DISCRIMINATIVE PERSISTENT HOMOLOGY OF BRAIN NETWORKS

Hyekyoung Lee [1,3,4], *Moo K. Chung* [4,6,7,*], *Hyejin Kang* [1, 3], *Bung-Nyun Kim* [2], *Dong Soo Lee* [1,5]

[1] Department of Nuclear Medicine, [2] Department of Neuropsychiatry,
[3] Institute of Radiation Medicine, Medical Research Center
[4] Department of Brain and Cognitive Sciences,
[5] WCU Department of Molecular Medicine and Biopharmaceutical Sciences,
Seoul National University, Seoul, Korea
[6] Department of Biostatistics and Medical Informatics
[7] Waisman Laboratory for Brain Imaging and Behavior
University of Wisconsin, Madison, WI 53706, USA

ABSTRACT

It is known that the brain network has small-world and scale-free topology, but the network structures drastically change depending on how to threshold a connectivity matrix. The exact threshold criterion is difficult to determine. In this paper, instead of trying to determine one fixed optimal threshold, we propose to look at the topological changes of brain network while increasing the threshold continuously . This process of continuously changing threshold level and looking at the resulting topological feature is related to the Rips filtration in persistent homology. The sequence of topological features obtained during the Rips filtration can be visualized and interpreted using barcode.

As an illustration, we apply the Rips filtration to construct the FDG-PET based functional brain networks out of 24 attention deficit hyperactivity disorder (ADHD) children, 26 autism spectrum disorder (ASD) children and 11 pediatric control subjects. We visually show the topological evolution of the brain networks using the barcode and perform statistical inference on the group differences. This is the first paper that deals with the persistence homology of the brain networks.

Index Terms— Brain Network, Thresholding, Persistent Homology, Rips Complex, Barcode

1. INTRODUCTION

The functional and anatomical connectivity studies of human brain have given us new understanding of the characteristics of brain, from microscale connectivity between single neurons to macroscale connectivity between regions of interest (ROI) in whole brain images. The connectivity matrix is an algebraic representation of the weighted brain network, which shows the relationship between all paired nodes. Since the interpretation of weighted graphs is somewhat complicated, we usually binarize the connectivity matrix into an adjacency matrix by thresholding the connectivity matrix.

So far the global topological characteristics of brain network, such as small-world network, has been mainly studied but recently, the local structure of the brain network, i.e., modularity, has started to draw attention [1]. However, most graph theoretic measures such as small-worldness and modularity can quantify only one aspect of the brain network at a fixed threshold. In this paper, we propose to look at the topological changes of the brain network for every possible thresholds, rather than trying to determine one fixed threshold that may not be optimal. We mathematically demonstrate that the changes of network structure when varying threshold of connectivity matrix can be exactly observed by finding the evolutionary history of the topological changes of Rips complex. Thus, we can borrow various algebraic topology tools such as barcodes and persistent diagrams for representing the change of topological features [2, 3]. Although the idea of persistent homology has already been applied to medical image analysis [4, 2], this is the first study modeling brain networks using the persistent homology.

The proposed method is applied to constructing functional brain networks with 97 regions of interest (ROIs) extracted from FDG-PET data for 24 attention-deficit hyperactivity disorder (ADHD), 26 autism spectrum disorder (ASD) and 11 pediatric control (PedCon). Numerical experiments show that their topological changes through varying threshold can be quantified and visualized through persistent homology and barcodes.

The correspondence should be send to M.K.C. (mkchung@wisc.edu). This work was supported by the WCU Grant from the government of Korea to M.K.C., by grant No. R31-2008-000-10103-0 from the WCU project of the MEST and the NRF, and by a grant (M103KV010016-08K2201-01610) from Brain Research Center of the 21st Century Frontier Research Program funded by the Ministry of Science and Technology.

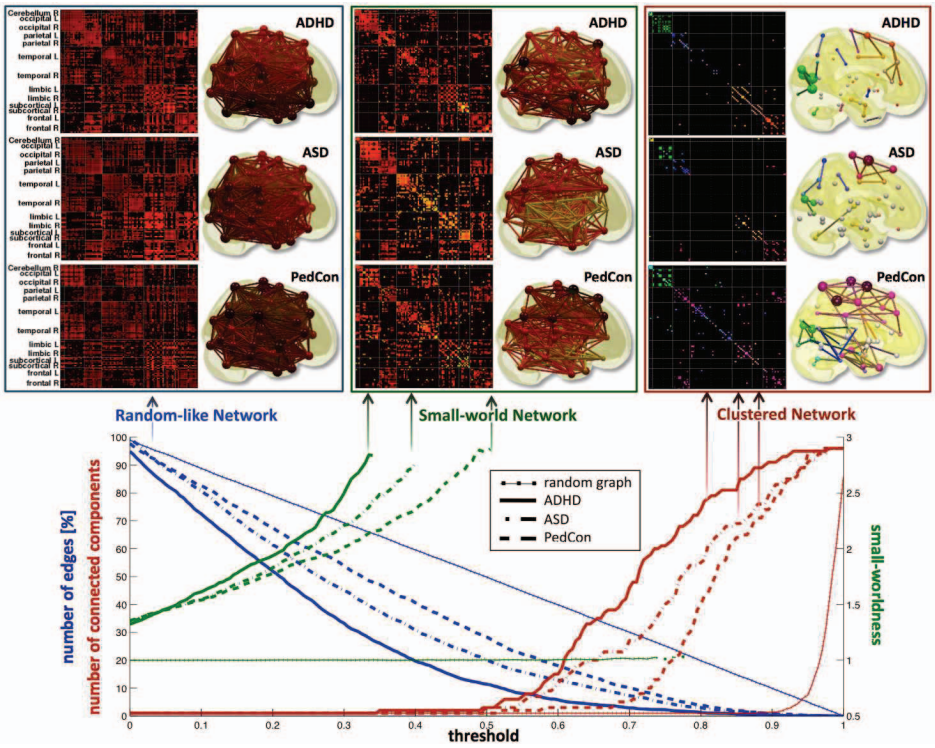

Fig. 1. How brain network changes for different thresholding. The blue, red and green lines in lower panel represent the number of edges, the number of connected components and small-worldness. The dot-solid, solid, dot-dashed and dashed lines in lower panel are for random graph, ADHD, ASD and PedCon networks, respectively. By varying the threshold, the brain network is changed to like random-like, small-world and clustered network. The colors in the clustered network indicates the clusters. The darker color in the adjacency matrices represents highly connected nodes.

2. THRESHOLDING CONNECTIVITY

In the usual connectivity analysis framework, the adjacency matrix is subsequently obtained by thresholding the correlation matrix of measurements selected brain regions. Finding the proper threshold is one of the most important issues in brain network modeling. Determining the threshold can be based on the statistical significance by the false discovery rate (FDR) or by fixing the graph metrics such as number of edges and nodes. However, these methods are fairly ad-hoc and everyone seem to use different thresholding techniques. This arbitrariness is demonstrated in Fig. 1, where the small-worldness (green lines), the number of edges (blue lines) and the number of connected components (red lines) substantially change depending on the threshold. According to the threshold, the resulting network graph has very different topological structures: random, small-world and clustered networks [5].

3. PERSISTENT HOMOLOGY

Instead of trying to determine one proper threshold that may not really proper, we decided to look at the over all change of topological structure over whole range of thresholds using persistent homology.

Consider a set $\boldsymbol{F}$ of point cloud data consisting of p points. We connect two points i and j by an edge if the distance $d(i,j)$ is smaller than ϵ. The generated graph is a *Rips complex* and denoted by $\text{Rips}(\boldsymbol{F},\epsilon)$. The topological informations of Rips complex are encoded into an algebraic form, known as a *Betti number*, where the 0-th Betti number β_0 counts the number of connected components in the graph. Fig. 2 (a)-(j) show a toy example of constructed Rips complex with different ϵ. The radius of circle around each dot is ϵ. If two dots are in the same circle, they are connected (red lines).

Since it can be easily seen that $\text{Rips}(\boldsymbol{F},t) \subset \text{Rips}(\boldsymbol{F},s)$ whenever $t \leq s$. Observing the topological transition by increasing the filtration value ϵ is called as a *Rips filtration* and is the main theme in *persistent homology*. During the filtration, the topological features such as the connected components is created and disappeared. These can be visualized by either using the persistent diagram or *barcode*. In the barcode, the vertical axis represents the Betti number and the horizontal axis corresponds to the filtration value. At $\epsilon = 0$, there is no connection and the number of connected components is

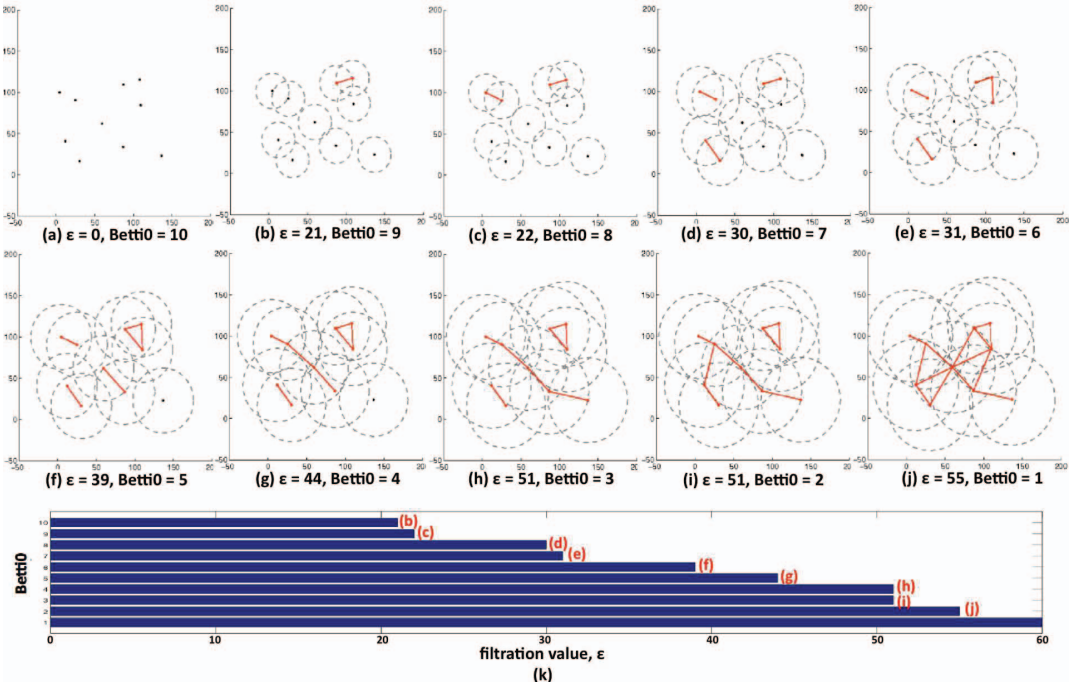

Fig. 2. Example of Rips filtration varying the filtration value ϵ (upper panels) and its barcode (the lower panel). The black dot represents the point cloud data which are connected (red line) when two nodes are in the same circle with radius ϵ. The lower panel is a barcode where the horizontal and vertical axes represent the filtration value and the connected components.

simply the number of nodes. Since there are 10 dots in Fig. 2 (k), the barcode starts at the height 10. At $\epsilon = 21$, two connected components are merged into a single connected component and one component disappears so the barcode stops at $\epsilon = 21$. In this way, all nodes are connected for sufficiently large filtration value ϵ and, finally, only one single connected component remains.

In the case of the brain network, we have p measurements $\boldsymbol{F} = \{\boldsymbol{f}_1, \ldots, \boldsymbol{f}_p\}$ obtained from p regions, which serve as point cloud to be connected. We assume that the measurements are centered and normalized. Thus, the correlation coefficient between $\boldsymbol{f}_i$ and $\boldsymbol{f}_j$ satisfies $\text{corr}(\boldsymbol{f}_i, \boldsymbol{f}_i) = \boldsymbol{f}_i^\top \boldsymbol{f}_i = 1$ for all $i = 1, \ldots, p$, and the correlation matrix is simply estimated by $\boldsymbol{\Sigma} = [\rho_{ij}] = [\boldsymbol{f}_i^\top \boldsymbol{f}_j]$.

The distance between two nodes i and j are defined not in the space where the nodes are residing but in the space where the measurements are defined. We denote this distance as $d(\boldsymbol{f}_i, \boldsymbol{f}_j)$ and link them if $d(\boldsymbol{f}_i, \boldsymbol{f}_j) < \epsilon$. One possible distance measure we can use is

$$d(\boldsymbol{f}_i, \boldsymbol{f}_j) = 1 - \text{corr}(\boldsymbol{f}_i, \boldsymbol{f}_j).$$

To simplify the problem, we consider only positive correlations and the Euclidean distances. The distance between $\boldsymbol{f}_i$ and $\boldsymbol{f}_j$ gets smaller as the correlation increases between them. We construct Rips($\boldsymbol{F}, \epsilon$) by connecting all nodes with the correlation larger than specific threshold ϵ. The constructed Rips complex is then exactly identical to the network connected by thresholding the correlation matrix. Therefore, the Rips filtration, a sequence of Rips complexes, is a more general framework than the usual connectivity matrix threshold method. The underlying topological change in the Rips filtration of the brain network is then encoded in the barcode.

4. RESULTS

The data consists of 24 ADHD, 26 ASD and 11 PedCon. PET images were preprocessed using Statistical Parametric Mapping (SPM) package. After spatial normalization to the standard template space, mean FDG uptake within 97 ROIs were extracted ROIs using Statistical Probabilistic Anatomical Map-Korean version (SPAM-K) as shown in Fig. 3. The values of FDG uptake were globally normalized to the individual's total gray matter mean count.

The barcodes of ADHD (red bars), ASD (green bars) and PedCon (blue bars) networks with the 0-th Betti number are shown in Fig. 4. Each bar is started and ended while each connected component is appeared and disappeared during the filtration. The bars of PedCon are merged faster into the last bar left while bars of ADHD and ASD are survived for a longer time. It might be due to common underconnectivity and local overconnectivity in ASD [6] and ADHD [7].

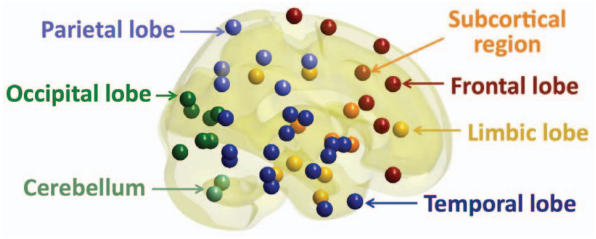

Fig. 3. Location of ROIs

The filtration values at which all connected components are created are identical so we only checked whether the death time of connected components are different between ADHD, ASD and PedCon networks using 1000 permutation test. The topological changes of connected components during filtration (threshold) are significantly different between ADHD-ASD ($p = 0.032$), ADHD-PedCon ($p = 0.030$) and ASD-PedCon ($p = 0.037$).

5. CONCLUSIONS

So far researchers are mainly concerned with the global characteristics of brain network such as small-worldness and scale-freeness. Such characteristics are one property of complex brain network at a certain threshold and do not completely characterize the network. By tabulating the topological changes for all possible threshold, we can obtain a more complete characterization of the network. We have shown that these characterization can be represented by barcodes in persistent homology. We have applied the proposed method in global characterization of ADHD, ASD and PedCon. The differences between ADHD and ASD groups are found in the local connectivity structures. However, since the algebraic topology approach is coordinate-free, we can't compare which parts of connected components are disappeared earlier or not. Finding the topological information combined with the location information of node is left as a future work.

6. REFERENCES

[1] P. Laurienti, C. Hugenschmidt, and S. Hayasaka, "Modularity maps reveal community structure in the resting human brain," *Nature Preceedings*, 2009.

[2] R.J. Adler, O. Bobrowski, M.S. Borman, E. Subag, and S. Weinberger, "Persistent homology for random fields and complexes," *ArXiv e-prints*, 2010.

[3] D. Horak, S. Maletić, and M. Rajković, "Persistent homology of complex networks," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2009, pp. P03034, 2009.

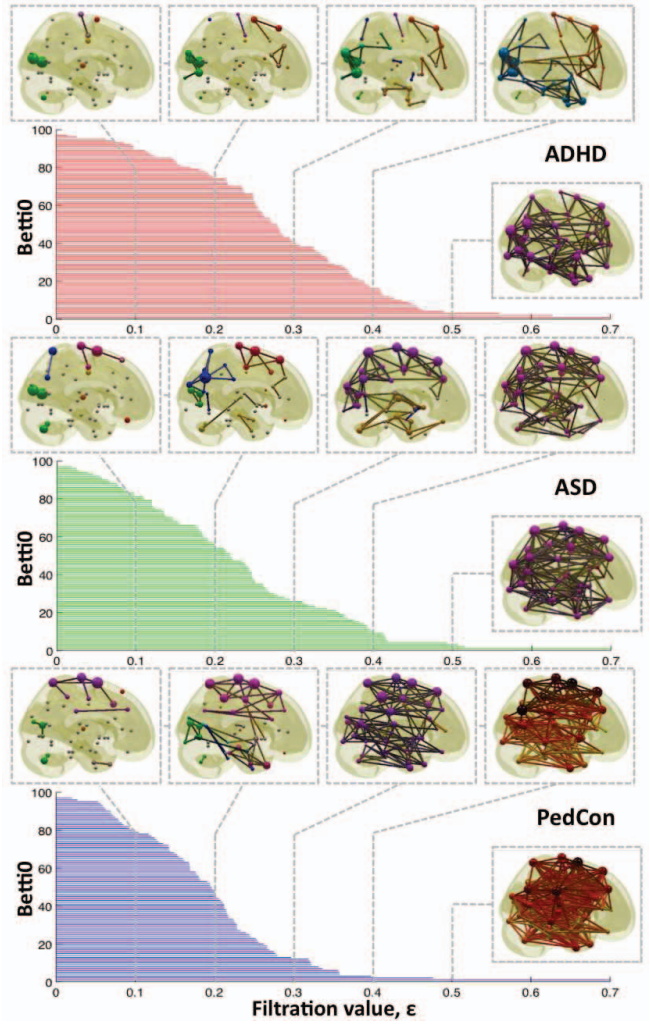

Fig. 4. Barcode of the 0-th Betti number.

[4] M.K. Chung, P. Bubenik, and P.T. Kim, "Persistence diagrams of cortical surface data," in *IPMI '09: Proceedings of the 21st International Conference on Information Processing in Medical Imaging*, 2009, pp. 386–397.

[5] J.M. Kleinberg, "Navigation in a small world," *Nature*, vol. 406, 2000.

[6] N.J. Minshew and D.L. Williams, "The new neurobiology of autism: Cortex, connectivity, and neuronal organization," *Arch Neurol.*, pp. 945–950, 2007.

[7] A. Cubillo and K. Rubia, "Neuroimaging of adult ADHD," *Expert Rev of Neurotherapeutics*, pp. 603–620, 2010.