The Waisman Laboratory for Brain Imaging and Behavior

University of Wisconsin SCHOOL OF MEDICINE AND PUBLIC HEALTH

Rapid Acceleration of Permutation Test via Transpositions

Moo K. Chung
University of Wisconsin-Madison

www.stat.wisc.edu/~mchung

NIBIB R01 EB022856

The method and code published in

Chung, M.K. et al. 2019. Rapid Acceleration of the permutation test via transpositions, *International Workshop on Connectomics in NeuroImaging*, Lecture Notes in Computer Science 11848:42-53.

http://pages.stat.wisc.edu/~mchung/papers/chung.2019.CNI.pdf

Matlab code

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

Given vector data x, y and the the number of permutation per_s the **permutation test** is done by

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

Given the the number of transpositions per_t, the **transposition test** is done by

```
stat_t=[];
for i=1:1000
   [stat, time] = test_transpose (x, y, per_t / 1000);
   stat_t=[stat_t; stat];
end
```

Standard 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

Permutation group of order *m*+*n*

Indicator function

Stat. at each permutation

Observation

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

Computing stat. at each permutation is the usual computational bottleneck

Limitation of permutation test

Serious computational bottleneck in large sample brain imaging studies.

1) Need to permute million voxels.
2) Compute the statistic for each permutation.

The proposed transposition test will bypass the computational bottleneck.

What is transposition?

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

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

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

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

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

How statistics change over a transposition

$$\mathbf{x} = (x_1, \cdots, x_m)$$

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

Mean and variance function changed incrementally

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

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

Online computation over transpositions

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

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

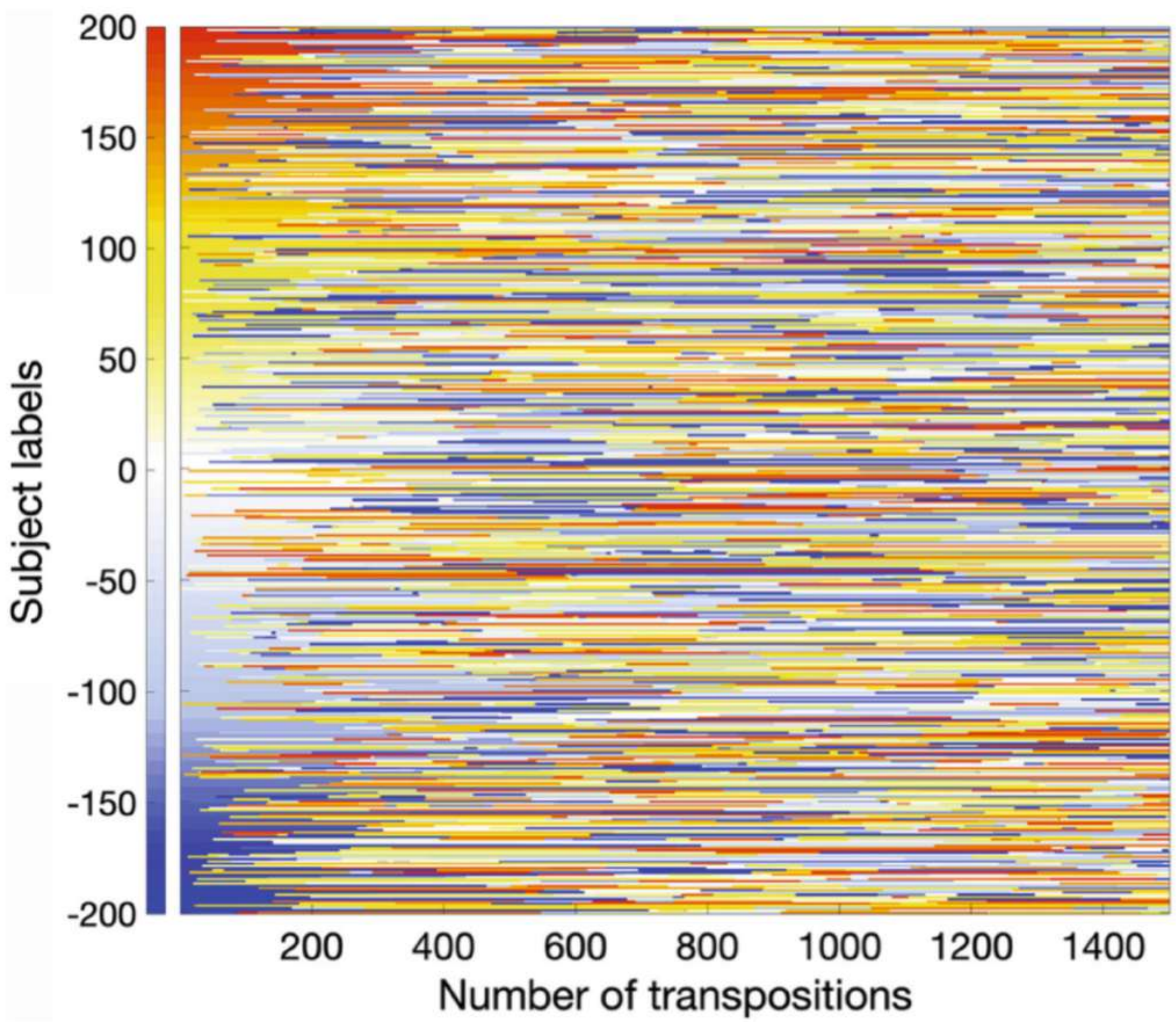
$$\nu(\pi_{ij}(\mathbf{x})) = \nu(\mathbf{x}) - x_i + y_j$$ O(2)

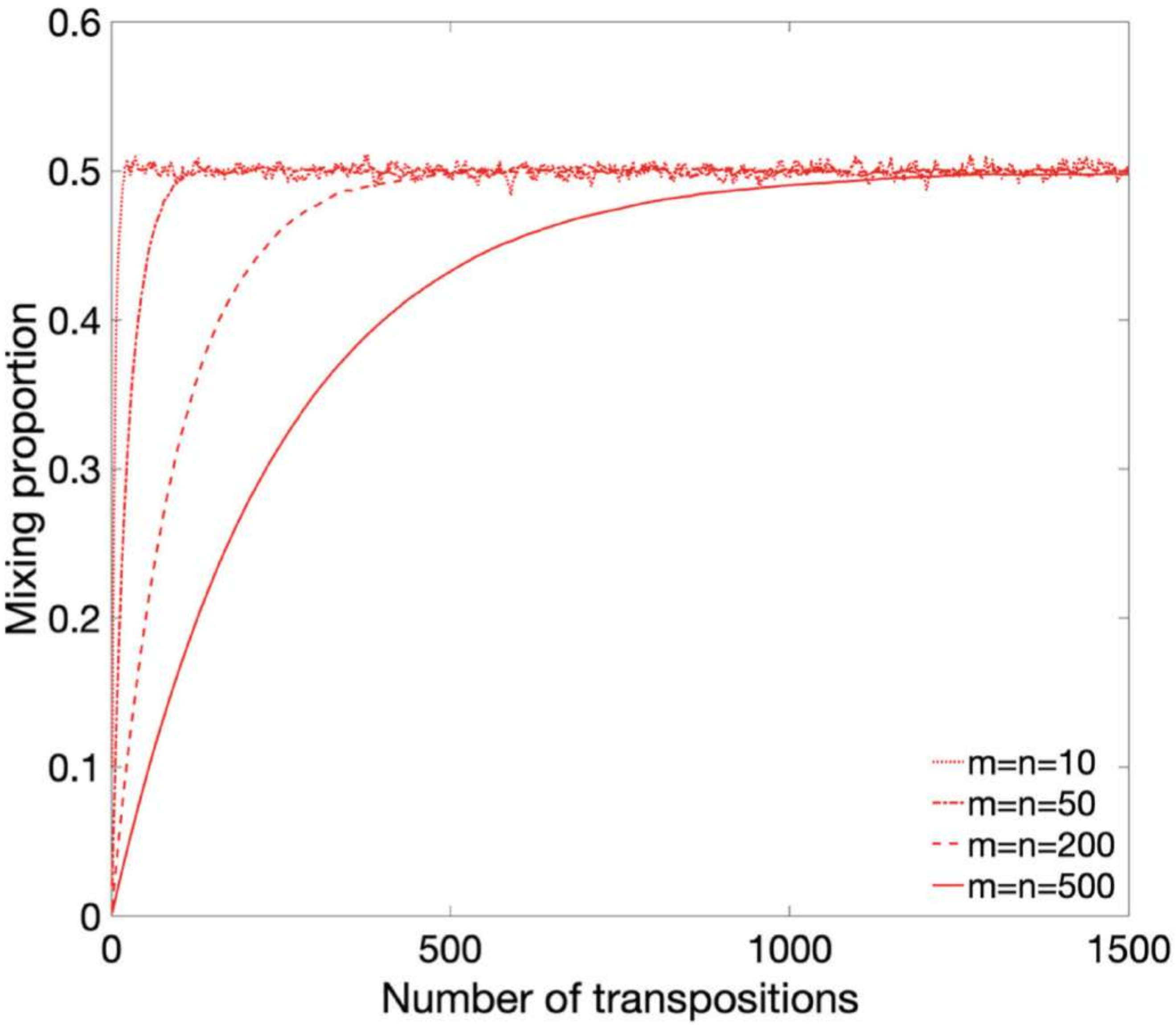
$$\omega(\pi_{ij}(\mathbf{x})) = \omega(\mathbf{x}) - x_i^2 + y_j^2 + \frac{\nu(\mathbf{x})^2 - \nu(\pi_{ij}(\mathbf{x}))^2}{m}$$ O(9)

T-stat computation per permutation O($4m$+$4n$+20)

T-stat computation per transpositions O(35)

Two groups mix quickly over transpositions

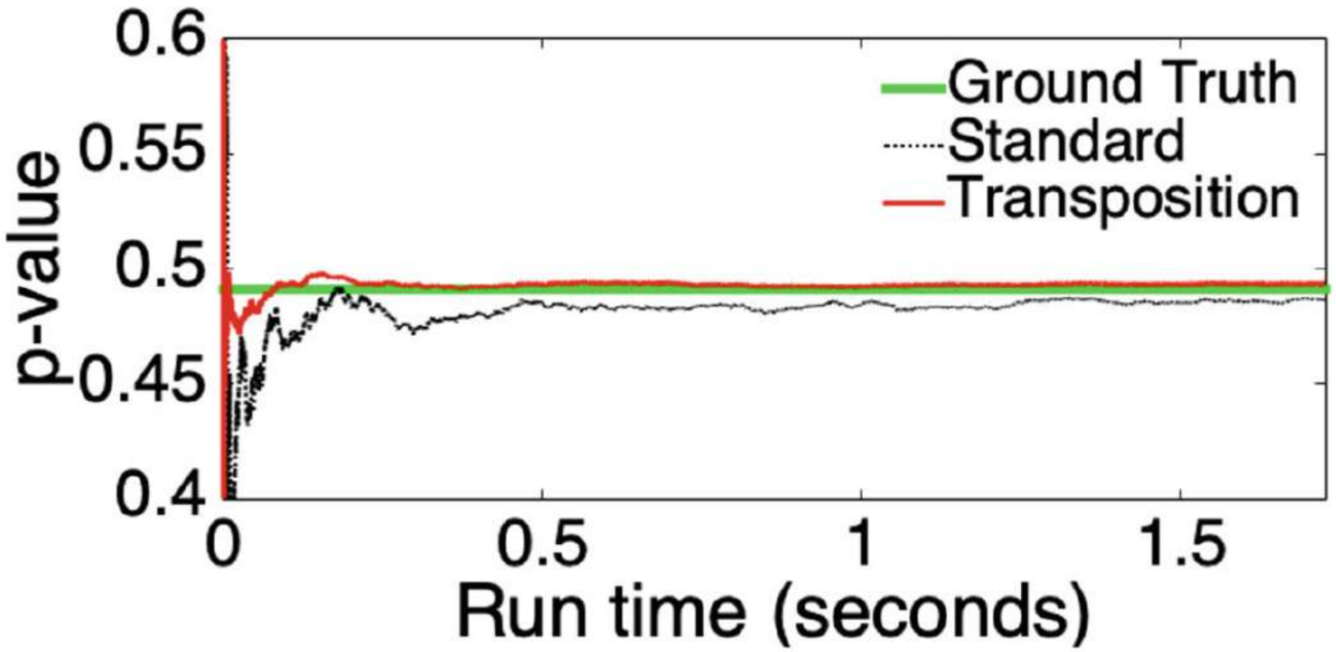


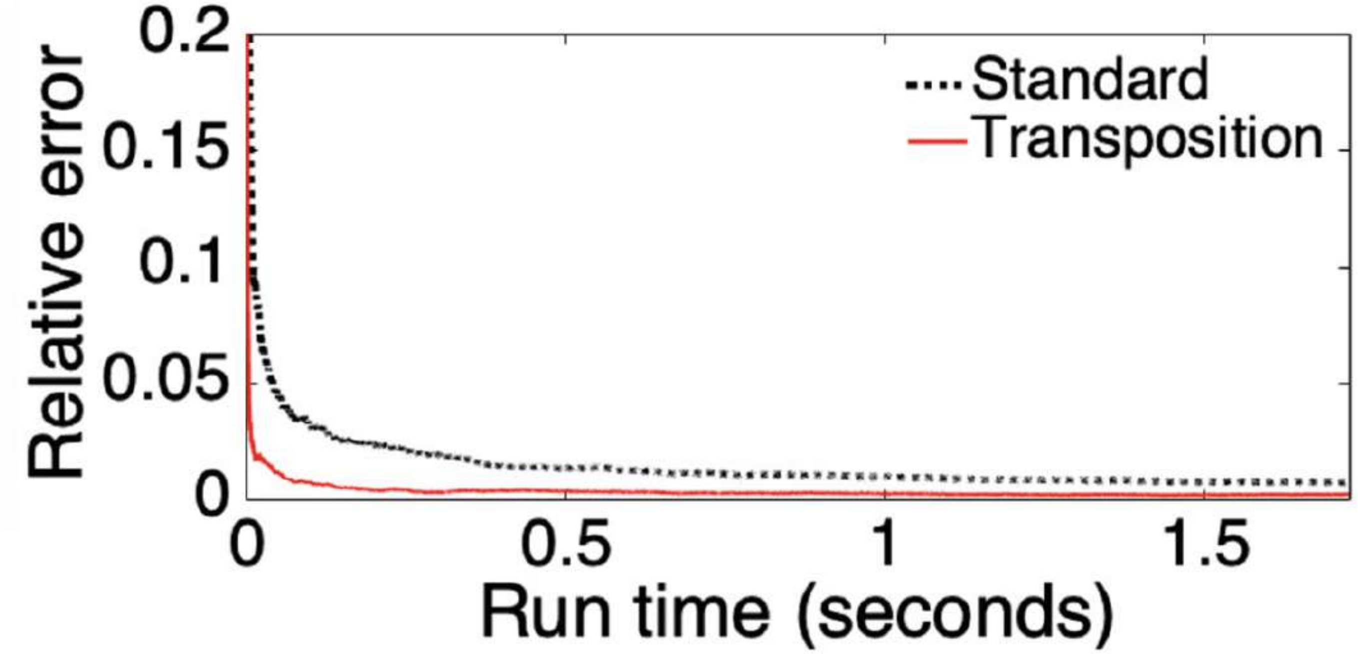


Convergence

m=n=10

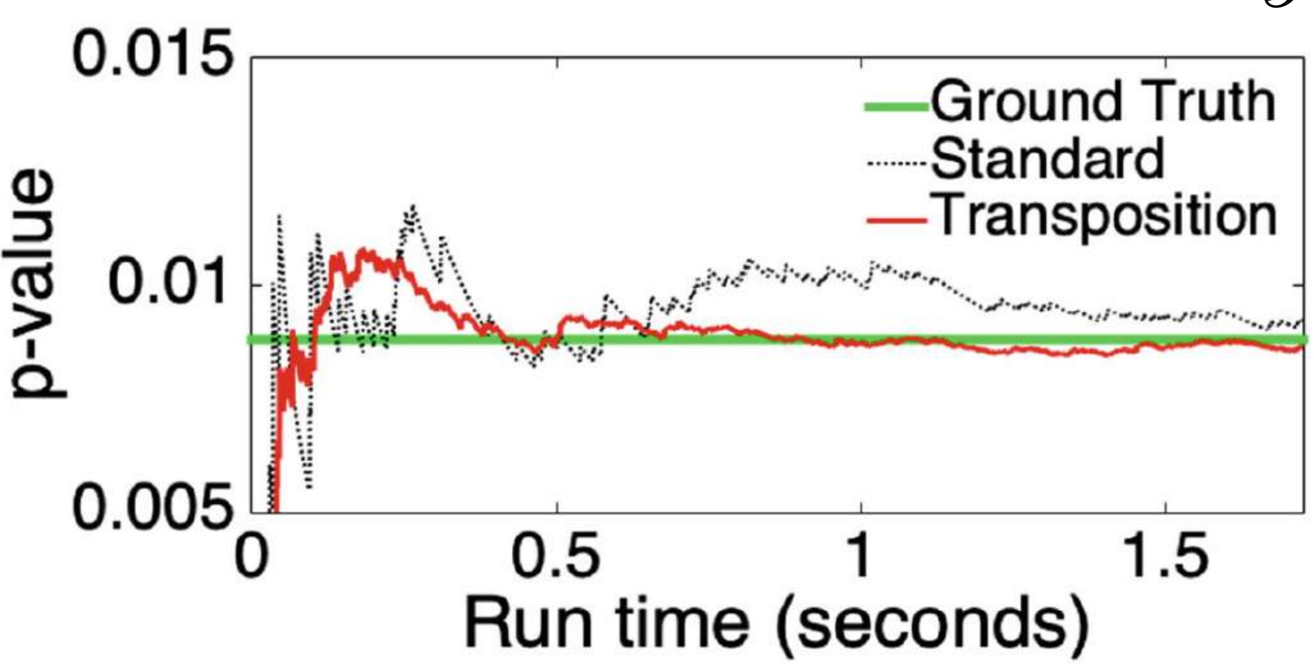
$$x_1, \cdots, x_m \sim 0.1 + Unif(0,1)$$
$$y_1, \cdots, y_n \sim Unif(0,1)$$

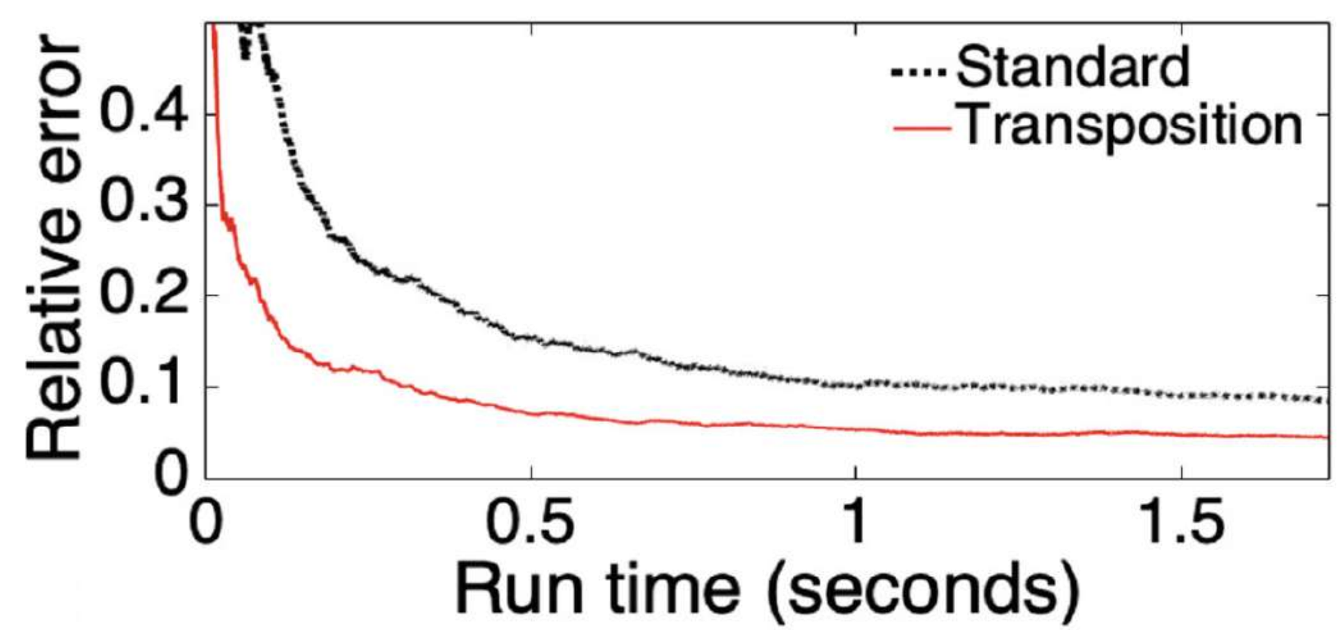




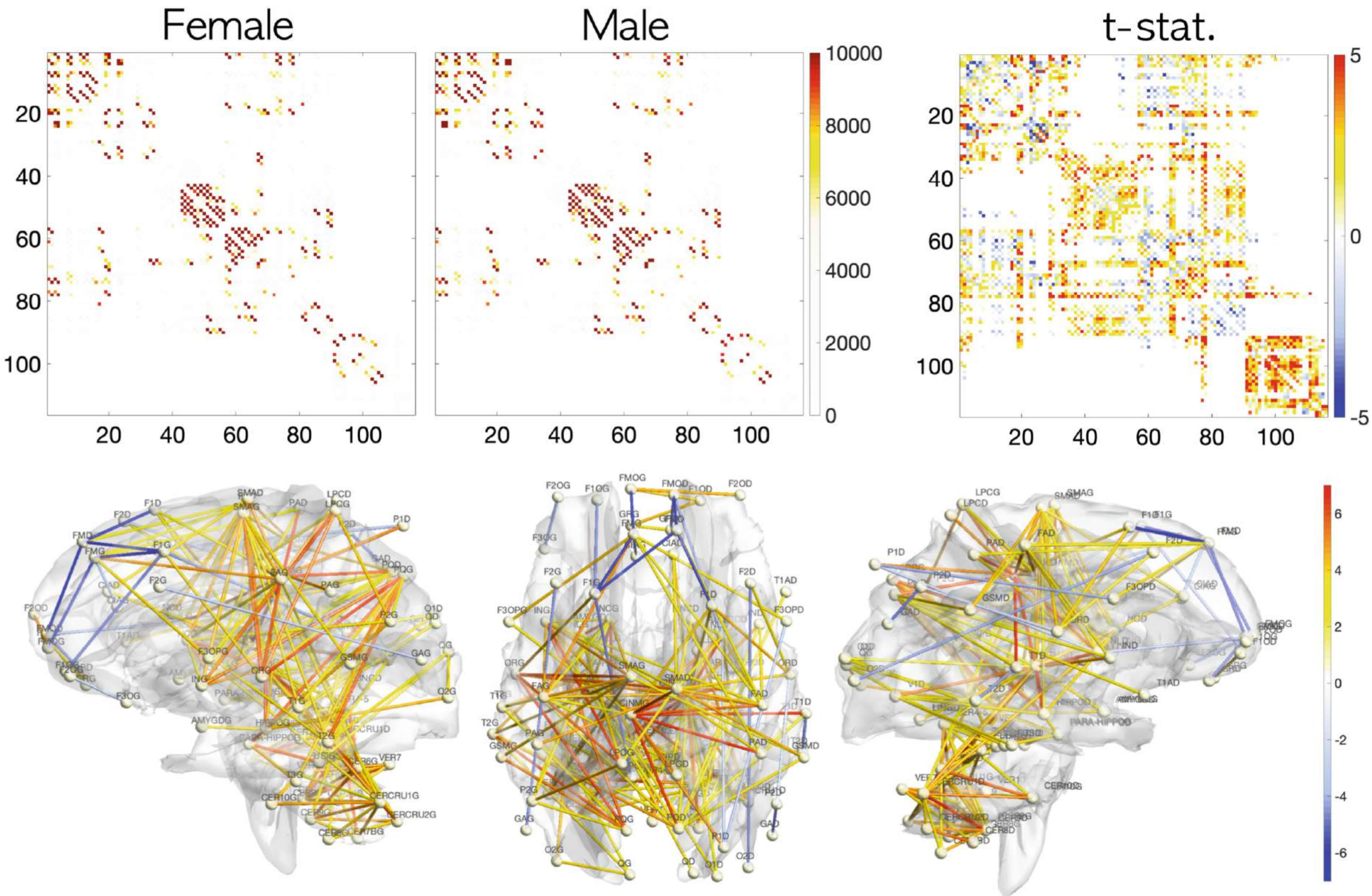
m=n=100

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





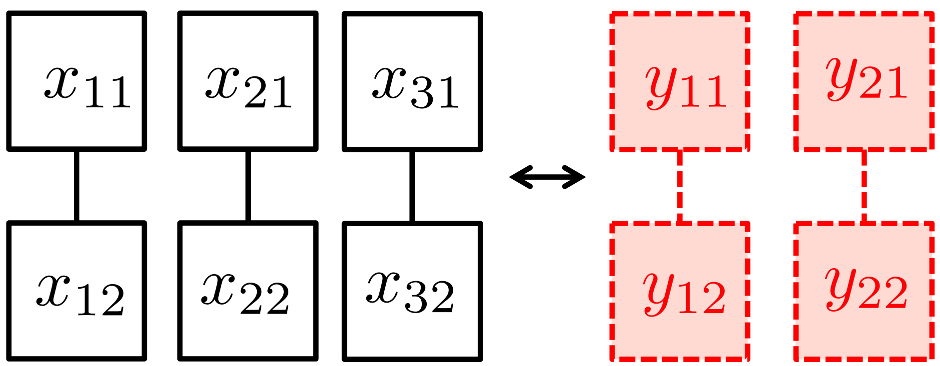
DTI study: structural connectivity difference



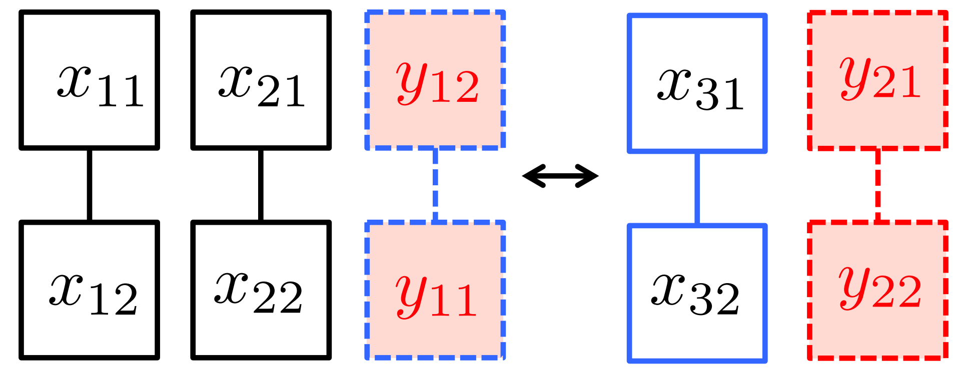
t-stat (202 females – 154 males)

More complex example: possible transpositions in twin data

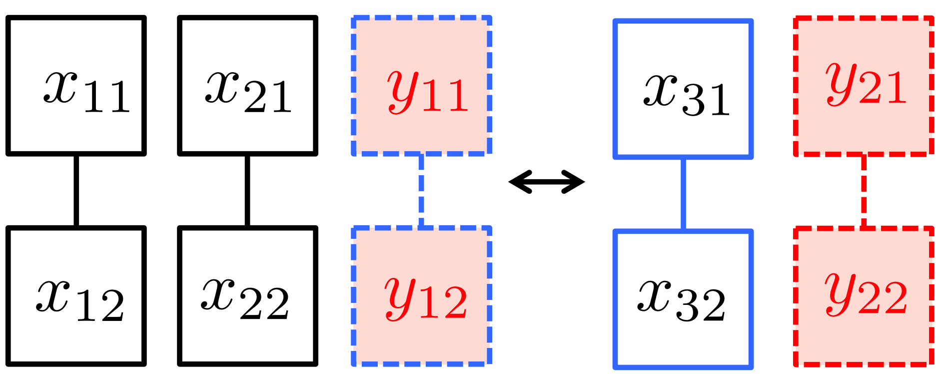
MZ-twins DZ-twins



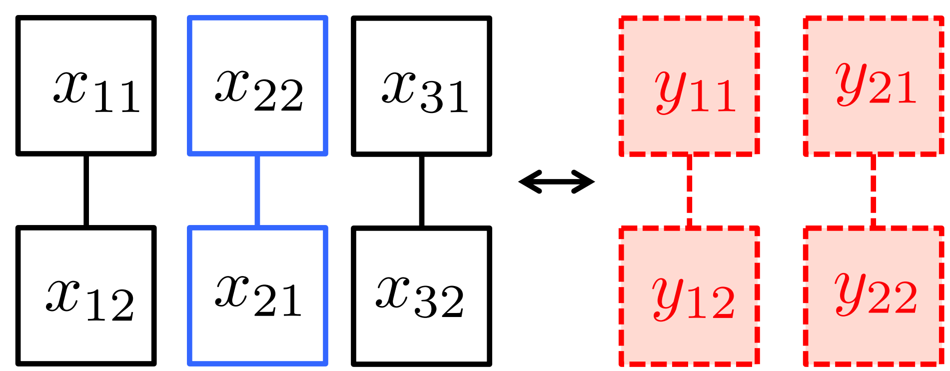
Permutation is the composition of two types of permutations



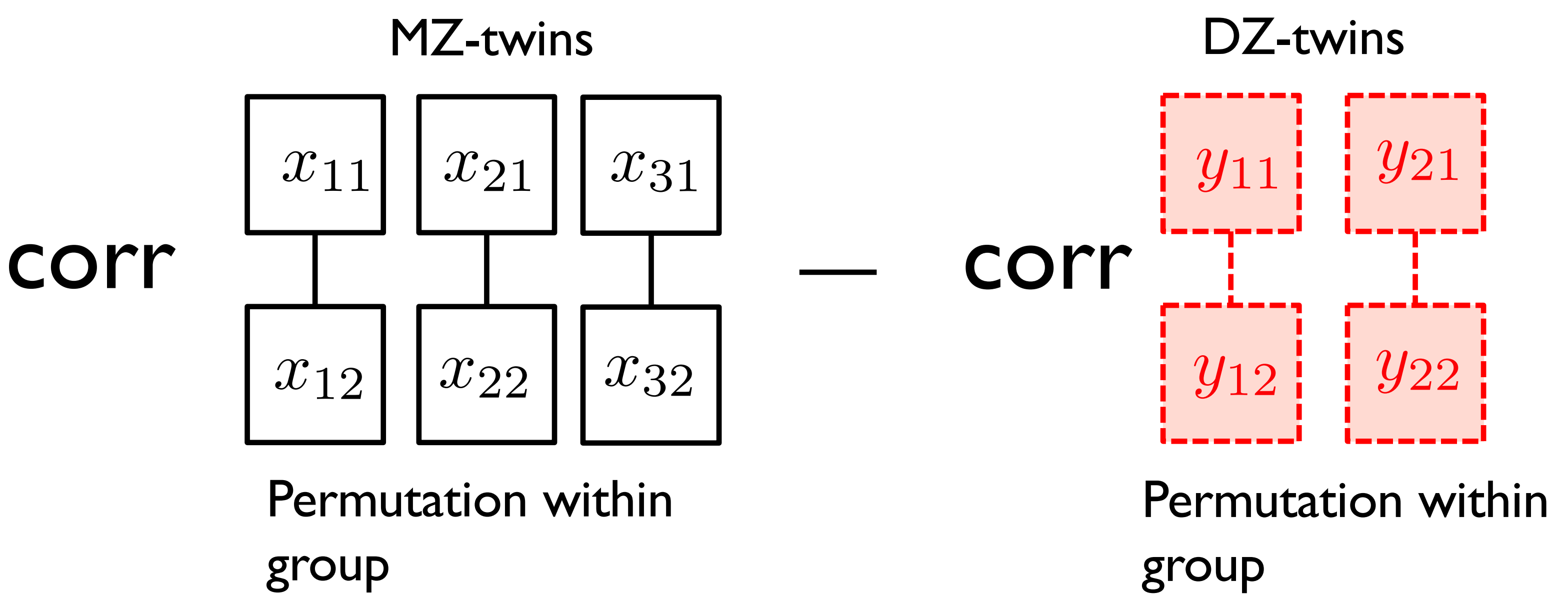
Permutation between groups



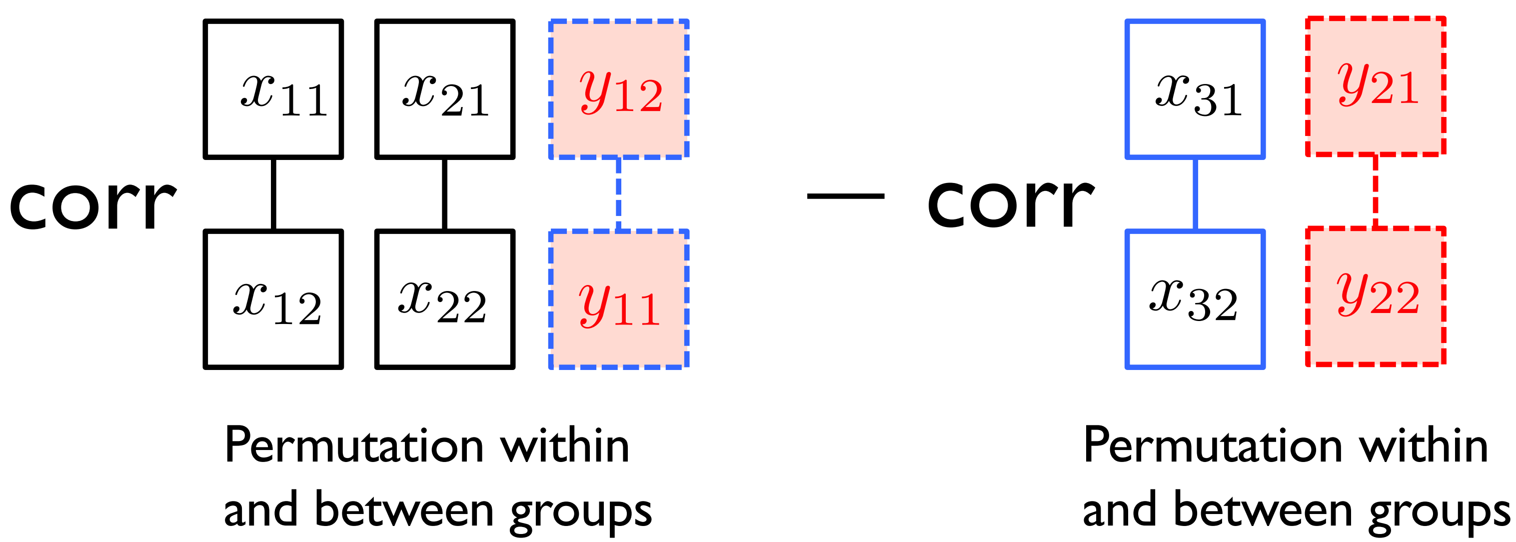
Permutation within twins



More complex example: twin correlation under transpositions



Observed correlation difference



Correlation under random transpositions within and between groups

p-value is computed by determining the empirical distribution of Correlation under random transpositions