Rapid Acceleration of the Permutation Test via Transpositions

Moo K. Chung[1]([✉]), Linhui Xie[2], Shih-Gu Huang[1], Yixian Wang[1],
Jingwen Yan[2], and Li Shen[3]

[1] University of Wisconsin, Madison, USA
mkchung@wisc.edu
[2] Indiana University-Purdue University Indianapolis, Indianapolis, USA
[3] University of Pennsylvania, Philadelphia, USA

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.

Keywords: Permutation test · Transposition test · Structural brain networks · Permutation group · Online statistics computation

1 Introduction

The permutation test is perhaps the most widely used nonparametric test procedure in sciences [8,19,21,24,27]. It is known as the exact test in statistics since the distribution of the test statistic under the null hypothesis can be exactly computed if we can calculate all the test statistics under every possible permutation. Unfortunately, generating every possible permutation for a large-sample network dataset is still extremely time consuming even for a modest sample size.

When the total number of permutations is large, various resampling techniques have been proposed to speed up the computation in the past [8,19,21,27]. In the resampling methods, only a small fraction of possible permutations is generated and the statistical significance is computed *approximately*. This approximate permutation test is the most widely used version of the permutation test. In most of brain imaging studies, 5,000–1,000,000 permutations are often used, which puts the total number of generated permutations usually less than a fraction of all possible permutations. In [27], 5,000 permutations out of possible

$\binom{27}{12} = 17,383,860$ permutations (0.029%) were used. In [21], 1 million permutations out of $\binom{40}{20}$ possible permutations (0.0007%) were generated using a super computer. In [18], 5,000 permutations out of possible $\binom{33}{10} = 92561040$ permutations (0.005%) were used.

To remedy the computational bottleneck, the tail regions of the distributions are often estimated using the extreme value theory [11,24]. One main tool in the extreme value theory is the use of generalized Pareto distribution in approximating the tail distributions. Unfortunately, without a prior information or model fit, it is difficult to even guess the shape of tails accurately. Recently, an exact topological inference approach with quadratic run time was proposed to combinatorially enumerate every possible permutation [8,9], but the method is limited to monotone network features and not applicable to more general settings.

In this paper, we propose a novel transposition test that is motived by the permutation test. The method is based on the concept of random transpositions that sequentially update the test statistic. Unlike the traditional permutation test that takes up to a few days on a computer, our method takes less than an hour and does not require large computer memory. As an application, the method is used in determining the statistical significance of the male and female differences in a large-sample structural brain network study.

2 Preliminary

The usual statistical test setting in brain imaging is a two-sample comparison [8,19,21]. Consider two ordered sets

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

The distance between $\mathbf{x}$ and $\mathbf{y}$ is measured by test statistic $f(\mathbf{x}, \mathbf{y})$ such as t-statistic or correlations. Under the null assumption of the equivalence of $\mathbf{x}$ and $\mathbf{y}$, elements in $\mathbf{x}$ and $\mathbf{y}$ are permutable. Consider combined ordered set $\mathbf{z} = (x_1, \cdots, x_m, y_1, \cdots, y_n)$ and its all possible permutations $\mathbb{S}_{m+n}$. Note $\mathbb{S}_{m+n}$ is a symmetric group of order $m + n$ with $(m + n)!$ possible permutations [14]. Since there is an isomorphism between $\mathbf{z}$ and integer set $\{1, 2, \cdots, m + n\}$, we will interchangeably use them when appropriate [17]. Permutation $\tau \in \mathbb{S}_{m+n}$ is often denoted as

$$\tau = \begin{pmatrix} x_1 & \cdots & x_m & y_1 & \cdots & y_n \\ \tau(x_1) & \cdots & \tau(x_m) & \tau(y_1) & \cdots & \tau(y_n) \end{pmatrix}$$

using a single combined sample notation in mathematical literature [10,14].

For instance, consider a permutation of $\{1, 2, 3, 4\}$ given by $\tau(1) = 4, \tau(2) = 2, \tau(3) = 1, \tau(4) = 3$. Since there are two cycles in the permutation, τ can be written in the cyclic form as $\tau = [2][1, 4, 3]$ indicating 2 is a cycle of length 1 $(2 \rightarrow 2)$ while 1, 3, 4 are a cycle of length 3 $(1 \rightarrow 4 \rightarrow 3 \rightarrow 1)$ [14]. A cycle of length 1 is simply ignored and the permutation can be written as $\tau = [1, 4, 3]$. If another permutation is given by $\pi(1) = 1, \pi(2) = 4, \pi(3) = 3, \pi(4) = 2$, the sequential application of π to τ is written as

$$\pi \cdot \tau = [1][3][2,4] \cdot [2][1,4,3] = [2,4] \cdot [1,4,3] = [1,2,4,3].$$

Let us split the permutation $\tau(\mathbf{z})$ into two groups with m and n elements

$$\tau(\mathbf{x}) = (\tau(x_1), \cdots, \tau(x_m)), \quad \tau(\mathbf{y}) = (\tau(y_1), \cdots, \tau(y_n)).$$

For test statistic f, the *exact p*-value for testing a one sided hypothesis is then given by the fraction

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

where $\mathcal{I}$ is an indicator function taking value 1 if the argument is true and 0 otherwise. In various brain imaging applications, computing statistic f for each permutation has been the main computational bottleneck [8,21].

If the test statistic f is a symmetric function such that $f(\mathbf{x}, \mathbf{y}) = f(\phi(\mathbf{x}), \psi(\mathbf{y}))$, where $\phi \in \mathbb{S}_m$ and $\psi \in \mathbb{S}_n$, then we only need to consider $\binom{m+n}{m}$ number of permutations in the denominator of (1), which reduces the number of possible permutations substantially. Still $\binom{m+n}{m}$ is an extremely large number and most computing systems including MATLAB/R cannot compute them exactly if the sample size is larger than 100 in each group. The total number of permutations when $m = n$ is given asymptotically by Stirling's formula [12]

$$\binom{2m}{m} \sim \frac{4^m}{\sqrt{\pi m}}.$$

The number of permutations *exponentially* increases as the sample size increases, and thus it is impractical to generate every possible permutation. In practice, up to hundreds of thousands of random permutations are generated using the uniform distribution on $\mathbb{S}_{m+n}$ with probability $1/\binom{m+n}{m}$.

3   Methods

Transpositions. Consider permutation π_{ij} that exchanges i-th and j-th elements between $\mathbf{x}$ and $\mathbf{y}$ and keeps all others fixed such that

$$\pi_{ij}(\mathbf{x}) = (x_1, \cdots, x_{i-1}, y_j, x_{i+1}, \cdots, x_m),$$
$$\pi_{ij}(\mathbf{y}) = (y_1, \cdots, y_{j-1}, x_i, y_{j+1}, \cdots, y_n).$$

Such a permutation is called the *transposition*, which is related to card shuffling problems [1,2,14]. Consider every possible sequence of transpositions applied to $\mathbf{x}$ and $\mathbf{y}$. If such sequence of transpositions covers every possible element in $\mathbb{S}_{m+n}$, we can perform the permutation test by sequentially transposing two elements at a time.

Theorem 1. *Any permutation in $\mathbb{S}_{m+n}$ can be reachable by a sequence of transpositions.*

Proof. Let $l = m + n$. Suppose $\tau \in \mathbb{S}_l$. For $x \in \{1, \cdots, l\}$, consider cycle

$$C_x = [x, \tau(x), \tau^2(x), \cdots, \tau^j(x)]$$

with $\tau^{j+1}(x) = x$ and $\tau^d(x) \neq x$ for $d \leq j$ [14]. Since we are dealing with a finite number of elements, such j always exists. If $\tau^c(x) = \tau^d(x)$ for some $c \leq d \leq j$, we have $\tau^{d-c}(x) = x$, thus all elements in the cycle C_x are distinct.

If C_x covers all the elements in $\{1, \cdots, l\}$ we proved the statement. If there is an element, say $y \in \{1, \cdots, l\}$, that is not covered by C_x, we construct a new cycle C_y. Cycles C_x and C_y must be disjoint. If not, we have $\tau^i(x) = \tau^j(y)$ and $y = \tau^{i-j}(x)$, which is in contradiction. Hence $\tau = C_x \cdot C_y$.

If $C_x \cdot C_y$ does not cover $\{1, \cdots, l\}$, we repeat the process until we exhaust all the elements in $\{1, \cdots, l\}$. Hence any permutation can be decomposed as a product of *disjoint* cycles. Then algebraic derivation can further show that cycle C_x can be decomposed as a product of 2-cycles

$$C_x = [x, \tau^j(x)] \cdot [x, \tau^{j-1}(x)] \cdots [x, \tau^2(x)] \cdot [x, \tau(x)].$$

A 2-cycle is simply a walk. Hence we proved τ is a sequence of walks.   □

From Theorem 1, any permutation can be reached by a sequence of transpositions. Thus, instead of performing uniform random sampling in $\mathbb{S}_{m+n}$, we will perform a sequence of random transpositions and compute the test statistic at each transposition. Over random transposition π_{ij}, the statistic changes from $f(\mathbf{x}, \mathbf{y})$ to $f(\pi_{ij}(\mathbf{x}), \pi_{ij}(\mathbf{y}))$. Instead of computing $f(\pi_{ij}(\mathbf{x}), \pi_{ij}(\mathbf{y}))$ directly, we will compute it from $f(\mathbf{x}, \mathbf{y})$ incrementally in *constant* run time by updating the value of $f(\mathbf{x}, \mathbf{y})$. Note, the statistics computation over transpositions is slightly different from the usual online computation where new data is added sequentially. Instead of adding the new data, the existing data is replaced.

Theorem 2. *If f is an algebraic function such as addition, subtraction, multiplication, division and integer exponents, there exists a function g such that*

$$f(\pi_{ij}(\mathbf{x}), \pi_{ij}(\mathbf{y})) = g(f(\mathbf{x}, \mathbf{y}), x_i, y_j), \tag{2}$$

where the computational complexity of g is constant.

The lengthy proof involves explicitly constructing an iterative formula for each algebraic operation so it will not be shown here. Often used statistics such as t-statistic and F-statistic are all algebraic functions. If we take computation involving *fractional exponents* as constant run time as well, then a much wider class of statistics such as correlations can all have constant run time. In this study, we will explicitly construct the t-statistic over transpositions that runs in constant run time. From this construction, it should be obvious that Theorem 2 should be applicable to other test statistics.

t-Statistic over a Transposition. Two sample t-statistic is a function of symmetric functions involving the mean and variance of $\mathbf{x}$ and $\mathbf{y}$. If we have the symmetric functions

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

the sample mean and variance of $\mathbf{x}$ are given by $\nu(\mathbf{x})/m$ and $\omega(\mathbf{x})/(m-1)$. We will determine how ν and ω change over a transposition.

Theorem 3. *Functions ν and ω are updated over transposition π_{ij} as*

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

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

Proof. The algebraic derivation follows applying the online computations of updating ν and ω twice. Suppose new data a and b is added to $\mathbf{x}' = (x_1, \cdots, x_{m-1})$ such that $\mathbf{x}_a = (\mathbf{x}', a)$ and $\mathbf{x}_b = (\mathbf{x}', b)$. Then we have

$$\nu(\mathbf{x}_b) = \nu(\mathbf{x}_a) - a + b.$$

Since ν is symmetric, by identifying $a = x_i$ and $b = y_j$, we obtain $\nu(\pi_{ij}(\mathbf{x})) = \nu(\mathbf{x}) - x_i + y_j$. An algebraic derivation can show that

$$\omega(\mathbf{x}_a) = \sum_{j=1}^{m-1} x_j^2 + a^2 - \frac{\nu(\mathbf{x}_a)^2}{m}, \quad \omega(\mathbf{x}_b) = \sum_{j=1}^{m-1} x_j^2 + b^2 - \frac{\nu(\mathbf{x}_b)^2}{m}. \qquad (3)$$

From (3), we obtain $\omega(\mathbf{x}_b) = \omega(\mathbf{x}_a) - a^2 + b^2 + \frac{\nu(\mathbf{x}_a)^2 - \nu(\mathbf{x}_b)^2}{m}$ and the result follows.   □

From Theorem 3, the two-sample t-statistic over a transposition is then computed as follows.

$$T(\pi_{ij}(\mathbf{x}), \pi_{ij}(\mathbf{y})) = \frac{\nu(\pi_{ij}(\mathbf{x}))/m - \nu(\pi_{ij}(\mathbf{y}))/n}{\sqrt{\frac{\omega(\pi_{ij}(\mathbf{x})) + \omega(\pi_{ij}(\mathbf{y}))}{m+n-2} \left(\frac{1}{m} + \frac{1}{n} \right)}}.$$

Computing two-sample t-statistic with m and n samples directly requires computing the sample means, which is m and n algebraic operations each. Then we need to compute the sample variances and pool them together, which requires $3m + 2$ and $3n + 2$ operations. Combining the numerator and denominator in t-statistic takes 16 operations. Thus, it takes total $3(m + n) + 20$ operations to compute the t-statistic per permutation. In comparison, it only takes 35 operations to computer t-statistic per transposition. In the case of $m = n = 200$, the proposed method can generate 1220 times more permutations compared to the standard permutation test.
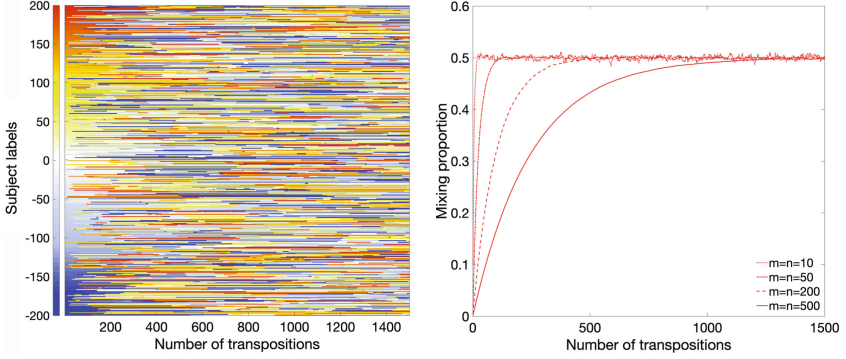
Fig. 1. Left: Mixing of subject labels over transpositions. Right: The estimated mixing proportion based on the average of 1000 simulations.

Reducing Mixing Time. Given $m = n$ elements in each group, the standard permutation test mixes half of elements in one group to the other. Thus, the mixing proportion is 0.5 on average. On the other, the transposition method mixes one element at a time, so the mixing is slow but it rapidly catches up. The rate of mixing can be formally measured by the *mixing time*, which is defined as the time until the transpositions are close to the uniform distribution in $\mathbb{S}_{m+n}$ in the variation distance sense [2,5]. Even though the transposition method does not guarantee the uniform distribution in $\mathbb{S}_{m+n}$ in the early stage of transpositions, the method converges to the uniform distribution quickly in $O((m+n)log(m+n))$ time [2,5]. This is demonstrated in Fig. 1.

Figure 1-left displays how the subject labels change over transpositions based on the sample sizes $m = n = 200$. The first group is indexed between 1 and 200 while the second group is indexed between -1 and -200. At each transposition, only two subjects are swapped. As the number of transpositions increases, subject labels rapidly mix up. Figure 1-right shows that how the mixing proportion converges to 0.5 based on the average of 1000 simulations. On average, about 1000 transpositions are enough to mix all the elements in the two groups uniformly. For far smaller sample sizes, to which most brain imaging studies belong, few hundreds transpositions are more than enough to mix the groups evenly.

To increase the rate of mixing further, we did not start with the original data **x** and **y** but started with a random permutation of **x** and **y**. This has the effect of starting with a completely mixed initial starting data. Then sequentially applied 5,000 random transpositions. This process is iteratively repeated. Thus, for every 5,000 random transpositions, one random permutation is intermixed. In real data, this process is repeated 10,000 times to generate 50 million random transpositions, which are intermixed with 10,000 permutations.

Multiple Comparisons. So far we have shown how test statistics change over transpositions. We now show how the multiple comparison corrected p-values are affected over transpositions. Suppose $\mathbf{x}(q)$ and $\mathbf{y}(q)$ are functional data on edge q in a brain network $\mathcal{M}$. Given statistic map $h(q) = f(\mathbf{x}(q), \mathbf{y}(q))$ at the edge level, the hypotheses of interests are given by

$$H_0 : h(q) = 0 \text{ for all } q \in \mathcal{M} \quad vs. \quad H_1 : h(q) > 0 \text{ for some } q \in \mathcal{M}. \quad (4)$$

Once the iterative algorithm for computing the test statistic is identified, the p-value for pointwise inference at each *fixed* q can be computed iteratively. At the k-th random transposition, the uncorrected p-value is given as p_k. Then p_{k+1} is computed from iterative formula

$$(k+1)p_{k+1} = kp_k + \mathcal{I}\Big(f(\pi_{ij}(\mathbf{x}), \pi_{ij}(\mathbf{y})) \geq f(\mathbf{x}, \mathbf{y})\Big), \quad (5)$$

where π_{ij} changes over random transpositions. Note the p-value for multiple comparisons over all q is given by

$$p\text{-value} = P\Big(\bigcup_{q \in \mathcal{M}} \{h(q) > c\}\Big) = 1 - P\Big(\bigcap_{q \in \mathcal{M}} \{h(q) \leq c\}\Big)$$
$$= 1 - P\Big(\sup_{q \in \mathcal{M}} h(q) \leq c\Big)$$

for some threshold c [25]. Thus, for multiple comparisons, the formula (5) changes to

$$(k+1)p_{k+1} = kp_k + \mathcal{I}\Big(\sup_{q \in \mathcal{M}} h(\pi_{ij}(\mathbf{x}(q)), \pi_{ij}(\mathbf{y}(q))) \geq \sup_{q \in \mathcal{M}} h(\mathbf{x}(q), \mathbf{y}(q))\Big).$$

For alternate hypothesis $H_1 : h(q) < 0$, a similar algorithm can be used for test statistic $\inf_{q \in \mathcal{M}} h(q)$.

Validation. The real data does not have the ground truth. Thus, we compared the proposed transposition method against the standard permutation test in random simulations with the ground truth. We simulated $x_1, \cdots, x_m \sim N(0, 1)$, standard normal distribution, and $y_1, \cdots, y_n \sim N(0.1, 1)$, which provides the *ground truth* in computing t-statistic and p-value. The use of t-statistic is the standard validation framework in many previous permutation test studies [6, 13, 19, 24]. The simulations were independently performed 100 times and their average was reported here. In both Simulations 1 and 2 below, we used the same model but different sample sizes.

Simulation 1 (small sample size). We used $m = n = 10$. There are exactly $\binom{20}{10} = 184756$ total permutations. The sample sizes are too small to differentiate the group difference. We obtained the t-statistic value of 0.0533, which corresponds to the exact p-value of 0.479 (Fig. 2a green line). We performed the standard permutation test with up to 10,000 permutations, which took 0.0926 s on average on a desktop computer. Within the same run time, we were able
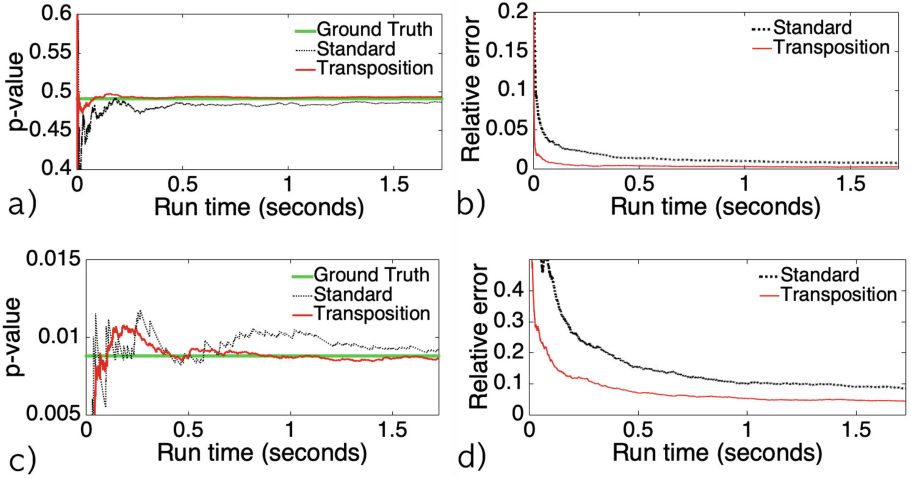
Fig. 2. (a, c) One representative simulation study showing faster convergence of the transposition method. The Gaussian distribution provides the exact ground truth. (b, d) The average relative error against the ground truth. The average of 100 independent simulations was plotted. (Color figure online)

to generate more than 1,220,000 transpositions. The transposition method uniformly converged faster than the standard permutation test due to 122 times more permutations the transposition method generated (Fig. 2b). The relative errors of the transposition method are about half the size of the standard method in most run time.

Simulation 2 (large sample size). We used $m = n = 100$. The sample sizes are big enough to differentiate the group difference. We obtained the t-statistic value of 2.39 and corresponding p-value of 0.0088, which are taken as the ground truth (Fig. 2c green line). We performed the standard permutation test with up to 1 million permutations, which took 173 s per simulation on average. With the same run time, the transposition was sequentially done about 125 million times. The transposition method uniformly converged faster than the standard permutation test through the whole run time (Fig. 2d). The performance results did not change much even if we performed more permutations over longer durations with different simulation parameters.

The computer code for performing the transposition test in the above simulation study is available at http://www.stat.wisc.edu/~mchung/transpositions. The MATLAB code is written in such a way that it accepts vector data. For brain connectivity matrices that are symmetric, vectorizing the upper triangle entries of connectivity matrices is necessary to reduce the computational time.
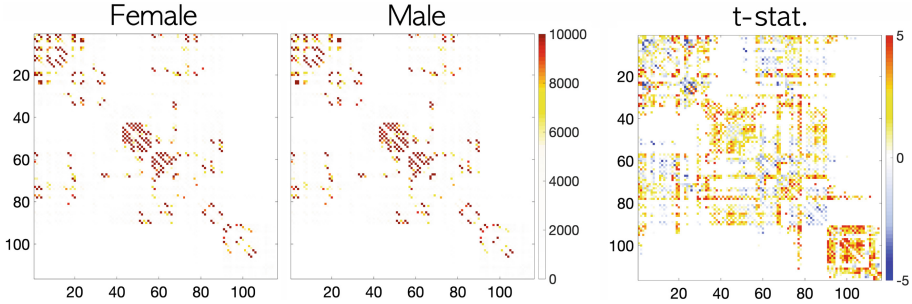
Fig. 3. Average connectivity matrices of females (left) and males (right) between 116 AAL parcellations. The two-sample t-statistic result (female $-$ male). Females have more structural connections between brain regions than males.

4   Application

**Subjects and Preprocessing.** Diffusion weighted imaging (DWI) data of 202 female and 154 male subjects (ages 29.2 ± 3.4) were obtained from the Human Connectome Project (HCP) [16]. The fiber orientation distribution functions were estimated and apparent fiber densities were exploited to produce reliable WM/GM/CSF volume maps [7,16]. Subsequently, random seeds on the basis of the voxel were selected to generate about initial 10 million streamlines per subject with the maximum fiber tract length at 250 mm and FA larger than 0.06 using MRtrix3 (http://www.mrtrix.org) [22,26]. The Spherical-Deconvolution Informed Filtering of Tractograms (SIFT2) technique making use of complete streamlines was subsequently applied to generate more biologically accurate brain connectivity, which results in about 1 million tracts per subject [20]. Non-linear diffeomorphic registration between subject images to the template was performed using ANTS [3,4]. Automated Anatomical Labeling (AAL) was used to parcellate the brain into 116 regions [23]. The subject-level connectivity matrices were constructed by counting the number of tracts connecting between brain regions.

Transposition Test. We are interested in testing and localizing the female and male differences in structural connectivity. Figure 3 displays the result of group averages and the two sample t-statistic (female $-$ male). Females have more structural connections between brain regions than males. Since the tract counts between brain regions do not follow normal distributions, assumption-free non-parametric procedures such as the permutation test are needed to determine the statistical significance of t-statistic accurately. We use the proposed transposition test by sequentially generating 50 million transpositions for 40 min on a desktop computer. For multiple comparisons correction, we counted the fraction of transpositions where the maximum t-statistic value over the whole connections is above the observed maximum t-statistic value. Any t-statistic value below

-4.05 and above 3.96 is statistically significant at 0.05 (Fig. 4). The statistically significant connections are shown in Fig. 5.
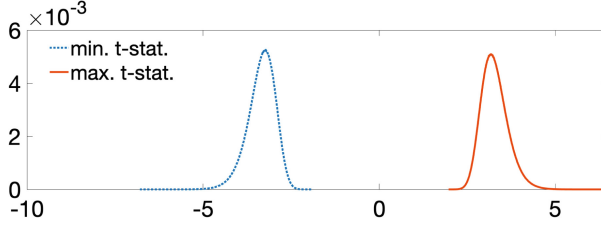
Fig. 4. The empirical distributions of minimum t-statistic (dotted blue) and maximum t-statistic (solid red), which do not follow well known statistical distributions. The proposed method is used to compute the multiple comparisons corrected p-value. (Color figure online)
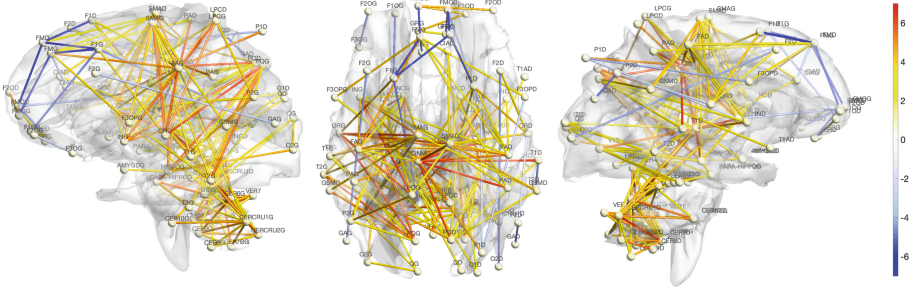
Fig. 5. t-statistic map (female $-$ male). Only the connections that are statistically significant (thresholded at -4.05 and 3.96) after multiple comparisons correction at 0.05 are shown. Females have more connections in most parts of the brain while males are more connected in the frontal regions of the brain.

Sex Difference in Connectivity. Females have far more connections in most parts of the brain while males have more connections in the frontal regions of the brain. Females also have more bilateral connections between the hemispheres. This indicates that females use the both sides of the brain while males only use one side of the brain. Females have more connections in the limbic structures that regulate emotions. Males have more connections between the back and frontal regions. Our findings are consistent with the previous structural connectivity study [15].

5   Discussion

Although we did not show here, it is also possible to construct incremental procedures for computing other test statistics such as F-statistic and Hotelling's T^2 statistic over random transpositions. These problems are left as future studies.

Compared to other approximate strategies for the permutation test, the proposed method is assumption and model free. The tail approximation method in [24] has parametric model assumptions to fit the tail regions, so the tail of the distribution needs to follow some specific pattern. On the other hand, the proposed method has no assumption on the distribution other than permutability between the groups and offers far more flexibility than [24].

We did not perform the comparisons between the methods in the real data since there is no ground truth. Thus the comparisons were done on the simulation, where the ground truths are exactly given.

References

1. Aldous, D.: Random walks on finite groups and rapidly mixing Markov chains. In: Azéma, J., Yor, M. (eds.) Séminaire de Probabilités XVII 1981/82. LNM, vol. 986, pp. 243–297. Springer, Heidelberg (1983). https://doi.org/10.1007/BFb0068322
2. Aldous, D., Diaconis, P.: Shuffling cards and stopping times. Am. Math. Monthly **93**, 333–348 (1986)
3. Avants, B., Epstein, C., Grossman, M., Gee, J.: Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. Med. Image Anal. **12**, 26–41 (2008)
4. Avants, B., Tustison, N., Song, G., Cook, P., Klein, A., Gee, J.: A reproducible evaluation of ANTs similarity metric performance in brain image registration. NeuroImage **54**, 2033–2044 (2011)
5. Berestycki, N., Schramm, O., Zeitouni, O.: Mixing times for random k-cycles and coalescence-fragmentation chains. Ann. Probab. **39**, 1815–1843 (2011)
6. Bullmore, E., Suckling, J., Overmeyer, S., Rabe-Hesketh, S., Taylor, E., Brammer, M.: Global, voxel, and cluster tests, by theory and permutation, for difference between two groups of structural MR images of the brain. IEEE Trans. Med. Imaging **18**, 32–42 (1999)
7. Christiaens, D., Reisert, M., Dhollander, T., Sunaert, S., Suetens, P., Maes, F.: Global tractography of multi-shell diffusion-weighted imaging data using a multitissue model. NeuroImage **123**, 89–101 (2015)
8. Chung, M.K., Luo, Z., Leow, A.D., Alexander, A.L., Davidson, R.J., Hill Goldsmith, H.: Exact combinatorial inference for brain images. In: Frangi, A.F., Schnabel, J.A., Davatzikos, C., Alberola-López, C., Fichtinger, G. (eds.) MICCAI 2018. LNCS, vol. 11070, pp. 629–637. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-00928-1_71
9. Chung, M.K., Villalta-Gil, V., Lee, H., Rathouz, P.J., Lahey, B.B., Zald, D.H.: Exact topological inference for paired brain networks via persistent homology. In: Niethammer, M., et al. (eds.) IPMI 2017. LNCS, vol. 10265, pp. 299–310. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-59050-9_24
10. Dummit, D., Foote, R.: Abstract Algebra. Wiley, Hoboken (2004)

11. Embrechts, P., Resnick, S., Samorodnitsky, G.: Extreme value theory as a risk management tool. North Am. Actuarial J. **3**, 30–41 (1999)
12. Feller, W.: An Introduction to Probability Theory and its Applications, vol. 2. Wiley, Hoboken (2008)
13. Hayasaka, S., Phan, K.L., Liberzon, I., Worsley, K.J., Nichols, T.E.: Nonstationary cluster-size inference with random field and permutation methods. Neuroimage **22**, 676–687 (2004)
14. Hungerford, T.: Algebra. Springer, New York (1980)
15. Ingalhalikar, M., et al.: Sex differences in the structural connectome of the human brain. Proc. Nat. Acad. Sci. **111**, 823–828 (2014)
16. Jeurissen, B., Tournier, J.D., Dhollander, T., Connelly, A., Sijbers, J.: Multi-tissue constrained spherical deconvolution for improved analysis of multi-shell diffusion MRI data. NeuroImage **103**, 411–426 (2014)
17. Kondor, R., Howard, A., Jebara, T.: Multi-object tracking with representations of the symmetric group. In: International Conference on Artificial Intelligence and Statistics (AISTATS), vol. 1, p. 5 (2007)
18. Lee, H., Kang, H., Chung, M., Lim, S., Kim, B.N., Lee, D.: Integrated multimodal network approach to PET and MRI based on multidimensional persistent homology. Hum. Brain Mapp. **38**, 1387–1402 (2017)
19. Nichols, T., Holmes, A.: Nonparametric permutation tests for functional neuroimaging: a primer with examples. Hum. Brain Mapp. **15**, 1–25 (2002)
20. Smith, R., Tournier, J.D., Calamante, F., Connelly, A.: SIFT2: enabling dense quantitative assessment of brain white matter connectivity using streamlines tractography. NeuroImage **119**, 338–351 (2015)
21. Thompson, P., et al.: Genetic influences on brain structure. Nat. Neurosci. **4**, 1253–1258 (2001)
22. Tournier, J., Calamante, F., Connelly, A., et al.: MRtrix: diffusion tractography in crossing fiber regions. Int. J. Imaging Syst. Technol. **22**, 53–66 (2012)
23. Tzourio-Mazoyer, N., et al.: Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. NeuroImage **15**, 273–289 (2002)
24. Winkler, A., Ridgway, G., Douaud, G., Nichols, T., Smith, S.: Faster permutation inference in brain imaging. NeuroImage **141**, 502–516 (2016)
25. Worsley, K., Marrett, S., Neelin, P., Vandal, A., Friston, K., Evans, A.: A unified statistical approach for determining significant signals in images of cerebral activation. Hum. Brain Mapp. **4**, 58–73 (1996)
26. Xie, L., et al.: Heritability estimation of reliable connectomic features. In: Wu, G., Rekik, I., Schirmer, M.D., Chung, A.W., Munsell, B. (eds.) CNI 2018. LNCS, vol. 11083, pp. 58–66. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-00755-3_7
27. Zalesky, A., et al.: Whole-brain anatomical networks: does the choice of nodes matter? NeuroImage **50**, 970–983 (2010)