The Waisman Laboratory for Brain Imaging and Behavior

University of Wisconsin SCHOOL OF MEDICINE AND PUBLIC HEALTH

Exact Topological Inference of the Resting-State Brain Network in Twins

Moo K. Chung

Department of Biostatistics and Medical Informatics

University of Wisconsin-Madison

www.stat.wisc.edu/~mchung

June18, 2019, Computational Topology

Abstract

A cycle in a brain network is a subset of a connected component with redundant additional connections. If there are many cycles in a connected component, the connected component is more densely connected. While the number of connected components represents the integration of the brain network, the number of cycles represents how strong the integration is. However, it is unclear how to perform statistical inference on the number of cycles in the brain network. In this lecture, we present a new Exact Topological Inference framework for determining the statistical significance of the number of cycles through the Kolmogorov-Smirnov (KS) distance, which was recently introduced to measure the similarity between networks across different filtration values using the zeroth Betti number. We show how to extend the method to the first Betti number. Using a twin imaging study, which provides biological ground truth, the methods are applied in determining if cycles are heritable network features in the resting-state functional brain networks of 217 twins. This talk is based on a paper of the same title: doi.org/10.1162/netn_a_00091. The MATLAB codes as well as the connectivity matrices used in the paper are freely available at www.stat.wisc.edu/~mchung/TDA.

Codes, data & lecture slides given in

www.stat.wisc.edu/~mchung/TDA

More codes & published brain imaging data given in

https://www.stat.wisc.edu/~mchung/software.html

Acknowledgement

Yixian Wang, Shih-Gu Huang, Andrey Grisenko, Ross Luo, Nagesh Adluru, Andrew Alexander, Richard Davidson, Hill Goldsmith
University of Wisconsin-Madison, USA

Yuan Wang *University of South Carolina*
Hyekyung Lee *Seoul National University*
Hernando Ombao *KAUST*

NIH grants: R01 EB022856, R01 MH101504, P30 HD003352, U54 HD09025

Full day course

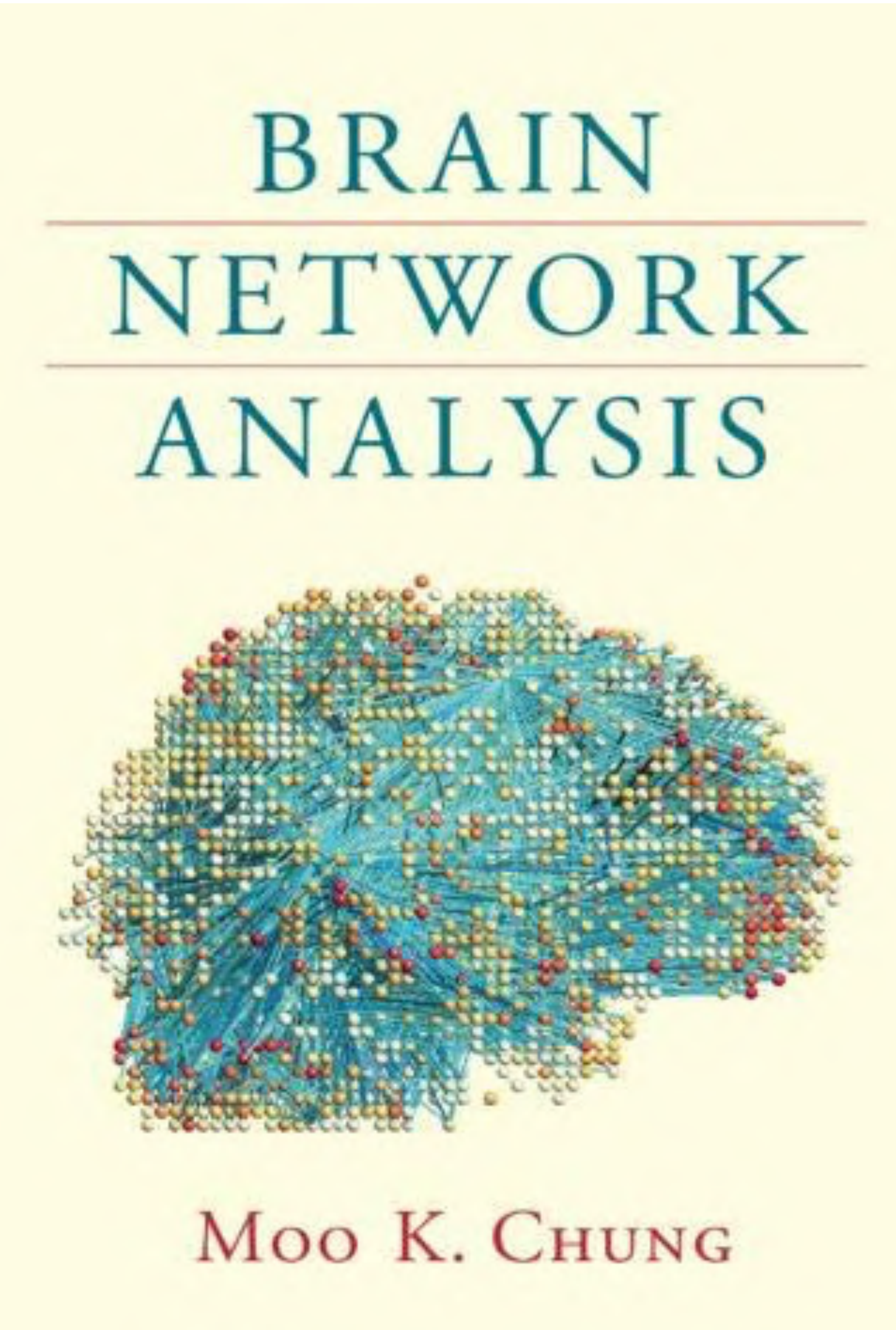
Topological and Object Oriented Data Analysis

International Biometric Conference (IBC2020)
COEX Seoul, Korea

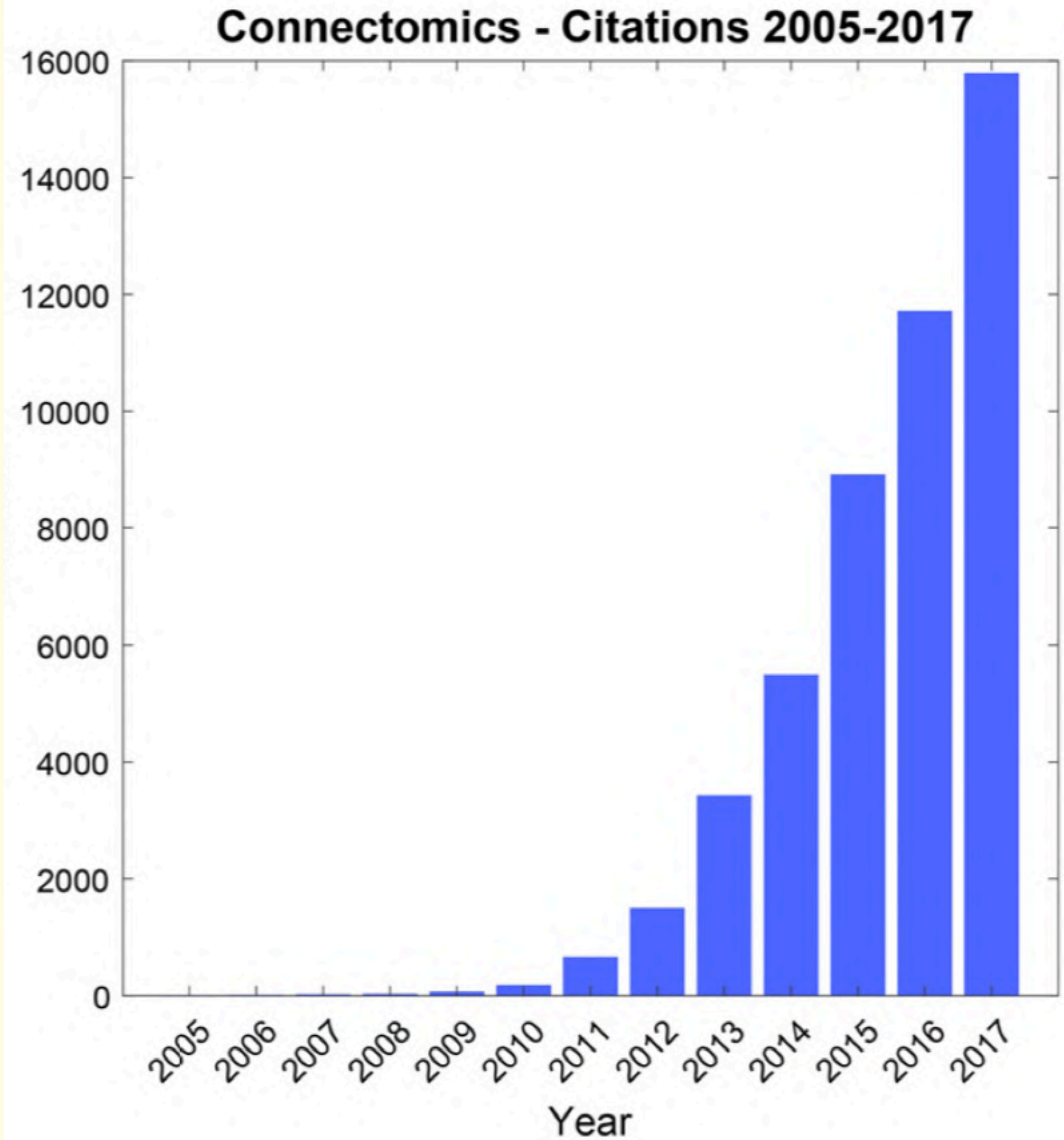
Sunday July 5, 2020

Steve Marron (UNC)
Yuan Wang (USC)
Moo K. Chung (UW-Madison)

http://www.tda-brain.com/teaching/ibc2020



Sporns & Bessett,2018
Network Neuroscience



Cambridge University Press
June 27, 2019

Motivation of this talk

There is a still huge gap between TDA theory to applications.

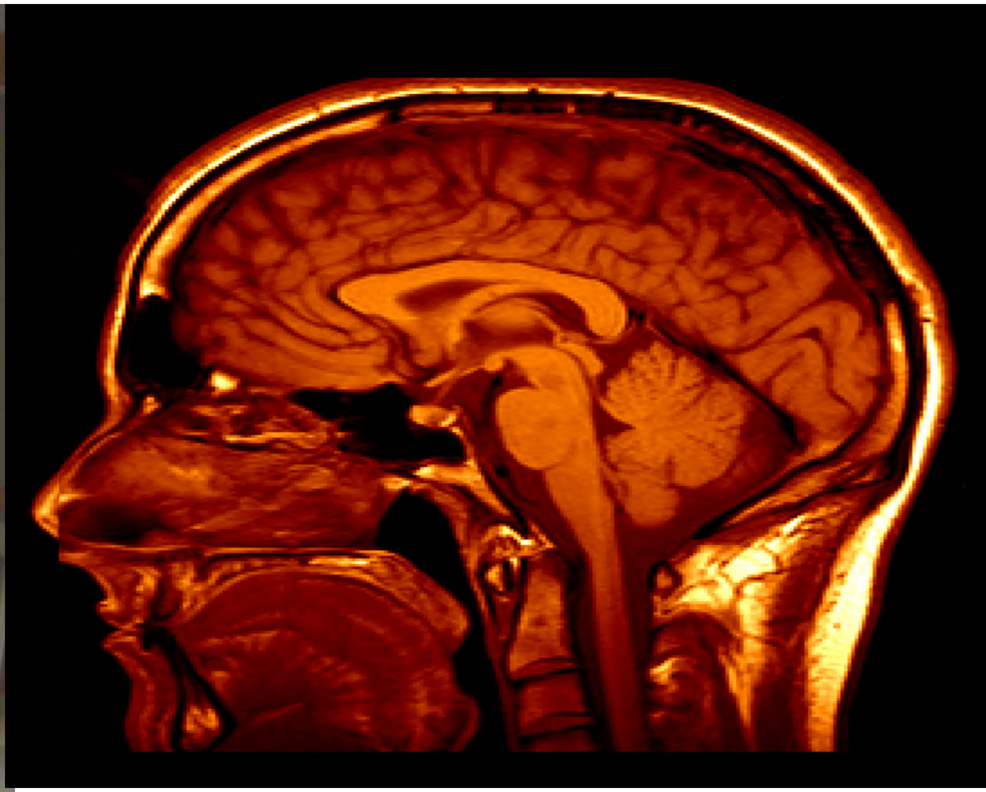
Theory → ? → Neuroimaging application

Must integrate multiple images: statistical problem

Previous works
& Preliminary

3T MRI research scanner in Madison

Structural MRI

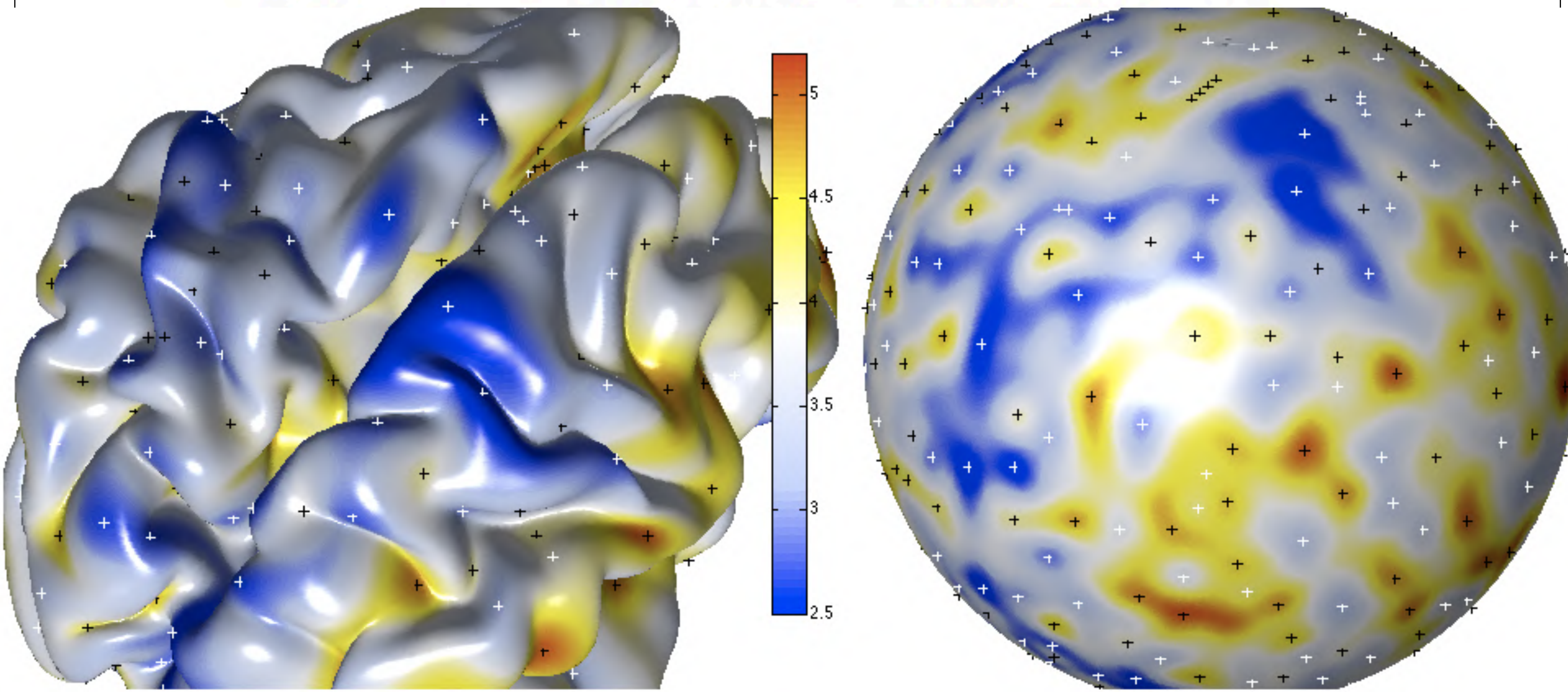


Functional MRI

Error using double
Requested
1083154800x1
(8.1GB) array exceeds
maximum array size.

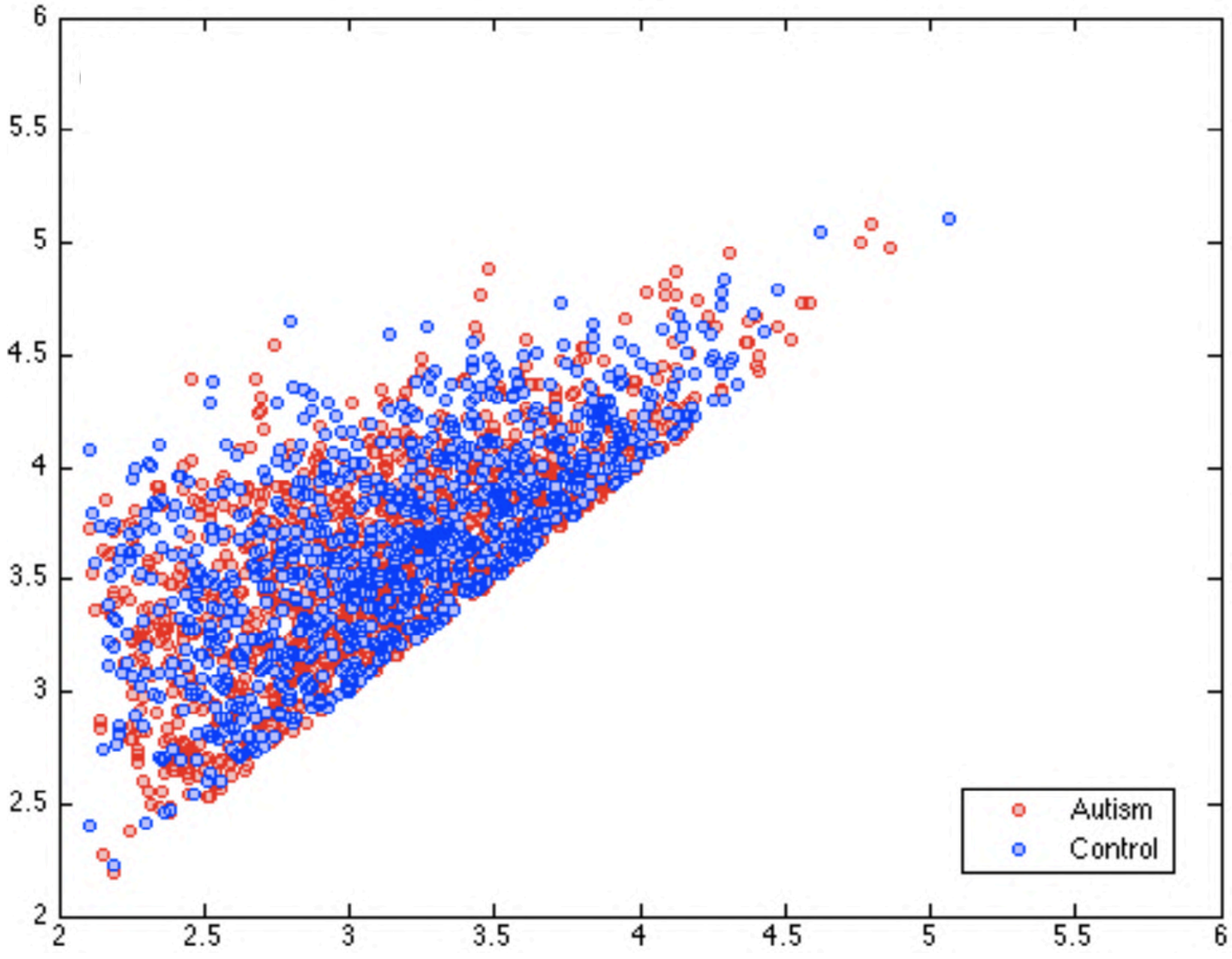
Persistence Diagrams of Cortical Surface Data

Moo K. Chung[1,2], Peter Bubenik[3], and Peter T. Kim[4]

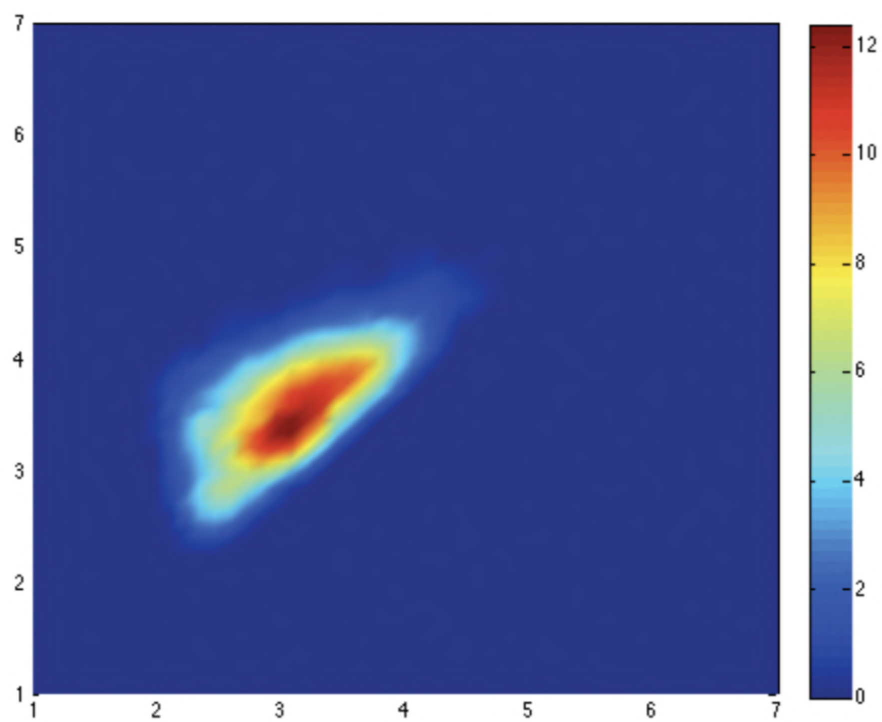


<u>Chung et al., 2009</u> Information Processing in Medical Imaging (IPMI)

Kernel density Estimation (uniform kernel)

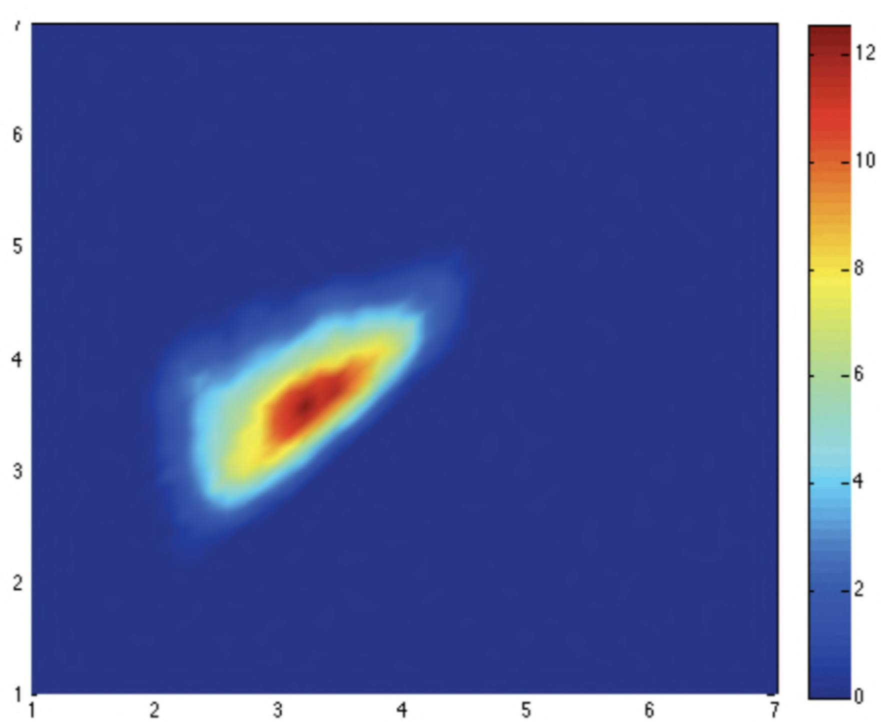


Autism



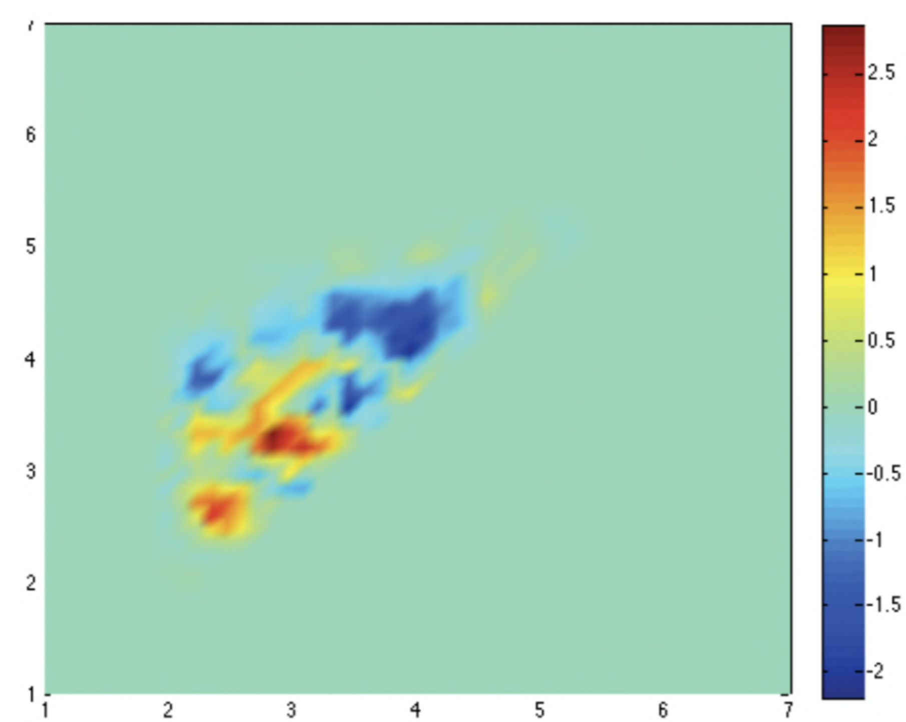
Average

Control



Average

Autism - Control



Difference in averages

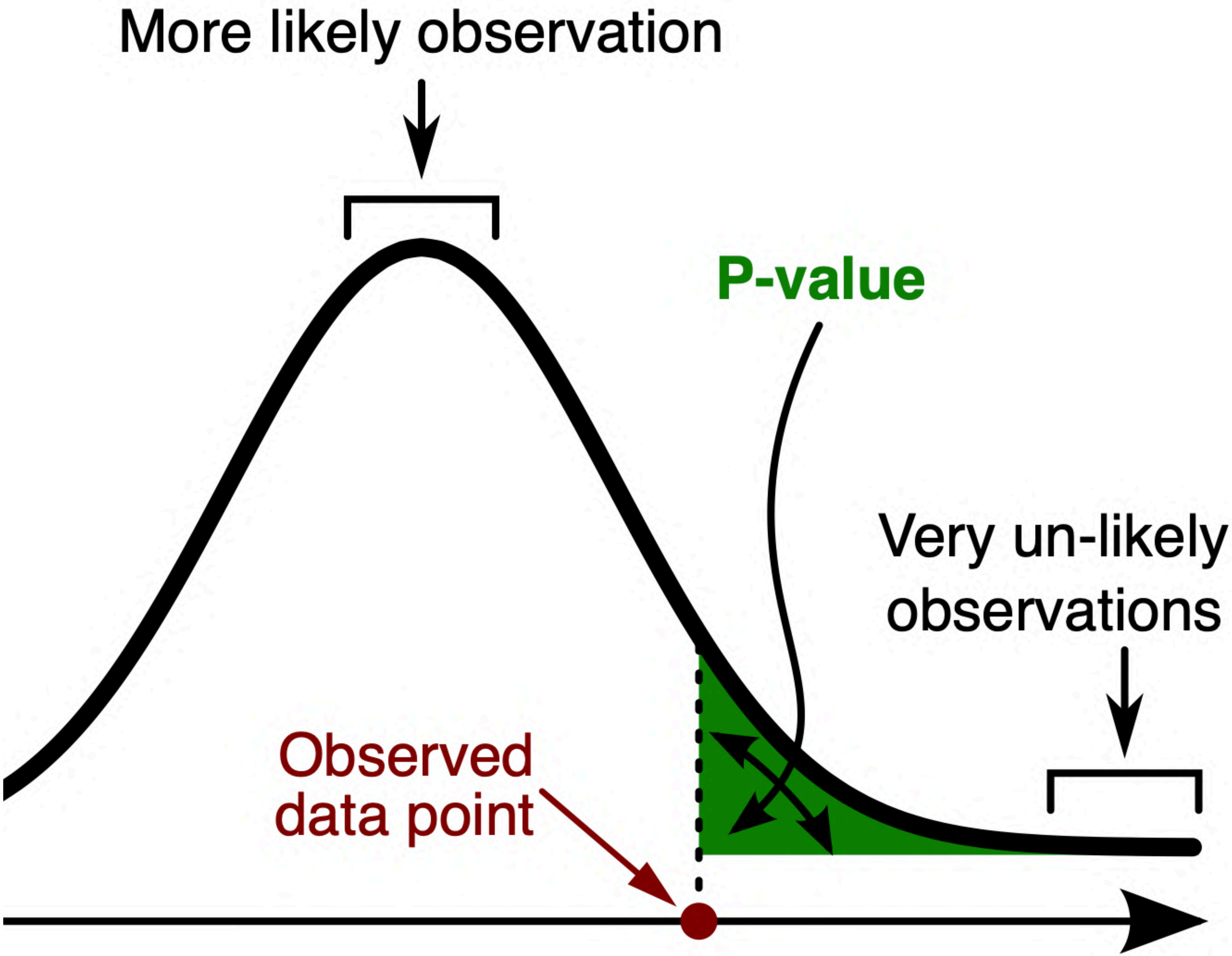
Permutation test

$$\mathbf{x} = (x_1, x_2, \cdots, x_m)$$
$$\mathbf{y} = (y_1, y_2, \cdots, y_n)$$
$$(\mathbf{x}, \mathbf{y}) = (x_1, \cdots, x_m, y_1, \cdots, y_n)$$

$$\pi(\mathbf{x}, \mathbf{y}) \in \mathbb{S}_{m+n}$$

Permutation group of order *m*+*n*

$$p\text{-value} = \frac{1}{(m+n)!} \sum_{\tau \in \mathbb{S}_{m+n}} \mathcal{I}\big(f(\tau(\mathbf{x}), \tau(\mathbf{y})) > f(\mathbf{x}, \mathbf{y})\big)$$


More likely observation
P-value
Very un-likely
observations
Observed
data point

Permutation test

Observation: $x=(x_1,x_2)=(1, 3)$, $y=(y_1,y_2)=(2, 4)$

Hypothesis: H_0: $x = y$ *vs.* H_1: $x > y$

Test stat: $f(x,y) = x_1 + x_2 - y_1 + y_2$

f large → H_1 is more likely

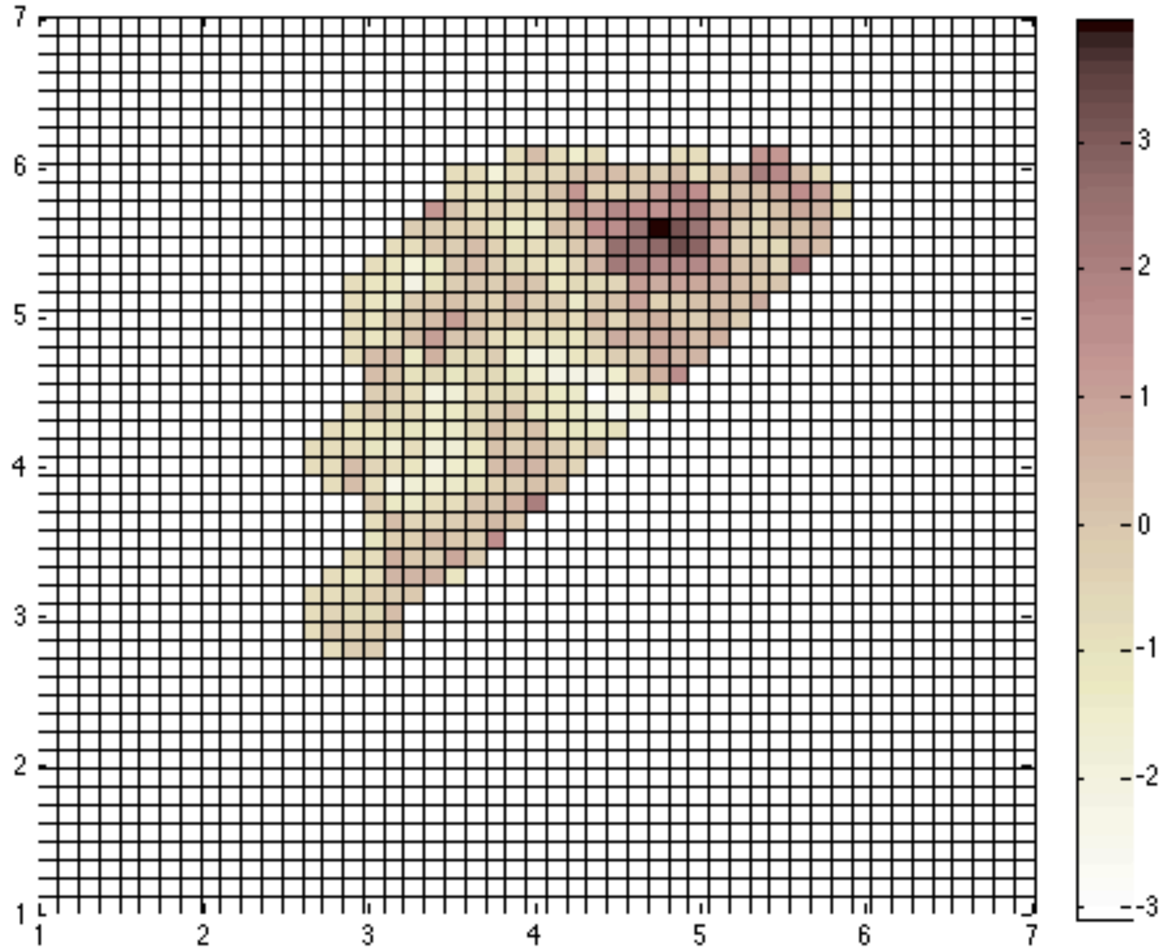
Permutations

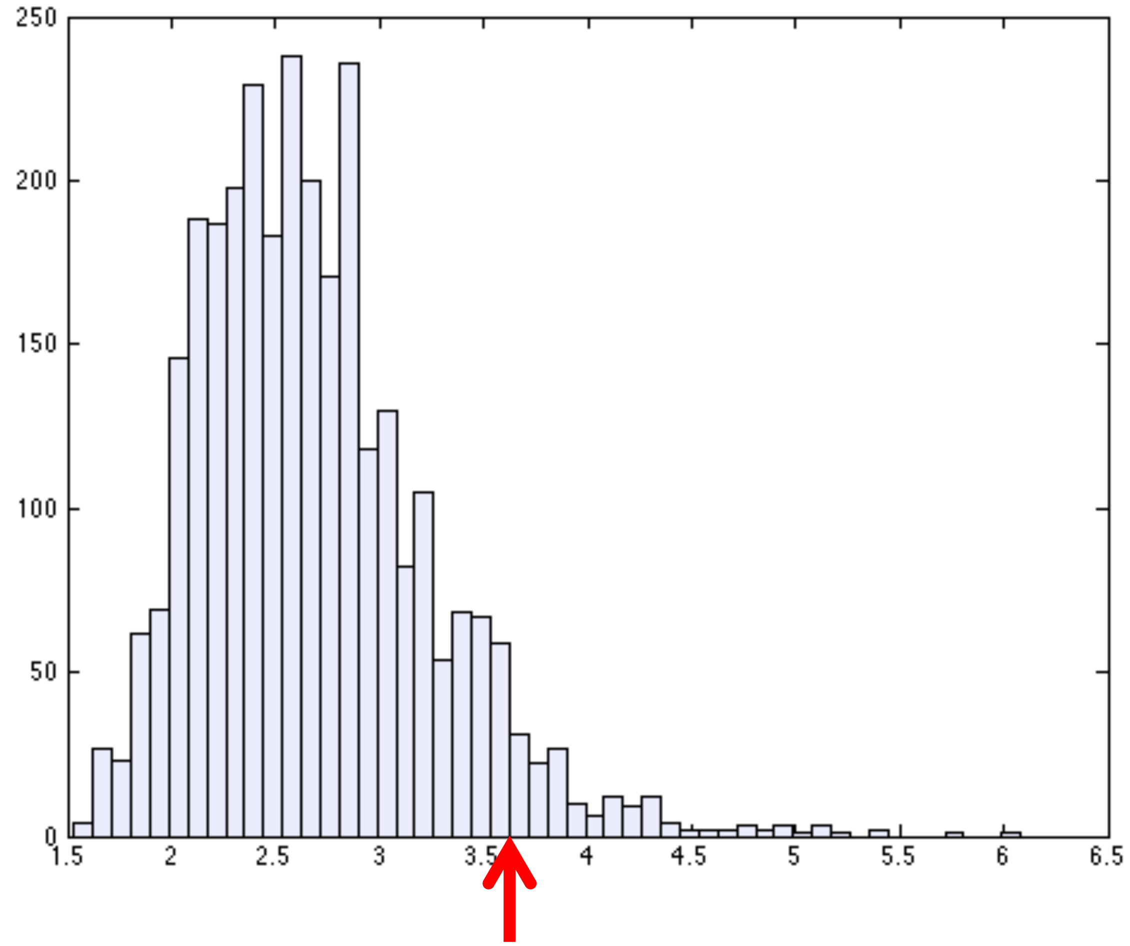
	(1,3)(2,4)	(2,4)(1,3)	(1,2)(3,4)	(3,4)(1,2)	(1,4)(3,2)	(3,2)(1,4)
f	-2	2	-4	4	0	0

p-value= **4/6**

We do not reject H_0

Permutation test on persistent diagrams





Max t = 3.9507

Min t = -3.0961

95 percentile = 3.6432

5 percentile = -4.0237

More pairings for the control subjects

= More cortical folding

History of permutation test

Fisher 1935, The Design of Experiment

$$\binom{8}{4} = 70$$

Thompson et al. 2001, Nature Neuroscience

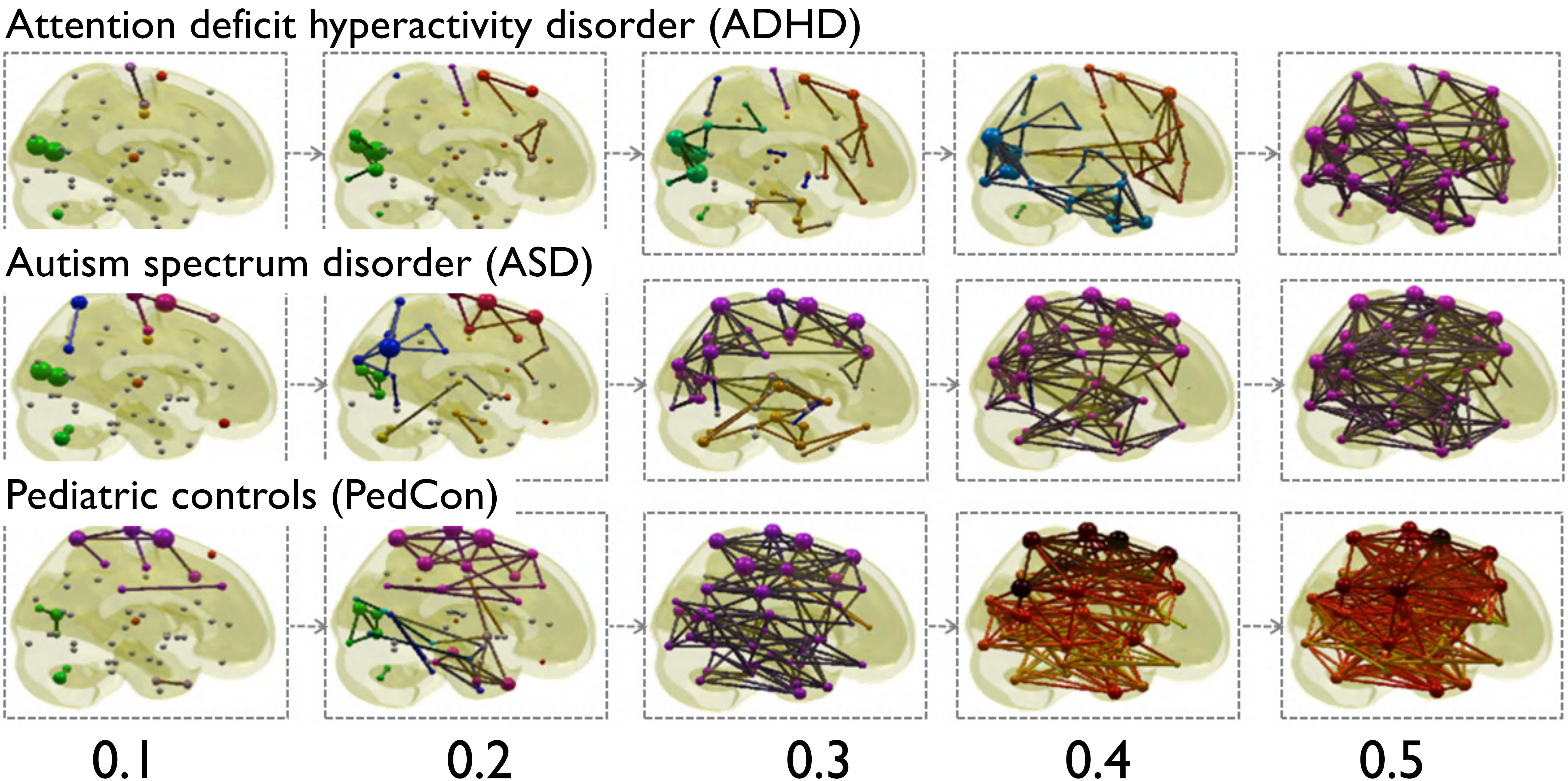
$$\binom{40}{20} = 1.34 \cdot 10^{11}$$

Supercomputer for 1million permutations

Nichols et al. 2002, Human Brain Mapping
4279 citations

$$\binom{6}{3} = 20$$

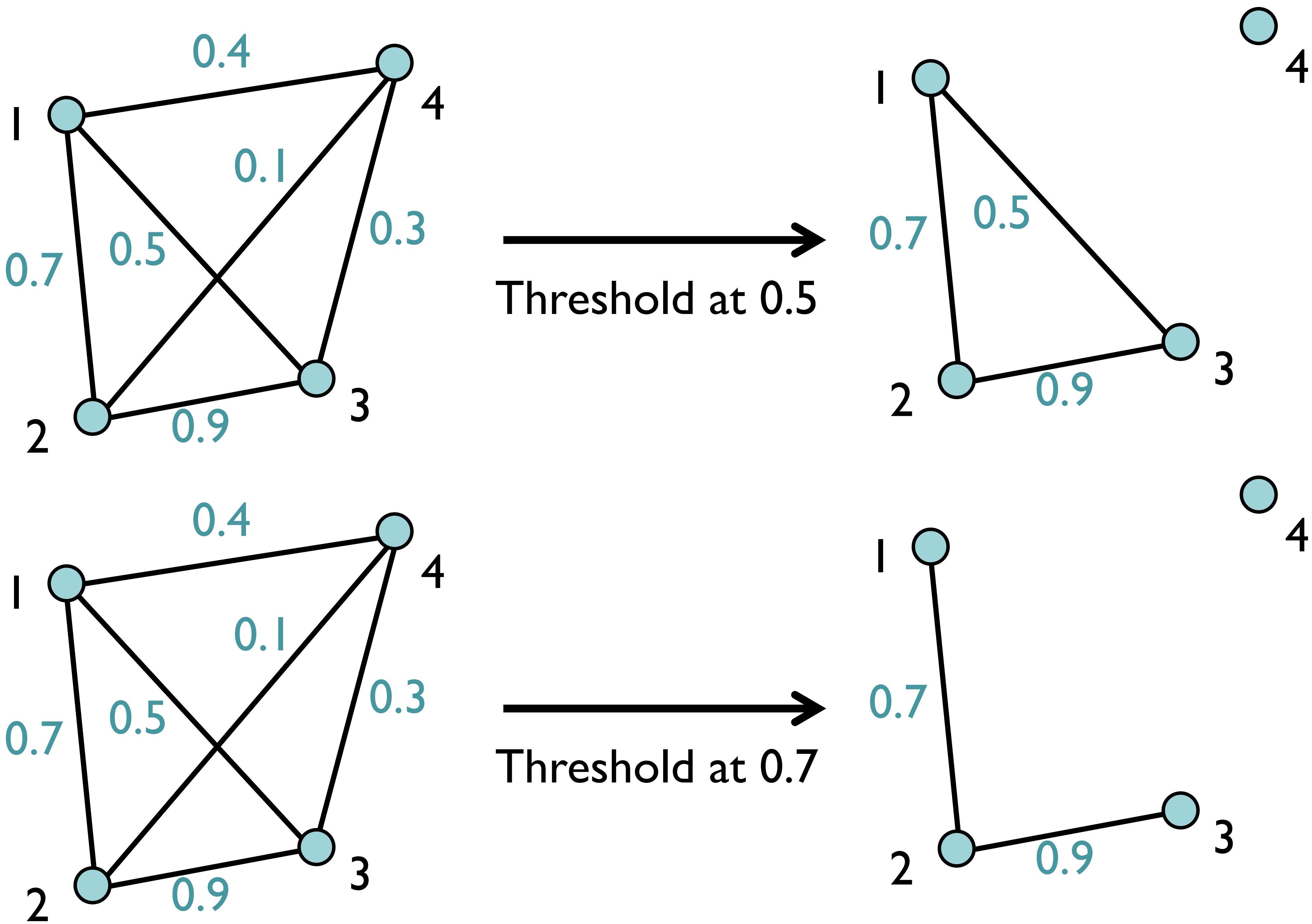
Graph filtration based network analysis



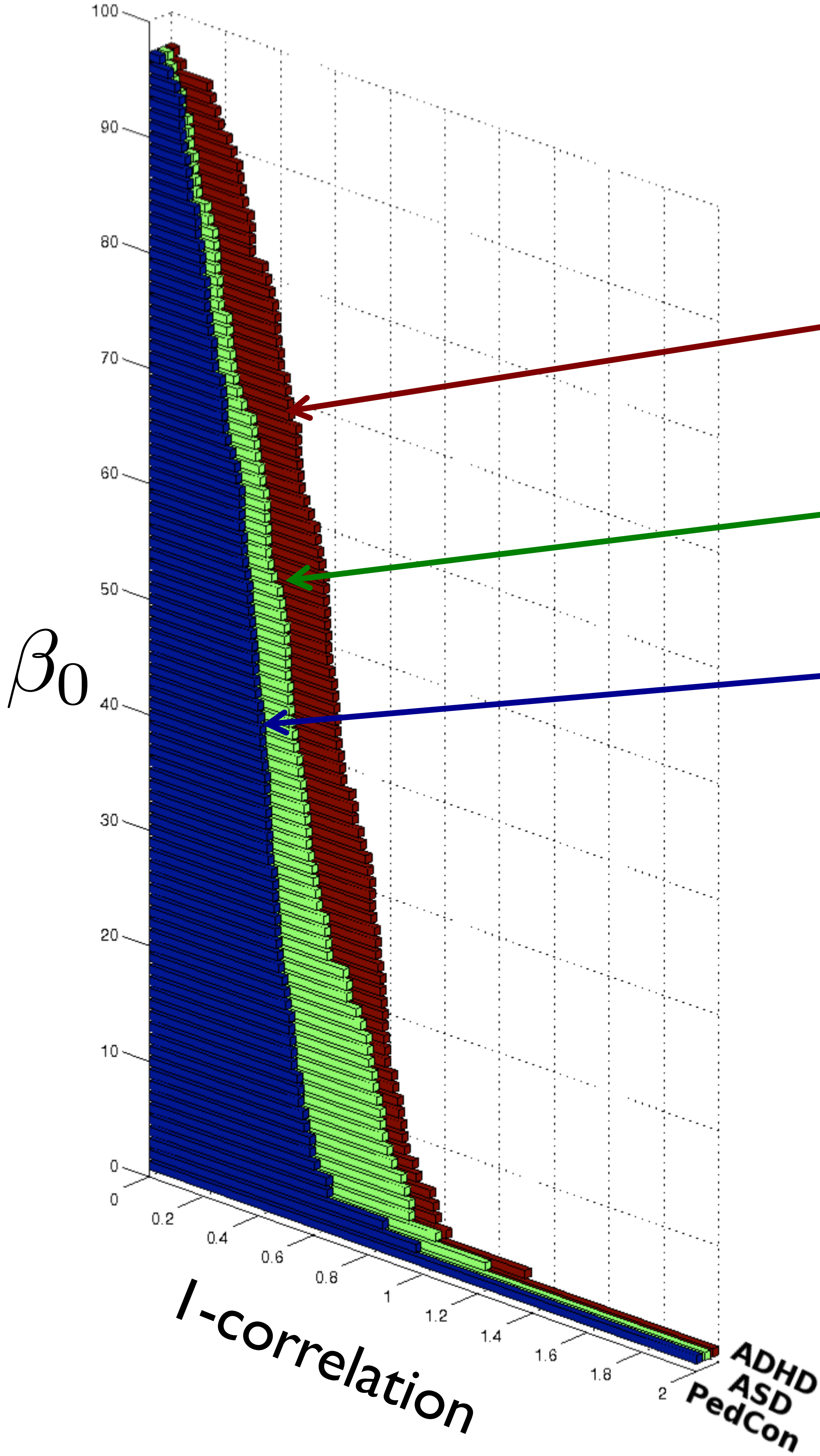
1-correlation

Lee. et al. 2012, IEEE Transactions on Medical Imaging

Graph theory based network analysis in year 2010



Betti-0 plot



24 attention deficit hyperactivity disorder (ADHD) children

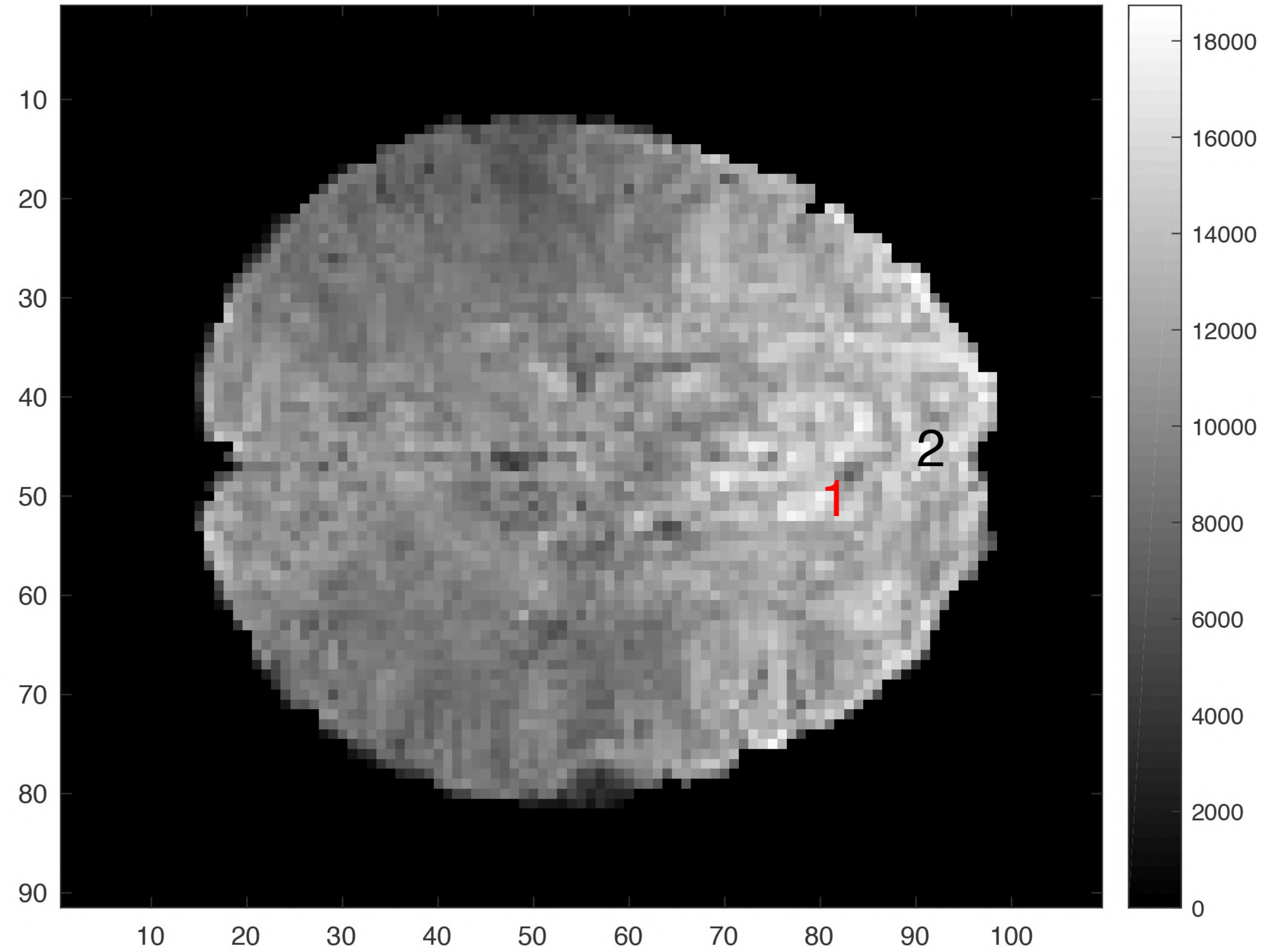
26 autism spectrum disorder (ASD) children

11 pediatric control subjects

The normal brain networks merges to a single component faster than other clinical populations.

Lee et al. 2010 ISBI

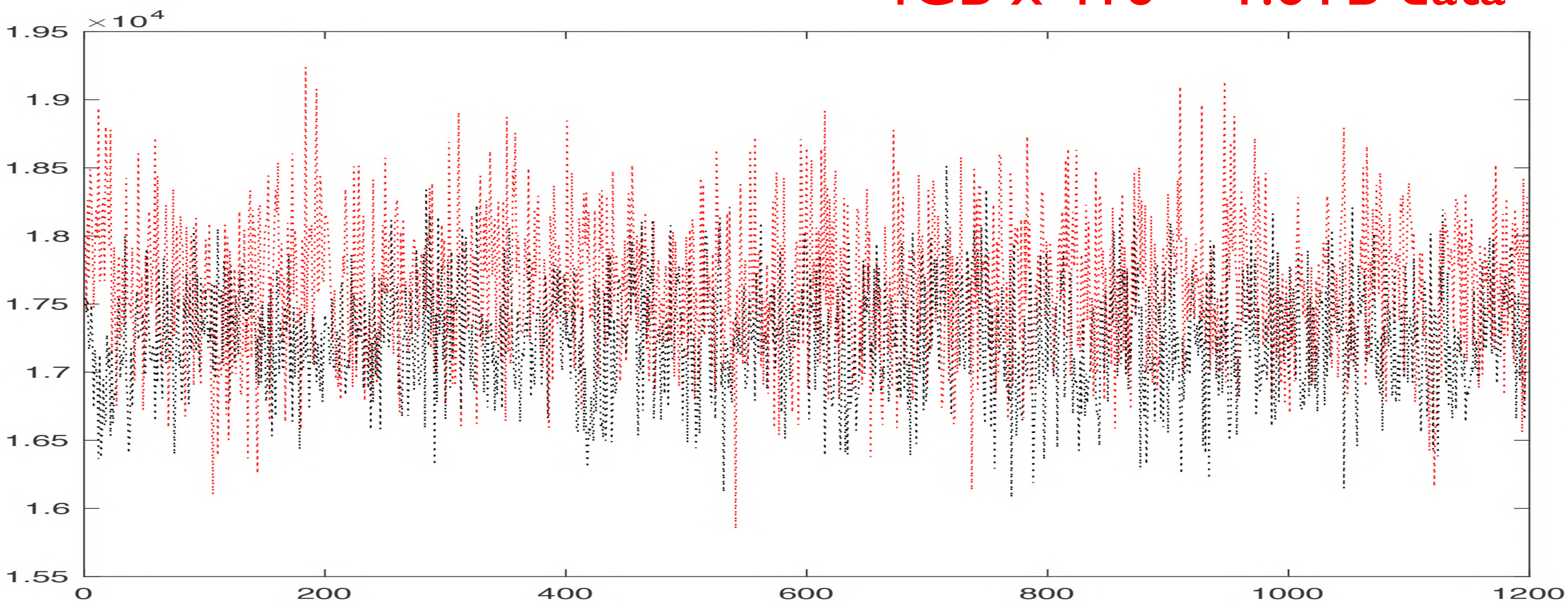
Resting-state functional magnetic resonance imaging (fMRI)



Time series with 1200 time points at 300000 voxels per subject measured over 14min 33 seconds inside MRI scanner

416 subjects
= 131 Monozygotic (MZ) twins
77 Dizygotic (DZ) twins

4GB x 416 = 1.6TB data



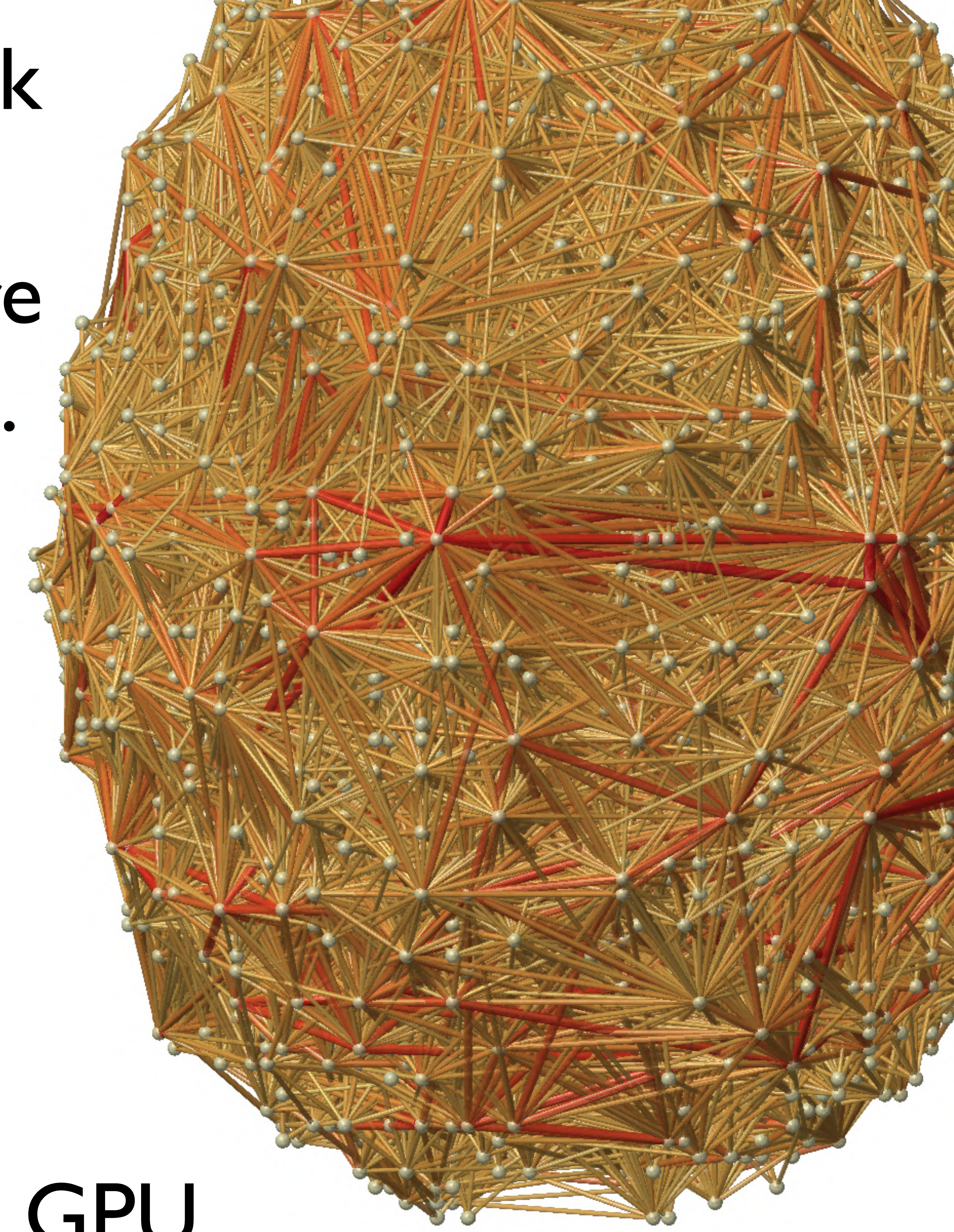
Permutation test impractical if sample size > 200

```
>> nchoosek(200,100)

Warning: Result may not be exact.
Coefficient is greater than
9.007199e+15 and is only accurate
to 15 digits

> In nchoosek (line 92)
ans =
  9.0549e+58
```

Dense brain network

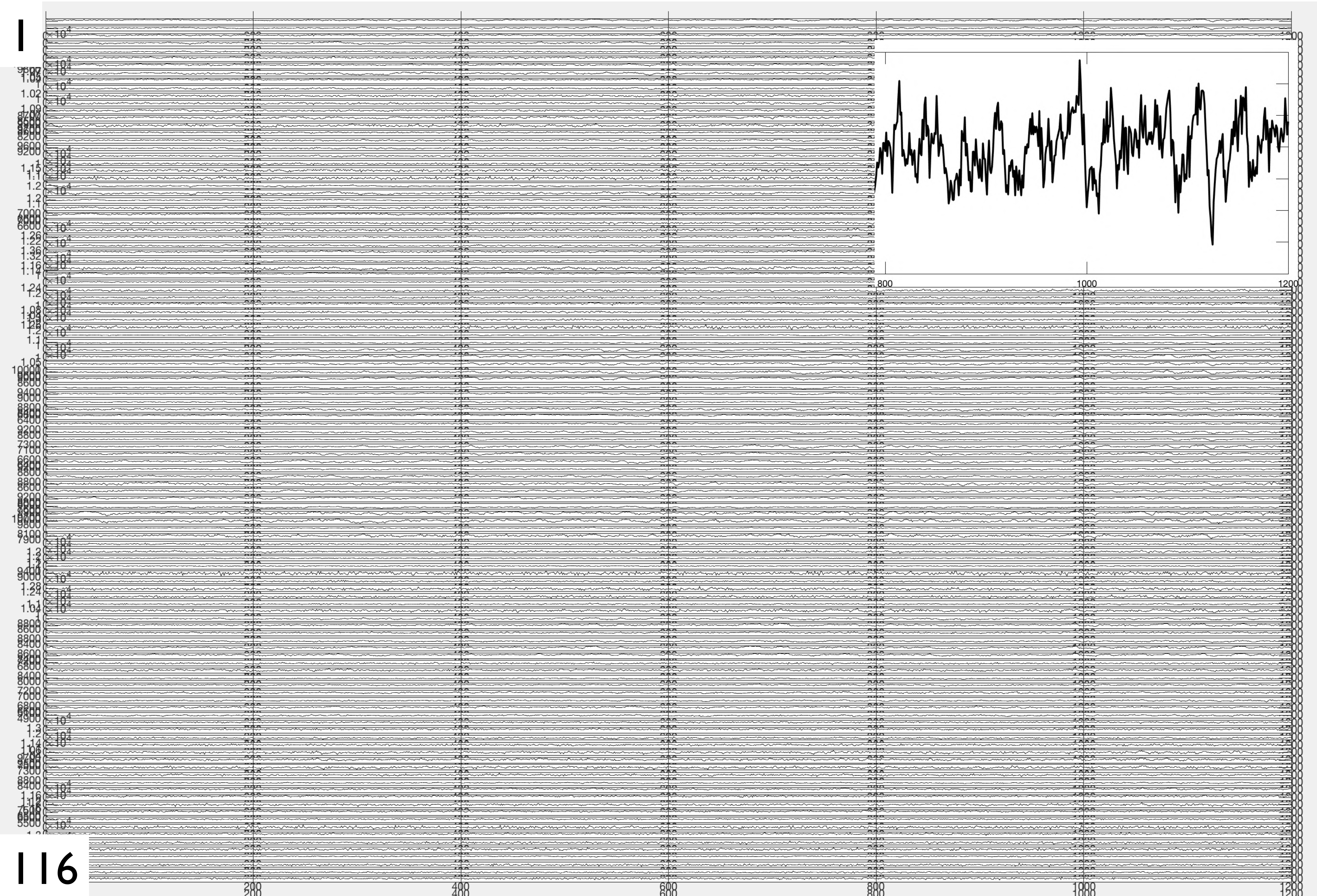
Brain network where each voxel is a node.

AMD vega64 GPU

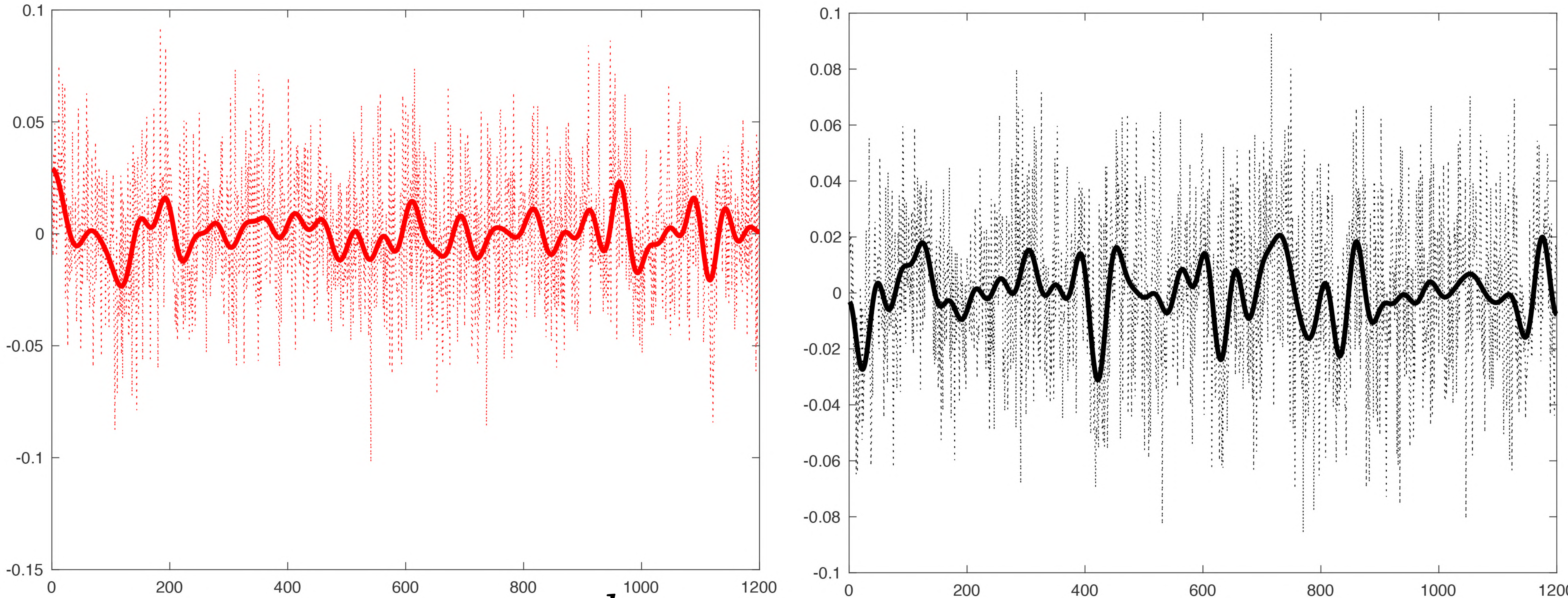


Time series averaged into 116 brain regions

116 time series at 1200 time points



Cosine Series Representation (*Wang 2018, Annals of Applied Stat.*)

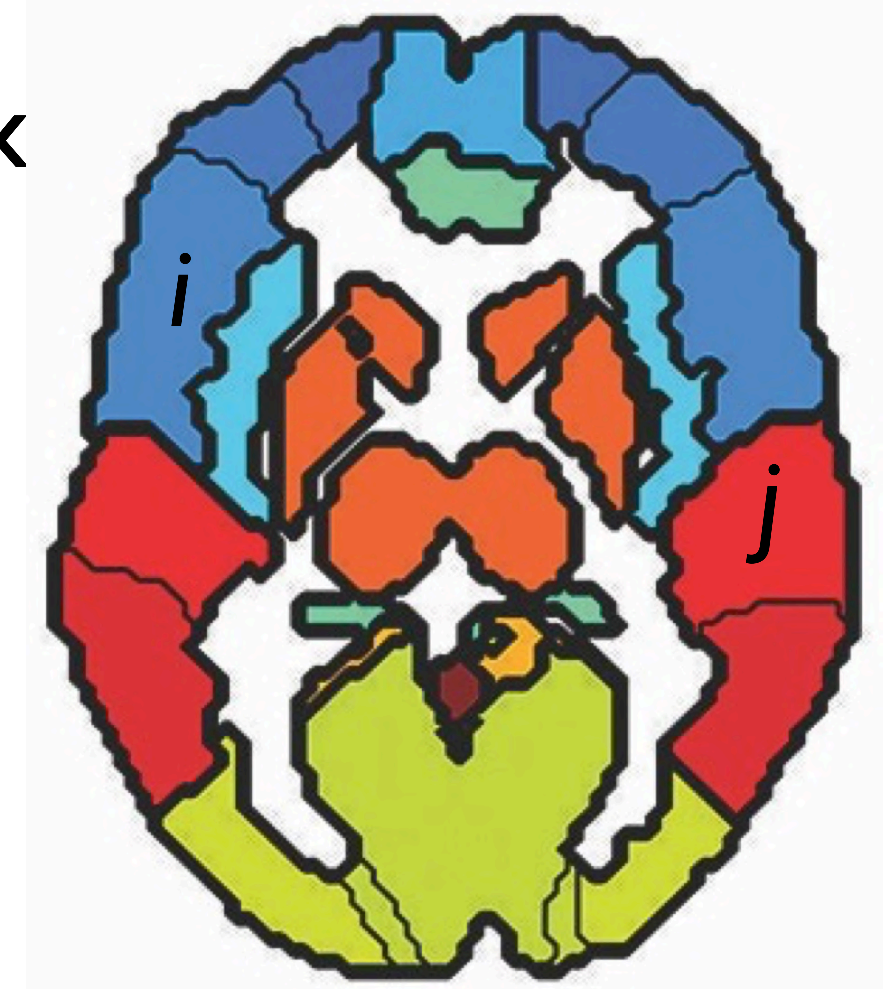


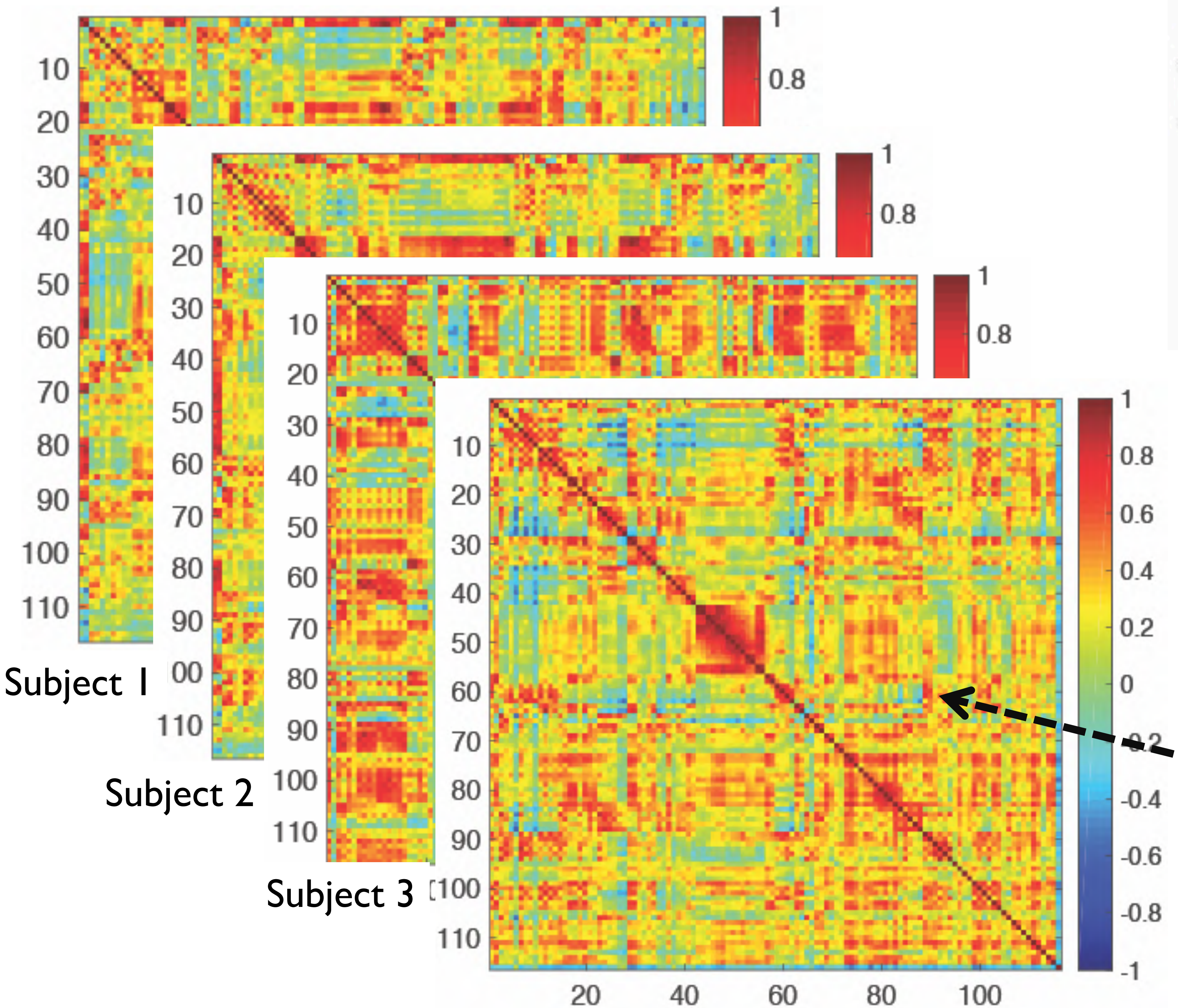
$$\zeta_i(t) = \sum_{l=0}^{k} d_{li}\psi_l(t),\ t \in [0,1]$$

$$\psi_0(t) = 1, \psi_l(t) = \sqrt{2}\cos(l\pi t)$$

120 features ⟶ $\mathbf{d}_i = (d_{0i}, d_{1i}, \cdots, d_{ki})$

Subject level brain connectivity matrix





$$c_{ij} = corr(\mathbf{d}_i, \mathbf{d}_j)$$

Correlation of Fourier coefficients

ACE model for twins

MZ-twins share 100% of genes
DZ-twins share 50% of genes

$$\rho_{\mathrm{MZ}} = A + C$$

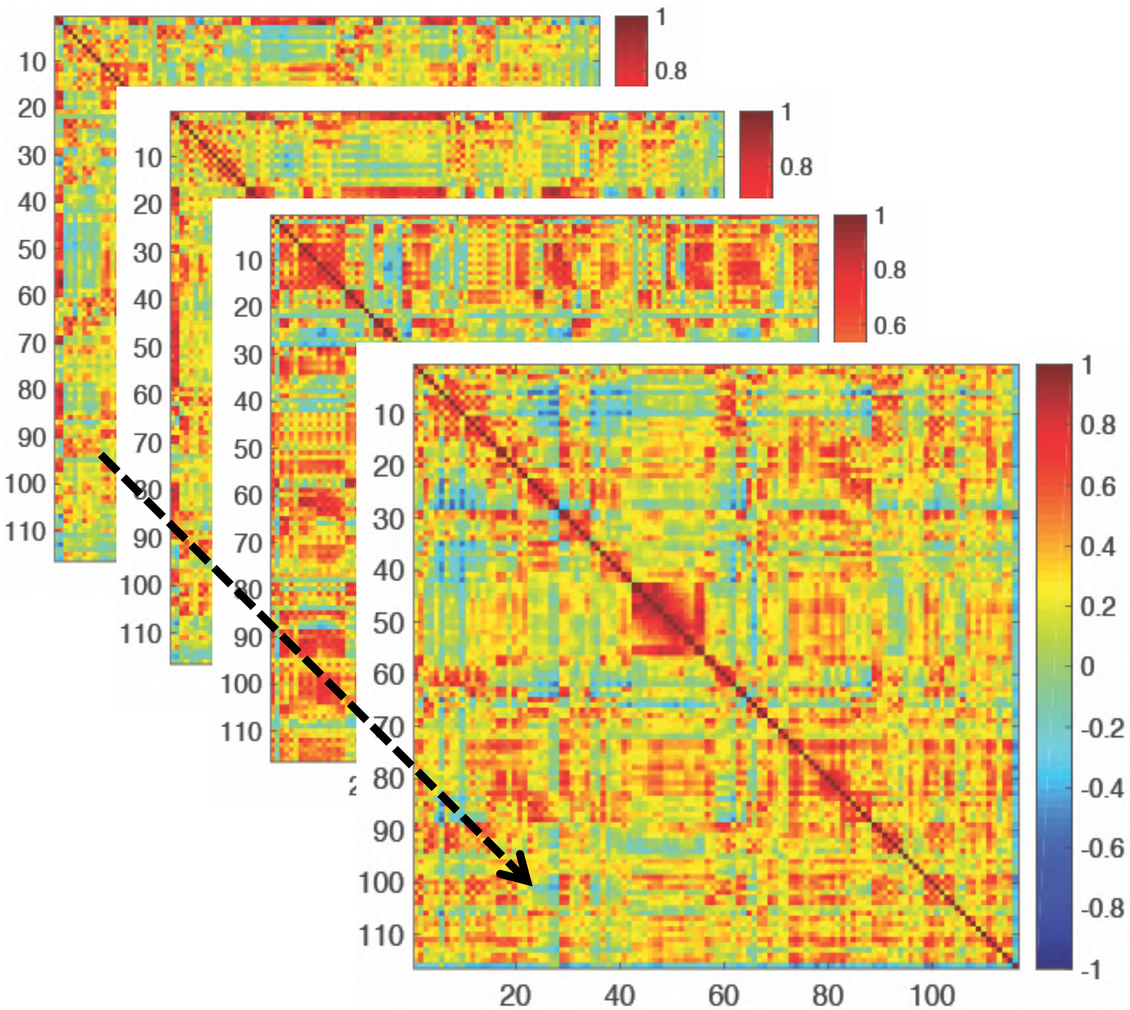
Twin correlation — Additive genetics — Common environment

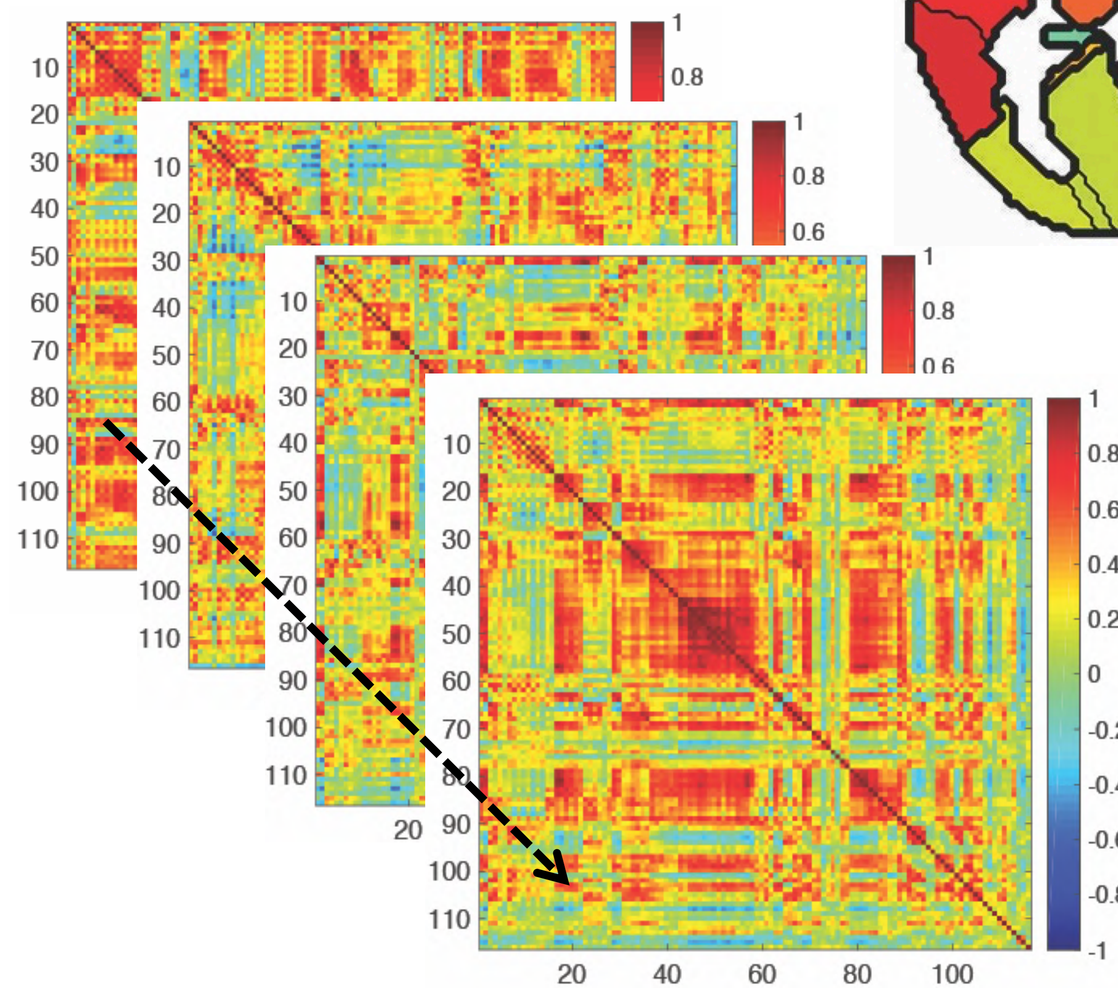
$$\rho_{\mathrm{DZ}} = A/2 + C$$

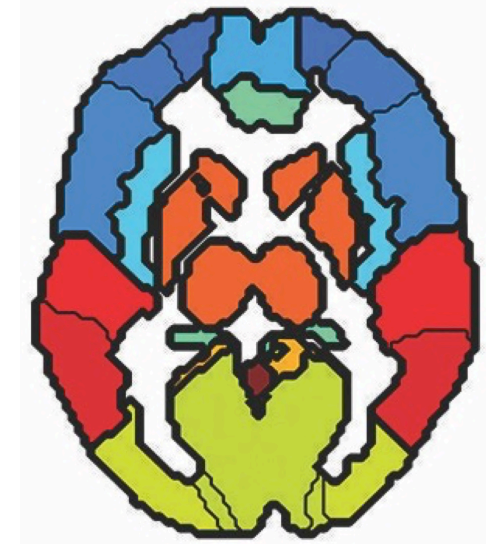
Falconer's formula for heritability index (HI)

$$\mathrm{HI} = A = 2(\rho_{\mathrm{MZ}} - \rho_{\mathrm{DZ}})$$

Correlation (group) of correlation (subject)





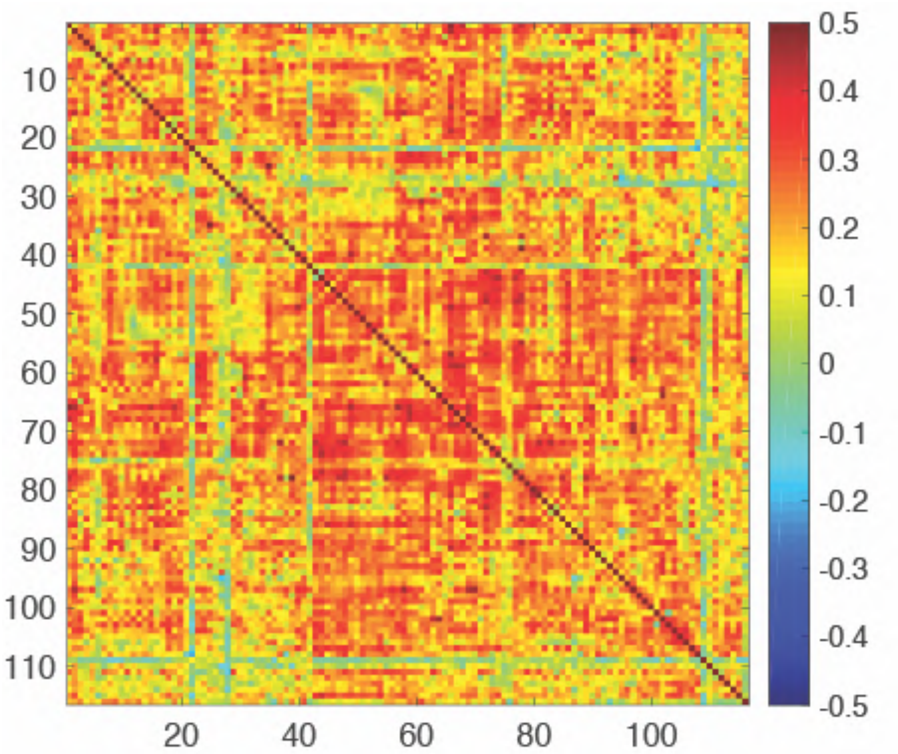


$$\mathbf{c}_{ij}^{1} = (c_{ij}^{11}, \cdots, c_{ij}^{1m}) \qquad \mathbf{c}_{ij}^{2} = (c_{ij}^{21}, \cdots, c_{ij}^{2m})$$

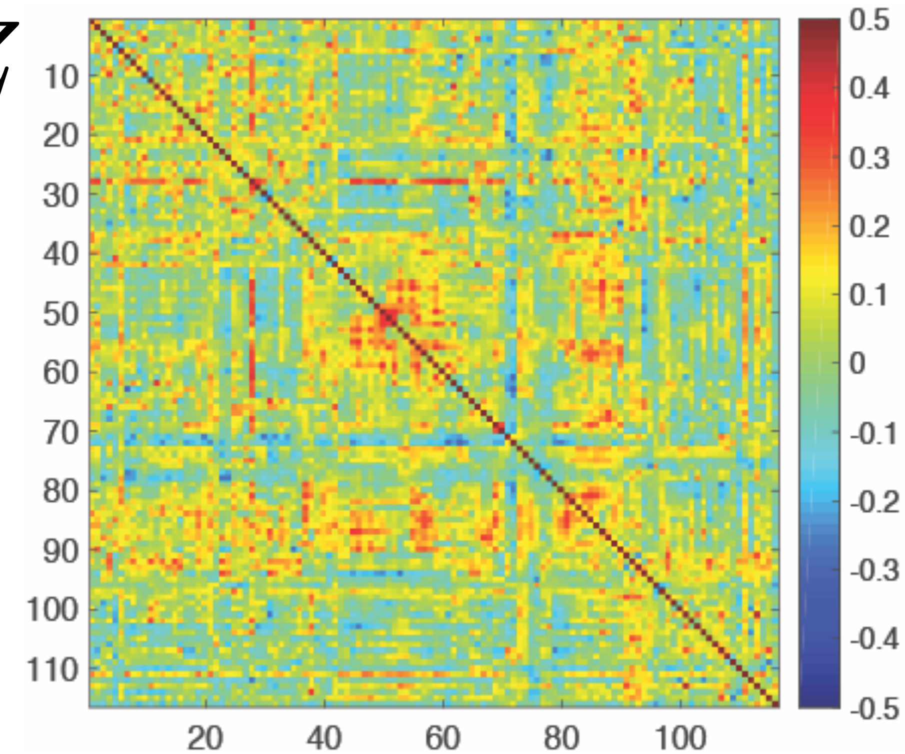
$$c_{ij}^{MZ} = corr(\mathbf{c}_{ij}^{1}, \mathbf{c}_{ij}^{2})$$

MZ- and DZ-twin correlation difference

c_{ij}^{MZ}

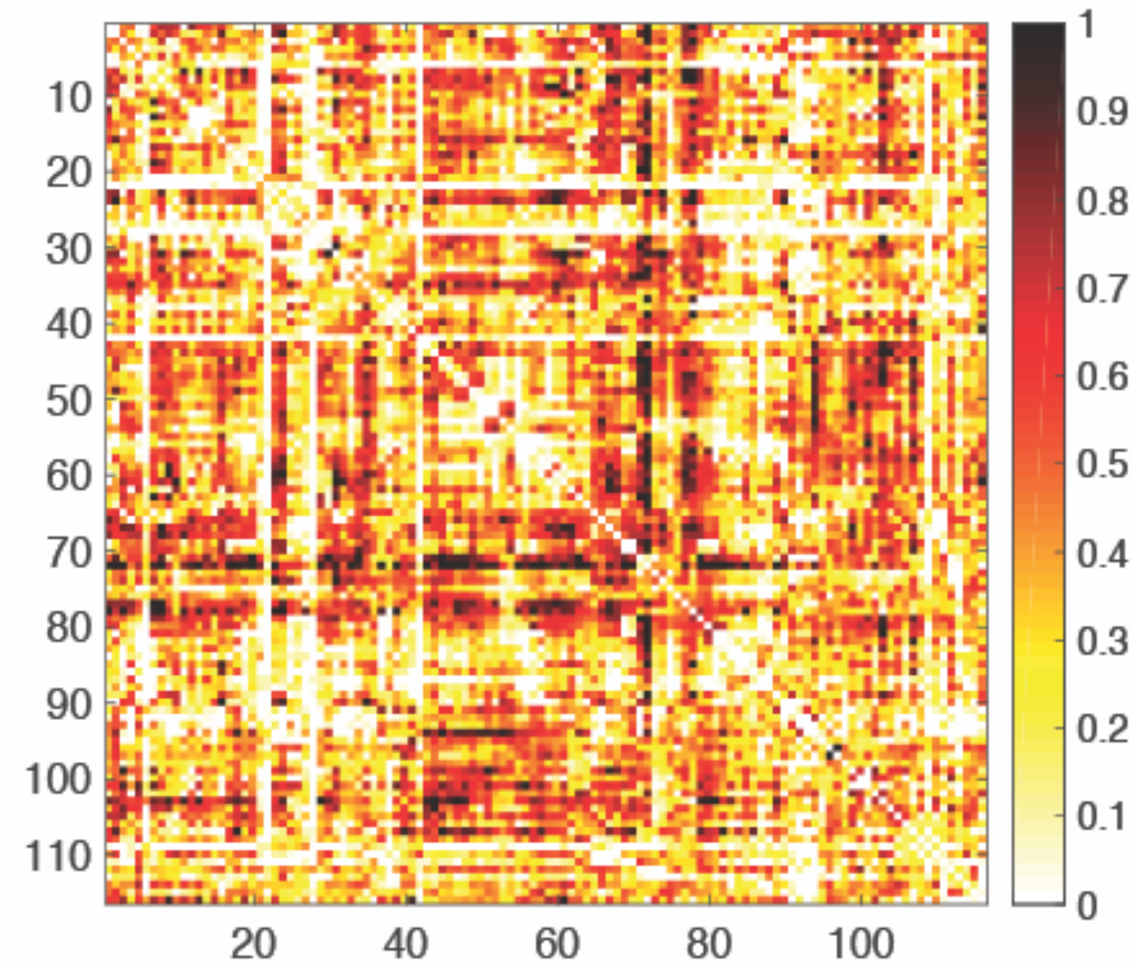


c_{ij}^{DZ}

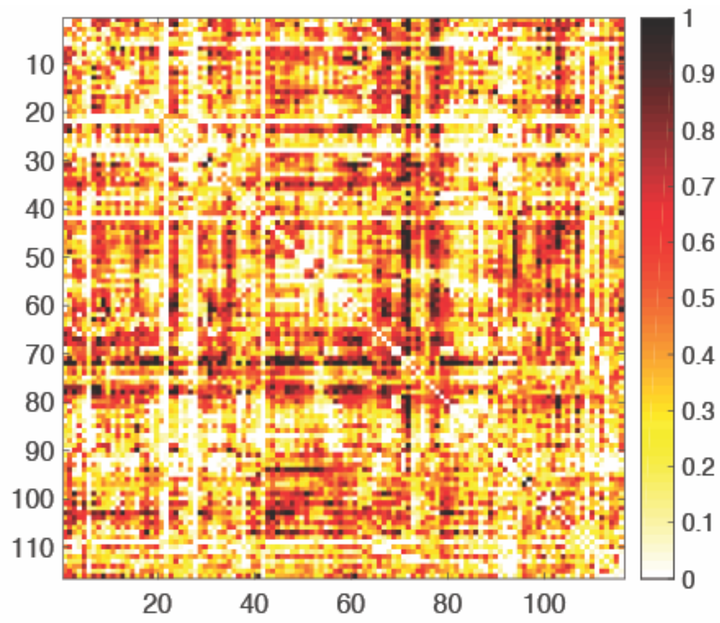


$$h_{ij} = 2(c_{ij}^{MZ} - c_{ij}^{CZ})$$

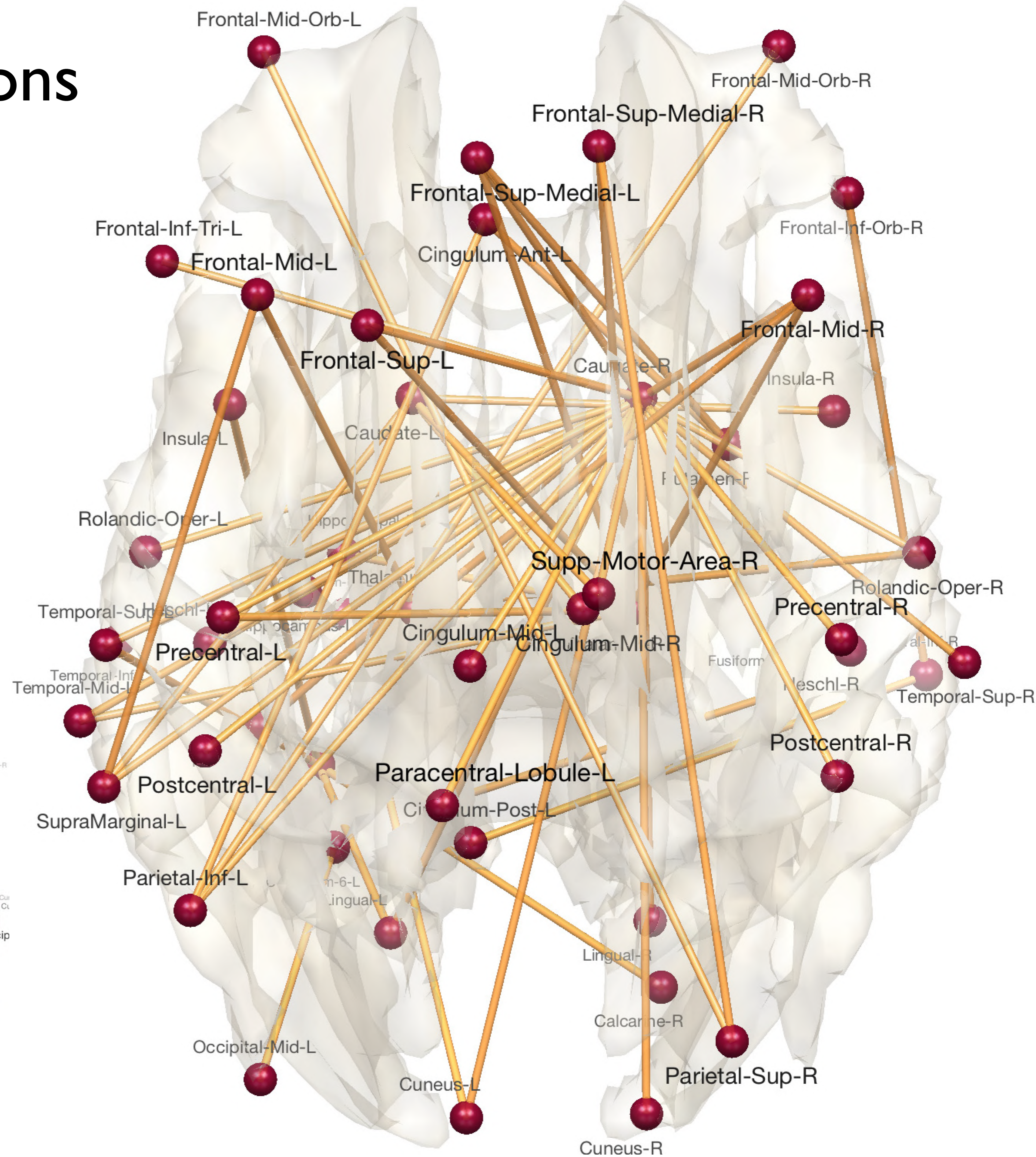
Heritability index = amount of genetic contribution



Heritable brain regions

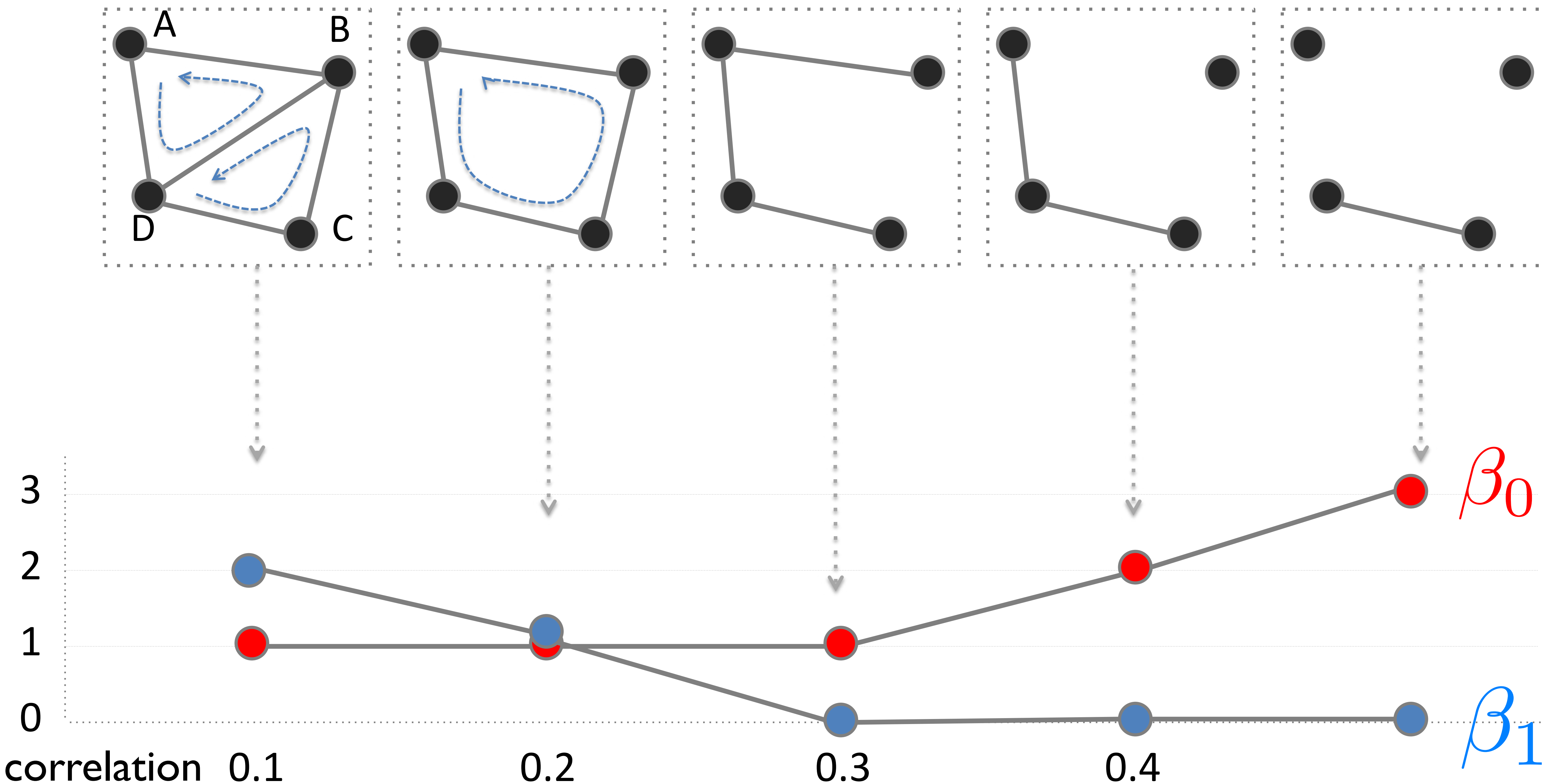


$$h_{ij} \geq 1$$



Statistical significance?

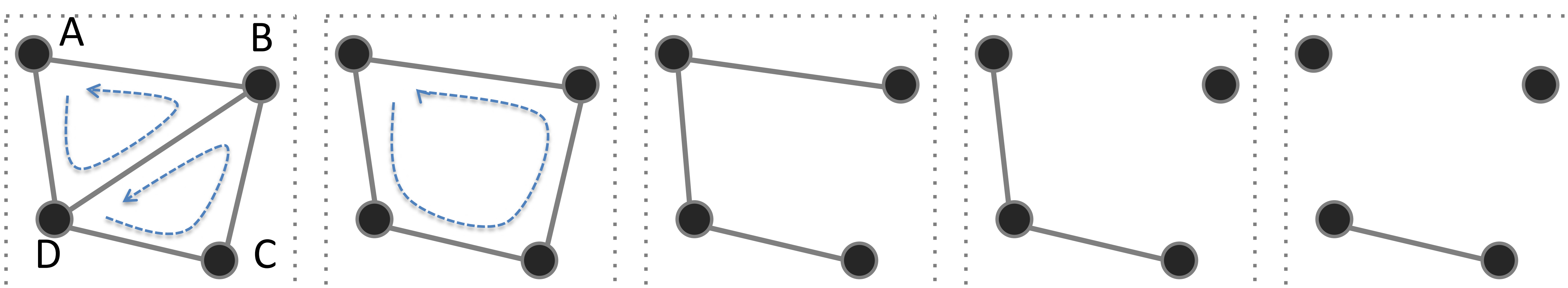
Betti-plots



Monotonicity:
<u>*Chung et al. 2019 Network Neuroscience*</u>

β_0 and β_1 are monotone over graph filtration.

Monotonicity of β_0: Deletion of edge increases the the number of connected components by at most 1. β_0 increases by 0 or 1.



β_0 and β_1 are monotone over graph filtration.

Monotonicity of β_1:

Euler characteristic:

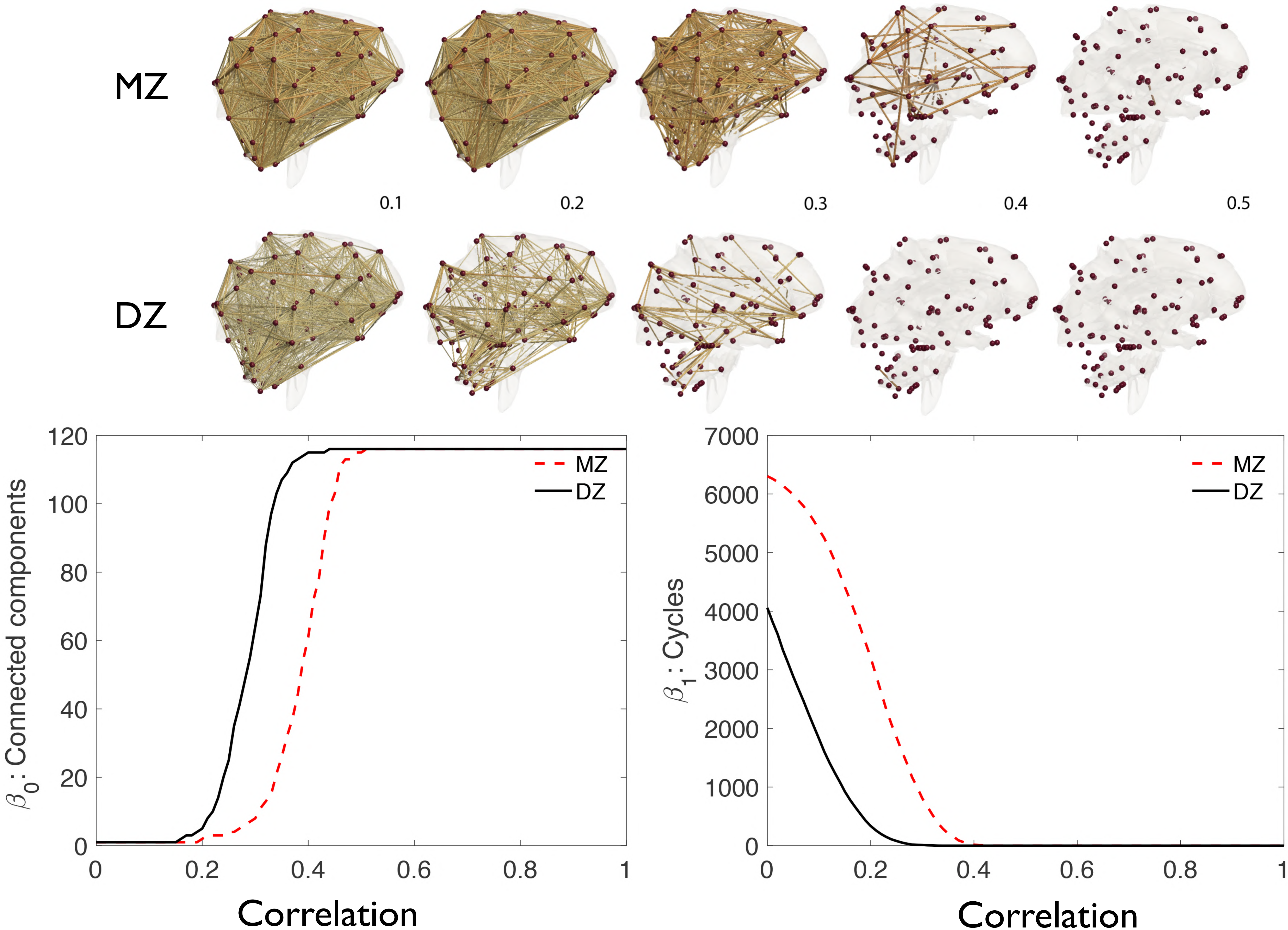
$$\chi = \beta_0 - \beta_1 = p - q$$

nodes edges

$$\beta_1 = \beta_0 - p + q$$

0, +1 fixed -1

Betti-plots on graph filtration



Exact Topological Inference (ETI)

Kolmogorov Smirnov (KS) distance

$$\mathbf{G}^1 = \{G^1_\lambda : 0 \leq \lambda \leq 1\}$$

$$\mathbf{G}^2 = \{G^2_\lambda : 0 \leq \lambda \leq 1\}$$

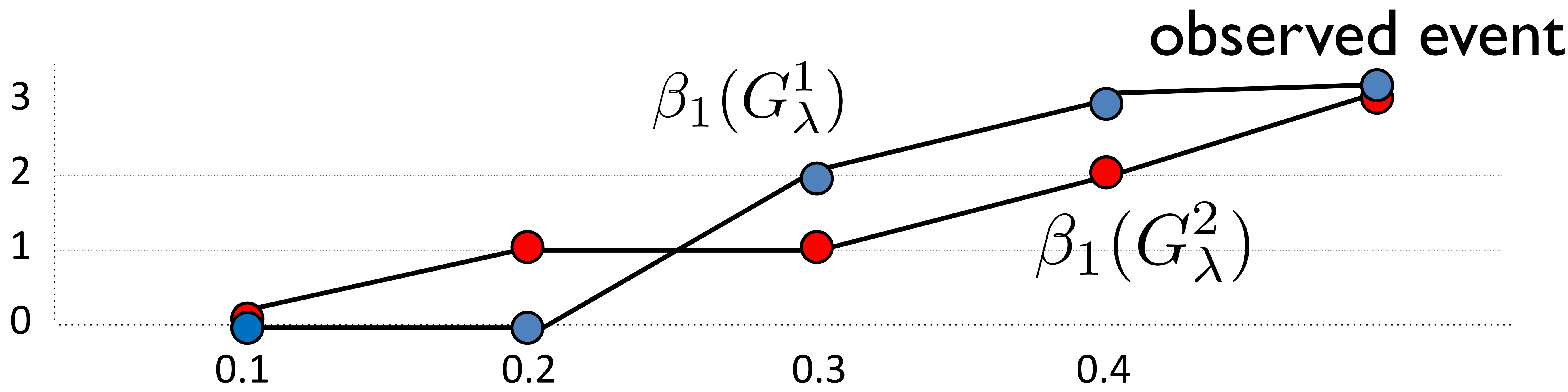
$$D(\mathbf{G}^1, \mathbf{G}^2) = \sup_{\lambda \in [0,1]} |\beta_i(G^1_\lambda) - \beta_i(G^2_\lambda)|$$

D satisfies all the axioms of metric except identity:

$$D(\mathbf{G}^1, \mathbf{G}^2) = 0 \not\longrightarrow \mathbf{G}^1 = \mathbf{G}^2$$

$$P(D(\mathbf{G}^1, \mathbf{G}^2) = 0) = 0$$

Inference on Betti-plots using KS-distance



Null hypothesis:

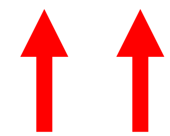
$$H_0 : \beta_1(G^1_\lambda) = \beta_1(G^2_\lambda) \text{ for all } \lambda$$

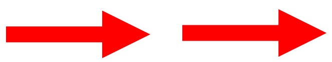
Need to determine the probability of observed event under the null hypothesis.

Under the null, generate every possible events (sample space) by permutations.

Permutation test on monotone features

Observed data: (1,3) (2,4)



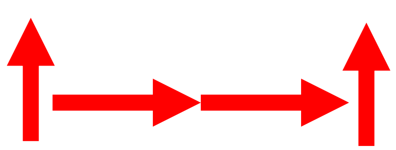


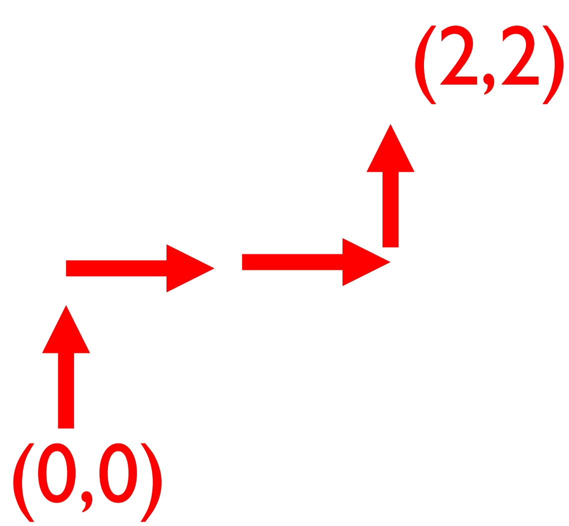
1) Combine features

1, 3, 2, 4

2) Permutation

3, 2, 4, 1



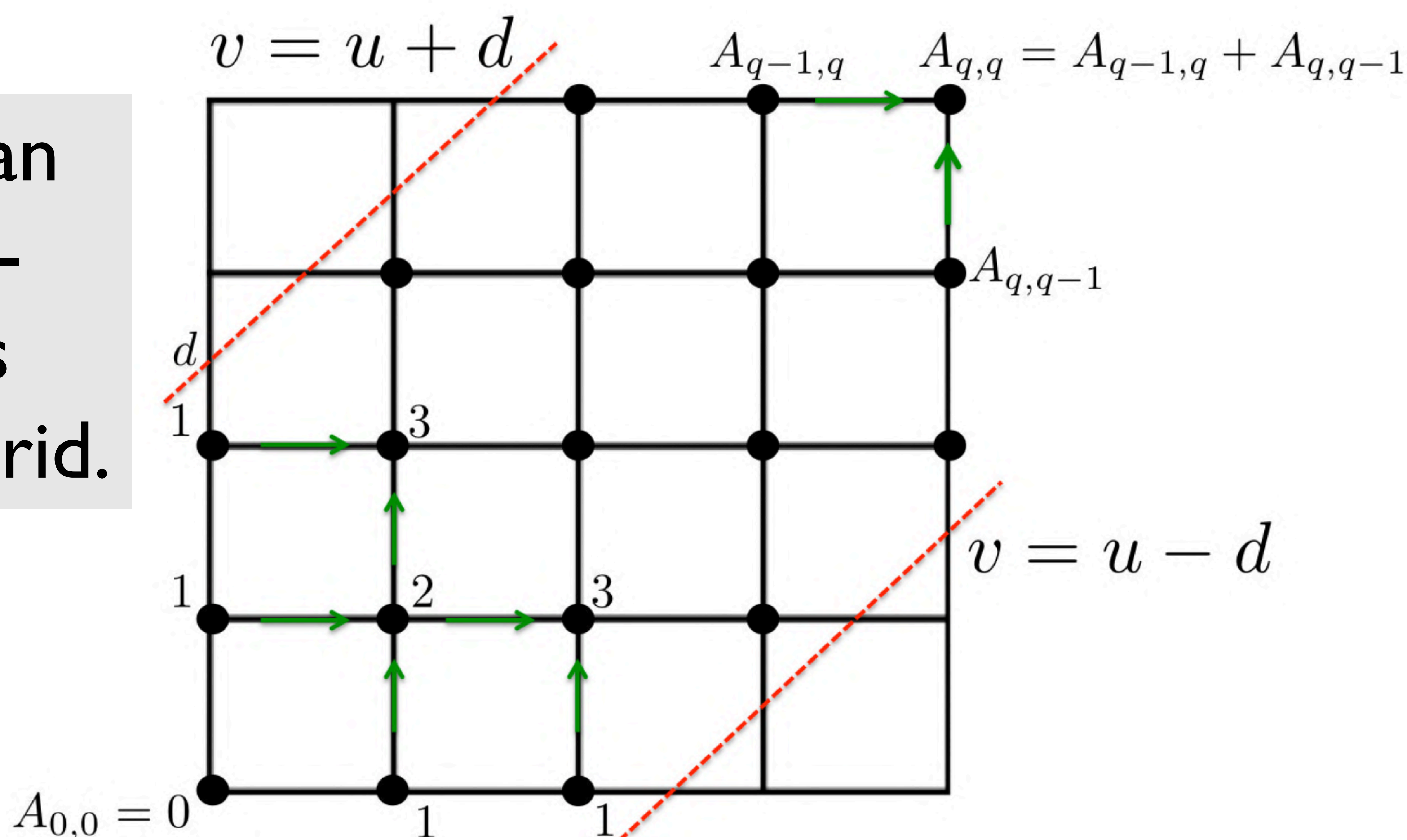


Exact Topological Inference

Theorem $$D_q = \sup_{1 \le j \le q} |\beta_i(G^1_{\lambda_j}) - \beta_i(G^2_{\lambda_j})|$$

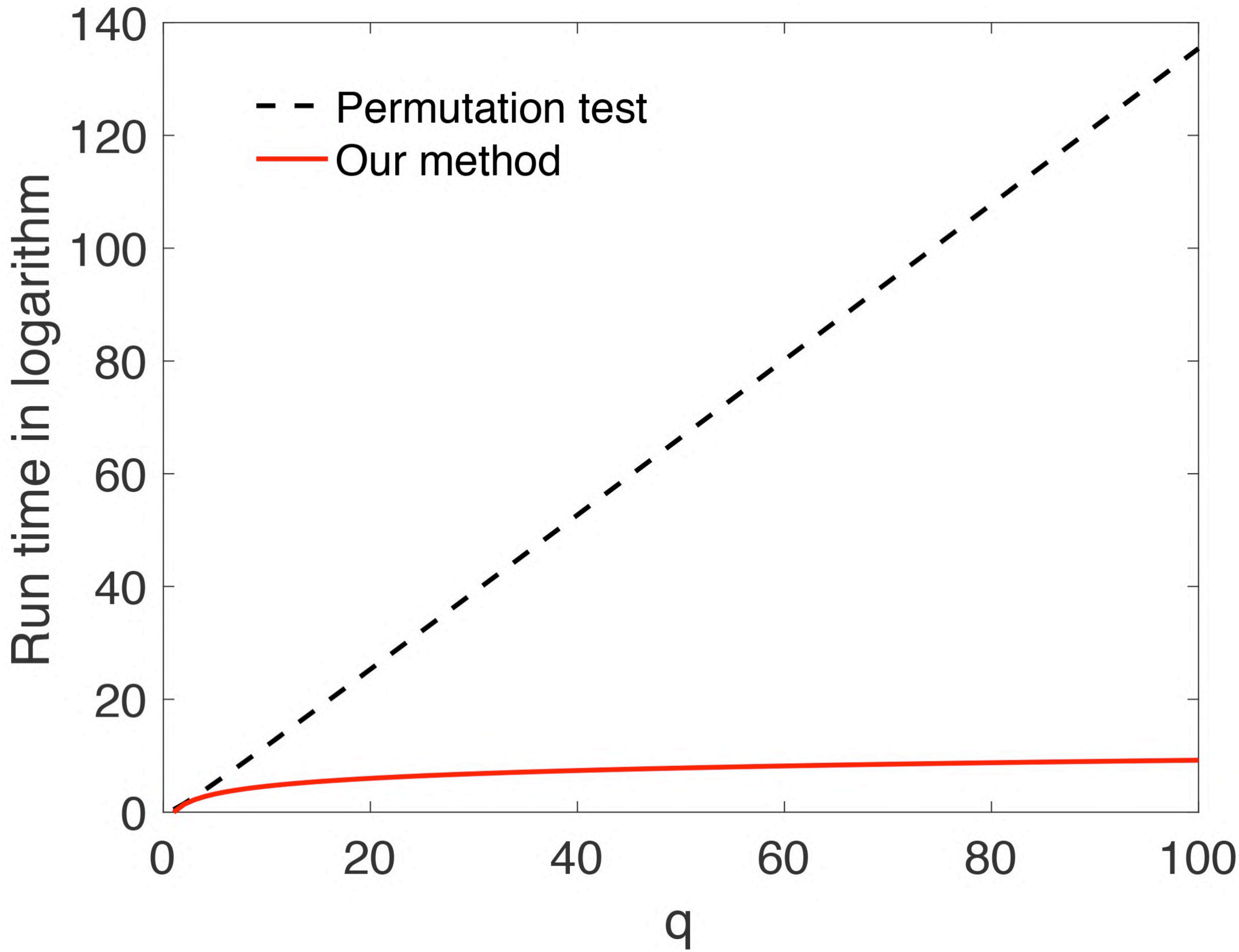
$$P(D_q \ge d) = 1 - \frac{A_{q,q}}{\binom{2q}{q}}$$

Permutations can be mapped one-to-one to walks on the square grid.



Chung et al. 2017 IPMI

Run time



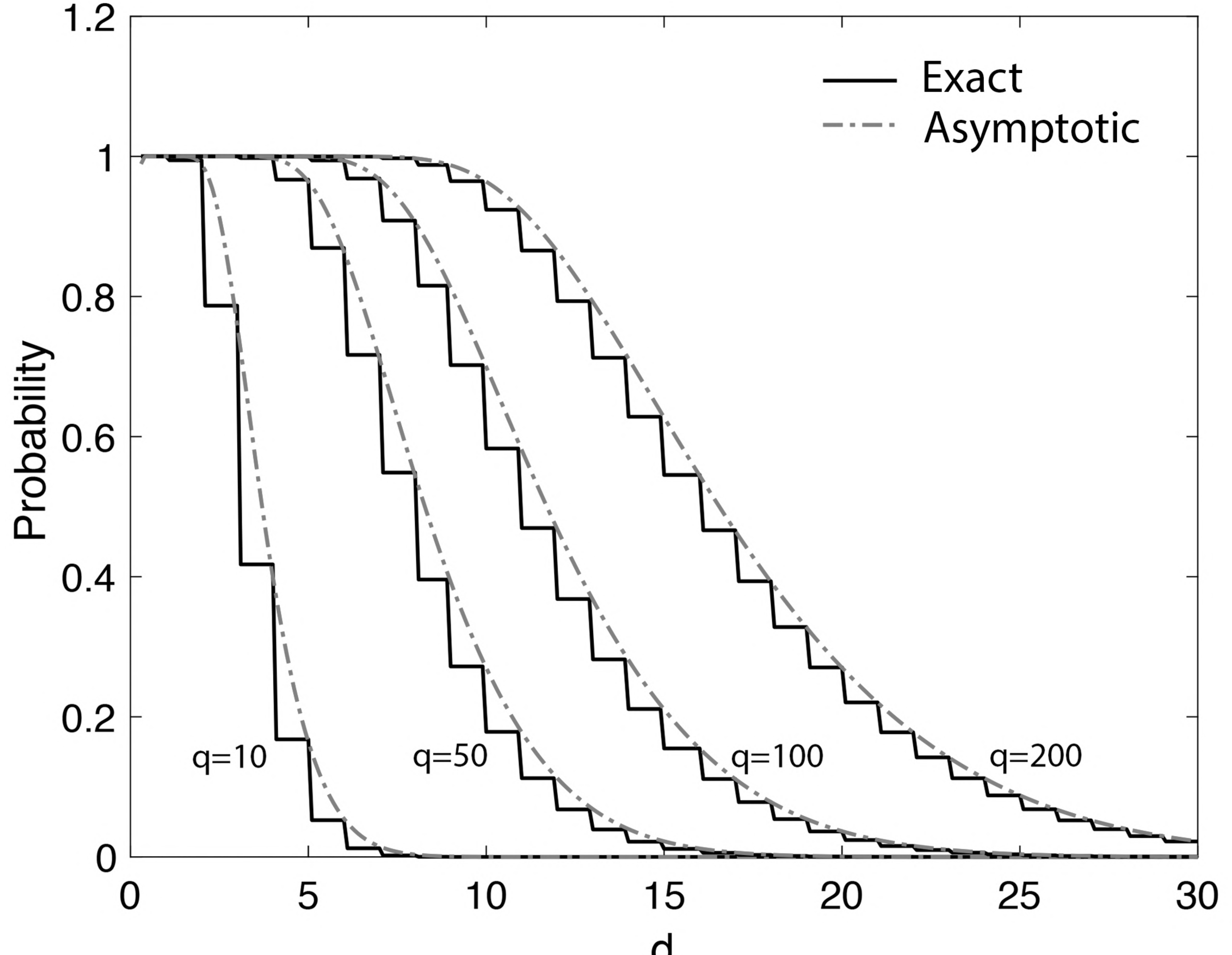
q = Number of edges

Permutation test impractical if sample size > 200

```
>> nchoosek(200,100)

Warning: Result may not be exact.
Coefficient is greater than
9.007199e+15 and is only accurate
to 15 digits

> In nchoosek (line 92)
ans =
  9.0549e+58
```

Asymptotic

$$\lim_{q\to\infty} P\left(D_q/\sqrt{2q} \geq d\right) = 2\sum_{i=1}^{\infty}(-1)^{i-1}e^{-2i^2d^2}$$



Validation via simulaiton

The purpose of (statistical) simulation is to generate synthetic data with the ground truth, where the performance of a method can be compared against existing methods.

Network simulation

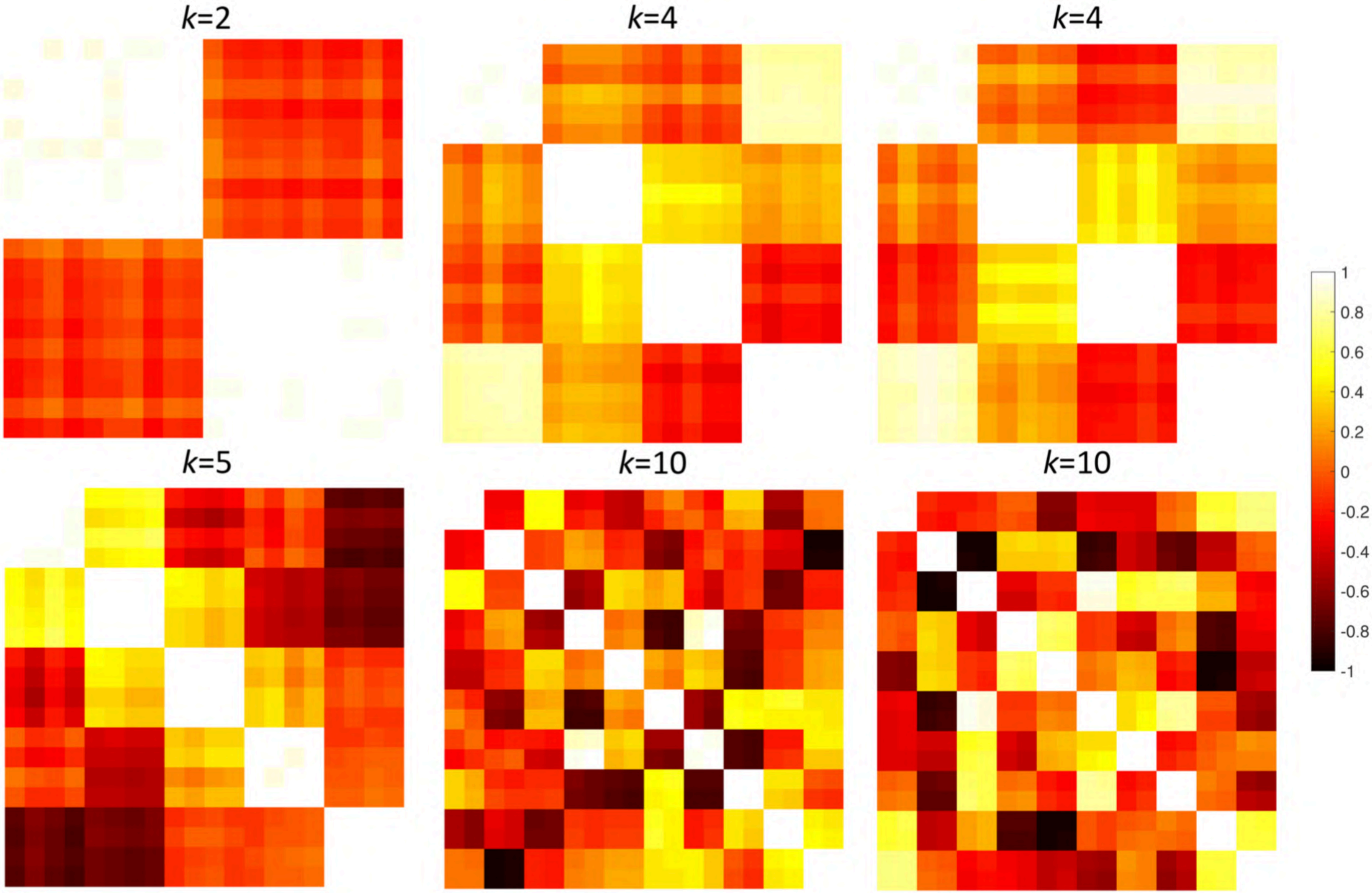
n x 1 data vector x_i at node *i*.

$$\mathbf{x}_i \sim N(0, I_n) \to C = (c_{ij}) = (corr(\mathbf{x}_i, \mathbf{x}_j))$$

$$\mathbb{E}C = I_n$$

Network with *k* modules

$$\begin{aligned}
\mathbf{y}_1, \cdots, \mathbf{y}_c &= \mathbf{x}_1 + N(0, \sigma^2 I_n) \\
\mathbf{y}_{c+1}, \cdots, \mathbf{y}_{2c} &= \mathbf{x}_{c+1} + N(0, \sigma^2 I_n) \\
&\vdots \\
\mathbf{y}_{c(k-1)+1}, \cdots, \mathbf{y}_{ck} &= \mathbf{x}_{c(k-1)+1} + N(0, \sigma^2 I_n)
\end{aligned}$$



k=2
k=4
k=4
k=5
k=10
k=10
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1

Matrix norm based distance

$$\mathcal{X}^1 = (V, w^1) \qquad \mathcal{X}^2 = (V, w^2)$$

$$D_l(\mathcal{X}^1, \mathcal{X}^2) = \Big(\sum_{i,j} |w^1_{ij} - w^2_{ij}|^l \Big)^{1/l}$$

$$D_\infty(\mathcal{X}^1, \mathcal{X}^2) = \max_{\forall i,j} |w^1_{ij} - w^2_{ij}|$$

Performance based on 100 simulations

False positives

Permutation test | ETI

p=20	L_1	L_2	L_∞	GH	KS (β_0)	KS (β_1)	Q
4 vs. 4	0.00	0.00	0.00	0.00	0.04	0.01	0.05
5 vs. 5	0.00	0.00	0.00	0.00	0.07	0.01	0.06
10 vs. 10	0.00	0.00	0.00	0.00	0.00	0.00	0.04
4 vs. 5	0.63	0.40	0.33	0.15	0.27	0.06	0.9
2 vs. 4	0.71	0.48	0.42	0.53	0.18	0.00	0.95
5 vs. 10	0.94	0.80	0.78	0.72	0.44	0.24	0.96

False negatives

Modularity

Performance based on 100 simulations

p=100	L_1	L_2	L_∞	GH	KS (β_0)	KS (β_1)	Q
4 vs. 4	0.00	0.00	0.00	0.00	0.26	0.54	0.03
5 vs. 5	0.00	0.00	0.00	0.00	0.14	0.43	0.05
10 vs. 10	0.00	0.00	0.00	0.00	0.05	0.05	0.05
4 vs. 5	0.51	0.37	0.35	0.16	0.11	0.00	0.93
2 vs. 4	0.66	0.45	0.57	0.61	0.03	0.00	0.91
5 vs. 10	0.94	0.86	0.79	0.72	0.11	0.00	0.98

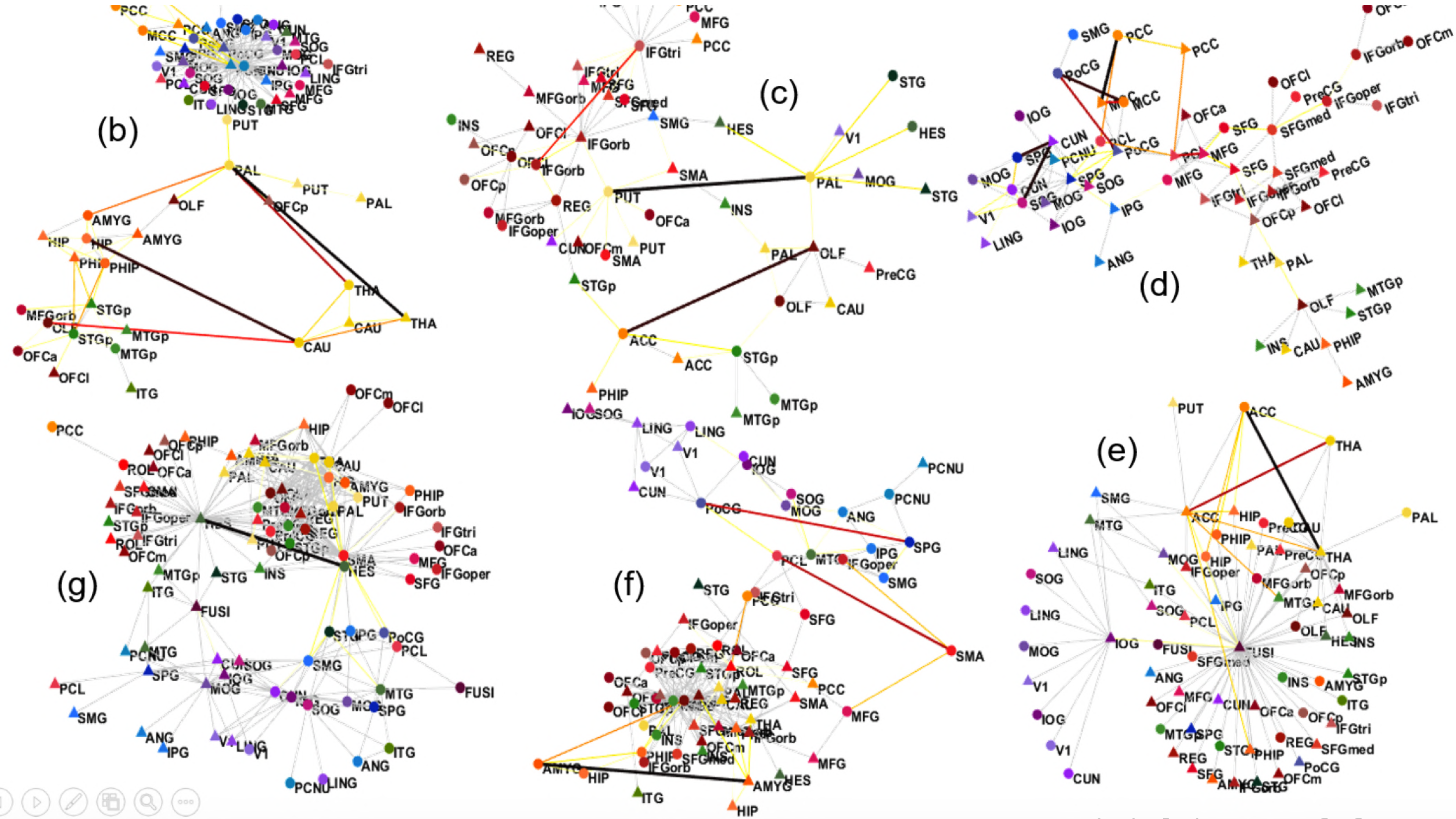
Performance based on 100 simulations

p= 500	L_1	L_2	L_∞	GH	KS (β_0)	KS (β_1)	Q
4 vs. 4	0.04	0.05	0.06	0.08	0.20	0.26	0.02
5 vs. 5	0.00	0.00	0.00	0.00	0.13	0.20	0.02
10 vs. 10	0.00	0.00	0.00	0.00	0.06	0.18	0.05
4 vs. 5	0.20	0.20	0.20	0.20	0.11	0.00	0.20
2 vs. 4	0.14	0.11	0.14	0.12	0.00	0.00	0.17
5 vs. 10	0.20	0.18	0.19	0.16	0.00	0.00	0.20

We need to come up with better *topology-aware* network distances!

What next?

Coidentification of cycles over multiple networks



Lee et al. 2019 MICCAI

Thank you!

Question? mkchung@wisc.edu