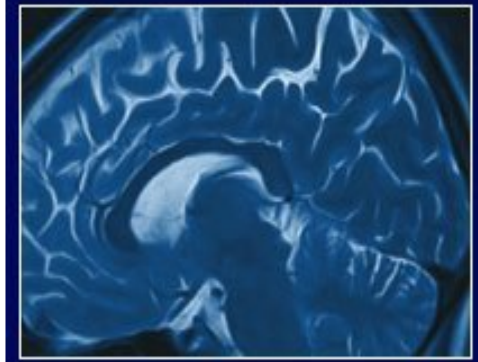
University of Wisconsin
SCHOOL OF MEDICINE AND PUBLIC HEALTH



The Waisman Laboratory for Brain Imaging and Behavior

February 1, 2016 Banff Station

Learning Large-scale Brain Networks for Twin fMRI

Moo K. Chung

Department of Biostatistics and Medical Informatics
Waisman Laboratory for Brain Imaging and Behavior
University of Wisconsin-Madison

www.stat.wisc.edu/~mchung

In collaboration with

Paul J. Rathouz
Department of Biostatistics and Medical Informatics
University of Wisconsin-Madison

Victoria G. Vilalta, David H. Zald
Department of Psychology
Vanderbilt University

Benjamin B. Lahey
Department of Public Health Sciences
University of Chicago

Functional MRI

11 monozygotic (MZ) pairs

14 dizygotic (DZ) pairs

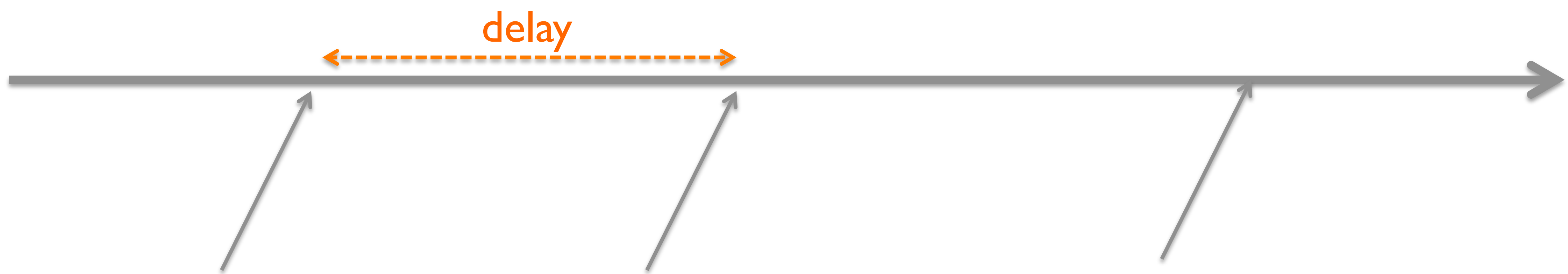
9 same-sex DZ pairs (5 male, 4 female)

5 different-sex DZ pairs

MZ-pair

DZ-pair

Reward experiment: 120 randomized trial



Cue for reward
$0, $1, $5

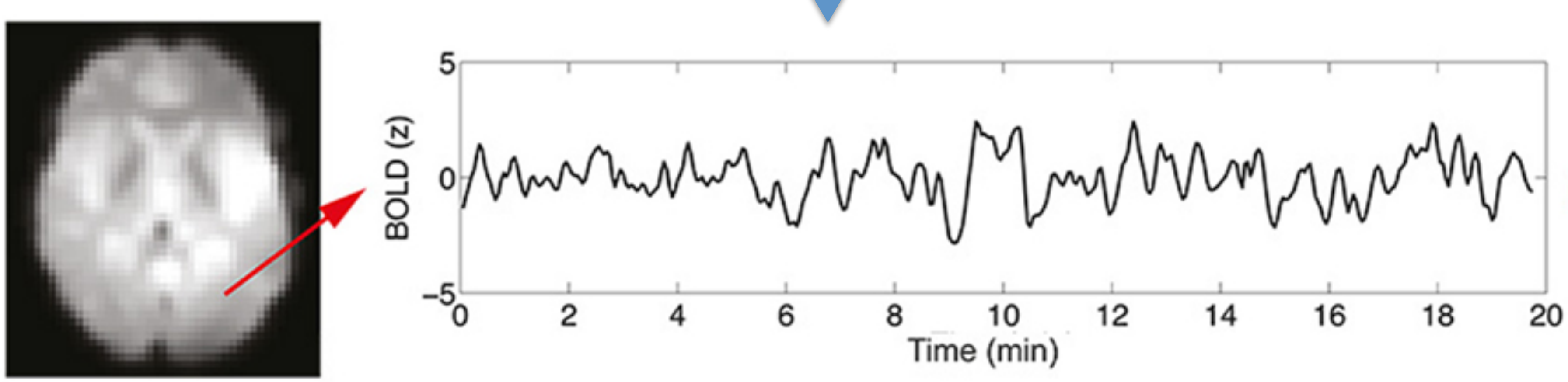
White star flashes
on screen.

Win if the response
button is pushed in time

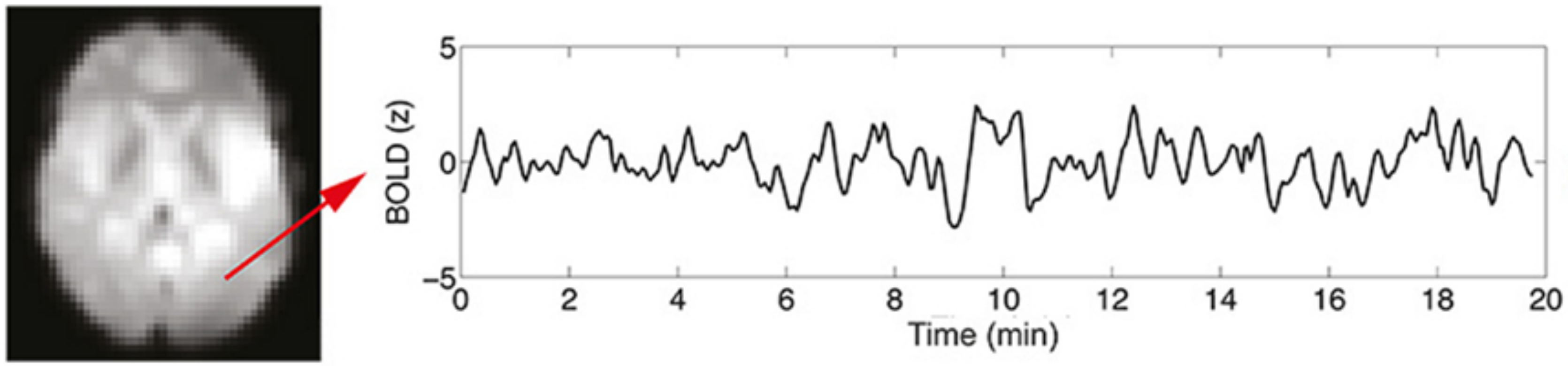
Feedback

9 conditions (each condition also includes a time and a dispersion derivative):

- delay for $0 trials
- delay for $1 trials
- delay for $5 trials
- reward $0
- reward $1
- reward $5
- miss $0
- miss $1
- miss $5



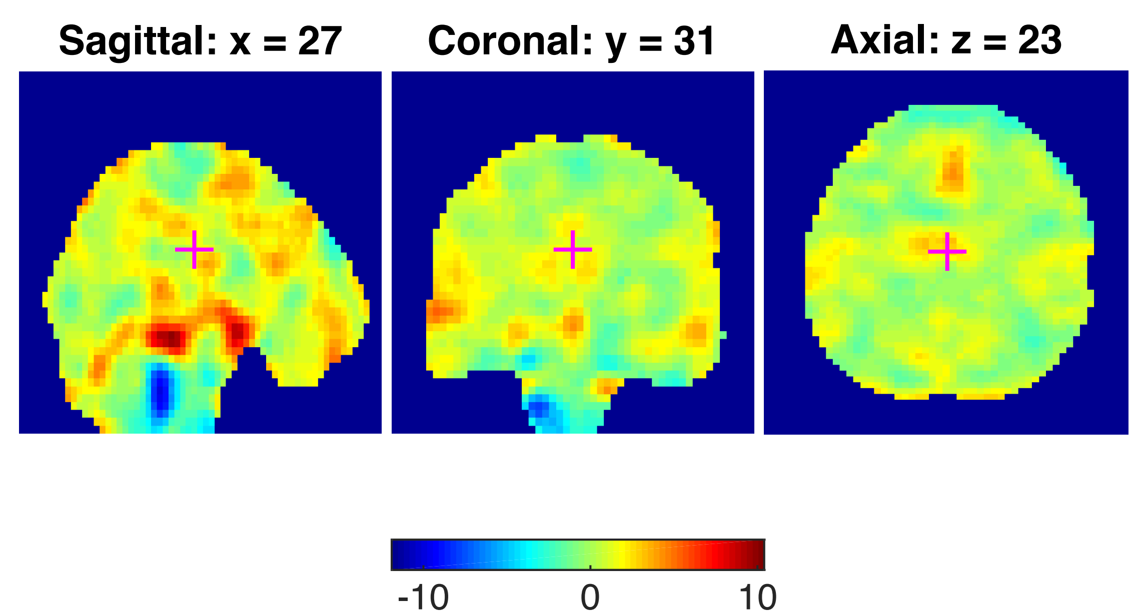
1^{st} level analysis: general linear model



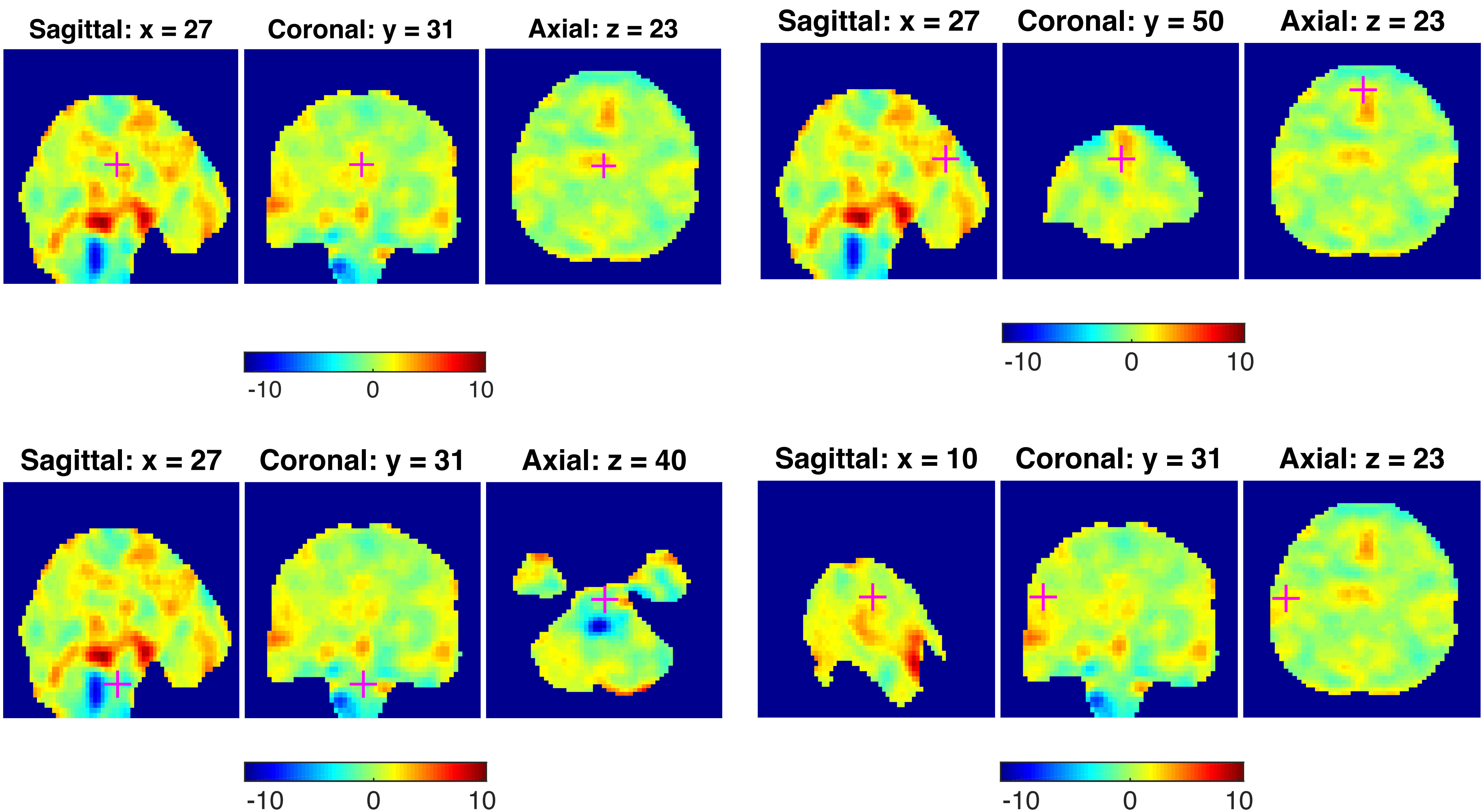
$$Y(v_i) = X\beta(v_i) + \varepsilon(v_i)$$

BOLD

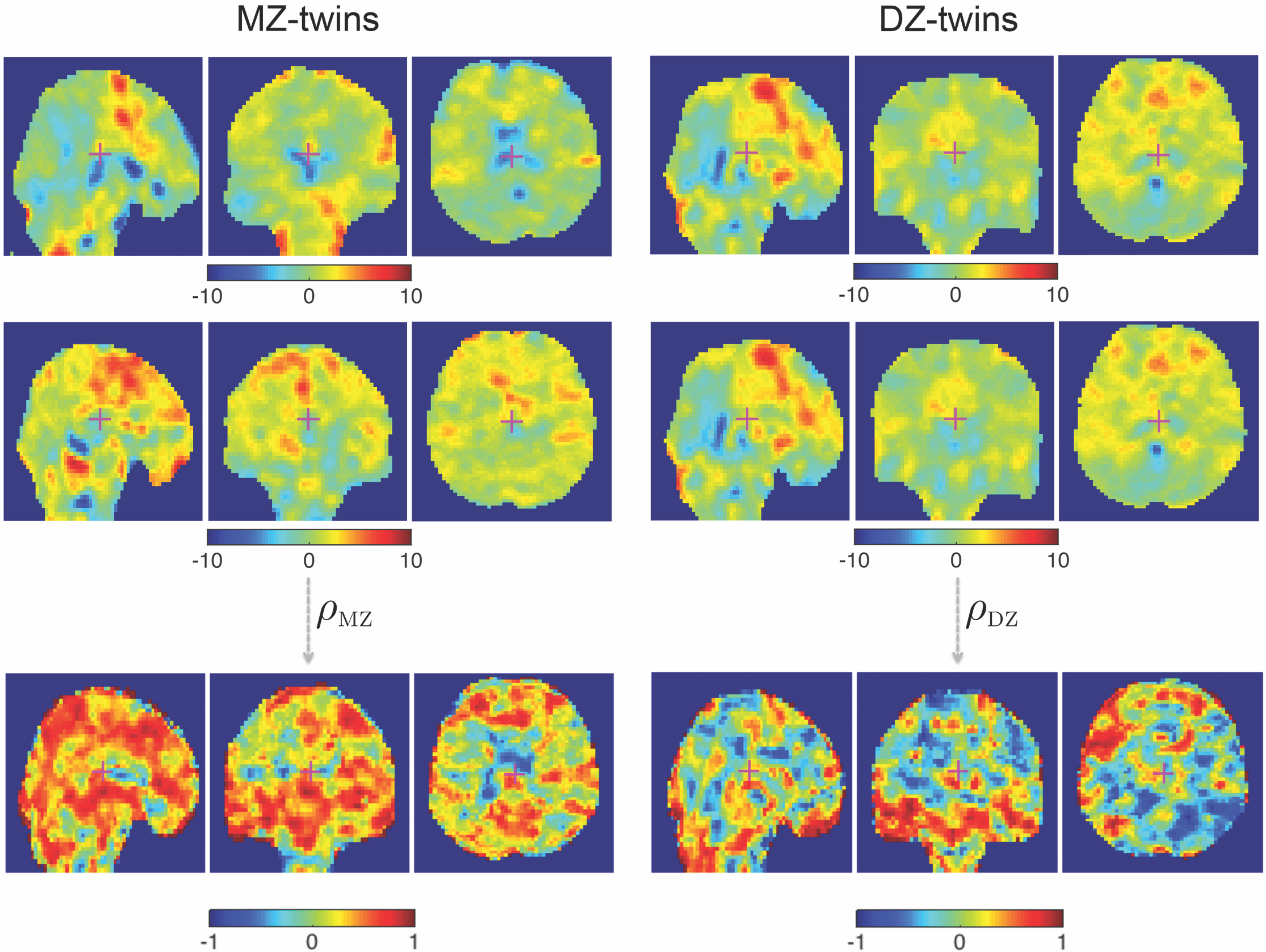
Contrast map: $c^T\beta(v_i)$



Contrast map: one subject



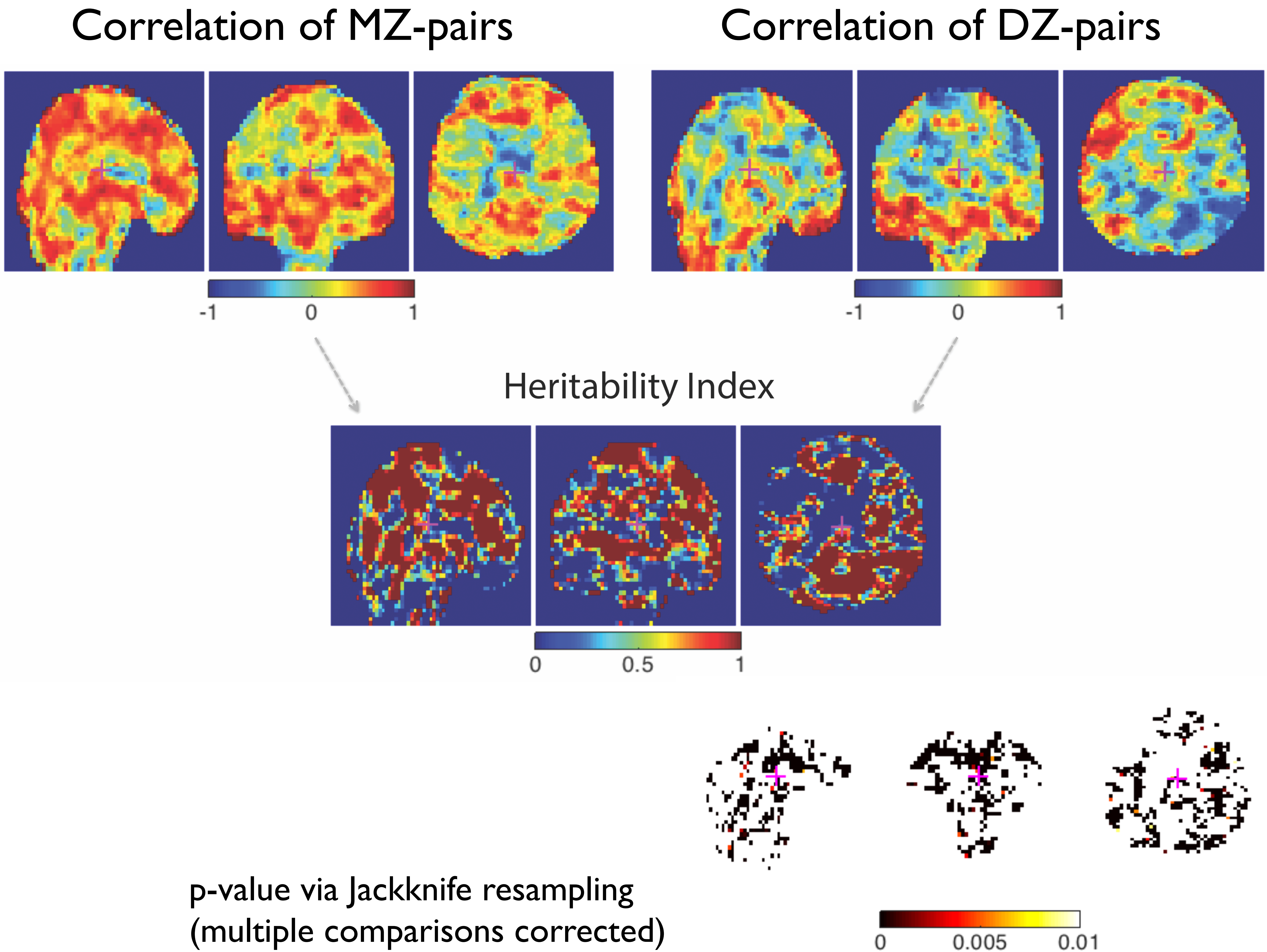
Heritability index


MZ-twins
DZ-twins
-10
0
10
ρ_{MZ}
ρ_{DZ}
-1
0
1

Heritability index (HI) determines the amount of variation due to genetic influence as percentage (%).

Falconer's formula at voxel v_i :

$$\mathrm{HI}(v_i) = 2\left[\rho_{\mathrm{MZ}}(v_i) - \rho_{\mathrm{DZ}}(v_i)\right]$$


Correlation of MZ-pairs
Correlation of DZ-pairs
-1
0
1
-1
0
1
Heritability Index
0
0.5
1
p-value via Jackknife resampling
(multiple comparisons corrected)
0
0.005
0.01

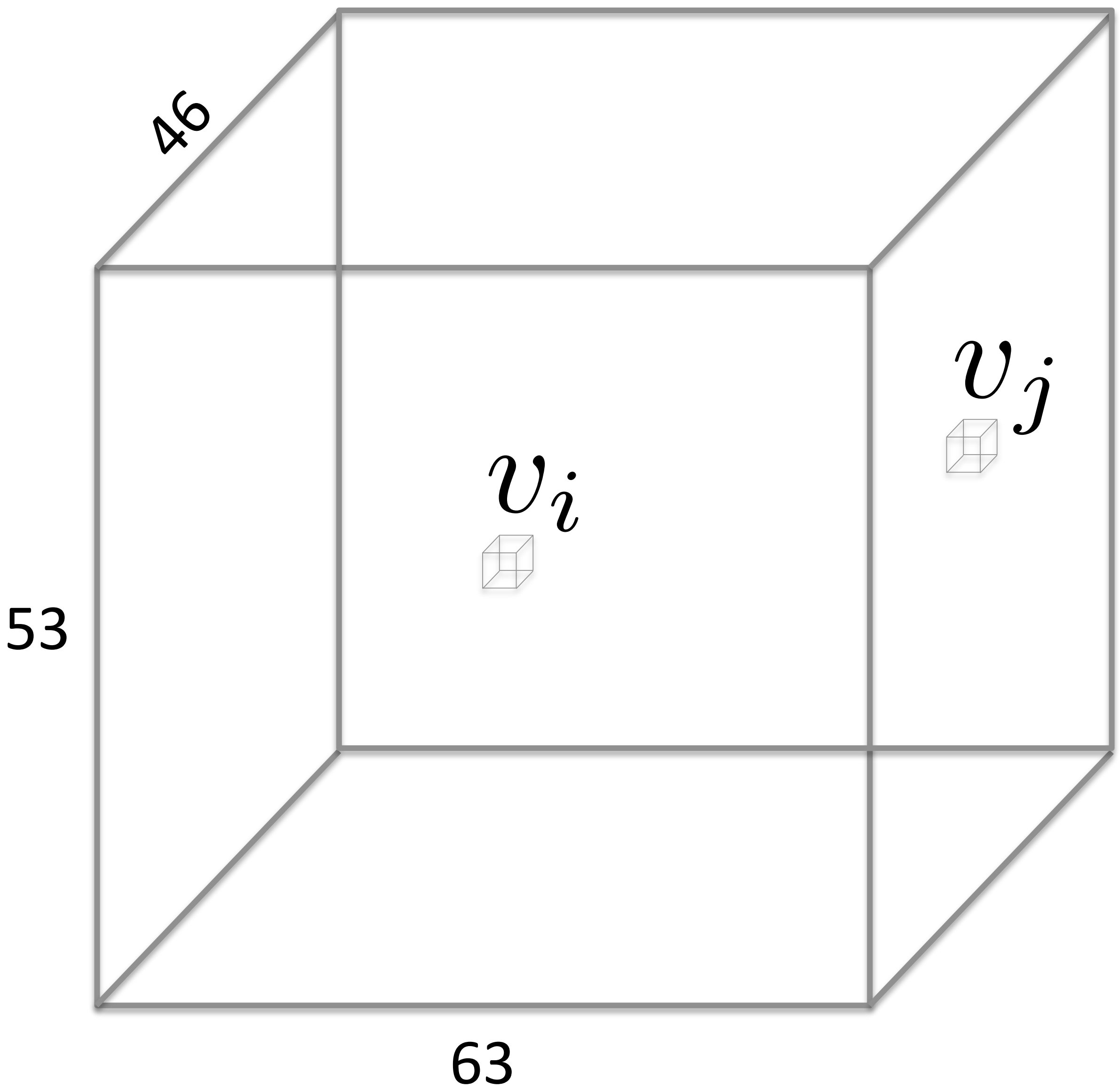
Three aims:

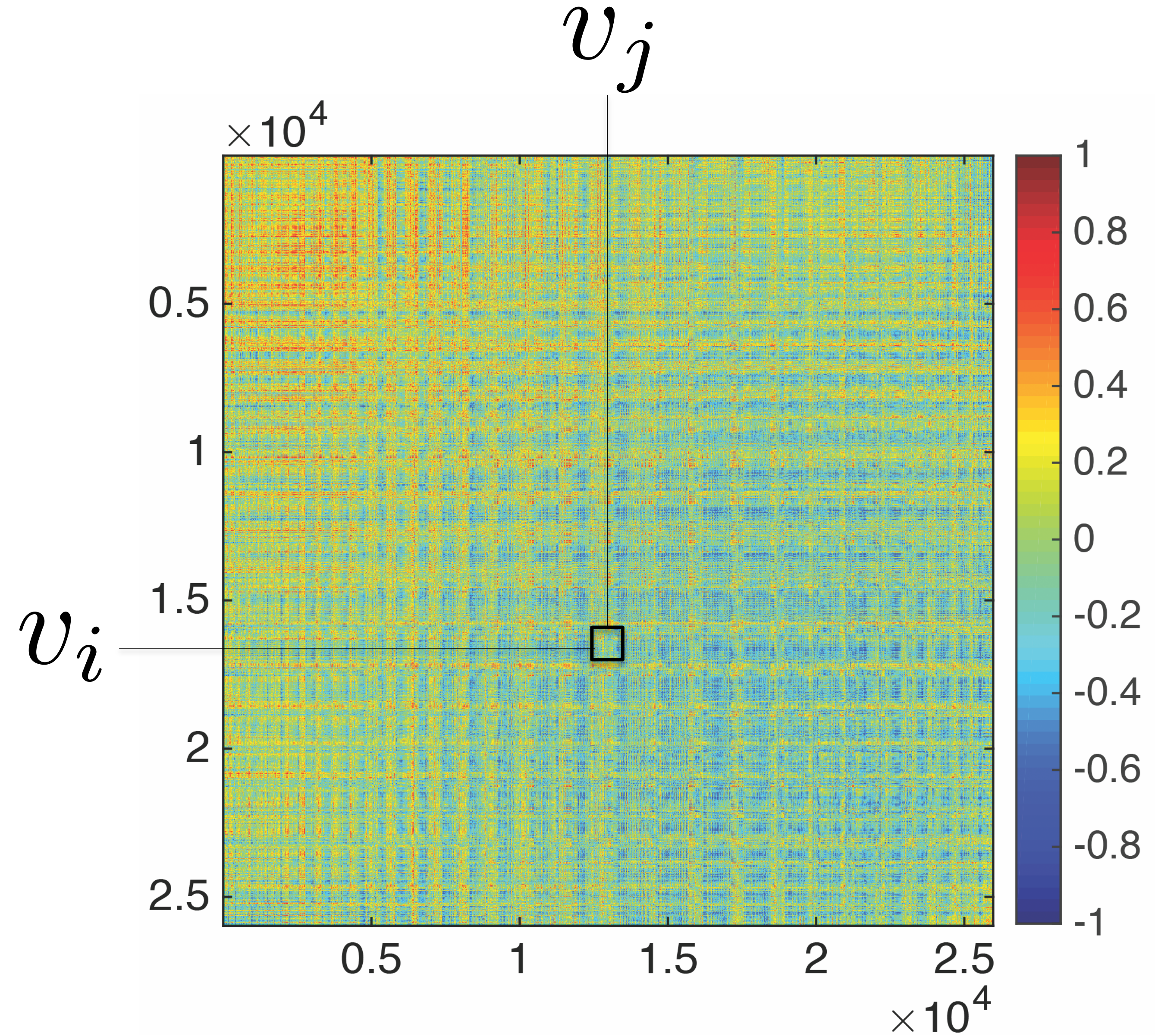
1) Construct a large-scale brain network

2) Compute HI of the large-scale network

3) Determine the statistical significance of HI

Large-scale connectivity analysis





p=(53 x 63 x 46) x (53 x 63 x 46)

Node set $V = \{v_1, v_2, \cdots, v_p\}$

(53 x 63 x 46) x (53 x 63 x 46) = 1.3 billion directed edges

Sparse Cross Correlations

Center and scaled data

n paired images in *p* voxels (*n* < *p*)

$x_k(v_i)$ *k*-th image intensity value at voxel v_i

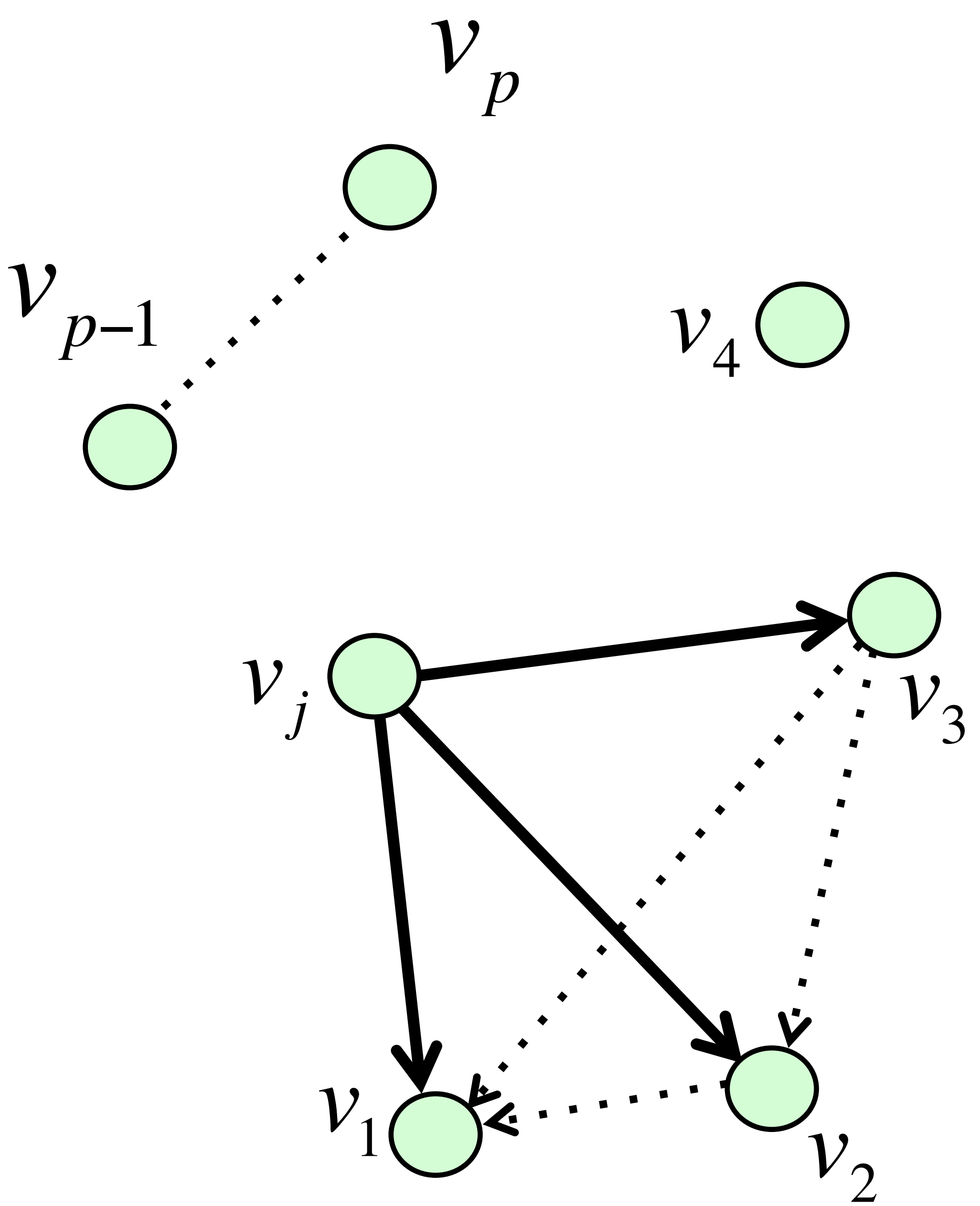
$y_k(v_i)$ *k*-th image intensity value at voxel v_i

$$\sum_{k=1}^{n} x_k(v_i) = \sum_{k=1}^{n} y_k(v_i) = 0$$

$$x = (x_1, \cdots, x_n)' \qquad y = (y_1, \cdots, y_n)'$$

$$\|x\|^2 = x'x = \|y\|^2 = y'y = 1$$

Massive regressions between nodes



$$y(v_j) = \sum_{i \neq j} \beta_{ij} x(v_i) + \varepsilon$$

$$y(v_j) = \beta_{ij} x(v_i) + \varepsilon$$

Least-squares estimation:

$$\hat{\beta}_{ij} = x'(v_i) y(v_j)$$

1.3 billion cross-correlations

Sparse cross-correlation network

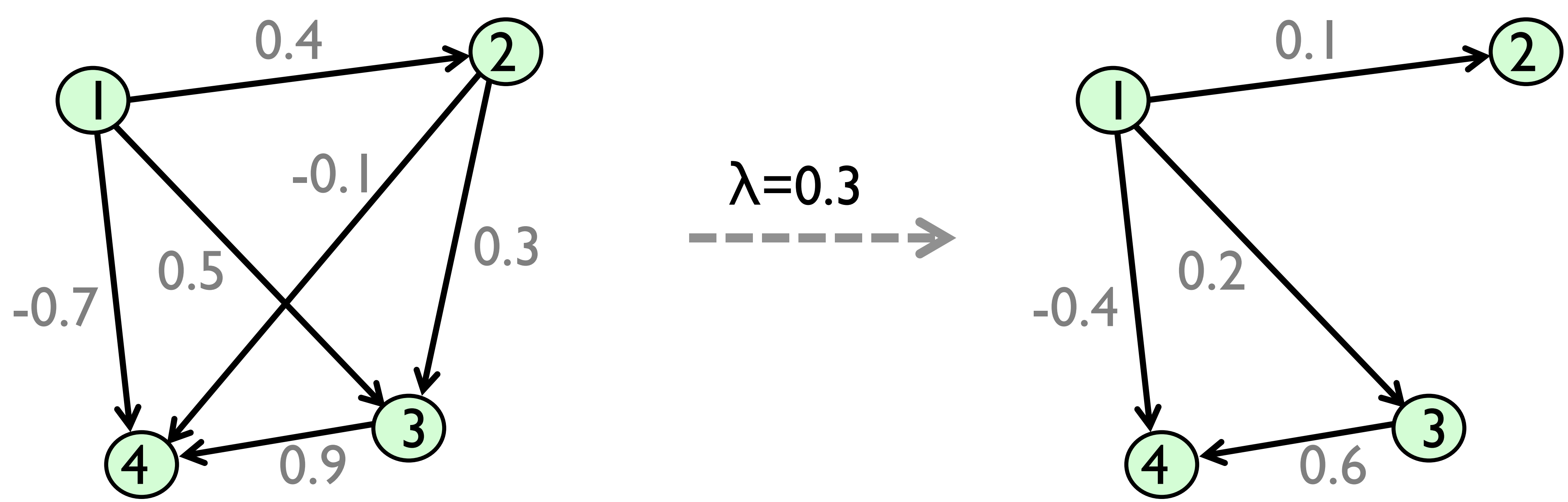
$$\hat{\beta} = \operatorname{argmin}_{\beta} \ \frac{1}{2}\sum_{i=1}^{p}\sum_{j=1}^{p}\left\| x(v_i) - \beta_{ij} y(v_j) \right\|^2 + \lambda \sum_{i,j} |\beta_{ij}|$$

Adjacency matrix

$$b_{ij}(\lambda) = \begin{cases} 1 \text{ if } \hat{\beta}_{ij} \neq 0 \\ 0 \text{ otherwise} \end{cases}$$

Sparse network

$$G(\lambda) = \{V, b(\lambda)\}$$



Soft-thresholding

LASSO
L1-optimization

$$b_{ij}(\lambda) = \begin{cases} 1 \text{ if } \hat{\beta}_{ij}(\lambda) \neq 0 \\ 0 \text{ otherwise} \end{cases}$$

$\longleftrightarrow$

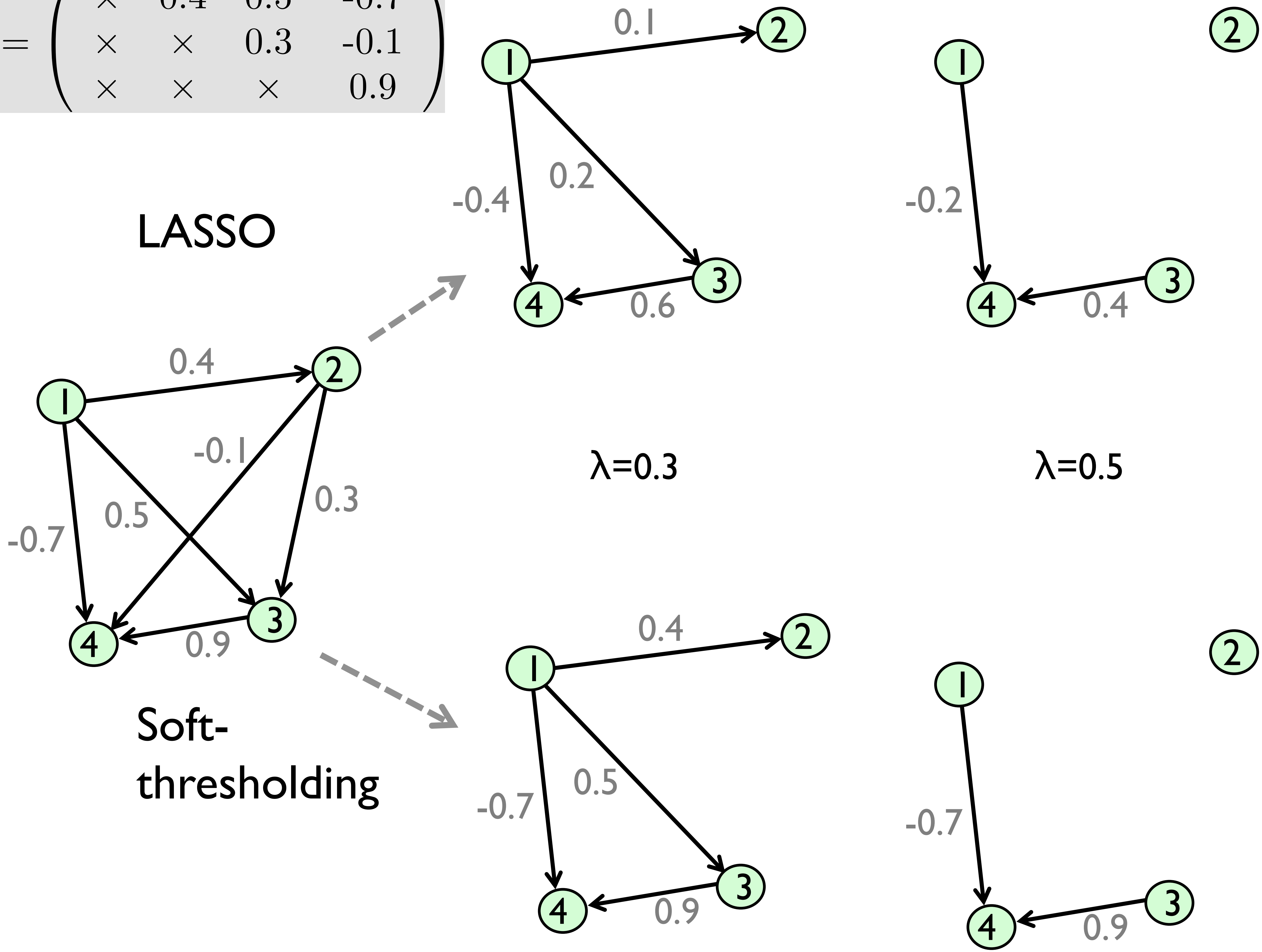
Soft-thresholding

$$b_{ij}(\lambda) = \begin{cases} 1 \text{ if } |x'(v_i)y(v_j)| > \lambda \\ 0 \text{ otherwise} \end{cases}$$

Threshold correlations at $\pm\lambda$ and get the identical sparse network.

Chung et al. 2015 IEEE Trans. Medical Imaging 34:1928-1939

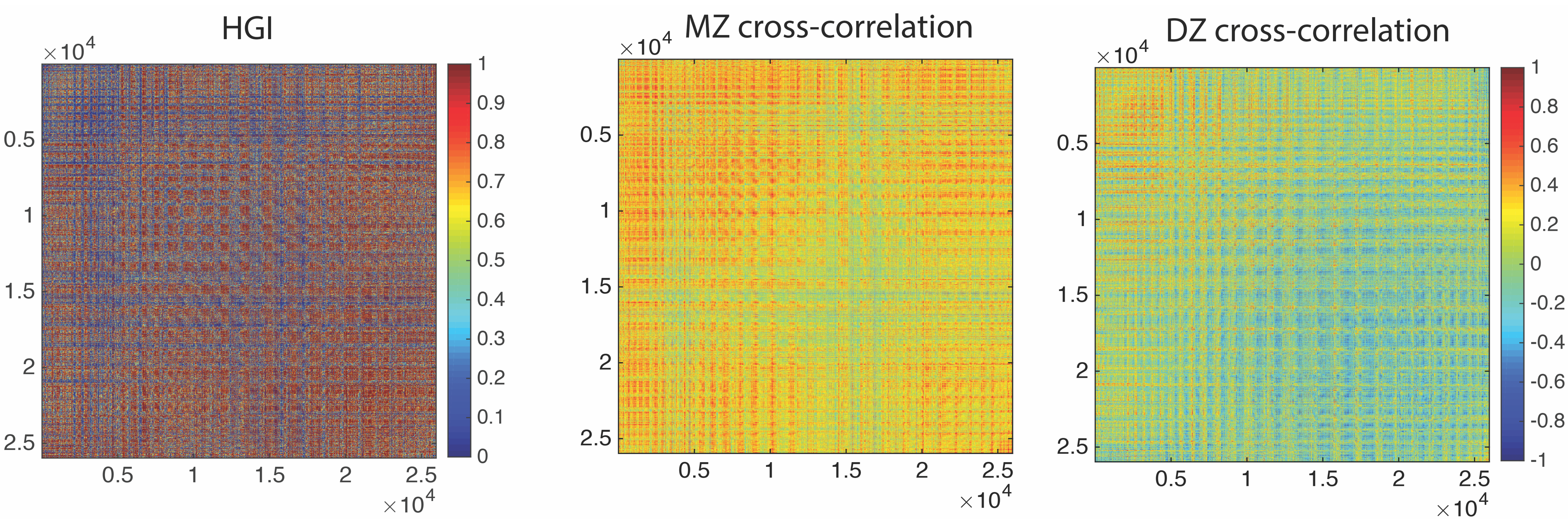
$$\mathbf{x}'\mathbf{y} = \begin{pmatrix} \times & 0.4 & 0.5 & -0.7 \\ \times & \times & 0.3 & -0.1 \\ \times & \times & \times & 0.9 \end{pmatrix}$$



Heritability graph index

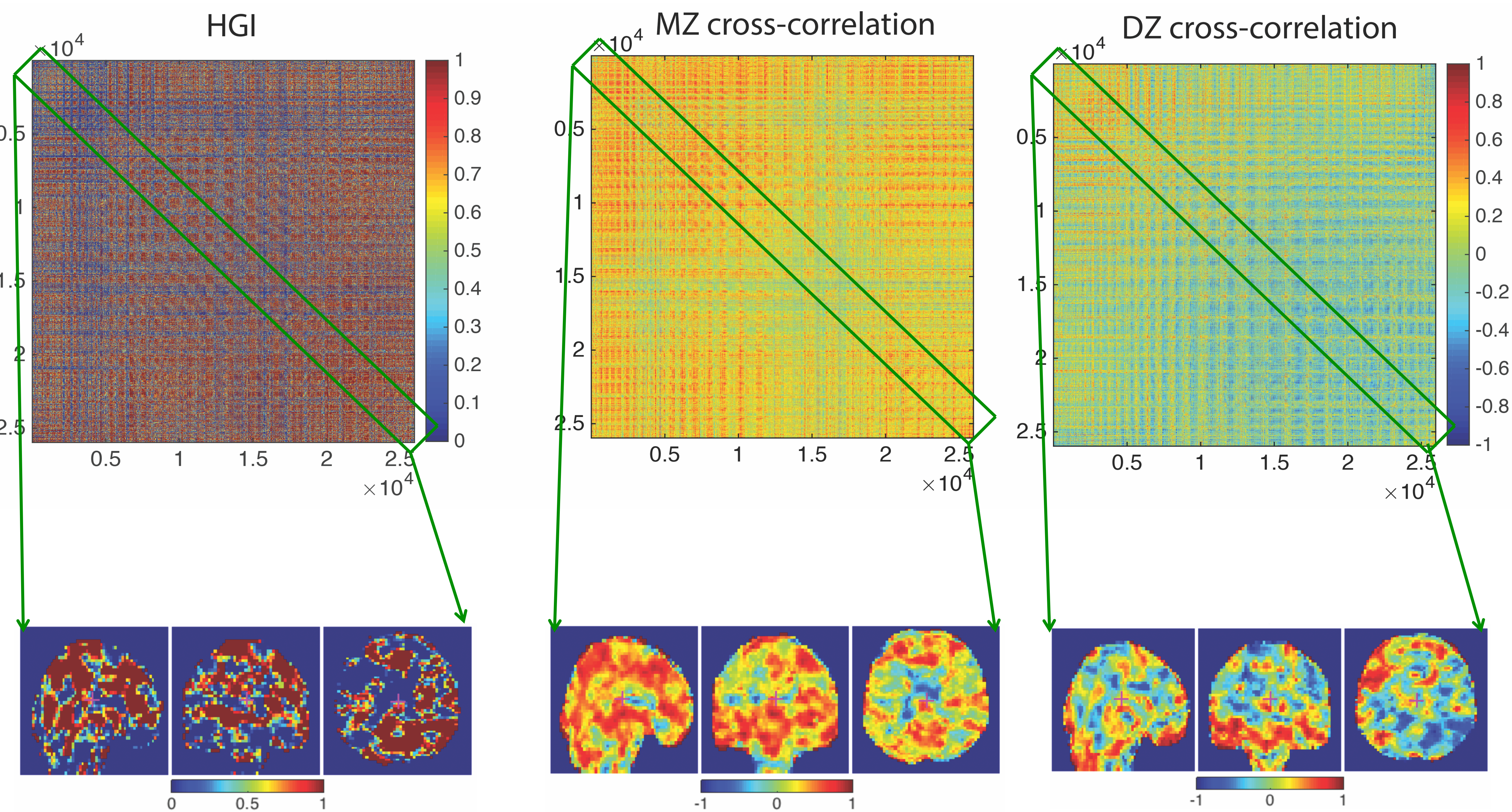
Heritability graph index (HGI):

$$\mathrm{HGI}(v_i, v_j) = 2\left[\rho_{\mathrm{MZ}}(v_i, v_j) - \rho_{\mathrm{DZ}}(v_i, v_j)\right]$$

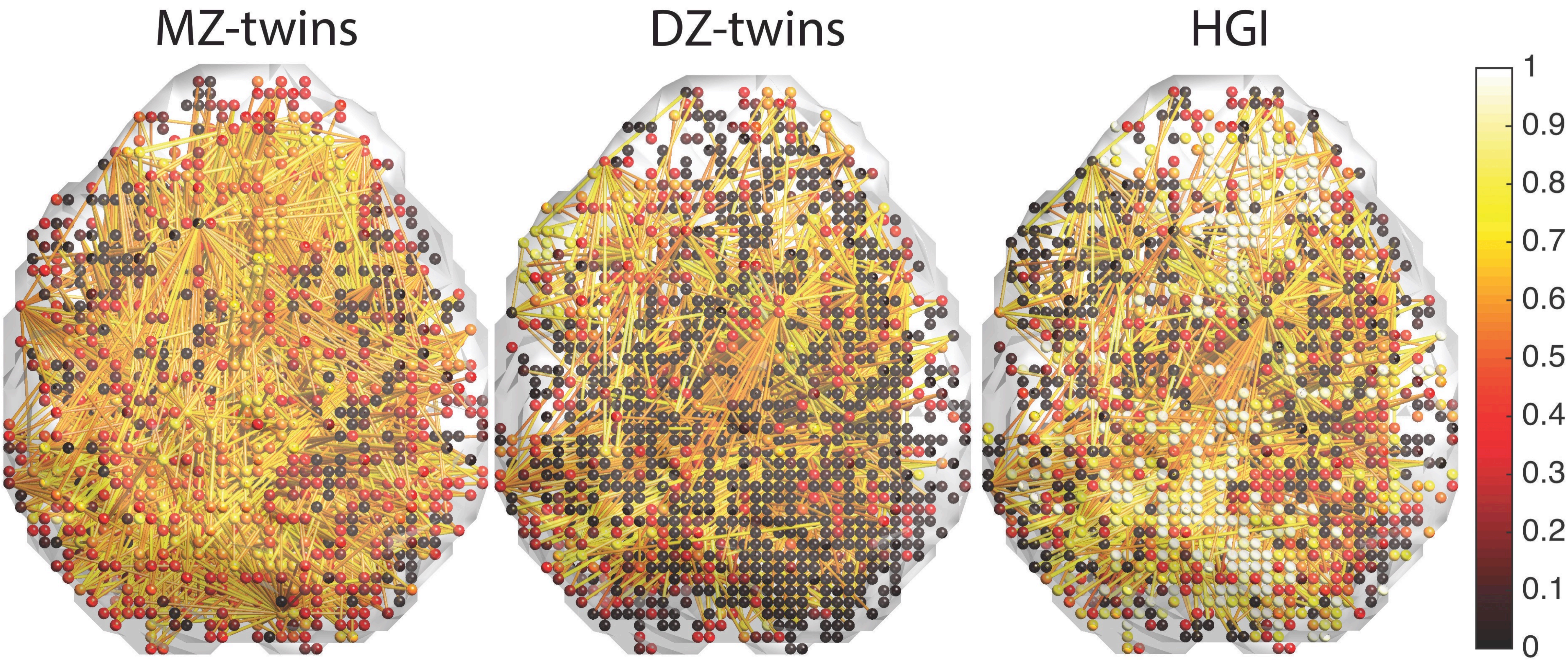


18 sec. computation per matrix
5.2GB per matrix, 3min to save to hard drive

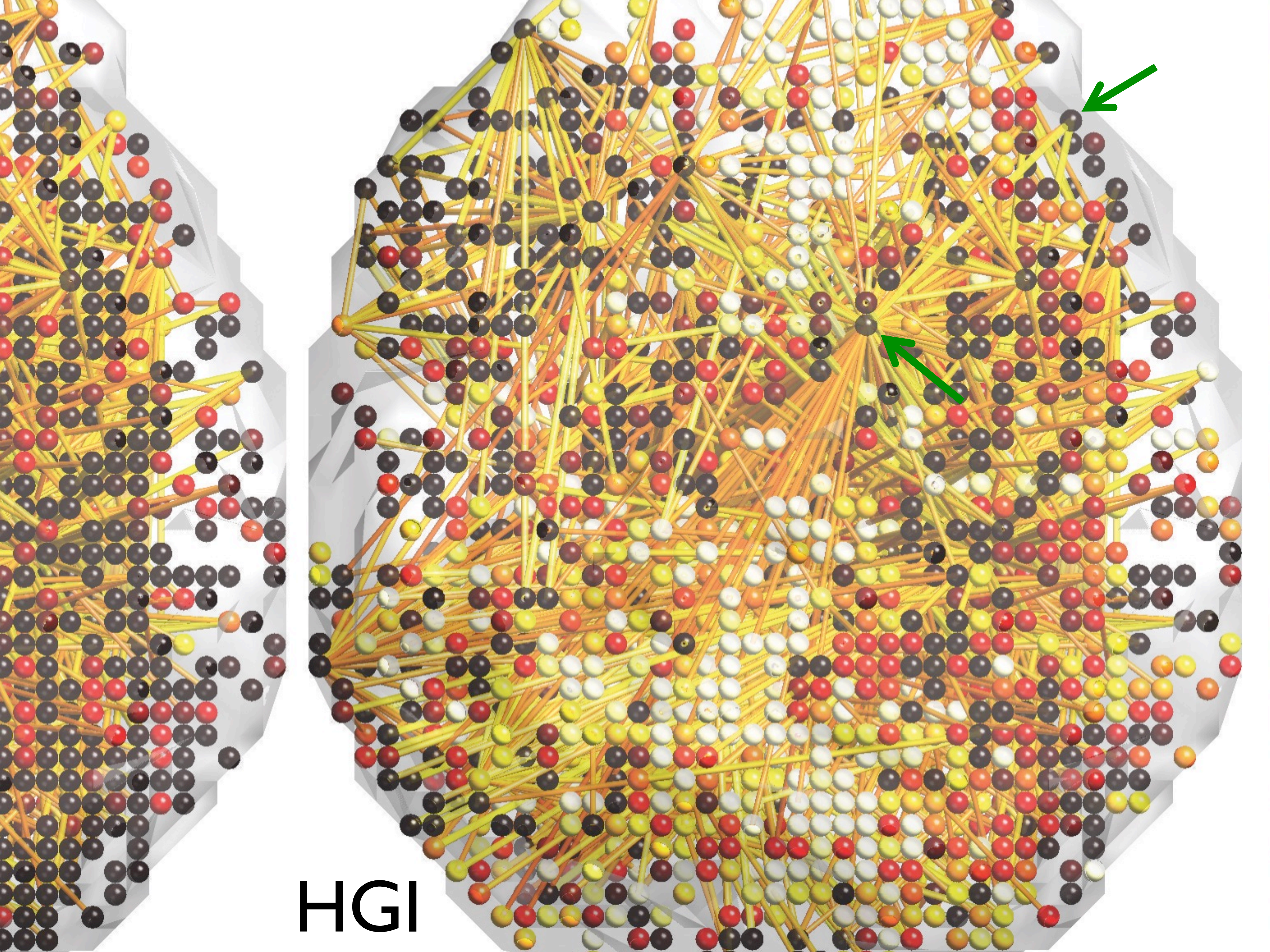
$$\mathrm{HGI}(v_i, v_i) = \mathrm{HI}(v_i)$$



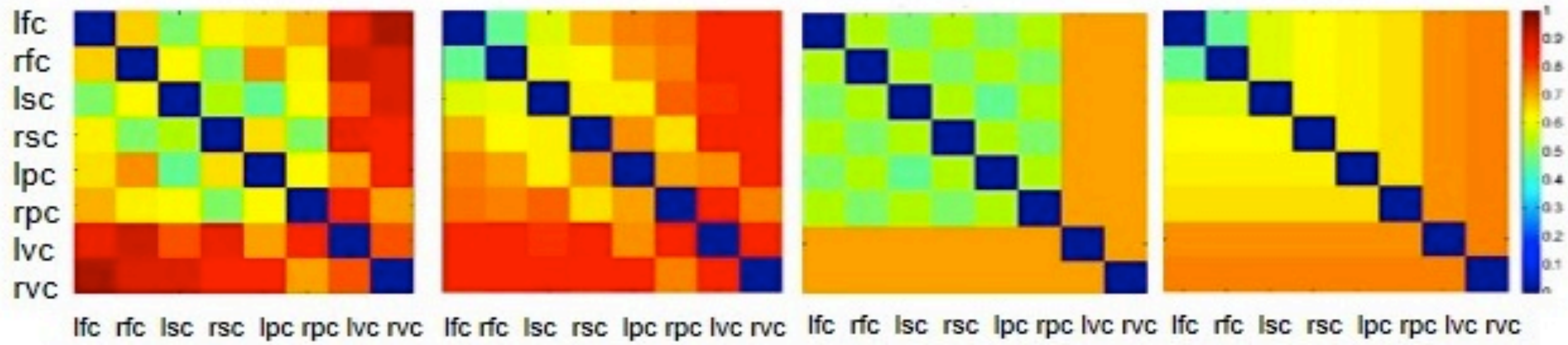
Heritability Graph Index (HGI) at sparse parameter 0.5



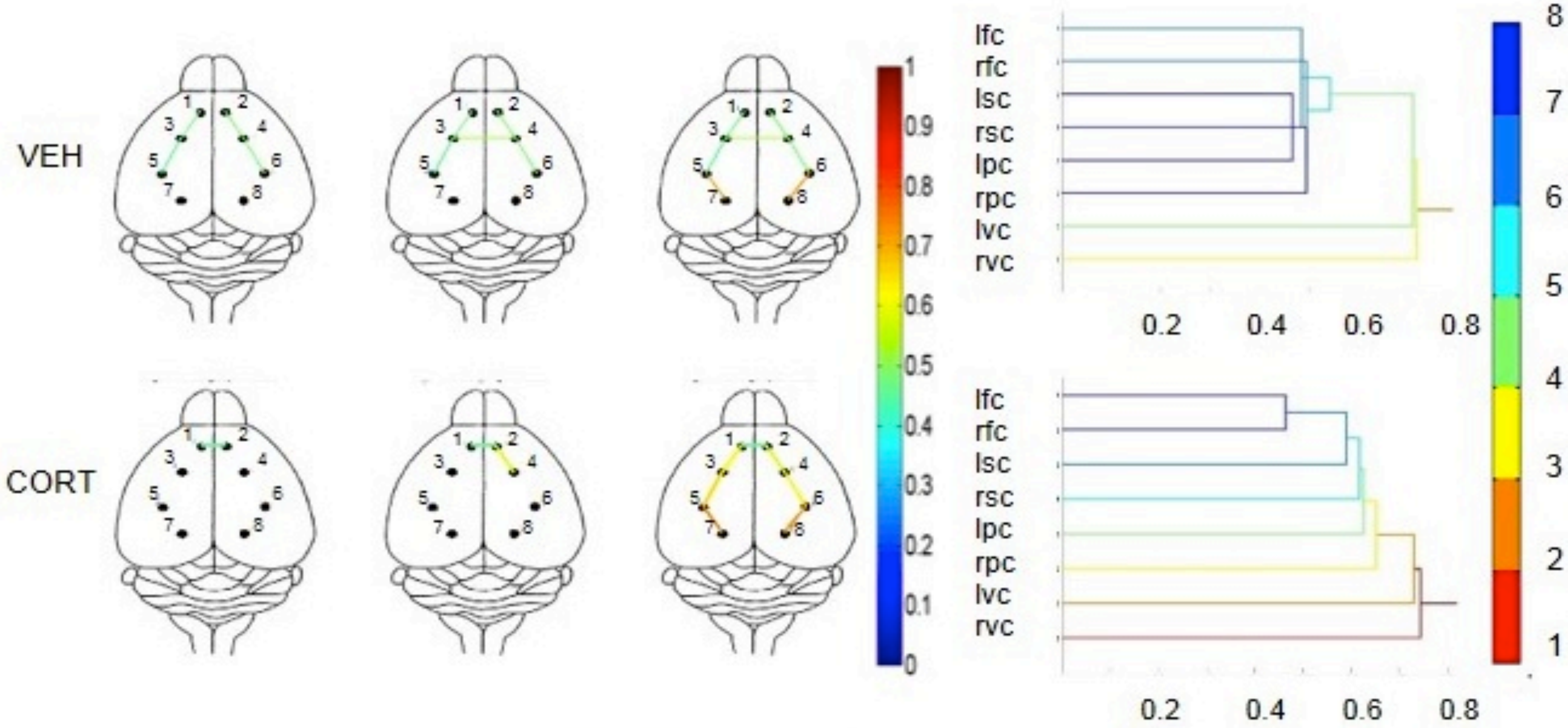
Each voxel is a network node!


HGI

Optimal sparse parameter?

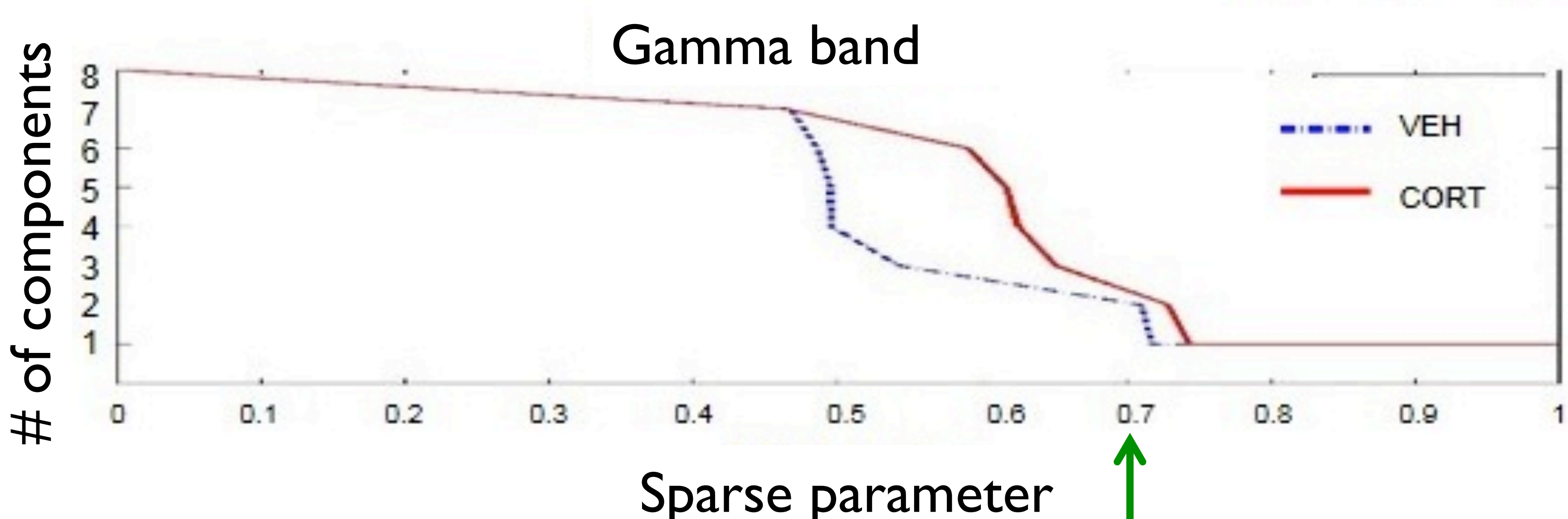


8 channel EEG mouse model of depression



~~Single optimal parameter is not useful!~~

~~Existing Multi-tresholding, Multi-scale methods~~

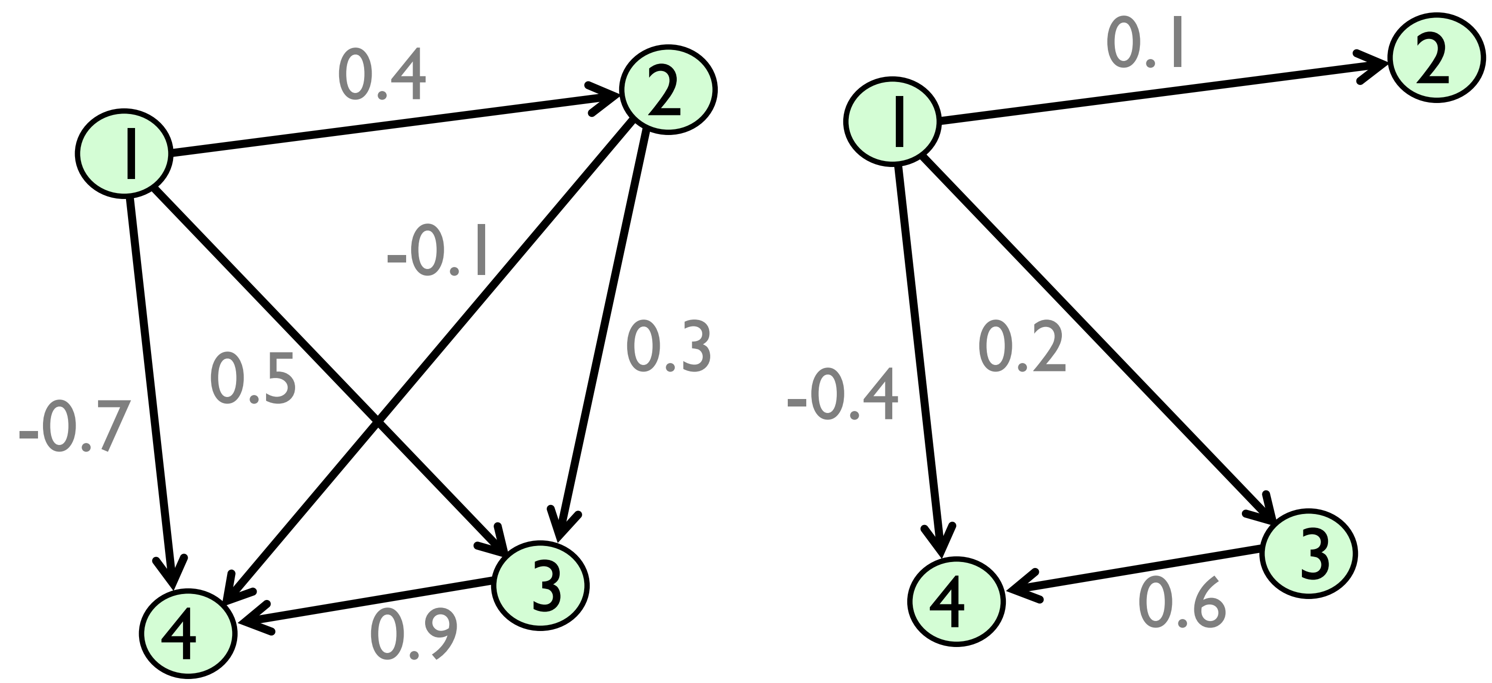


Khalid et al. 2014, NeuroImage 101:351-363

Graph filtration / persistent homology

Analyze the infinite collection of networks

$$\{G(\lambda), \lambda \in \mathrm{R}^{+}\}$$



$$G(0) \supset G(0.3)$$

Graph filtration

$$G(\lambda_1) \supset G(\lambda_2) \supset G(\lambda_3) \supset \cdots$$

for $\lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \cdots$

Lee et al. 2011
MICCAI 6892:302-309

Topological Inference

$H_0 : \mathrm{HGI}(\lambda) = 0$ for all $\lambda \geq 0$

vs.

$H_1 : \mathrm{HGI}(\lambda) \neq 0$ for some $\lambda \geq 0$

$H_0 : \rho_{MZ}(\lambda) = \rho_{DZ}(\lambda)$ for all $\lambda \geq 0$

vs.

$H_1 : \rho_{MZ}(\lambda) \geq \rho_{DZ}(\lambda)$ for some $\lambda \geq 0$

$H_0 : G_{MZ}(\lambda) = G_{DZ}(\lambda)$ for all $\lambda \geq 0$

vs.

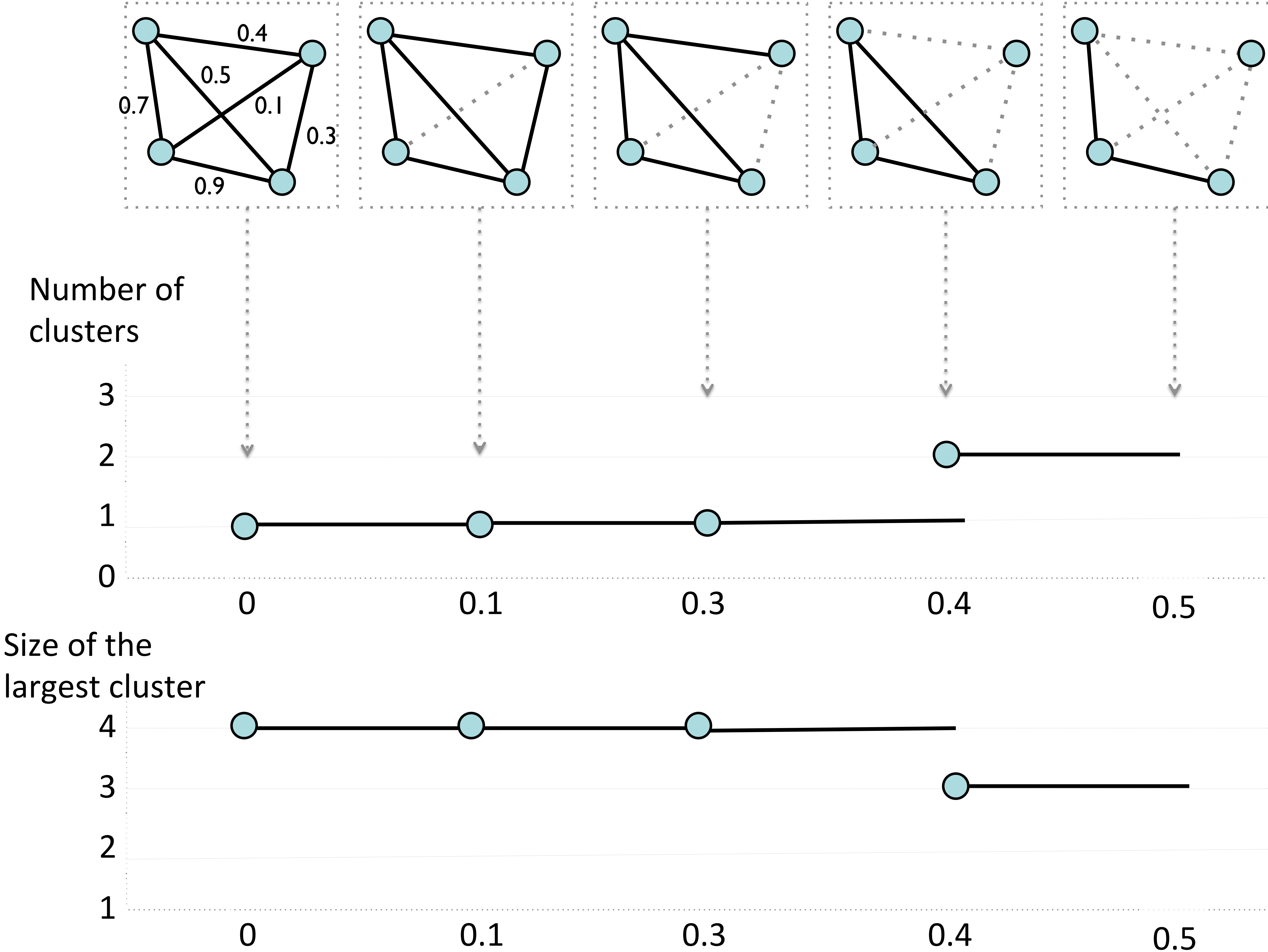
$H_1 : G_{MZ}(\lambda) \neq G_{DZ}(\lambda)$ for some $\lambda \geq 0$

$H_0 : f(G_{MZ})(\lambda) = f(G_{DZ})(\lambda)$ for all $\lambda \geq 0$

vs.

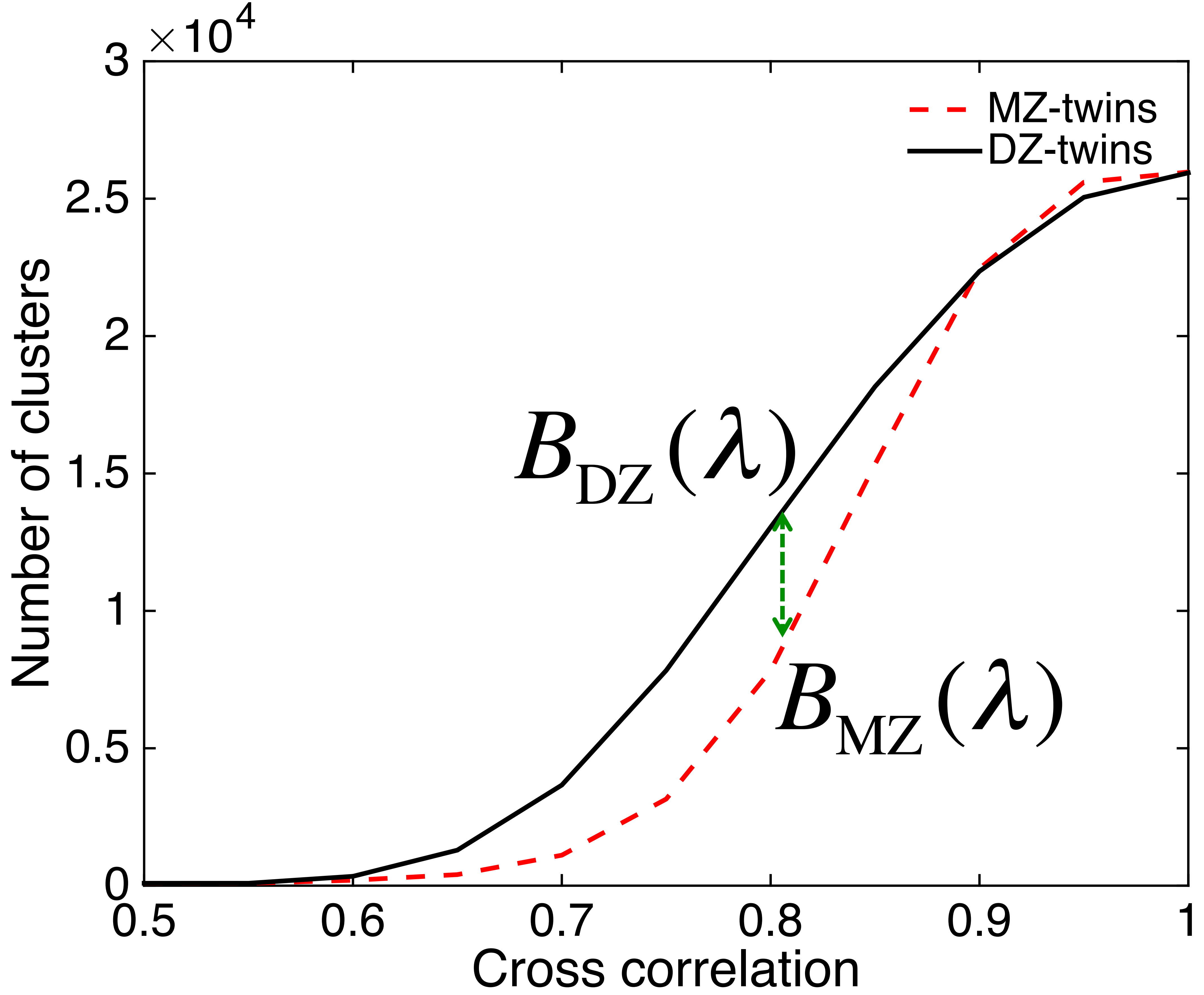
$H_1 : f(G_{MZ})(\lambda) \neq (G_{DZ})(\lambda)$ for some $\lambda \geq 0$

with monotonic graph function $f(G)(\lambda_1) \leq f(G)(\lambda_2) \leq f(G)(\lambda_3) \leq \ldots$

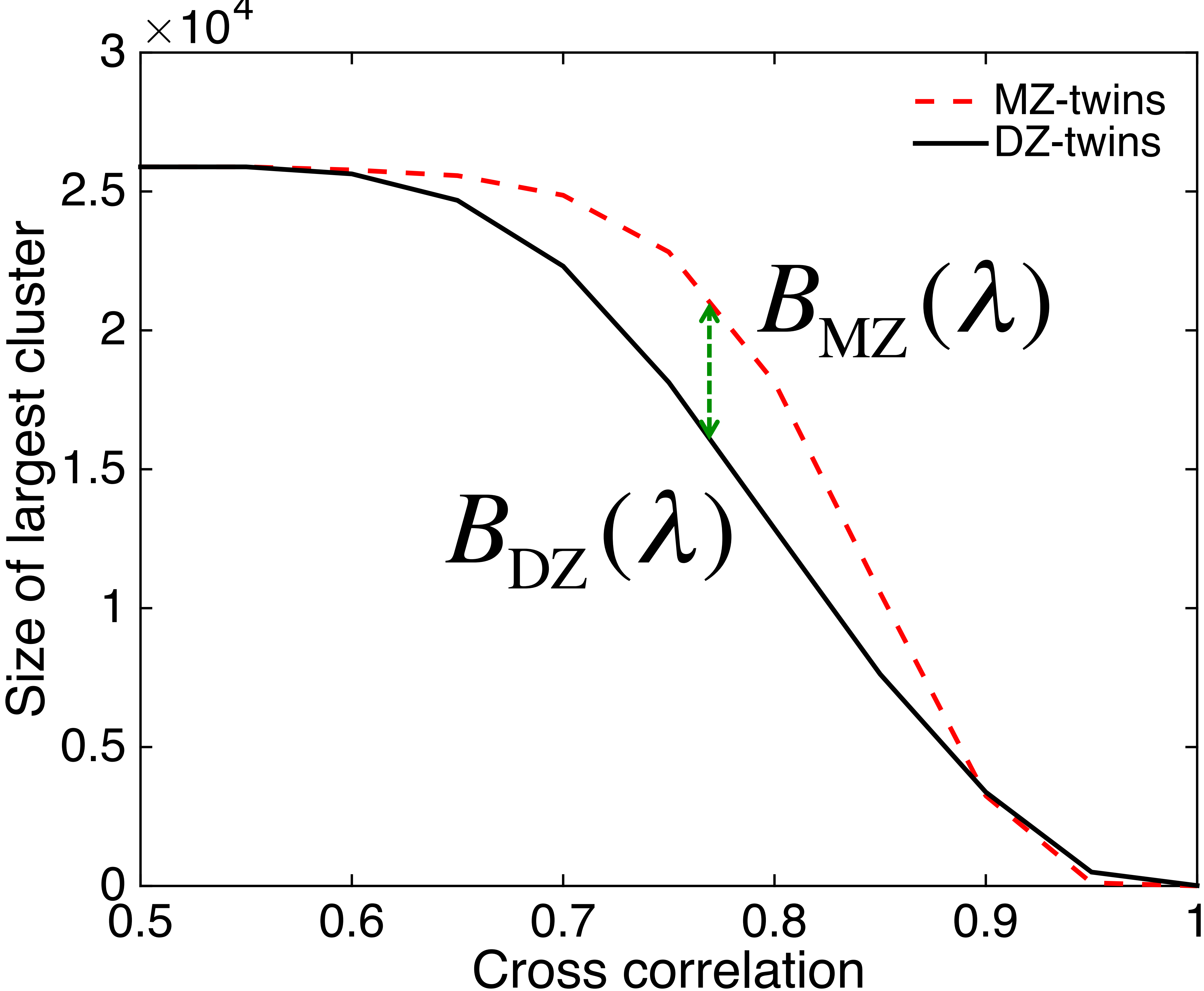

0.4
0.5
0.7
0.1
0.3
0.9
Number of clusters
3
2
1
0
0
0.1
0.3
0.4
0.5
Size of the largest cluster
4
3
2
1
0
0.1
0.3
0.4
0.5

Betti plots

$$D = \min_{\lambda} \left| B_{MZ}(\lambda) - B_{DZ}(\lambda) \right|$$



p-value < 0.0002



p-value < 0.0001

$$p\text{-value} = P\left(\max_{\lambda} \left| B_{\mathrm{MZ}}(\lambda) - B_{\mathrm{DZ}}(\lambda) \right| \geq \sqrt{2(p-1)}d\right) \approx 2e^{-2d^2} + \ldots$$

Chung et al. 2015 arxiv:509.04771

Discussion

Multiscale methods are usually better than monoscale methods. Infinite scale methods will be even better.

Thresholding feature/connectivity is not a bad idea.

The method can be easily generalized to other sparse models: full LASSO with LARS.

Thank you