The Waisman Laboratory
for Brain Imaging and Behavior

University of Wisconsin
SCHOOL OF MEDICINE
AND PUBLIC HEALTH

Rapid Acceleration of the Permutation Test via Transpositions

Moo K. Chung

University of Wisconsin-Madison
www.stat.wisc.edu/~mchung

October 13, 2019 CNI workshop 13:37-13:49

Abstract

The permutation test is an often used test procedure for determining statistical significance in brain network studies. Unfortunately, generating every possible permutation for large-scale brain imaging datasets such as HCP and ADNI with hundreds of subjects is not practical. Many previous attempts at speeding up the permutation test rely on various approximation strategies such as estimating the tail distribution with known parametric distributions. In this study, we propose the novel transposition test that exploits the underlying algebraic structure of the permutation group. The method is applied to a large number of diffusion tensor images in localizing the regions of the brain network differences.

Acknowledgement

Shih-Gu Huang, Yixian Wang

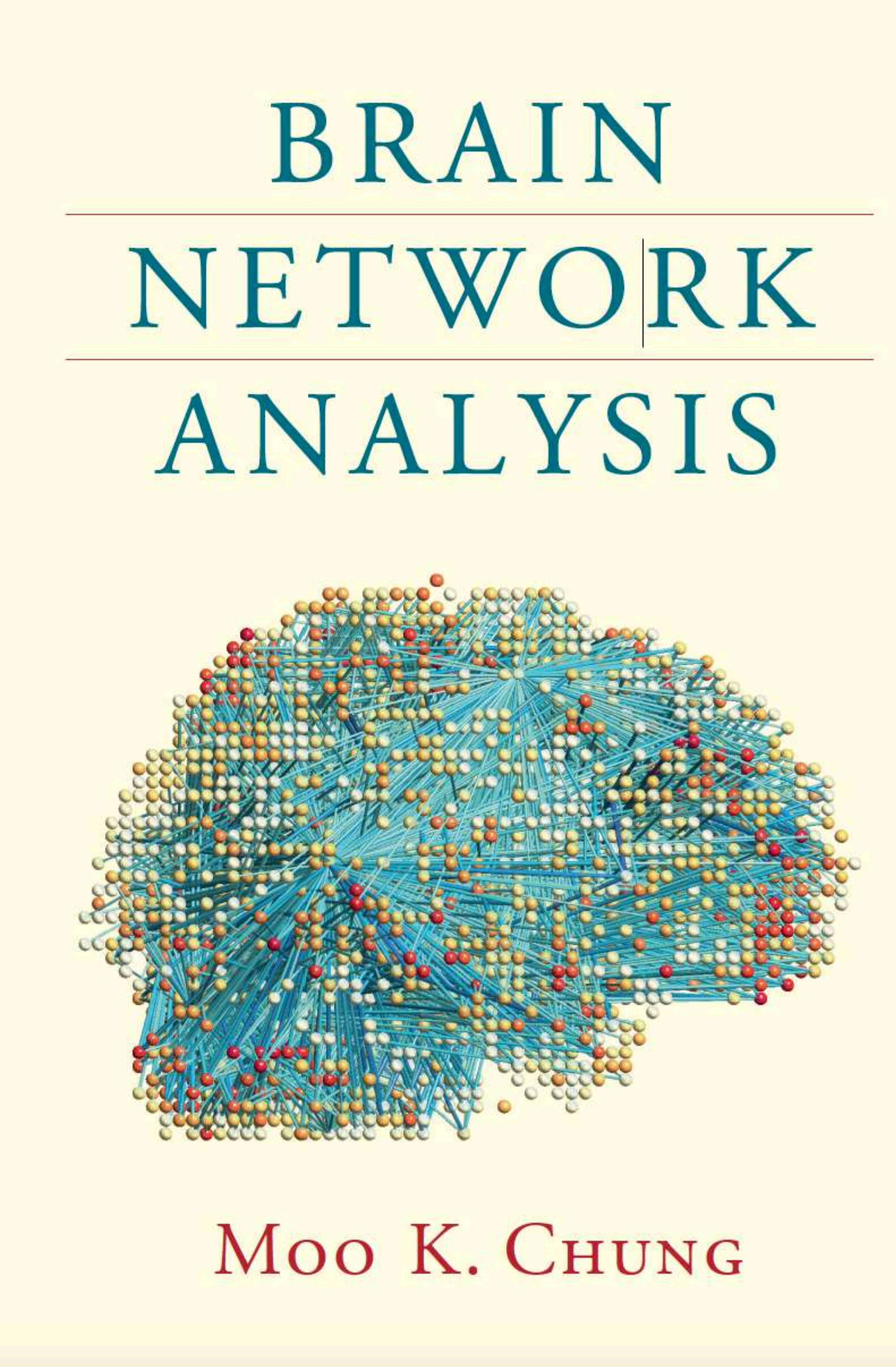
University of Wisconsin-Madison, USA

Linhui Xie, JingwenYan

Indiana University-Perdue University, USA

Li Shen *University of Pennsylvania, USA*

Grants NIH R01 EB022856, R01 EB022574, NSF IIS 1837964



New book!

Published August 3, 2019
Cambridge University Press

NIH funded
Postdoctoral Research
Associate Position
at University of
Wisconsin-Madison
2020-2022

ISBI 2020 Workshop April 4, 2020

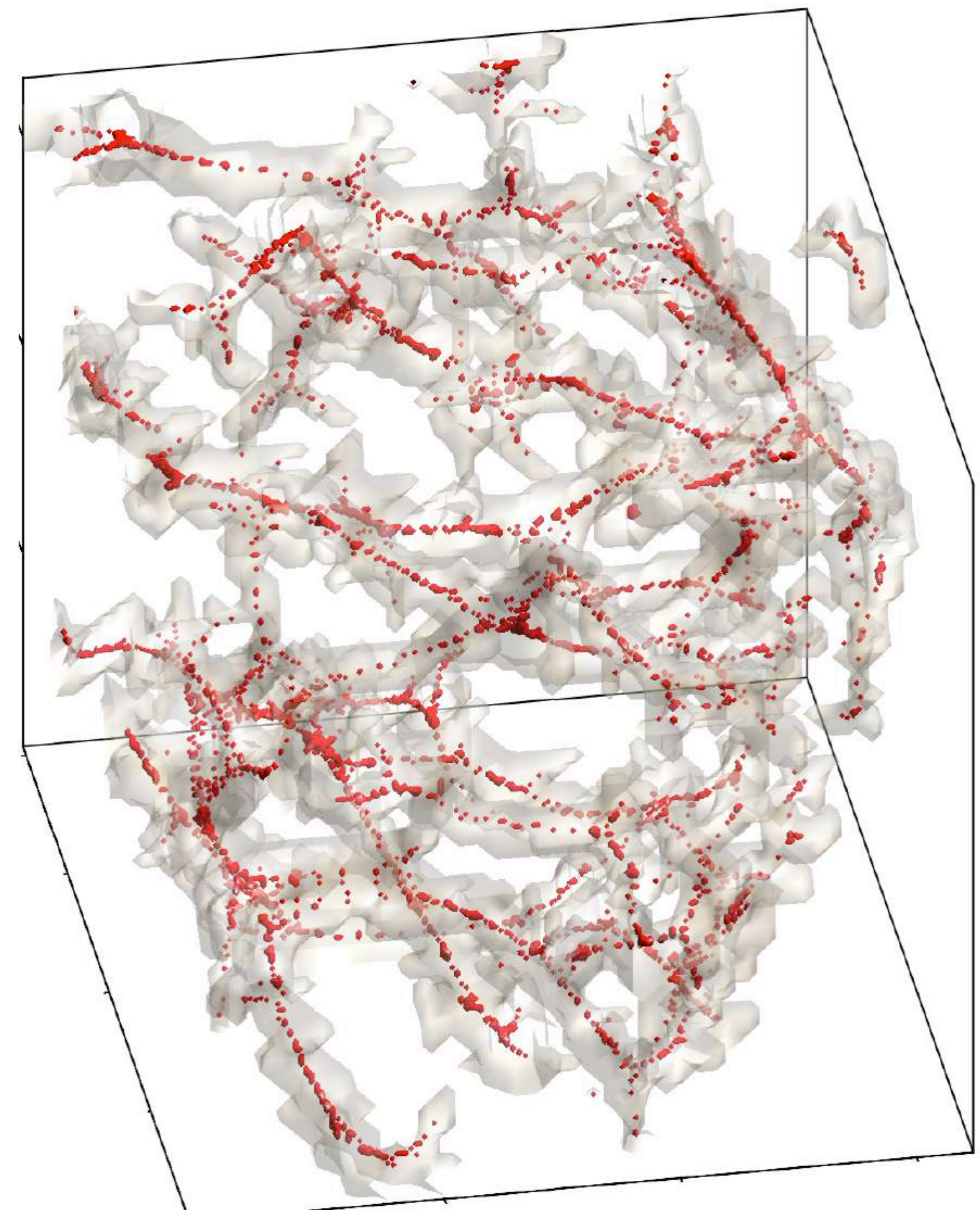
Interaction of Topology and Geometry

Organizers:

Joseph Reinhardt
University of Iowa

Moo K. Chung
Univ. of Wisconsin-Madison

Six Invited speakers
Best paper awards



What is 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*

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

Computational bottleneck

History of permutation test

Fisher 1935, The Design of Experiment

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

Thompson et al. 2001, Nature Neuroscience

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

Nichols et al. 2002, Human Brain Mapping

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

Google scholar 434,000 papers.

Limitation of permutation test

Serious computational bottleneck

1) Need to permute million voxels.
2) Very slow: Exponential run time

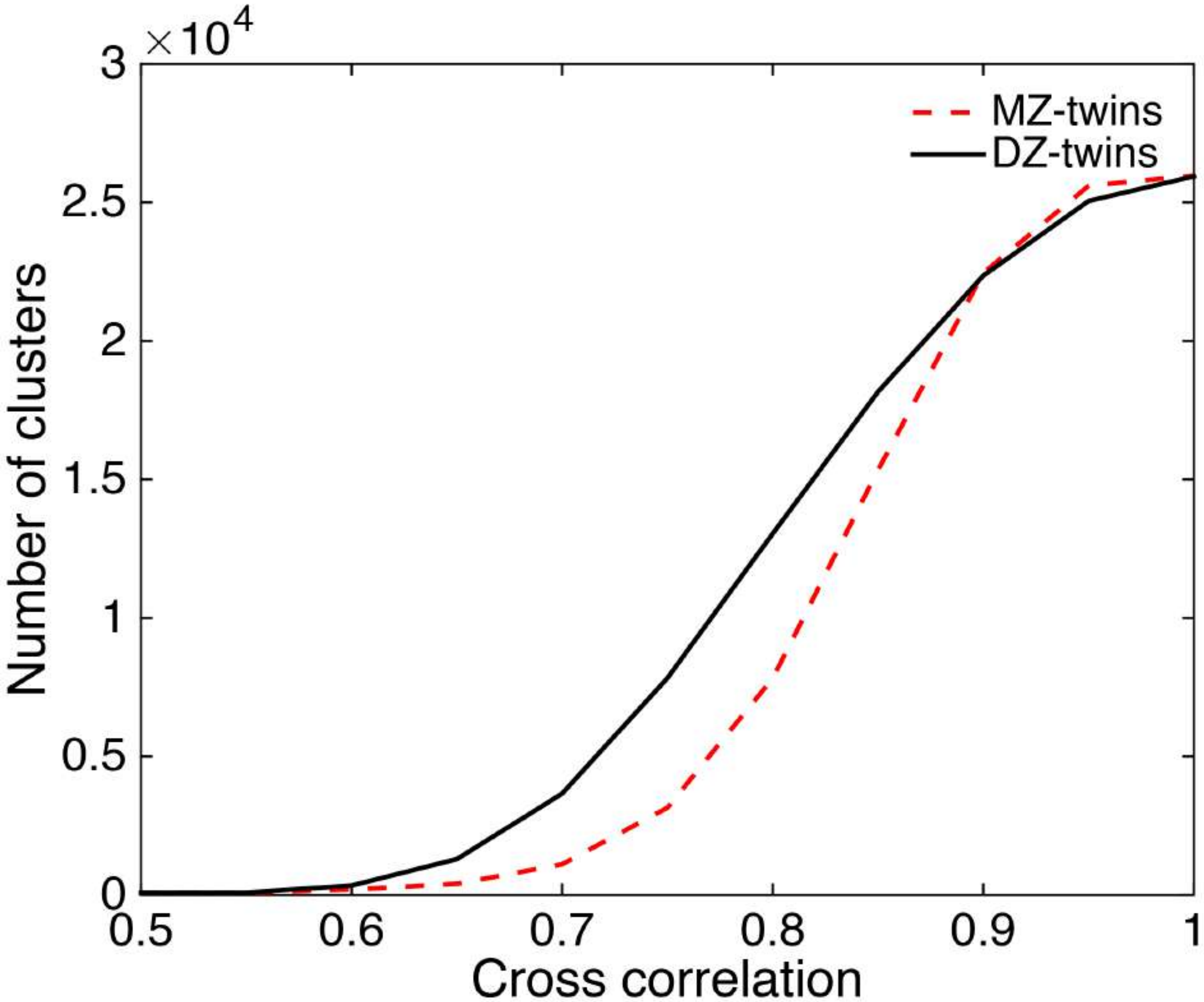
Thompson *et al*. (2001) used supercomputer:
1million permutations for

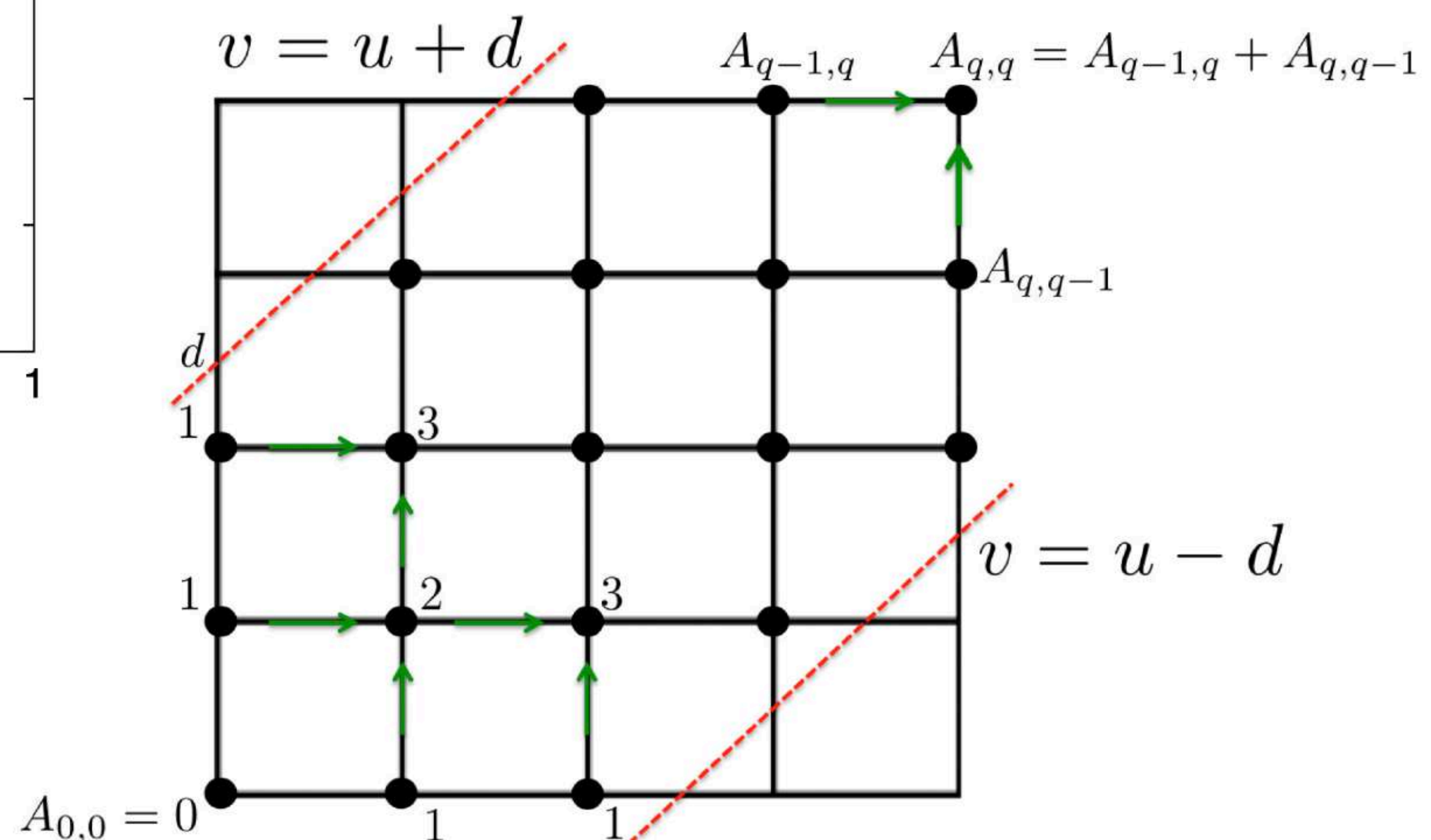
$$\binom{40}{20} = 1.34 \cdot 10^{11}$$ hundred billion

Pervious method: Exact topological inference (ETI)

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

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





Chung et al. 2017 IPMI

Random transposition on the permutation group

$$\mathbf{x} = (x_1, x_2, \cdots, x_{i-1}, x_i, x_{i+1}, \cdots, x_m)$$

transpose *i*-th and *j*-th data

$$\mathbf{y} = (y_1, y_2, \cdots, y_{j-1}, y_j, y_{j+1}, \cdots, y_n)$$

$$\pi_{ij}(\mathbf{x}) = (x_1, x_2, \cdots, x_{i-1}, y_j, x_{i+1}, \cdots, x_m)$$

$$\pi_{ij}(\mathbf{y}) = (y_1, y_2, \cdots, y_{j-1}, x_i, y_{j+1}, \cdots, y_n)$$

Theorem 1: If the test statistic is algebraic, there exists a function g such that

$$f(\pi_{ij}(\mathbf{x}), \pi_{ij}(\mathbf{y})) = g(f(\mathbf{x}, \mathbf{y}), x_i, y_i)$$

where computational complexity of g is constant.

Theorem 2: Any permutation in $\mathbb{S}_{m+n}$ can be reachable by a sequence of transpositions.

Chung et al. 2019 CNI

Online computation for t-stat.

$$\nu(\mathbf{x}) = \sum_{j=1}^{m} x_j, \qquad \omega(\mathbf{x}) = \sum_{j=1}^{m} \left(x_j - \frac{\nu(\mathbf{x})}{m} \right)^2$$

O(m) O($3m$+2)

$$\nu(\pi_{ij}(\mathbf{x})) = \nu(\mathbf{x}) - x_i + y_j$$ O(2)

$$\omega(\pi_{ij}(\mathbf{x})) = \omega(\mathbf{x}) - x_i^2 + y_j^2 + \frac{\nu(\mathbf{x})^2 - \nu(\pi_{ij}(\mathbf{x}))^2}{m}$$

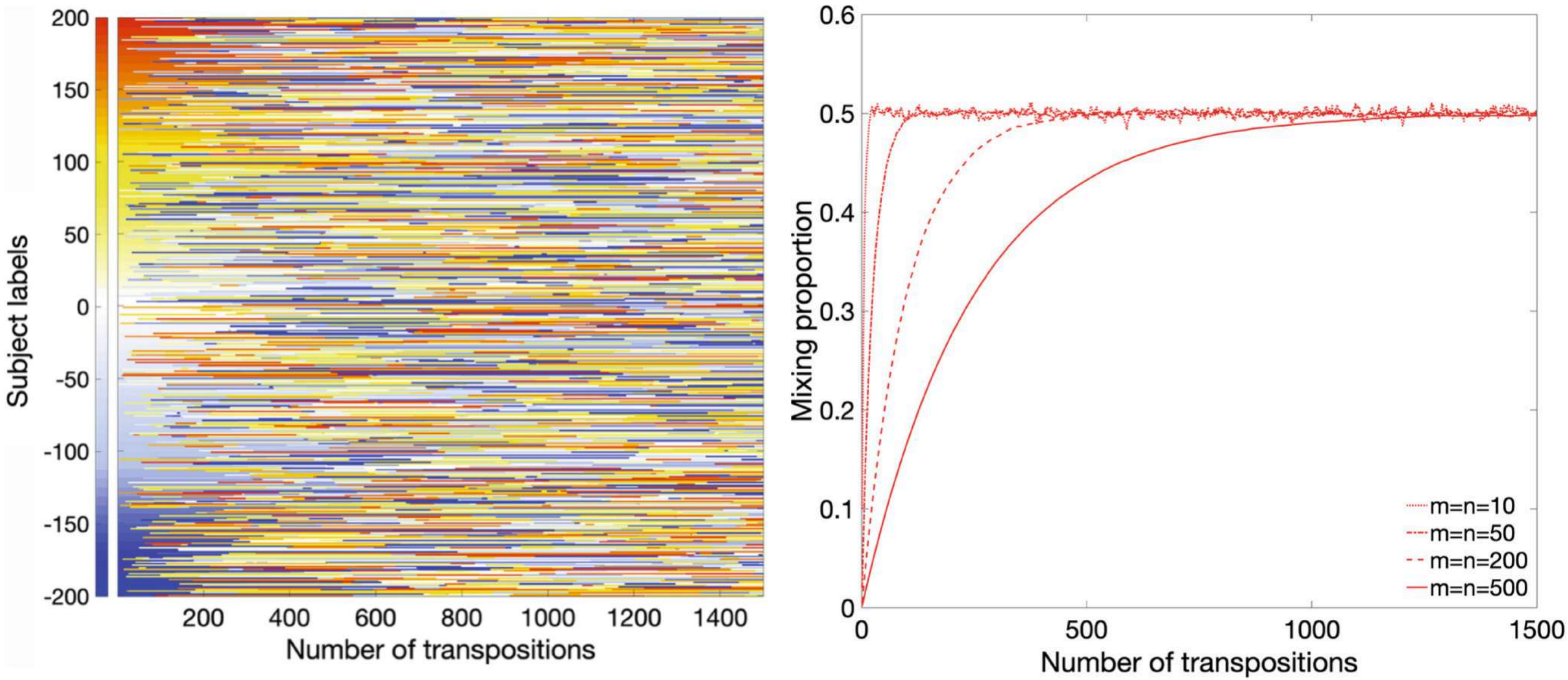
O(9)

T-stat computation per permutation

Permutation test: O($4m$+$4n$+20)

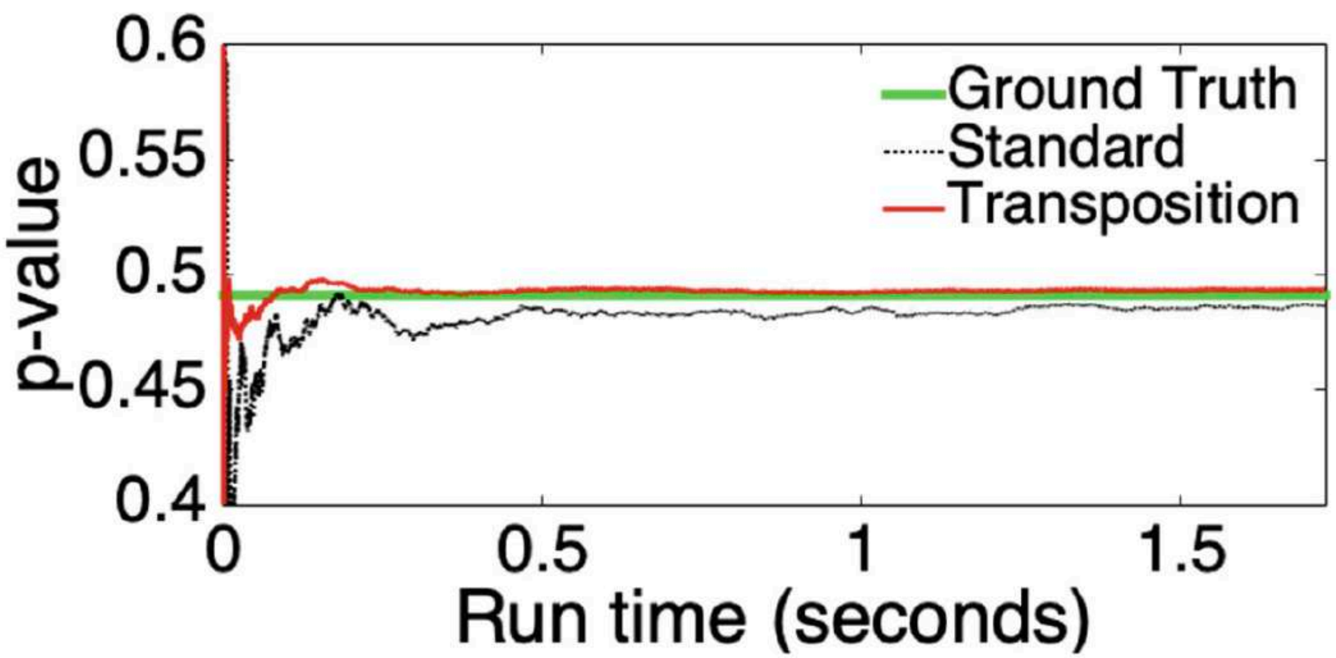
Transposition test: O(35)

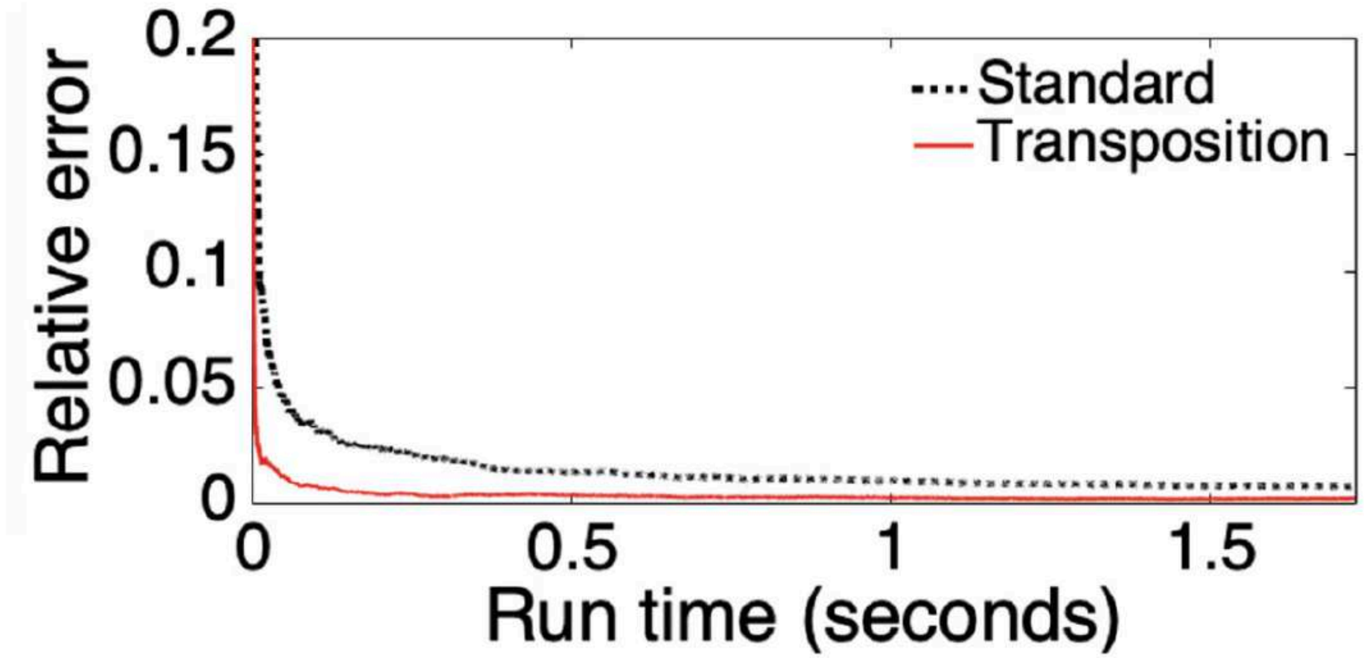
Simulation: Mixing proportion over transpositions



m=*n*=10

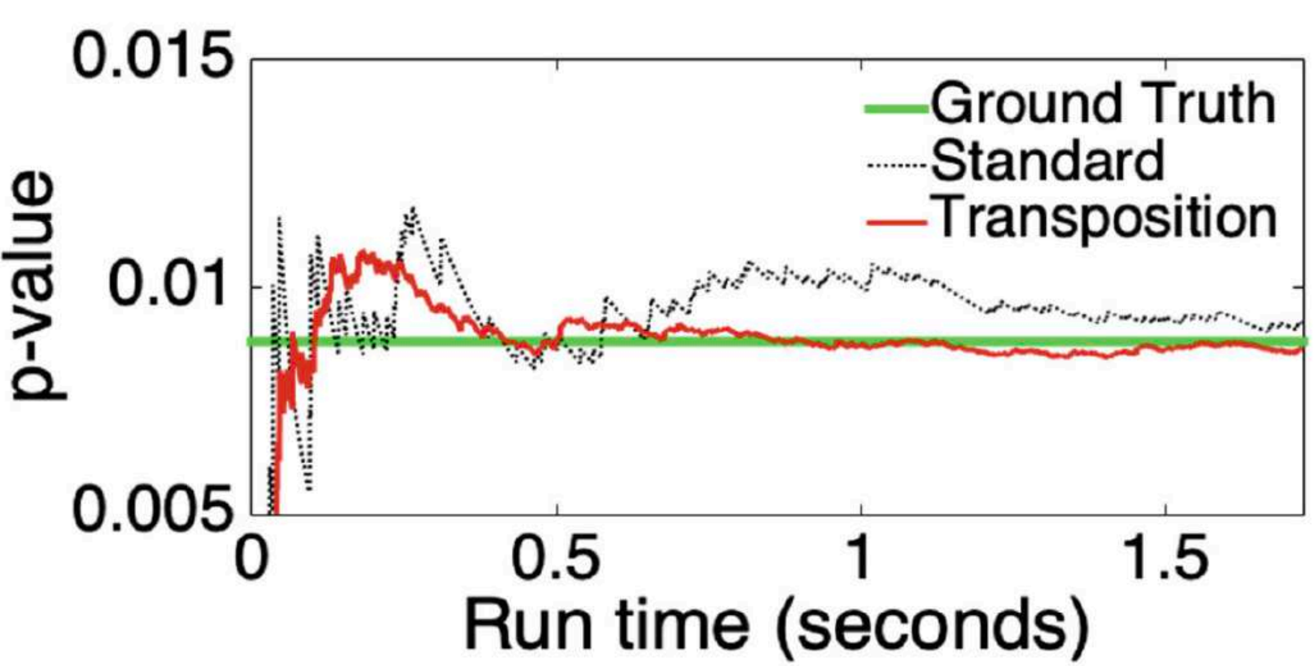
$$x_1, \cdots, x_m \sim 0.1 + Unif(0,1)$$
$$y_1, \cdots, y_n \sim Unif(0,1)$$

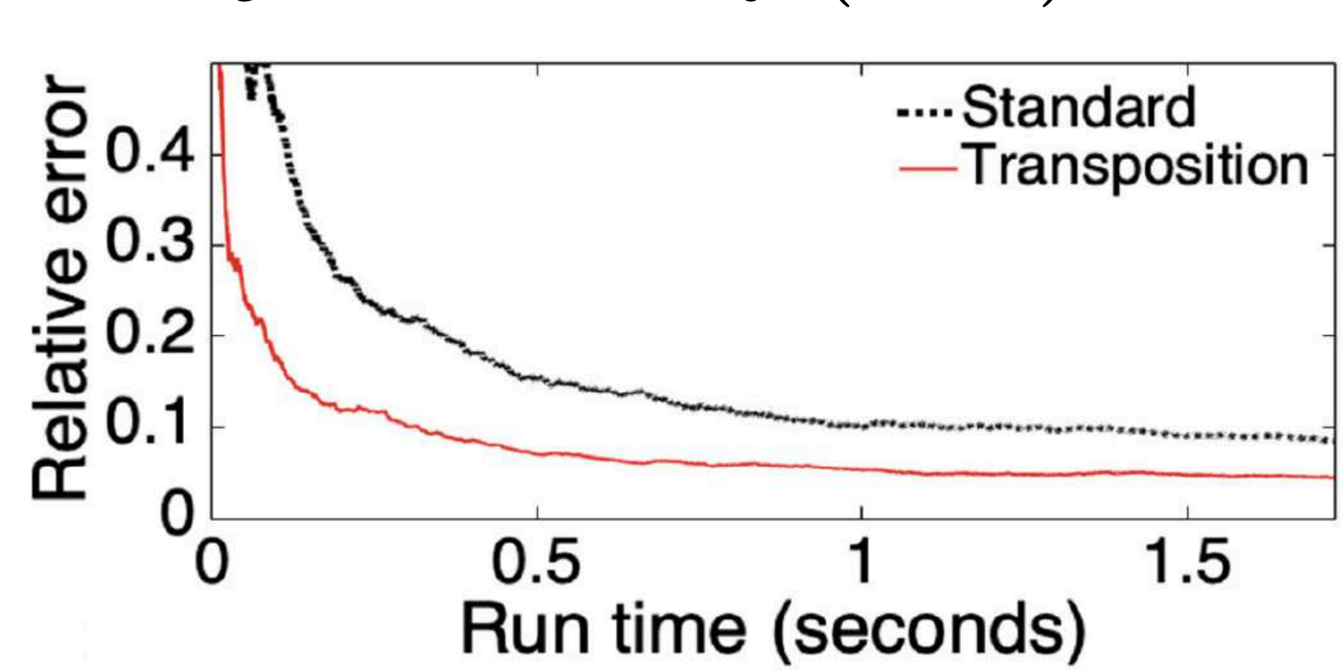




m=*n*=100

$$x_1, \cdots, x_m \sim 0.1 + Unif(0,1)$$
$$y_1, \cdots, y_n \sim Unif(0,1)$$





Matlab code

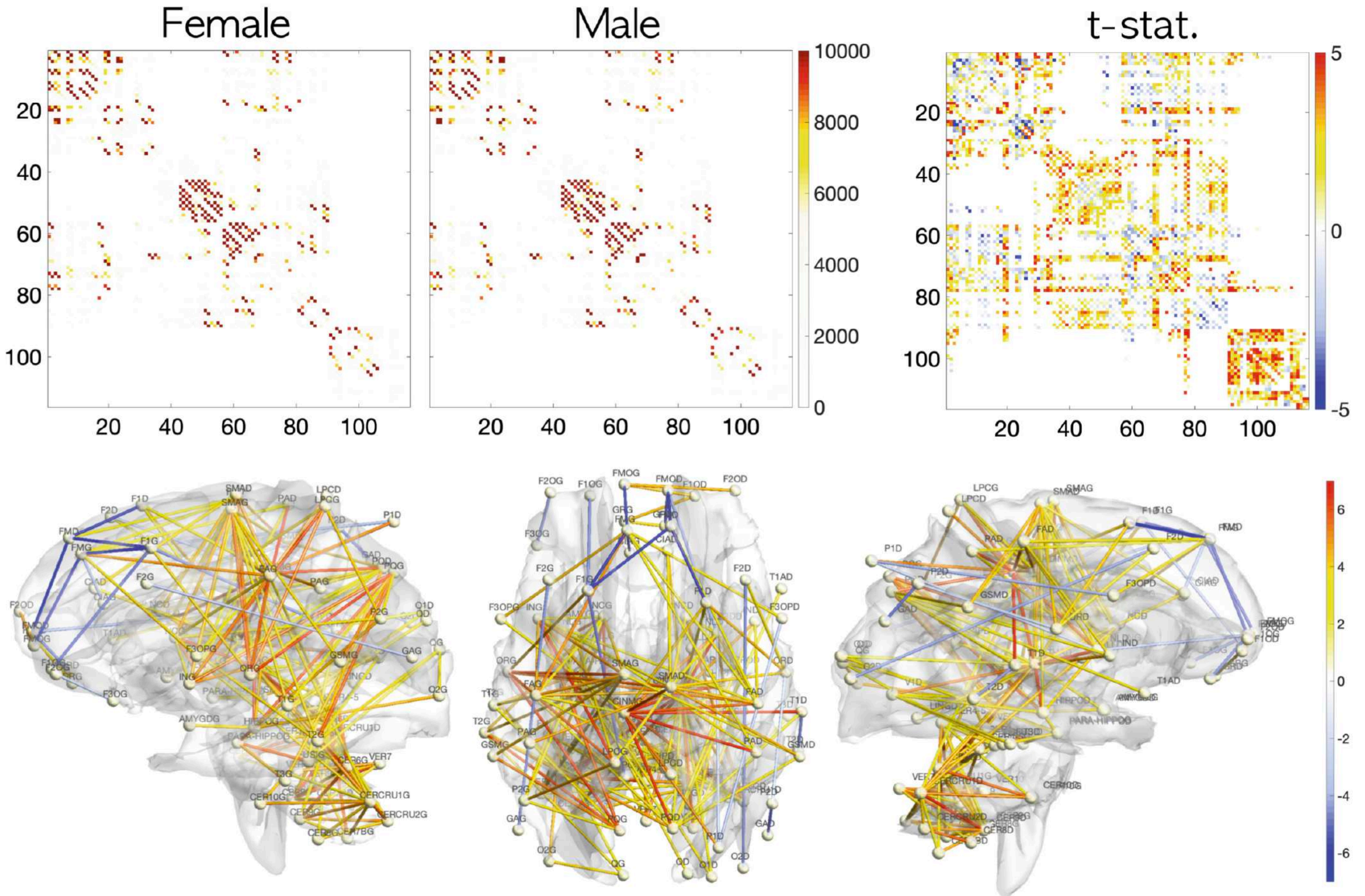
http://www.stat.wisc.edu/~mchung/transpositions

Permutation test

```
[stat_s, time_s] = test_permute (x , y, per_s)
```

Transposition test

```
stat_t=[];
for i=1:10000
    [stat, time] = test_transpose (x, y, per_t / 10000);
    stat_t=[stat_t; stat];
    time_t=time_t + time;
end
```



t-stat (202 females – 154 males)

Thank you!

Question? mkchung@wisc.edu