Annual Meeting of the Organization for Human Brain Mapping (OHBM) 2022

Wasserstein Graph Clustering in Determining the Genetic Contribution of State Changes in rs-fMRI

Abstract No: 2536

Authors:
Moo K. Chung[1], Shih-Gu Huang[1], Ian C. Carroll[1], Vince D. Calhoun[2], H. Hill Goldsmith[1]

Institutions:
[1]University of Wisconsin, Madison, WI, USA
[2]Tri-institutional Center for Translational Research in Neuroimaging and Data Science (TReNDS), Georgia State, Georgia Tech, Emory Georgia State University, Georgia, USA

Introduction:

We present the novel Wasserstein graph clustering method for networks (Anand, 2021). The Wasserstein clustering penalizes the topological discrepancy between 0D and 1D homological features. The Wasserstein clustering outperforms the widely used k-means clustering. We applied the method in more accurately determining the genetic contribution of the dynamically changing state spaces in rs-fMRI.

Methods:

Subjects: Rs-fMRI were collected on a 3TMRI scanner (Discovery MR750, GE) with a 32-channel RF head coil array. The functional scans were processed with AFNI (Burghy, 2016). 479 healthy subjects consisting of 231 males and 248 females ranging in age from 13 to 25 years. The dataset also consists of 132 monozygotic (MZ) twin pairs and 93 same-sex dizygotic (DZ) twin pairs.

Preprocessing: We parcellated the brain volume into 116 non-overlapping brain regions (Tzourio, 2002). The fMRI data were averaged across voxels within each parcellation resulting in 116 average fMRI signals per subject. The averaged fMRI signal is further temporally smoothed using heat kernel expansion (Huang, 2019). The smoothed fMRI are correlated across parcellations form dynamically changing correlation matrices of size 116 by 116 at 295 time points.

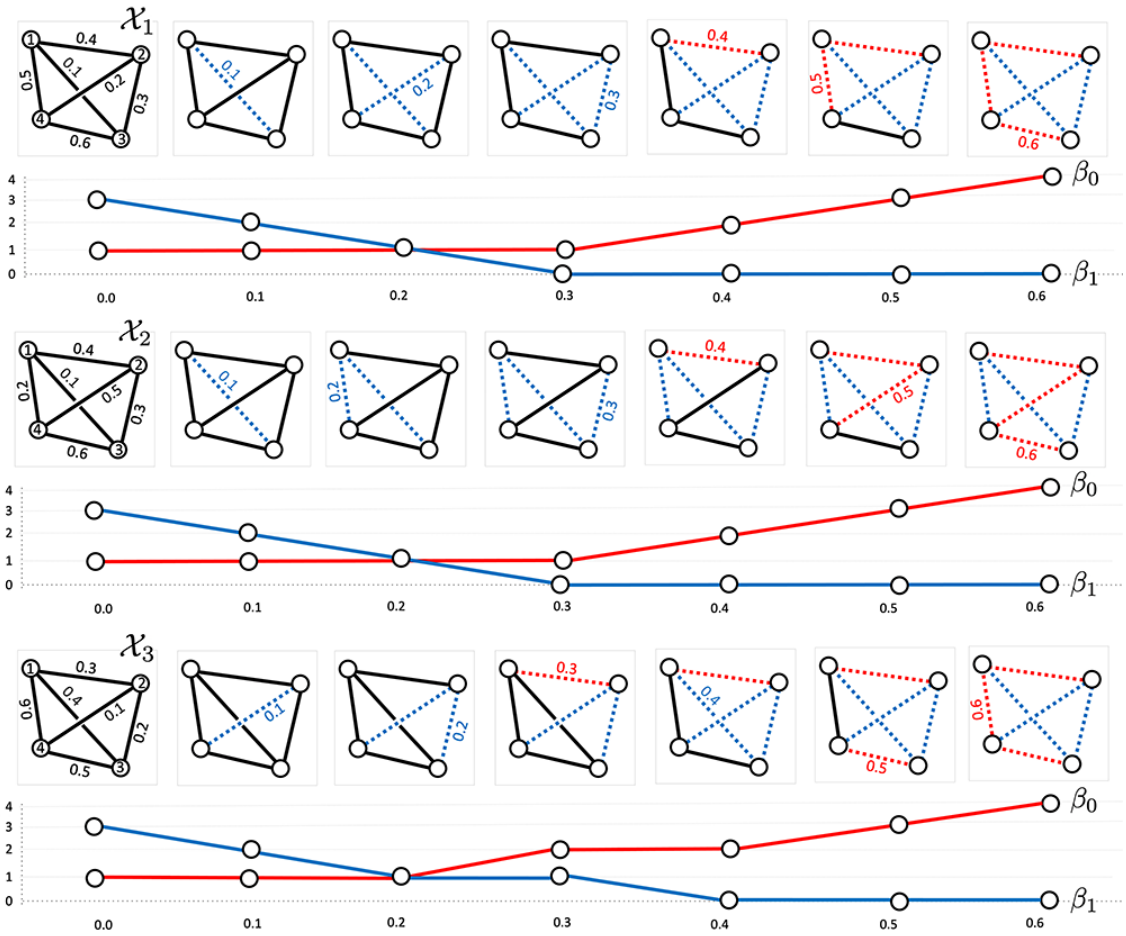

Figure 1-(a). Betti number plots over graph filtration. When each edge is removed from the smallest to the largest edge weights, the number of connected components (β0) increases while the number of cycles (β1) decreases. The edge weights corresponding to the birth of connecteced components and the death of cycles are used in computing the topological discrepancy through the Wasserstein distance.

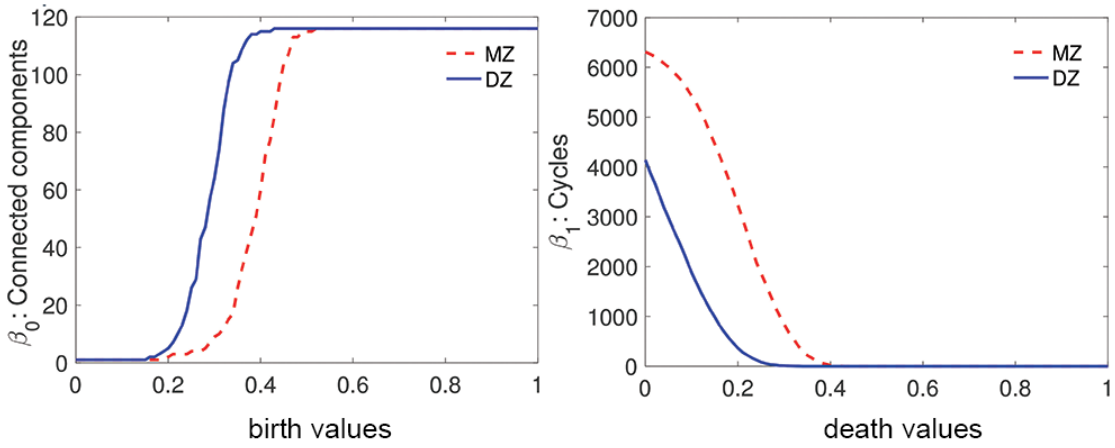

Figure 1-(b). Change of Betti numbers over birth and death values for MZ- and DZ-twins.The β0- and β1-curves and corresponding birth and death values are heritable topological features.

Graph filtration: Graph filtration was performed on correlation matrices to extract persistent homological features by thresholding one edge at a time (Figure 1-a). Betti numbers β_0 and β_1, which measures the number of connected components and cycles respectively, are monotone over the filtration (Chung, 2019). The edge weights that increase β_0 or decrease β_1 are called the birth or death values respectively (Songdechakraiwut, 2021).

Wasserstein graph clustering: The Wasserstein distance is given by the smallest sum of squared distance between birth and death values. The optimization is solved as a linear assignment problem. The Wasserstein distance is used to cluster a collection of graphs by minimizing the within-cluster distance. The minimization is done as the two-step optimization like the expectation maximization (Bishop, 2006). Since the algorithm depends on the initial estimation of the cluster center, the clustering is run 10 times with different initial seeds.

Validation: For testing false positives (no cluster difference), k-means clustering incorrectly clustered with 1.00 +/- 0.04 accuracy while the Wasserstein clustering clustered with 0.53 +/- 0.08 accuracy. For testing false negatives (Figure 2-a), k-means clustering only able to correct cluster with 0.83 +/- 0.16 accuracy while the Wasserstein clustering achieved remarkable with 0.96 +/- 0.10 accuracy.

Results:

We clustered time varying time correlation matrices at the group level. The optimal number of clusters is determined to be 3 through the elbow method (Huang, 2020). The ratio of within-cluster to between-cluster distance is 0.034 +/- 0.012 for Wasserstein clustering and 0.202 +/- 0.047 for k-means clustering (Figure 2-b). The six times smaller ratio for the Wasserstein clustering demonstrates the better model fit. Figure 2-c displays the topological average brain networks within each cluster. The averaging is done through the Wasserstein distance, which provides 6 times shaper averaging performance compared to the Euclidean distance.

We determined the extent of the genetic influence on the estimated state spaces. We computed heritability index through the MZ- and DZ-twin correlations (Chung, 2019). Figure 2-c displays the estimated heritability index. The heritability of the first state is characterized by lateralization of the hemisphere connections. The heritability of the second state is characterized by asymmetric front and back connections (Chen, 2008). The method provides far more accurate and stable heritability index map than previous attempts (Chung, 2019).

Conclusions:

Through the proposed novel persistent homological clustering method, we were able to estimate accurate heritability index map of changing state patterns in rs-fMRI.

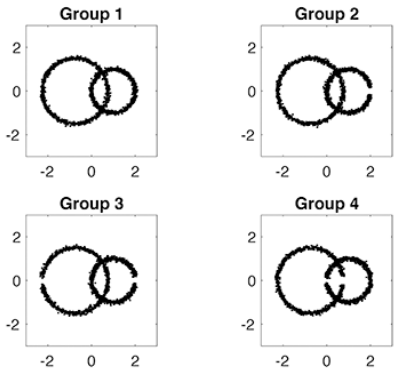

Figure 2-(a). Simulation study showing topologically different patterns across 4 groups.

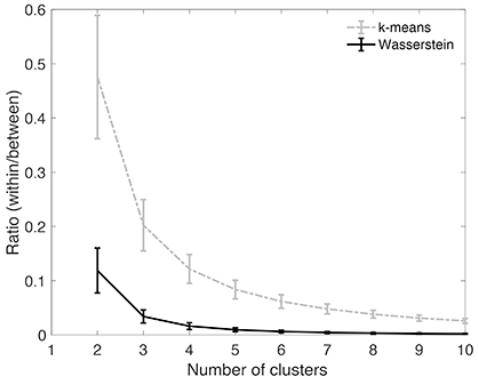

Figure 2-(b). The plot of ratio of within-cluster to between cluster dsitance.

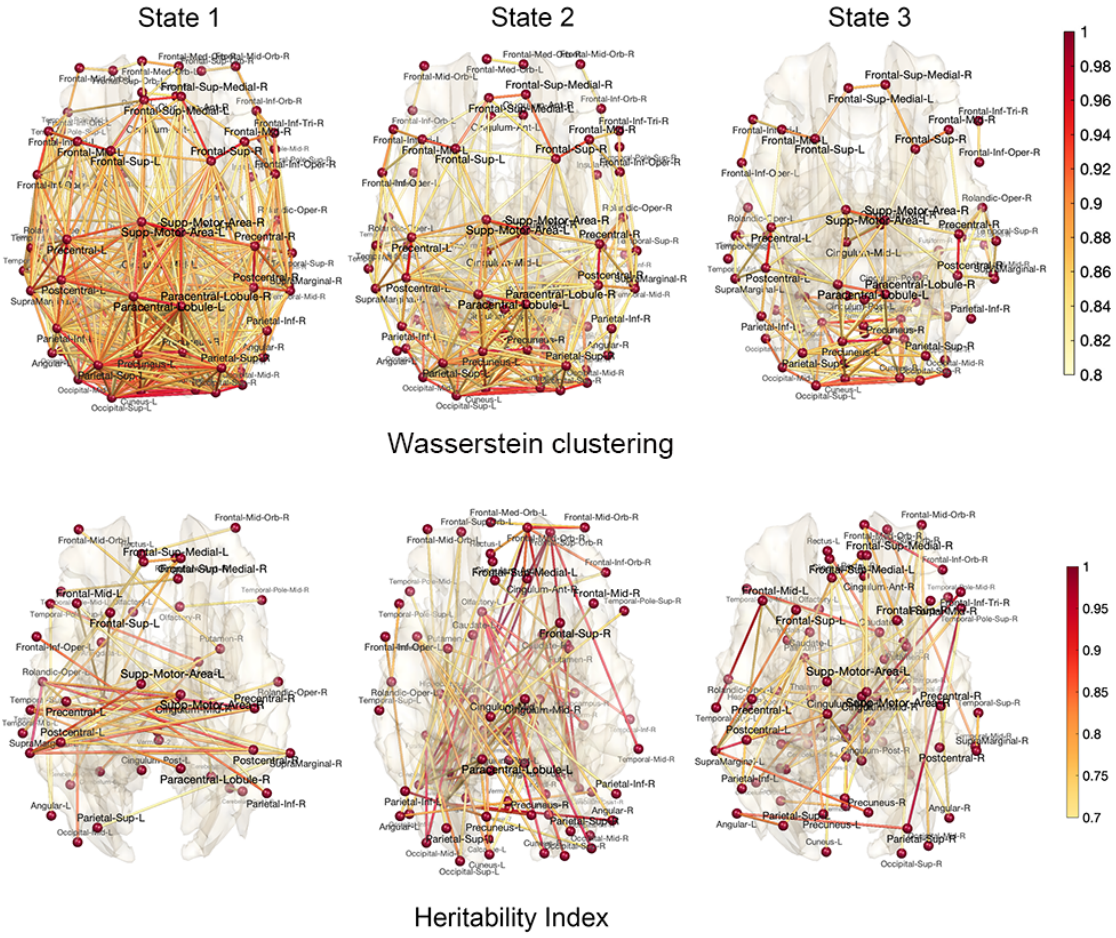

Figure 2-(c). Estimated state spaces using the proposed Wasserstein graph clustering. The heritability index is computed for each state showing the extensive genetic contribution of dynamically changing states than previously determined using the baseline k-means clustering

References:

Anand, D.V. (2021) Hodge-Laplacian of brain networks and its application to modeling cycles. arXiv:2110.14599

Bishop, C. (2006) Pattern recognition and machine learning. springer

Burghy, C. (2016) Experience-driven differences in childhood cortisol predict affect-relevant brain function and coping in adolescent Monozygotic twins. Scientific Reports 6, 37081

Chen, C.C. (2009) Forward and backward connections in the brain: A DCM study of functional asymmetries. NeuroImage 45:453-462

Chung, M.K. (2020) Statistical analysis of dynamic functional brain networks in twins, arXiv 1911.02731

Chung, M.K. (2019) Statistical inference on the number of cycles in brain networks. IEEE International Symposium on Biomedical Imaging (ISBI) 113-116

Huang, S.-G. (2020) Statistical model for dynamically-changing correlation matrices with application to brain connectivity. Journal of Neuroscience Methods 331:108480

Huang, S.-G. (2019) Dynamic functional connectivity using heat kernel. IEEE Data Science Workshop (DSW), 222-226

Songdechakraiwut, T. (2021) Topological learning and its application to multimodal brain network integration, Medical Image Computing and Computer Assisted Intervention (MICCAI), LNCS 12902:166-176

Tzourio-Mazoyer, M. (2002) Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain, NeuroImage, vol. 15, pp. 273–289