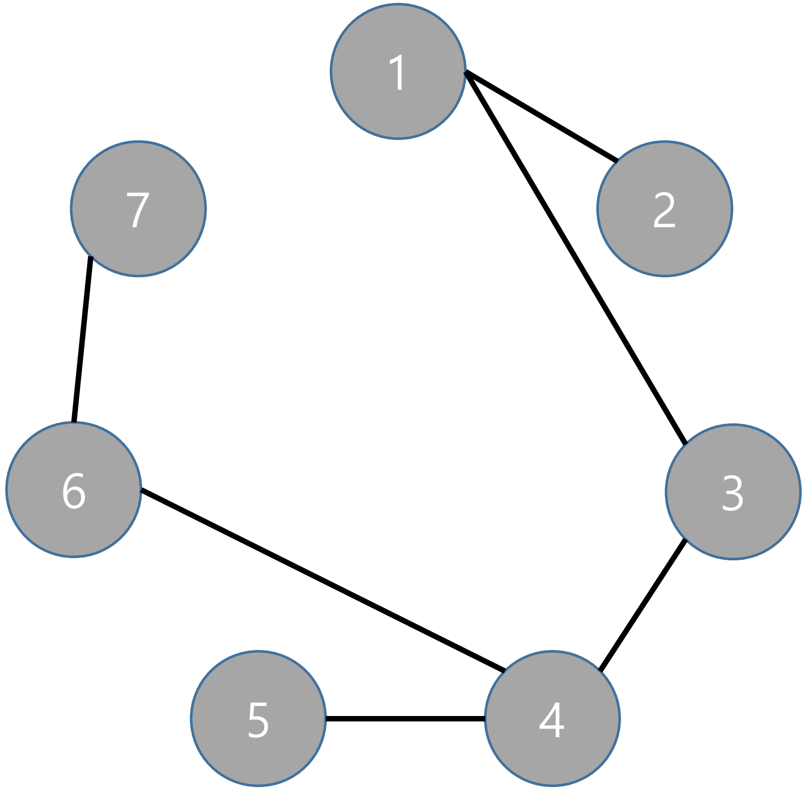
Precision Matrix



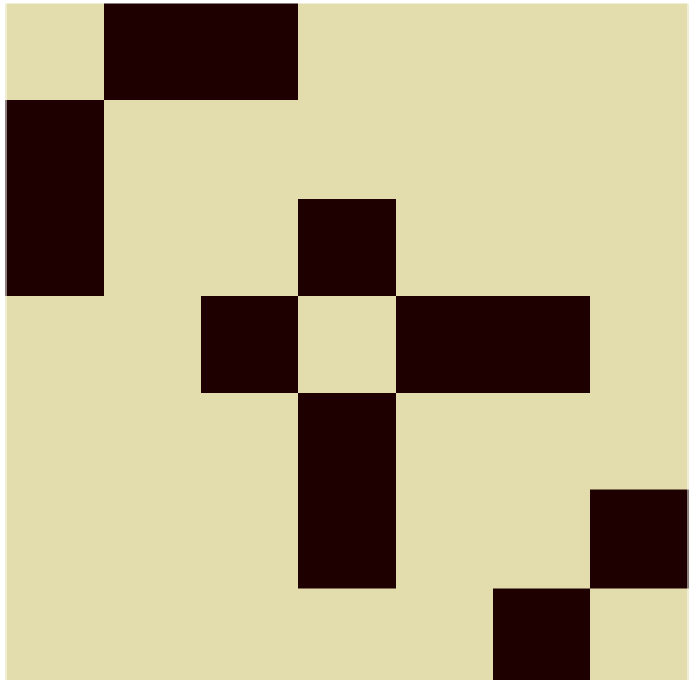
$$\Theta = \sum_{\ell=1}^{n} \frac{1}{\lambda_\ell} V_\ell V_\ell^T$$

Graphical Model Selection

Estimate conditional independence represented by undirected graphical models



Undirected graphical model

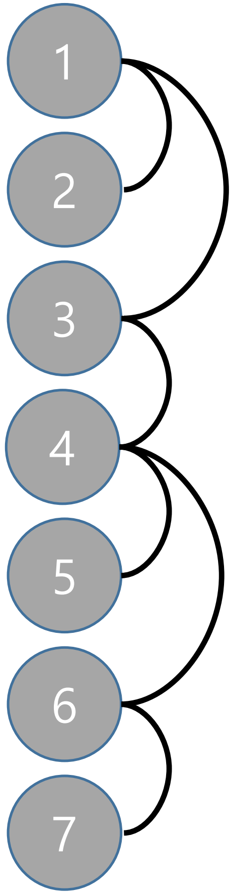


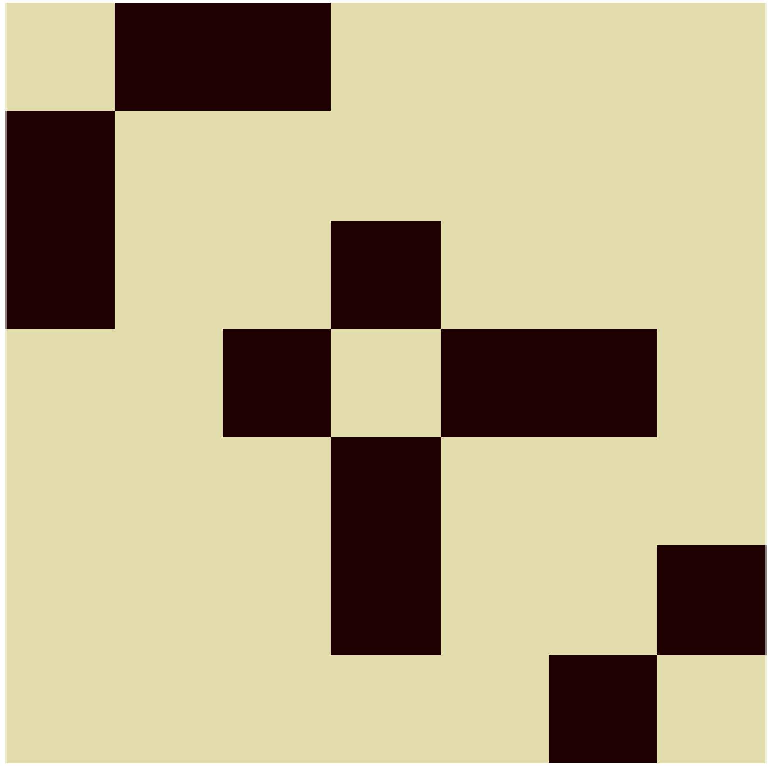
Precision matrix

Latent variables Graphical Model Selection

Challenges:

1. Latent (unobserved) variables affecting conditional independence. (Dense/noise precision matrix)

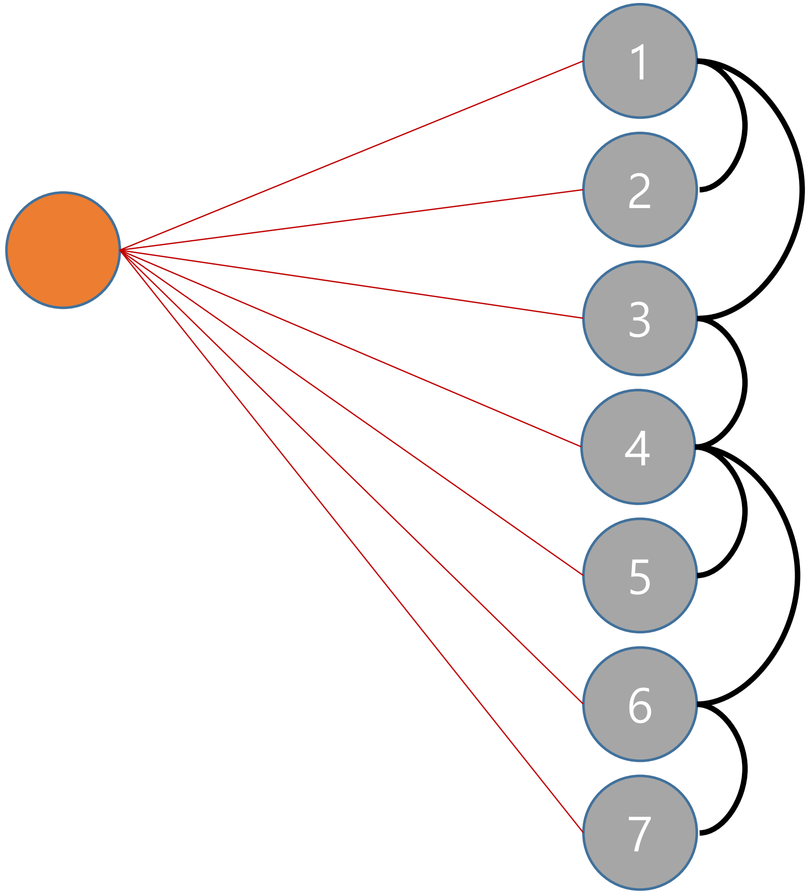


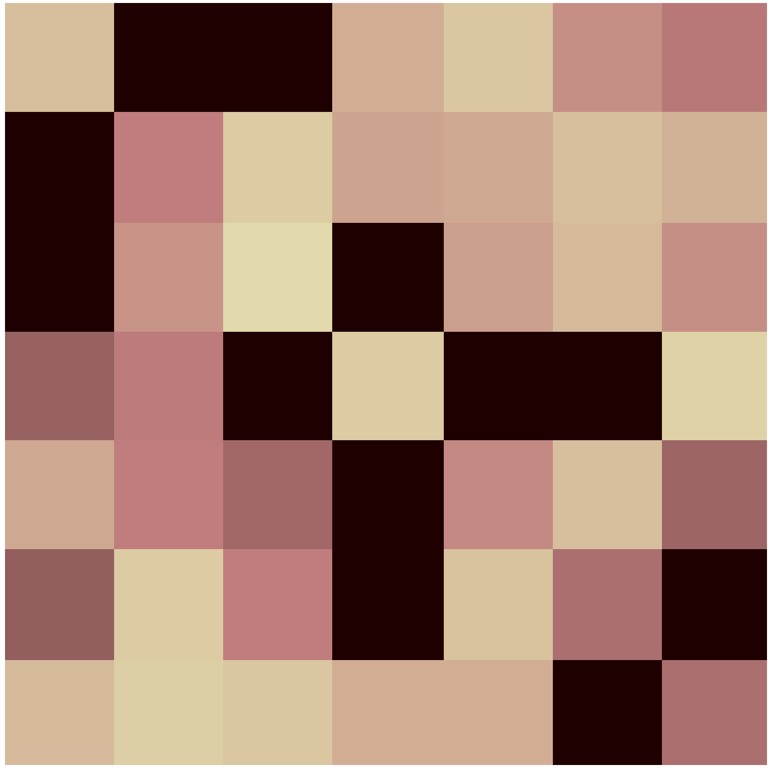


Latent variables Graphical Model Selection

Challenges:

1. **Latent (unobserved) variables** affecting conditional independence. (Dense/noisy precision matrix)

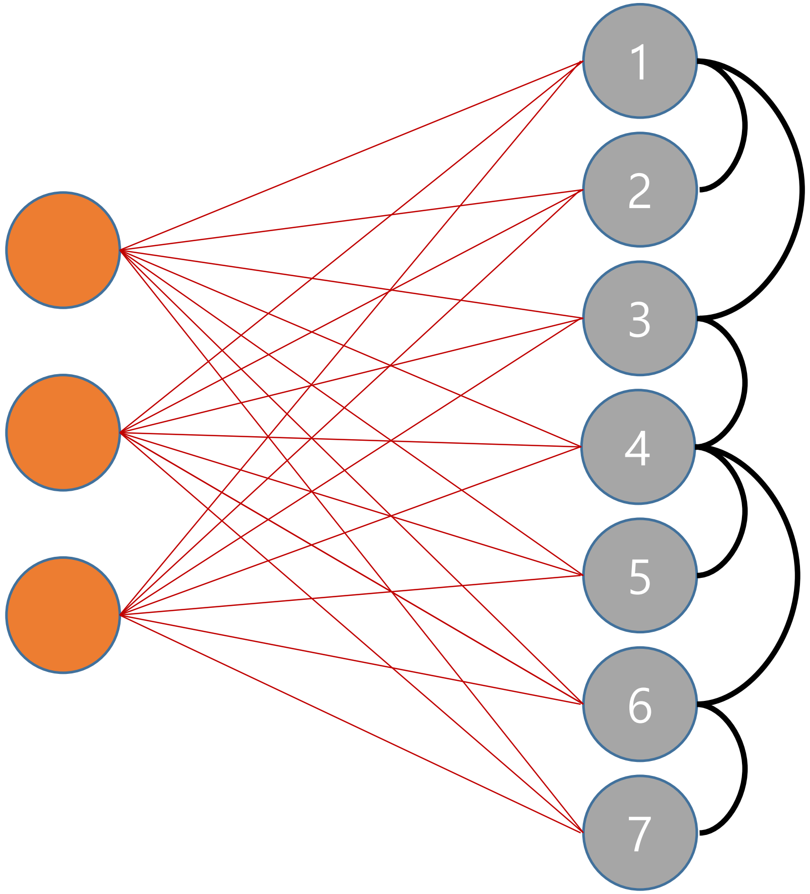


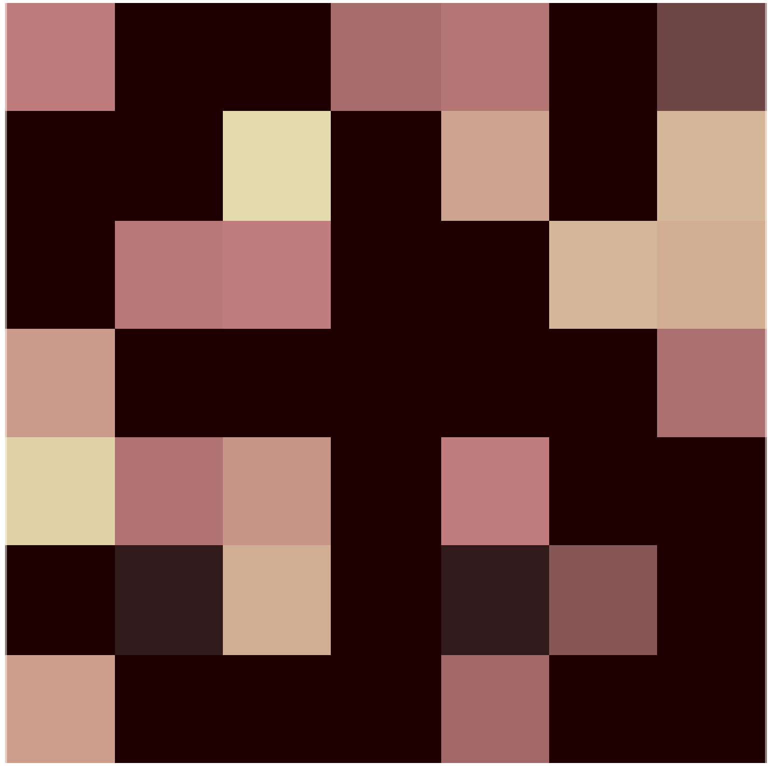


Latent variables Graphical Model Selection

Challenges:

1. **Latent (unobserved) variables** affecting conditional independence (Dense/noisy precision matrix).
2. **The number of Latent variables** is not known.



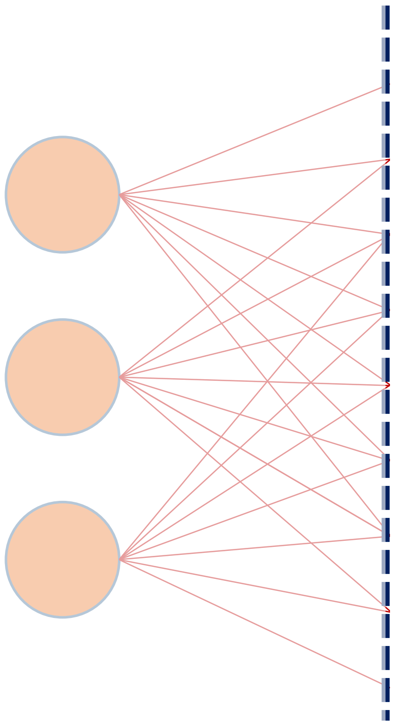


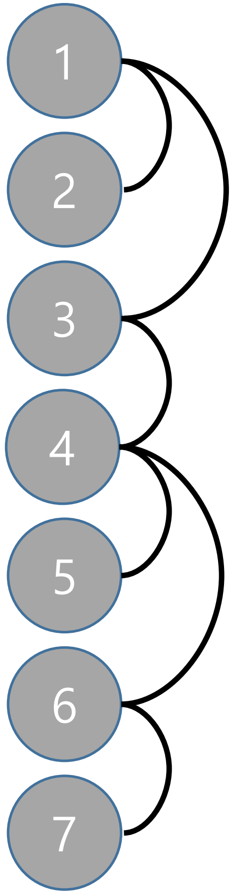
Latent variables Graphical Model Selection

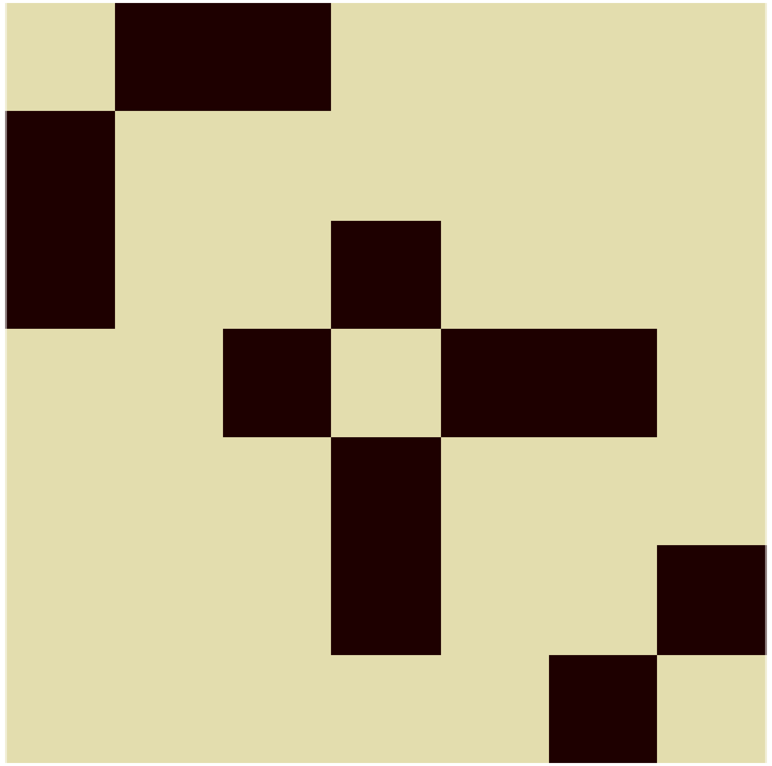
Challenges:

1. **Latent (unobserved) variables** affecting conditional independence. (Dense/noisy precision matrix)
2. **The number of Latent variables** is not known

Our Solution: Remove the effect from latent variables using Harmonic Analysis on Graphs





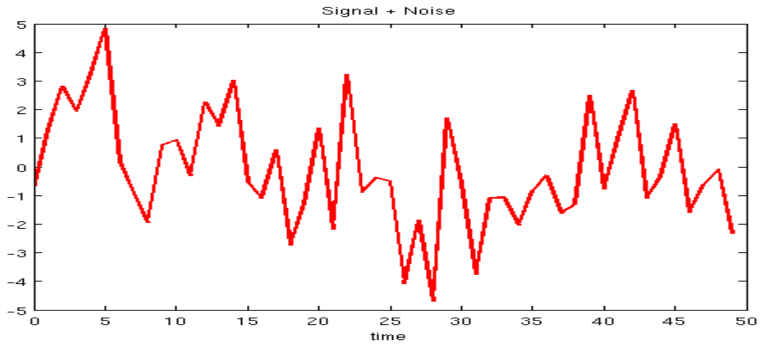


Harmonic Analysis of Latent Variable Graphical Models

Multi-scale Analysis of Precision Matrix

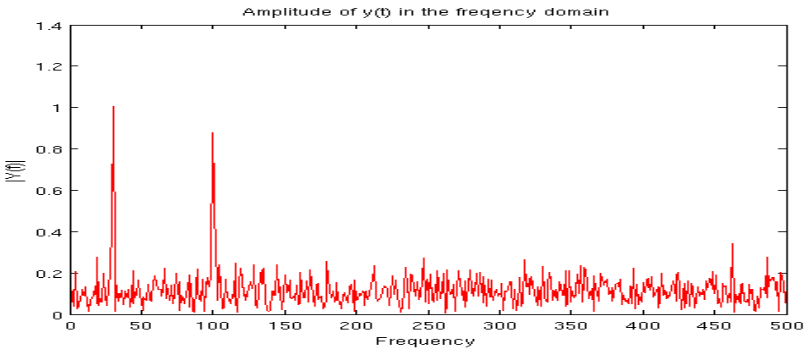
Invertible transform (e.g., Fourier transform) for novel perspective

Original Space



invertible

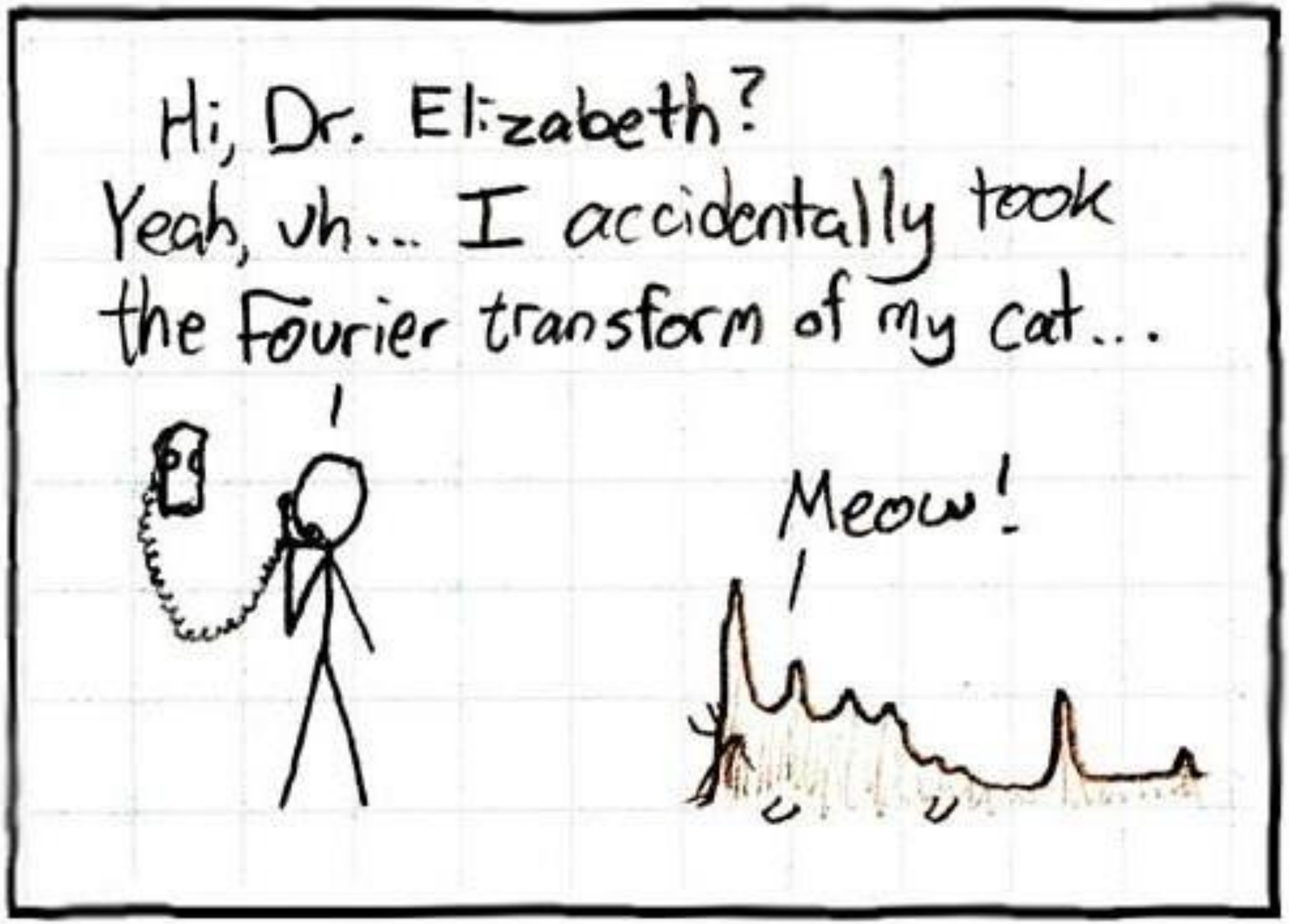
Dual Space



Harmonic Analysis of Latent Variable Graphical Models

Multi-scale Analysis of Precision Matrix

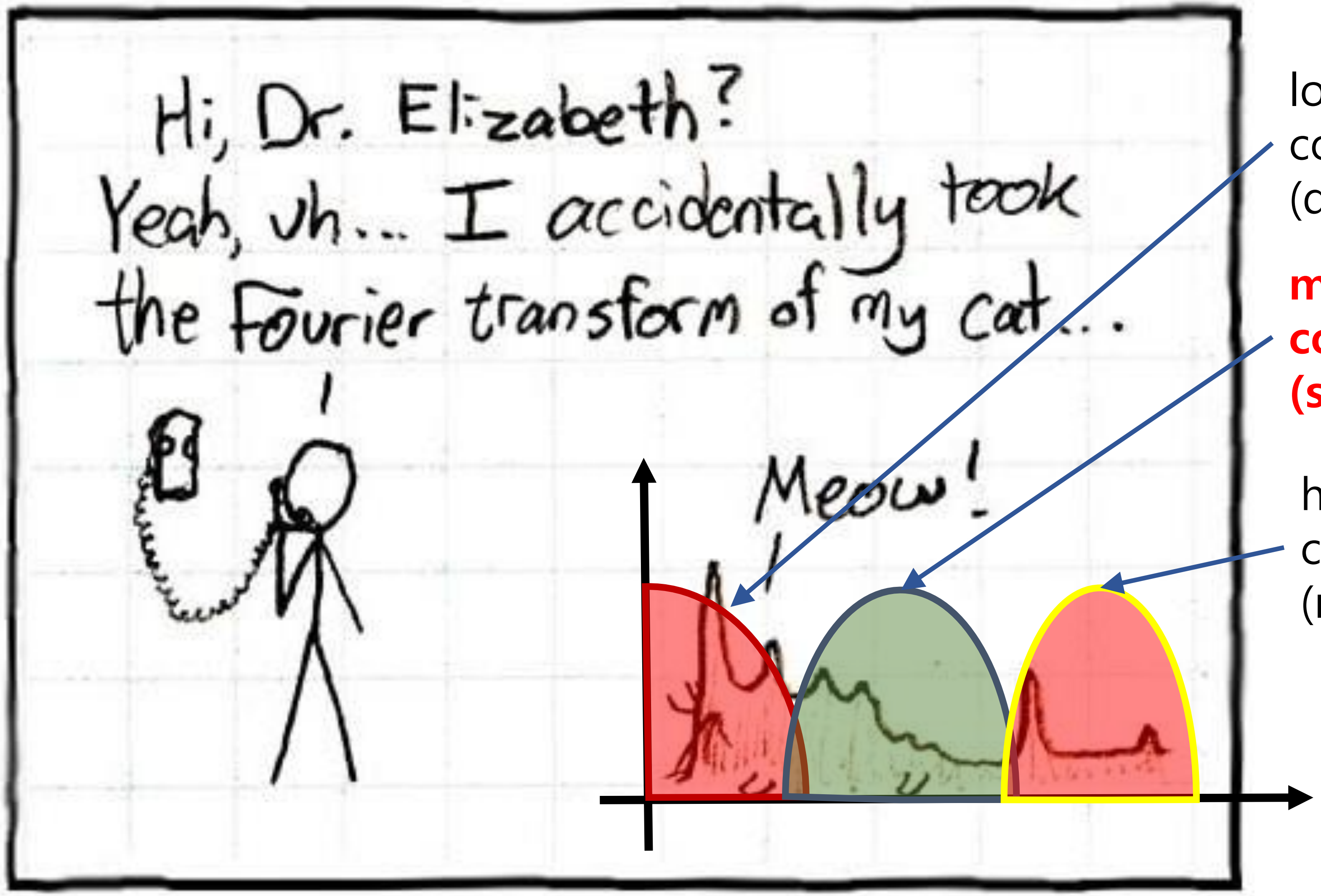
Focusing on specific band in the frequency domain: **WAVELET (on graphs)**



Harmonic Analysis of Latent Variable Graphical Models

Multi-scale Analysis of Precision Matrix

Focusing on specific band in the frequency domain: **WAVELET (on graphs)**



low-frequency components (dense part)

middle band components (sparse part)

high-frequency components (noise)

Harmonic Analysis of Latent Variable Graphical Models

Multi-scale Analysis of Precision Matrix

Basis function to derive multi-scale view of Θ

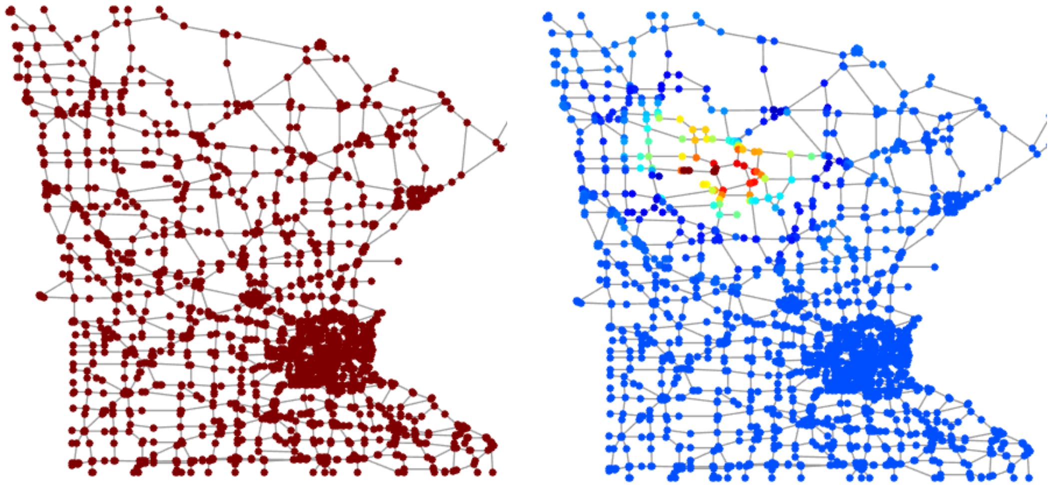
$$\psi_{\ell,s}(i,j) = g(s\sigma_\ell)V_\ell^*(i)V_\ell(j), \forall \ell \in \{1,\ldots,n\}$$

$g()$: kernel function (band-pass filter)

Spectral graph wavelet (Hammond et al, 2011)

$$\psi_{s,n}(m) = \sum_{\ell=0}^{N-1} g(s\lambda_l)\chi_\ell^*(n)\chi_\ell(m)$$

χ: basis from spectral graph theory



Our Estimation (at scale s)

$$\tilde{\Theta} = \sum_{\ell=1}^{n} \sigma_\ell g^2(s\sigma_\ell)V_\ell V_\ell^T$$

Harmonic Analysis of Latent Variable Graphical Models

Estimating the optimal scale

Objective function in the original space:

$$\max_{s\geq 0} \ \mathrm{tr}(\tilde{\Theta}\Theta^{-1}) - \mathrm{logdet}(\tilde{\Theta}\Theta^{-1}) - n + \gamma|\tilde{\Theta}|_1$$

$$\text{subject to } \tilde{\Theta} = \sum_{\ell=1}^{n} \sigma_\ell g^2(s\sigma_\ell) V_\ell V_\ell^T$$

Our objective function in the dual space:

$$\max_{s\geq 0} \sum_{\ell=1}^{n} \lambda_\ell K(s,\sigma_\ell) - \sum_{\ell=1}^{n} \log(\lambda_\ell K(s,\sigma_\ell)) - n \ + \gamma \sum_{i=1}^{n}\sum_{j=1}^{n} \left| \sum_{\ell=1}^{n} K(s,\sigma_\ell) X_\ell(i,j) \right|$$

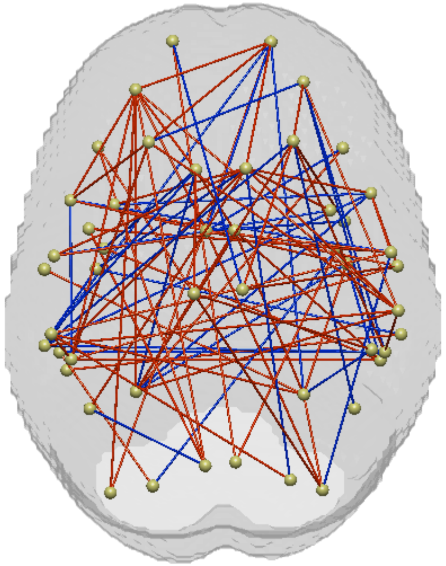
where $X_\ell = V_\ell V_\ell^T$, $K(s,\sigma_\ell) := \sigma_\ell g^2(s\sigma_\ell)$

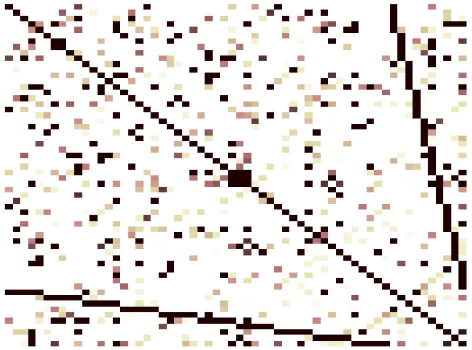
Experimental Result

Statistical dependency estimation on synthetic brain connectivity

n_o = 60 (i.e., 50 pathways, 10 covariates) / $\boldsymbol{n_h}$ **= 5**

1) Precision Matrix



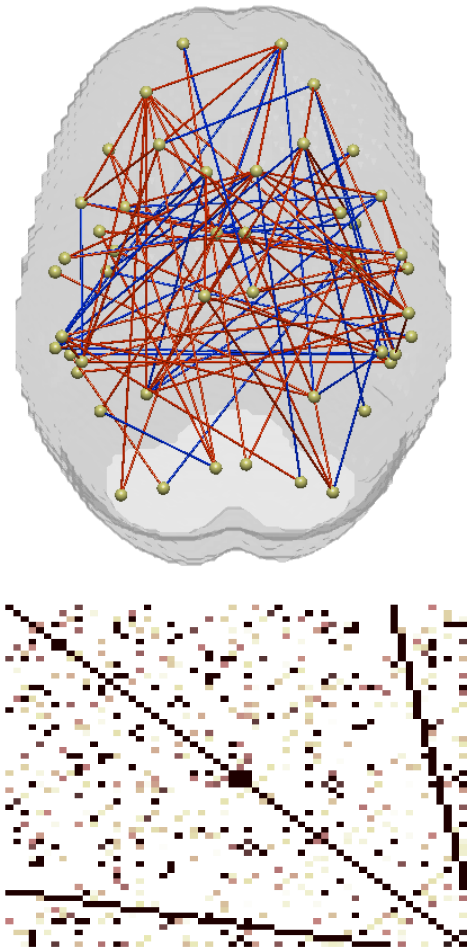


Experimental Result

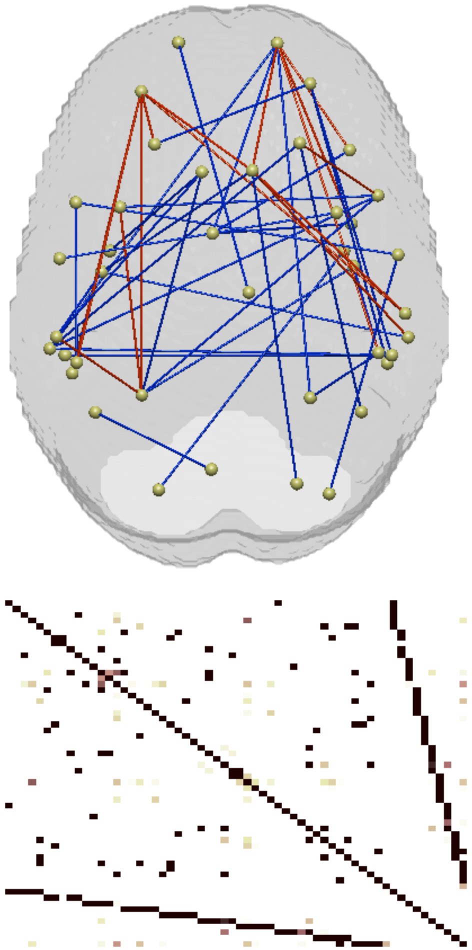
Statistical dependency estimation on synthetic brain connectivity

n_o = 60 (i.e., 50 pathways, 10 covariates) / $\boldsymbol{n_h = 5}$

1) Precision Matrix



2) Graphical Lasso

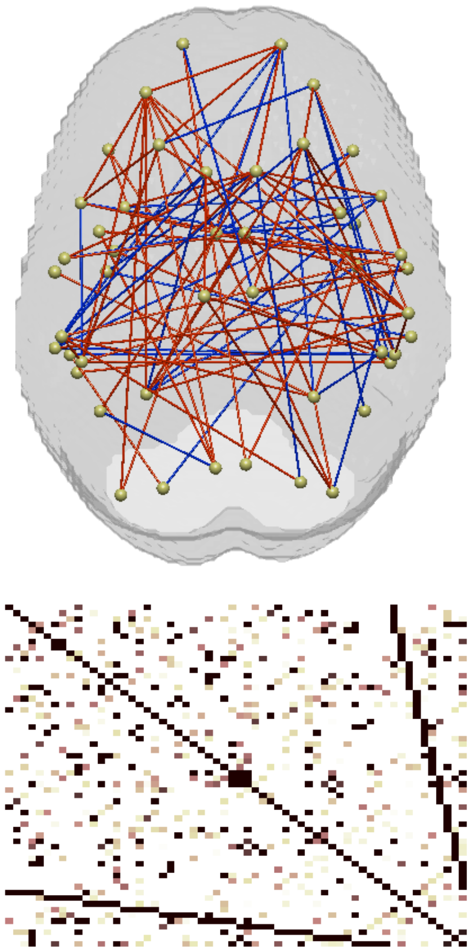


Experimental Result

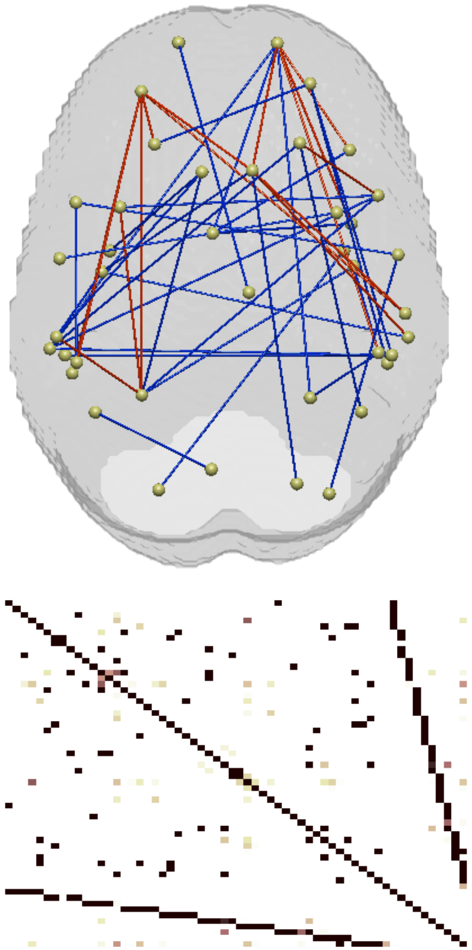
Statistical dependency estimation on synthetic brain connectivity

n_o = 60 (i.e., 50 pathways, 10 covariates) / $\boldsymbol{n_h}$ **= 5**

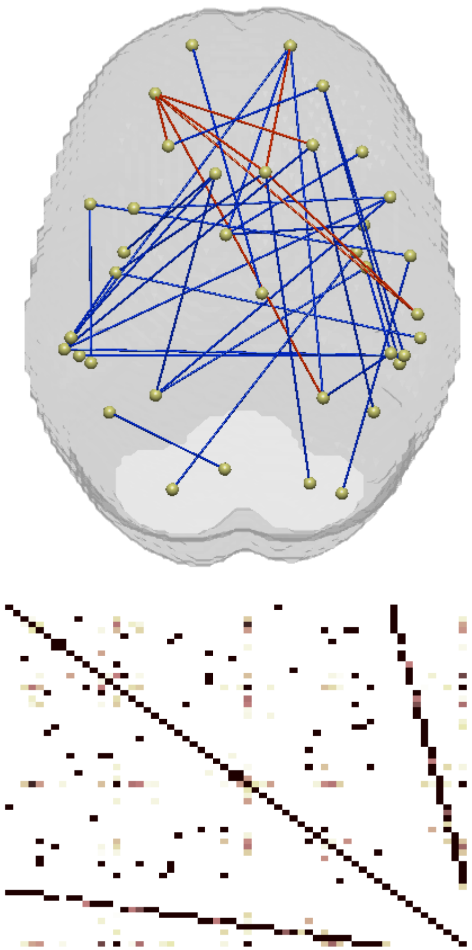
1) Precision Matrix



2) Graphical Lasso



3) Chandrasekaran et al

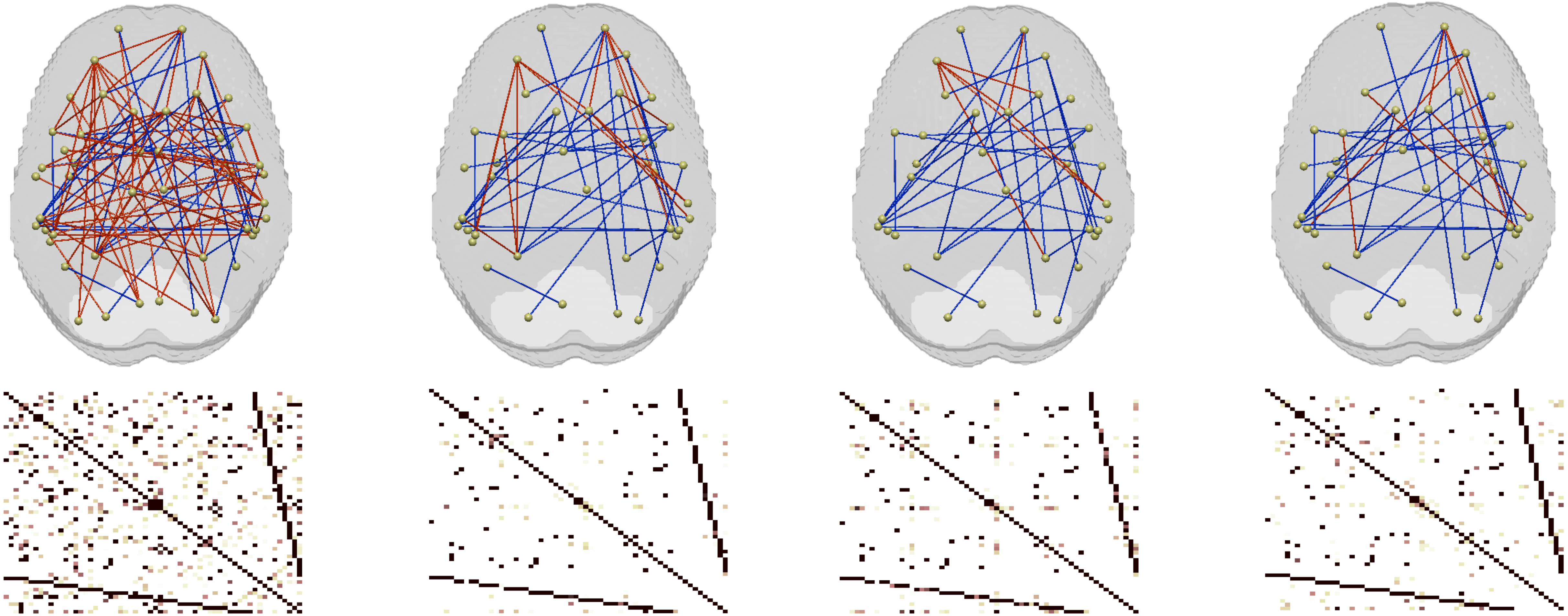


Experimental Result

Statistical dependency estimation on synthetic brain connectivity

n_o = 60 (i.e., 50 pathways, 10 covariates) / $\boldsymbol{n_h = 5}$

1) Precision Matrix 2) Graphical Lasso 3) Chandrasekaran et al 4) Ours

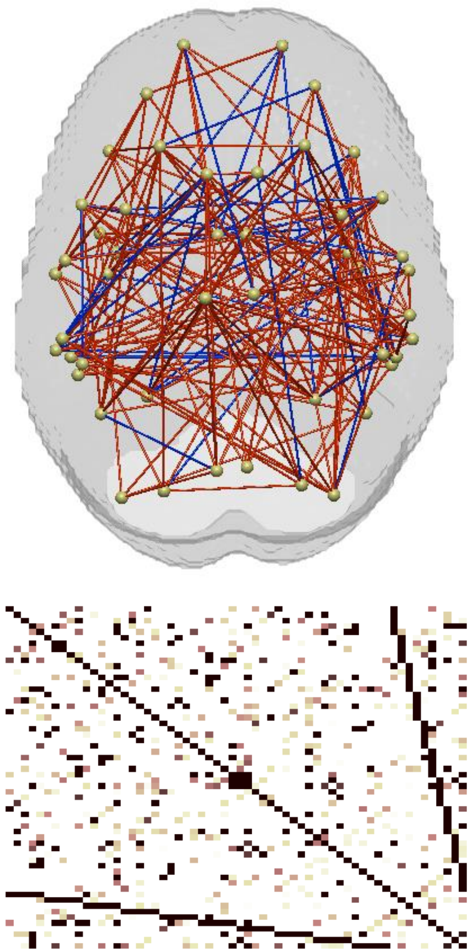


Experimental Result

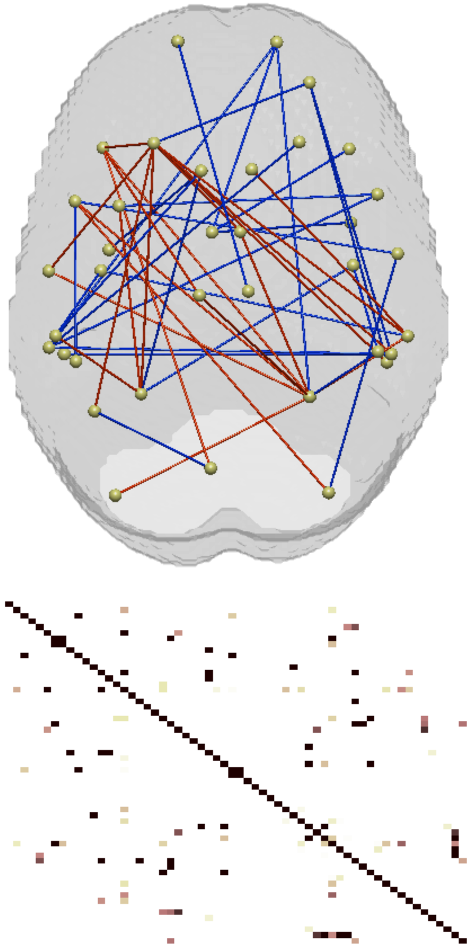
Statistical dependency estimation on synthetic brain connectivity

n_o = 60 (i.e., 50 pathways, 10 covariates) / $\boldsymbol{n_h}$ **= 10**

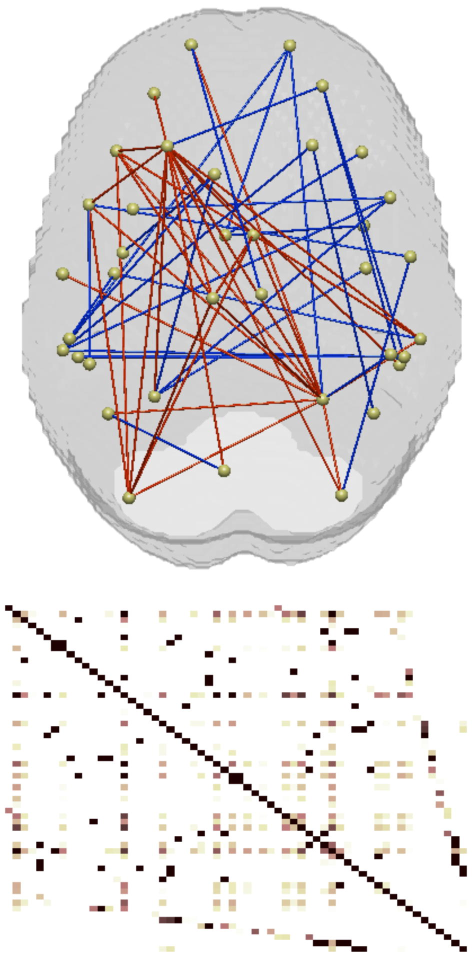
1) Precision Matrix



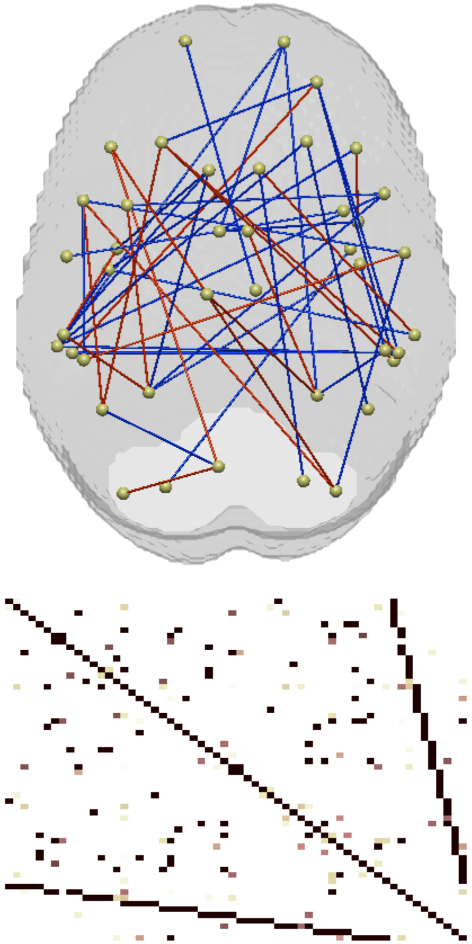
2) Graphical Lasso



3) Chandrasekaran et al



4) Ours

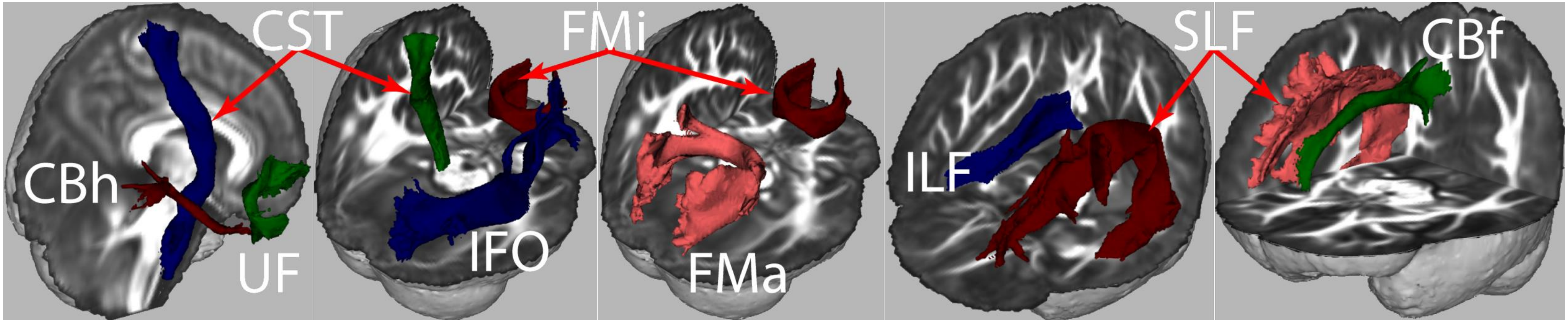


Experimental Result (HCP)

Statistical dependency estimation on HCP dataset

n_o = 60 (i.e., 17 pathways, 22covariates)

Pathways



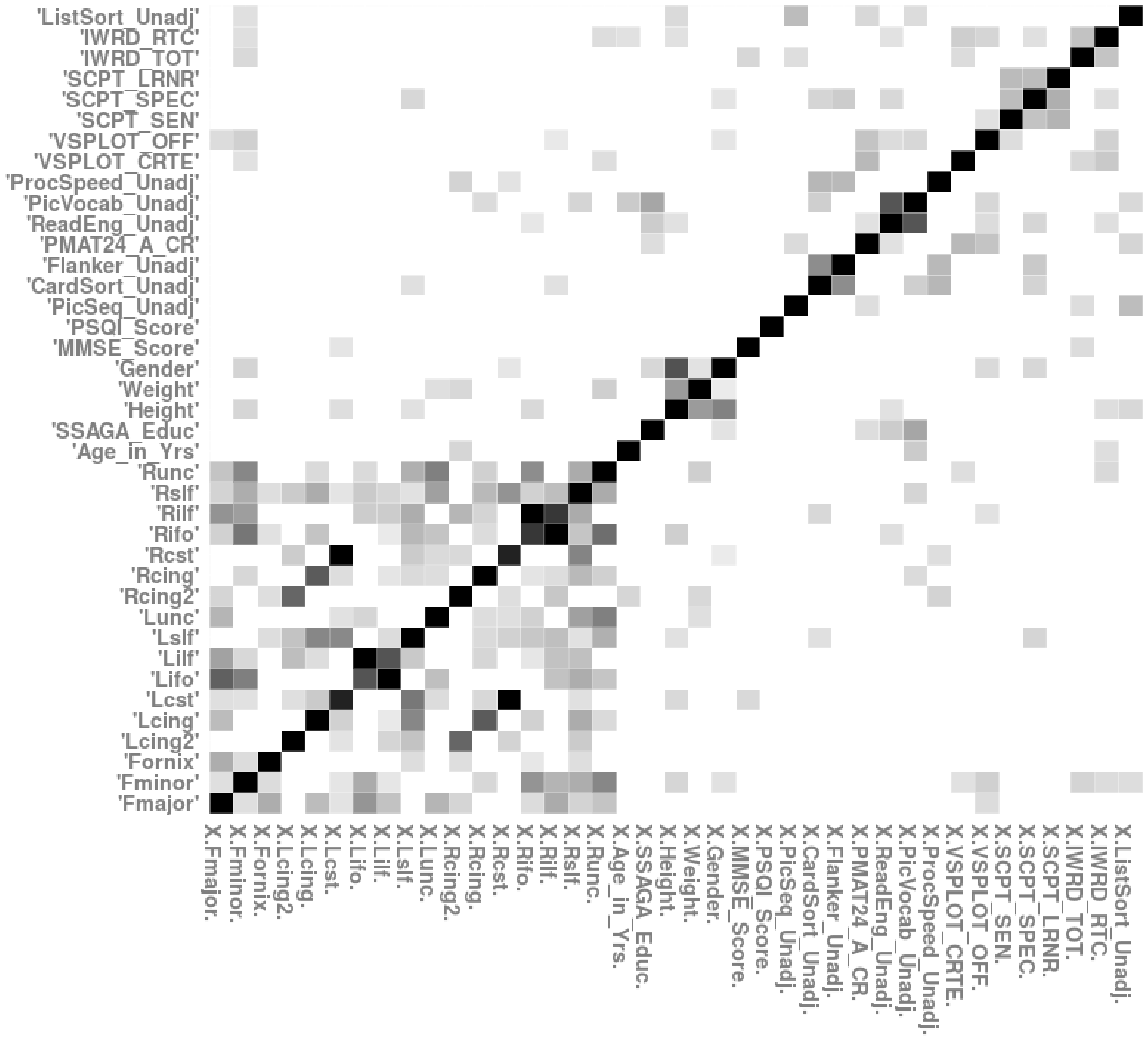
Covariates:

- Demographics, Physical health, Memory, Cognition, etc.

Experimental Result (HCP)

Statistical dependency estimation on HCP dataset

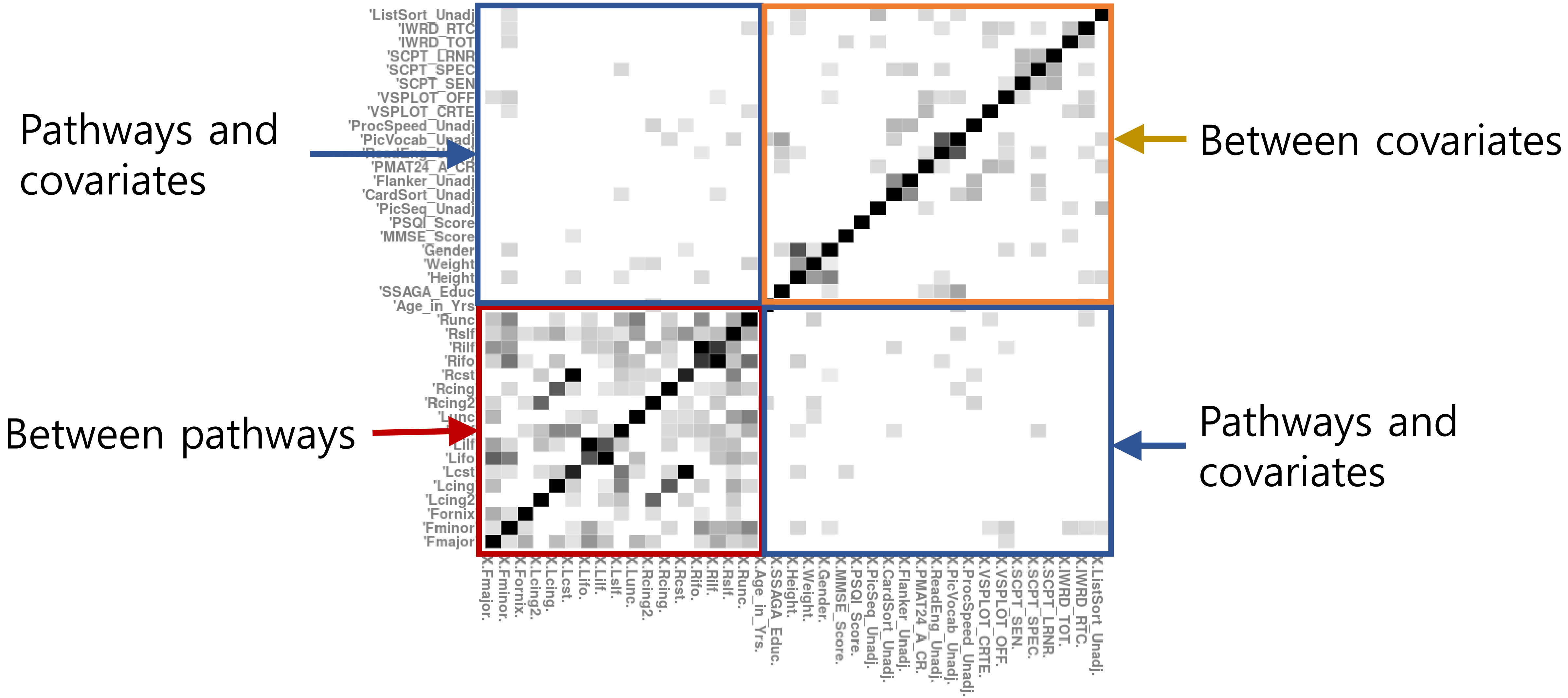
n_o = 60 (i.e., 17 pathways, 22covariates) / unknown latent variables



Experimental Result (HCP)

Statistical dependency estimation on HCP dataset

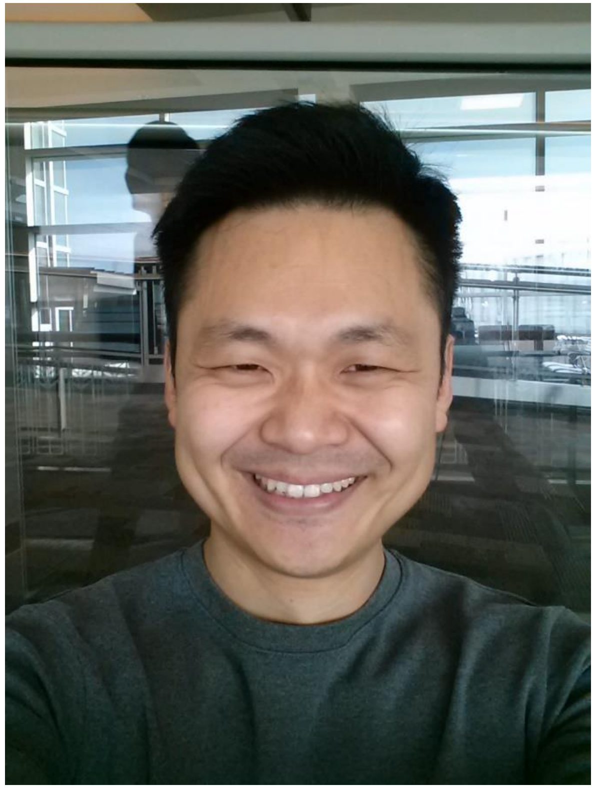
n_o = 39 (i.e., 17 pathways, 22covariates) / unknown latent variables



Find us at our poster for details!

Latent Variable Graphical Model Selection using Harmonic Analysis: Applications to the Human Connectome Project (HCP)

Tue. at 4:45pm (poster #20)



Won Hwa Kim

Hyunwoo J. Kim