



Exact Combinatorial Inference for Brain Images

Moo K. Chung¹(✉), Zhan Luo¹, Alex D. Leow², Andrew L. Alexander¹,
Richard J. Davidson¹, and H. Hill Goldsmith¹

¹ University of Wisconsin, Madison, USA
mkchung@wisc.edu

² University of Illinois, Chicago, USA

Abstract. The permutation test is known as the exact test procedure in statistics. However, often it is not exact in practice and only an approximate method since only a small fraction of every possible permutation is generated. Even for a small sample size, it often requires to generate tens of thousands permutations, which can be a serious computational bottleneck. In this paper, we propose a novel combinatorial inference procedure that enumerates all possible permutations combinatorially without any resampling. The proposed method is validated against the standard permutation test in simulation studies with the ground truth. The method is further applied in twin DTI study in determining the genetic contribution of the minimum spanning tree of the structural brain connectivity.

1 Introduction

The permutation test is perhaps the most widely used nonparametric test procedure in sciences [1, 7, 10]. 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 possible values of the test statistic under every possible permutation. Unfortunately, generating every possible permutation for whole images is still extremely time consuming even for modest sample size.

When the total number of permutations are too large, various resampling techniques have been proposed to speed up the computation in the past. In the resampling methods, only a small fraction of possible permutations are 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, 5000–1000000 permutations are often used, which puts the total number of generated permutations usually less than 1% of all possible permutations. In [10], 5000 permutations are out of possible $\binom{27}{12} = 17383860$ permutations (2.9%) were used. In [7], for instance, 1 million permutations out of $\binom{40}{20}$ possible permutations (0.07%) were generated using a super computer.

In this paper, we propose a novel *combinatorial inference* procedure that enumerates all possible permutations combinatorially and simply avoids resampling that is slow and approximate. Unlike the permutation test that takes few

hours to few days in a desktop, our exact procedure takes few seconds. Recently combinatorial approaches for statistical inference are emerging as an powerful alternative to existing statistical methods [1, 5]. Neykov et al. proposed a combinatorial technique for graphical models. However, their approach still relies on bootstrapping, which is another resampling technique and still approximate [5]. Chung et al. proposed another combinatorial approach for brain networks but their method is limited to integer-valued graph features [1]. Our main contributions of this paper are as follows.

(1) A new combinatorial approach for the permutation test that does not require resampling. While the permutation tests require exponential run time and approximate [1], our combinatorial approach requires $\mathcal{O}(n^2)$ run time and exact. (2) Showing that the proposed method is a more sensitive and powerful alternative to the existing permutation test through extensive simulation studies. (3) A new formulation for testing the brain network differences by using the minimum spanning tree differences. The proposed framework is applied to a twin DTI study in determining the heritability of the structural brain network.

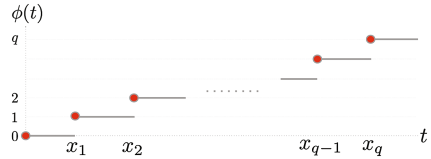


Fig. 1. Monotone sequence $x_1 < x_2 < \dots < x_q$ is mapped to integers between 1 and q without any gap via $\phi(t)$.

2 Exact Combinatorial Inference

The method in this paper extends our previous *exact topological inference* [1], which is limited to integer-valued monotone functions from graphs. Through Theorem 1, we extend the method to any arbitrary monotone function.

Definition 1. For any sets G_1 and G_2 satisfying $G_1 \subset G_2$, function f is strictly monotone if it satisfies $f(G_1) < f(G_2)$. \subset denotes the strict subset relation.

Theorem 1. Let f be a monotone function on the nested set $G_1 \subset G_2 \subset \dots \subset G_q$. Then there exists a nondecreasing function ϕ such that $\phi \circ f(G_j) = j$.

Proof. We prove the statement by actually constructing such a function. Function ϕ is constructed as follows. Let $x_j = f(G_j)$. Then obviously $x_1 < x_2 < \dots < x_q$. Define an increasing step function ϕ such that

$$\phi(t) = 0 \text{ if } t < x_1, \quad \phi(t) = j \text{ if } x_j \leq t < x_{j+1}, \quad \phi(t) = q \text{ if } x_q \leq t.$$

The step function ϕ is illustrated in Fig. 1. Then it is straightforward to see that $\phi \circ f(G_j) = j$ for all $1 \leq j \leq q$. Further ϕ is nondecreasing. \square

Consider two nested sets $F_1 \subset F_2 \subset \dots \subset F_q$ and $G_1 \subset G_2 \subset \dots \subset G_q$. We are interested in testing the null hypothesis H_0 of the equivalence of two monotone functions defined on the nested sets:

$$f(F_1) < f(F_2) < \dots < f(F_q) \quad \text{vs.} \quad g(G_1) < g(G_2) < \dots < g(G_q).$$

We have nondecreasing functions ϕ and ψ on $f(F_j)$ and $g(G_j)$ respectively that satisfies the condition Theorem 1. We use pseudo-metric

$$D_q = \max_t |\phi(t) - \psi(t)|$$

as a test statistic that measures the similarity between two monotone functions. The use of maximum removes the problem of multiple comparisons. The distribution of D_q can be determined by combinatorially.

Theorem 2. $P(D_q \geq d) = 1 - \frac{A_{q,q}}{\binom{2q}{q}}$, where $A_{u,v}$ satisfies $A_{u,v} = A_{u-1,v} + A_{u,v-1}$ with the boundary condition $A_{0,v} = A_{u,0} = 1$ within band $|u - v| < d$ and initial condition $A_{0,0} = 0$ for $u, v \geq 1$.

Proof. Let $x_j = \phi \circ f(F_j)$ and $y_j = \phi \circ g(G_j)$. From Theorem 1, the sequences x_1, \dots, x_q and y_1, \dots, y_q are monotone and integer-valued between 1 and q without any gap. Perform the permutation test on the sorted sequences. If we identify each x_j as moving one grid to right and y_j as moving one grid to up, each permutation is mapped to a walk from $(0, 0)$ to (q, q) . There are total $\binom{2q}{q}$ number of such paths and each permutation can be mapped to a walk uniquely.

Note $\max_{1 \leq j \leq q} |x_j - y_j| < d$ if and only if $|x_j - y_j| < d$ for all $1 \leq j \leq q$. Let $A_{q,q}$ be the total number of paths within $|x - y| < d$ (Fig. 2 for an illustration). Then it follows that $A_{q,q}$ is iteratively given as $A_{u,v} = A_{u-1,v} + A_{u,v-1}$ with $A_{0,0} = 0, A_{0,v} = A_{u,0} = 1$, within $|u - v| < d$. Thus $P(D_q < d) = \frac{A_{q,q}}{\binom{2q}{q}}$. \square

For example, $P(D_3 \geq 2)$ is computed sequentially as follows (Fig. 2). We start with the bottom left corner $A_{0,0} = 0$ and move right or up toward the upper corner. $A_{1,0} = 1, A_{0,1} = 1 \rightarrow A_{1,1} = A_{1,0} + A_{0,1} \rightarrow \dots \rightarrow A_{3,3} = A_{3,2} + A_{2,3} = 8$. The probability is then $P(D_3 \geq 2) = 1 - 8/\binom{6}{3} = 0.6$. The computational complexity of the combinatorial inference is $\mathcal{O}(q^2)$ for computing $A_{q,q}$ in the grid while the permutation test is exponential.

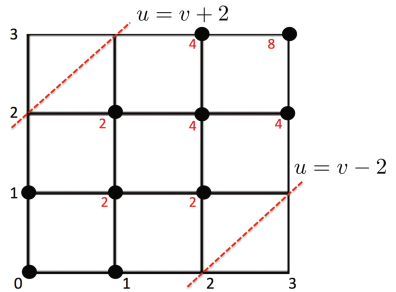


Fig. 2. In this example, $A_{u,v}$ is computed within the boundary (dotted red line) from $(0,0)$ to $(3,3)$.

3 Inference on Minimum Spanning Trees

As a specific example of how to apply the method, we show how to test for shape differences in minimum spanning trees (MST) of graphs. MST are often used in speeding up computation and simplifying complex graphs as simpler trees [9]. We et al. used MST in edge-based segmentation of lesion in brain MRI [9]. Stam et al. used MST as an unbiased skeleton representation of complex brain networks [6]. Existing statistical inference methods on MST rely on using graph theory features on MST such as the average path length. Since the probability

distribution of such features are often not well known, the permutation test is frequently used, which is not necessarily exact or effective. Here, we apply the proposed combinatorial inference for testing the shape differences of MST.

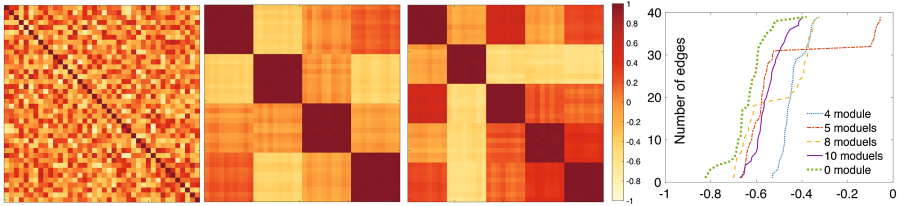


Fig. 3. Randomly simulated correlation matrices with 0, 4 and 5 modules. The plot shows the number of nodes over the largest edge weights added into MST construction during Kruskal’s algorithm for 4, 5, 8, 10 and 0 modules.

For a graph with p nodes, MST is often constructed using Kruskal’s algorithm, which is a greedy algorithm with runtime $\mathcal{O}(p \log p)$. The algorithm starts with an edge with the smallest weight. Then add an edge with the next smallest weight. This sequential process continues while avoiding a loop and generates a spanning tree with the smallest total edge weights. Thus, the edge weights in MST correspond to the order, in which the edges are added in the construction of MST. Let M_1 and M_2 be the MST corresponding to $p \times p$ connectivity matrices C_1 and C_2 . We are interested in testing hypotheses

$$H_0 : M_1 = M_2 \quad \text{vs.} \quad H_1 : M_1 \neq M_2.$$

The statistic for testing H_0 vs. H_1 is as follows. Since there are p nodes, there are $p-1$ edge weights in MST. Let $w_1^1 < w_2^1 < \dots < w_{p-1}^1$ and $w_1^2 < w_2^2 < \dots < w_{p-1}^2$ be the sorted edge weights in M_1 and M_2 respectively. These correspond to the order MST are constructed in Kruskal’s algorithm. w_j^1 and w_j^2 are edge weights obtained in the j -th iteration of Kruskal’s algorithm. Let ϕ and ψ be monotone step functions that map the the edge weights obtained in the j -th iteration to integer j , i.e., $\phi(w_j^1) = j$, $\psi(w_j^2) = j$. ϕ and ψ can be interpreted as the number of nodes added into MST in the j -th iteration.

4 Validation and Comparisons

For validation and comparisons, we simulated the random graphs with the ground truth. We used $p = 40$ nodes and $n = 10$ images, which makes possible permutations to be exactly $\binom{10+10}{10} = 184756$ making the permutation test manageable. The data matrix $X_{n \times p} = (x_{ij}) = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p)$ is simulated as standard normal in each component, i.e., $x_{ij} \sim N(0, 1)$. or equivalently each column is multivariate normal $\mathbf{x}_j \sim N(0, I)$ with identity matrix as the covariance.

Table 1. Simulation results given in terms of p -values. In the case of no network differences (0 vs. 0 and 4 vs. 4), higher p -values are better. In the case of network differences (4 vs. 5, 4 vs. 8 and 5 vs. 10), smaller p -values are better.

| | Combinatorial | Permute 0.1% | Permute 0.5% | Permute 1% |
|----------|---------------|---------------|---------------|---------------|
| 0 vs. 0 | 0.831 ± 0.187 | 0.746 ± 0.196 | 0.745 ± 0.195 | 0.744 ± 0.196 |
| 4 vs. 4 | 0.456 ± 0.321 | 0.958 ± 0.075 | 0.958 ± 0.073 | 0.958 ± 0.073 |
| 4 vs. 5 | 0.038 ± 0.126 | 0.381 ± 0.311 | 0.377 ± 0.311 | 0.378 ± 0.311 |
| 4 vs. 8 | 0.053 ± 0.138 | 0.410 ± 0.309 | 0.411 ± 0.306 | 0.411 ± 0.306 |
| 5 vs. 10 | 0.060 ± 0.126 | 0.391 ± 0.283 | 0.395 ± 0.284 | 0.395 ± 0.283 |

Let $Y = (y_{ij}) = (\mathbf{y}_1, \dots, \mathbf{y}_p) = X$. So far, there is no statistical dependency between nodes in Y . We add the following block modular structure to Y . We assume there are $k = 4, 5, 8, 10, 40$ modules and each module consists of $c = p/k = 10, 8, 5, 4, 1$ number of nodes. Then for the i -th node in the j -th module, we simulate

$$\mathbf{y}_{c(j-1)+i} = \mathbf{x}_{c(j-1)+1} + N(0, \sigma I) \quad \text{for } 1 \leq i \leq c, 1 \leq j \leq k \quad (1)$$

with $\sigma = 0.1$. Subsequently, the connectivity matrix $C = (c_{ij})$ is given by $c_{ij} = \text{corr}(\mathbf{y}_i, \mathbf{y}_j)$. This introduces the block modular structure in the correlation network (Fig. 3). For 40 modules, each module consists of just 1 node, which is basically a network with 0 module.

Using (1), we simulated random networks with 4, 5, 8, 10 and 0 modules. For each network, we obtained MST and computed the distance D between networks. We computed the p -value using the combinatorial method. In comparison, we performed the permutation tests by permuting the group labels and generating 0.1, 0.5 and 1% of every possible permutation. The procedures are repeated 100 times and the average results are reported in Table 1.

In the case of no network differences (0 vs. 0 and 4 vs. 4), higher p -values are better. The combinatorial method and the permutation tests all performed well for no network difference. In the case of network differences (4 vs. 5, 4 vs. 8 and 5 vs. 10), smaller p -values are better. The combinatorial method performed far superior to the permutation tests. None of the permutation tests detected modular structure differences. The proposed combinatorial approach on MST seems to be far more sensitive in detecting modular structures. The performance of the permutation test does not improve even when we sample 10% of all possible permutations. The permutation test doesn't converge rapidly with increased samples. The codes for performing exact combinatorial inference as well as simulations can be obtained from <http://www.stat.wisc.edu/~mchung/twins>.

5 Application to Twin DTI Study

Subjects. The method is applied to 111 twin pairs of diffusion weighted images (DWI). Participants were part of the Wisconsin Twin Project [2]. 58 monozy-

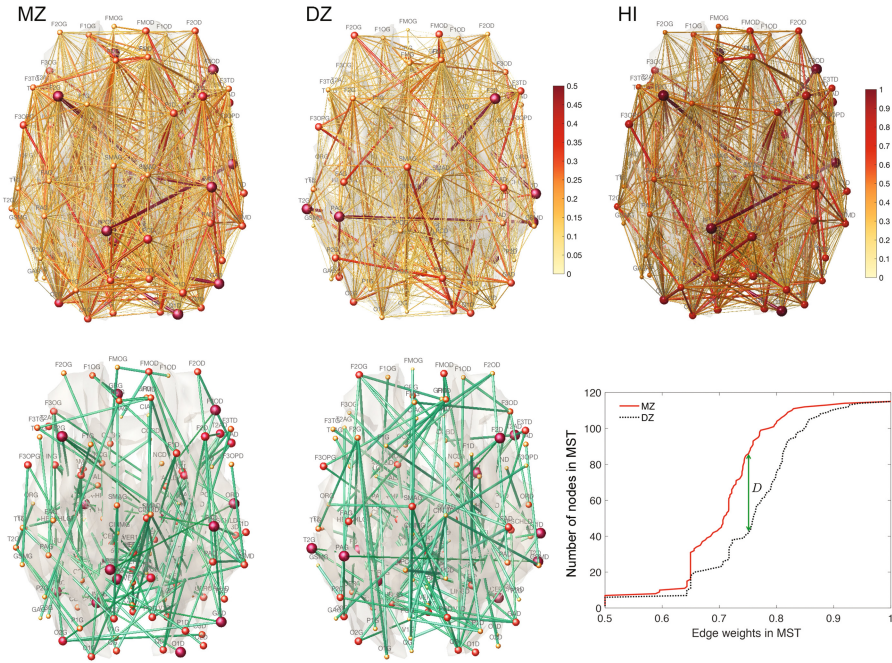


Fig. 4. Top: Correlation network of MZ- and DZ-twins and heritability index (HI). Bottom: Minimum spanning trees (MST) constructed using Kruskal’s algorithm on 1-correlation. Plot: The number of added nodes is plotted over the largest edge weights of MST for MZ- (solid red) and DZ-twins (dotted black) during the MST construction. The pseudo-metric D is 46 at edge weight 0.75 (corresponding to correlation 0.25).

gotic (MZ) and 53 same-sex dizygotic (DZ) twins were used in the analysis. We are interested in knowing the extent of the genetic influence on the structural brain network of these participants and determining its statistical significance between MZ- and DZ-twins. Twins were scanned in a 3.0 Tesla GE Discovery MR750 scanner with a 32-channel receive-only head coil. Diffusion tensor imaging (DTI) was performed using a three-shell diffusion-weighted, spin-echo, echo-planar imaging sequence. A total of 6 non-DWI ($b = 0$ s-mm²) and 63 DWI with non-collinear diffusion encoding directions were collected at $b = 500, 800, 2000$ (9, 18, 36 directions). Other parameters were TR/TE = 8575/76.6 ms; parallel imaging; flip angle = 90°; isotropic 2 mm resolution (128 × 128 matrix with 256 mm field-of-view).

Image Processing. FSL were used to correct for eddy current related distortions, head motion and field inhomogeneity. Estimation of the diffusion tensors at each voxel was performed using non-linear tensor estimation in CAMINO. DTI-TK was used for constructing the study-specific template [3]. Spatial normalization was performed for tensor-based white matter alignment using a non-parametric diffeomorphic registration method [11]. Each subject’s tractography was constructed in the study template space using the streamline method, and tracts

were terminated at FA-value less than 0.2 and deflection angle greater than 60° [4]. Anatomical Automatic Labeling (AAL) with $p = 116$ parcellations were used to construct $p \times p$ structural connectivity matrix that counts the number of white matter fiber tracts between the parcellations [8].

Exact Combinatorial Inference. From the individual structural connectivity matrices, we computed pairwise twin correlations in each group using Spearman's correlation. The resulting twin correlations matrices C_{MZ} and C_{DZ} (edge weights in Fig. 4) were used to compute the heritability index (HI) through Falconer's formula, which determines the amount of variation due to genetic influence in a population: $HI = 2(C_{MZ} - C_{DZ})$ [1]. Although HI provides quantitative measure of heritability, it is not clear if it is statistically significant. We tested the significance of HI by testing the equality of C_{MZ} and C_{DZ} . We used $1 - C_{MZ}$ and $1 - C_{DZ}$ as edge weights in finding MST using Kruskal's algorithm. This is equivalent to using C_{MZ} and C_{DZ} as edge weights in finding the *maximum spanning trees*. Figure 4 plot shows how the number of nodes increase as the edges are added into the MST construction. At the same edge weights, MZ-twins are more connected than DZ-twins in MST. This implies MZ-twins are connected less in lower correlations and connected more in higher correlations.

Results. At edge weight 0.75, which is the maximum gap and corresponding to correlation 0.25, the observed distance D was 46. The corresponding p -value was computed as $P(D \geq 46) = 1.57 \times 10^{-8}$. The localized regions of brain that genetically contribute the most can also be identified by identifying the nodes of connections around edge weight 0.75 (0.75 ± 0.2). The following AAL regions are identified as the region of statistically significant MST differences: Frontal-Mid-L, Frontal-Mid-R, Frontal-Inf-Oper-R, Rolandic-Oper-R, Olfactory-L, Frontal-Sup-Medial-L, Frontal-Sup-Medial-R, Occipital-Inf-L, SupraMarginal-R, Precuneus-R, Caudate-L, Putamen-L, Temporal-Pole-Sup-L, Temporal-Pole-Sup-R, Temporal-Pole-Mid-R, Cerebelum-Crus2-R, Cerebelum-8-R, Vermis-8 (Fig. 5). The identified frontal and temporal regions are overlapping with the previous MRI-based twin study [7].

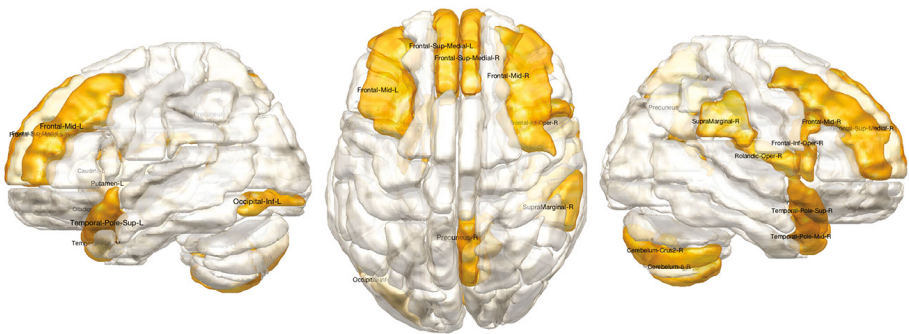


Fig. 5. Regions corresponding to the maximum difference between MST of MZ- and DZ-twins.

6 Conclusion and Discussion

We presented the novel exact combinatorial inference method that outperforms the traditional permutation test. The main innovation of the method is that it works for any monotone features. Given any two sets of measurements, all it requires is to sort them and we can apply the method. Thus, the method can be applied to wide variety of applications. For this study, we have shown how to apply the method in testing the shape differences in MST of the structural brain networks. The method was further utilized in localizing brain regions influencing such differences.

In graphs, there are many monotone functions including the number of connected components, total node degrees and the sorted eigenvalues of graph Laplacians. These monotone functions can all be used in the proposed combinatorial inference. The proposed method is also equally applicable to wide variety of monotonically decreasing graph features such as the largest connected components [1]. If $\phi \circ f$ is monotonically decreasing, $-\phi \circ f$ is monotonically increasing, thus the same method is applicable to decreasing functions. The applications of other features are left for future studies.

Acknowledgements. This work was supported by NIH grants R01 EB022856, R01 MH101504, P30 HD003352, U54 HD09025, UL1 TR002373. We thank Nagesh Adluru, Yuan Wang, Andrey Gritsenko (University of Wisconsin-Madison), Hyekeyoung Lee (Seoul National University), Zachery Morrissey (University of Illinois-Chicago) and Sourabh Palande, Bei Wang (University of Utah) for valuable supports.

References

1. 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. (ed.) IPMI 2017. LNCS, vol. 10265, pp. 299–310. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-59050-9_24
2. Goldsmith, H.H., Lemery-Chalfant, K., Schmidt, N.L., Arneson, C.L., Schmidt, C.K.: Longitudinal analyses of affect, temperament, and childhood psychopathology. *Twin Res. Hum. Genet.* **10**, 118–126 (2007)
3. Hanson, J.L., Adluru, N., Chung, M.K., Alexander, A.L., Davidson, R.J., Pollak, S.D.: Early neglect is associated with alterations in white matter integrity and cognitive functioning. *Child Develop.* **84**, 1566–1578 (2013)
4. Lazar, M., et al.: White matter tractography using tensor deflection. *Hum. Brain Mapp.* **18**, 306–321 (2003)
5. Neykov, M., Lu, J., Liu, H.: Combinatorial inference for graphical models. arXiv preprint [arXiv:1608.03045](https://arxiv.org/abs/1608.03045) (2016)
6. Stam, C.J., Tewarie, P., Van Dellan, E., van Straaten, E.C.W., Hillebrand, A., Van Mieghem, P.: The trees and the forest: Characterization of complex brain networks with minimum spanning trees. *Int. J. Psychophysiol.* **92**, 129–138 (2014)
7. Thompson, P.M., et al.: Genetic influences on brain structure. *Nat. Neurosci.* **4**, 1253–1258 (2001)

8. 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)
9. Wu, Z., Leahy, R.: An optimal graph theoretic approach to data clustering: theory and its application to image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **15**, 1101–1113 (1993)
10. Zalesky, A., et al.: Whole-brain anatomical networks: does the choice of nodes matter? *NeuroImage* **50**, 970–983 (2010)
11. Zhang, H., Yushkevich, P.A., Alexander, D.C., Gee, J.C.: Deformable registration of diffusion tensor MR images with explicit orientation optimization. *Med. Image Anal.* **10**, 764–785 (2006)