ABNORMAL HOLE DETECTION IN BRAIN CONNECTIVITY BY KERNEL DENSITY OF PERSISTENCE DIAGRAM AND HODGE LAPLACIAN

Hyekyoung Lee[a] *Moo K. Chung*[f] *Hyejin Kang*[b] *Hongyoon Choi*[e] *Yu Kyeong Kim*[c] *Dong Soo Lee*[a,b,d]

[a] Seoul National University Hospital, [b] Seoul National University,
[c] Seoul National University Boramae Medical Center, Seoul,
[d] Korea Brain Research Institute, Daegu, [e] Cheonan Public Health Center, Chungnam,
Republic of Korea [f] University of Wisconsin, Madison, WI 53706 USA

ABSTRACT

Community and rich-club detection are a well-known method to extract functionally specialized subnetwork in brain connectivity analysis. They find densely connected subregions with large modularity or high degree in brain connectivity studies. However, densely connected nodes are not the only representation of network shape. In this study, we propose a new method to extract abnormal holes, which are another representation of network shape. While densely connected component characterizes network's efficiency, abnormal holes characterize inefficiency. The proposed method differs from the existing hole detection in two respects. One is to use Hodge Laplacian to obtain a harmonic hole in the linear combination of edges, rather than a subset of edges. The other is to use the kernel density estimation of persistence diagram of random networks to determine the significance of a hole, rather than using the persistence of a hole. We applied the proposed method to find the abnormality of metabolic connectivity in the FDG PET data of ADNI. We found that, as AD severely progressed, the brain network had more abnormal holes. The localized holes showed how inefficient the structure of brain network became as the disease progressed.

Index Terms— Hole, Brain connectivity, Alzheimer's disease, Kernel density estimation, Hodge Laplacian

1. INTRODUCTION

Functional specialization has been often studied in brain imaging analysis. By introducing the concept of network, the research interest of brain imaging analysis has been extended from functional specialization of local regions to the functional integration in the whole brain connectivity [1]. Although the complex graph theoretic measures such as small-worldness and modularity have been widely used for finding the global characteristics of brain connectivity, we still want to know which parts of brain are related to the specific cognitive function or neurodegenerative disease. In this sense, community and rich-club detection have been proposed to extract functionally specialized subnetwork [2, 3]. They find densely connected subregions with large modularity or high degree in the whole brain connectivity. Those regions are usually considered to be related with the specific cognitive function, and weakened by the disease progression. However, it is rarely mentioned how the shape of brain network is changed after the deterioration of connections in the regions.

In this paper, we propose a hole detection method to find the shape change during the disease progression. Hole is also a connected component, but there is a path that any node in the hole can return to itself, and the path consists of at least four consecutive edges. Like a connected component, a hole is also the fundamental shape descriptor defined in algebraic topology [4]. While densely connected component represents the network's local efficiency, the hole with many consecutive edges characterizes the network's inefficiency. The concept of hole defined in topological data analysis has been already used for brain network analysis [5, 6, 7]. In the topological data analysis, a persistent hole that lasts for a long threshold is considered as a signal of network and a hole with short duration of threshold is considered as a noise. In this sense, the existing hole detection methods find the significant hole using the persistence of the hole. However, we determine the significance of hole using the probability map of persistence diagram because not only the persistent hole, but also the hole that appears and disappears at abnormal thresholds are impor-

Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf. This work is supported by Basic Science Research Program through the National Research Foundation (NRF) (No.2013R1A1A2064593 and No.2016R1D1A1B03935463), NRF Grant funded by MSIP of Korea (No.2015M3C7A1028926 and No.2017M3C7A1048079), NRF grant funded by the Korean Government (No. 2016R1D1A1A02937497, No.2017R1A5A1015626, and No.2011-0030815), and NIH grant EB022856.

tant to determine the shape of brain network. The kernel density estimation is used for obtaining the probability map of persistence diagram.

The proposed method finds a harmonic hole represented by the weighted sum of edges based on Hodge Laplacian [8, 5], while the existing method usually finds a hole in the binary representation of edges, i.e., a subset of edges [6, 7, 9]. The weights of the harmonic hole are proportional to the contribution of edges to the hole. Moreover, while the existing method can not find the unique representation of a hole, i.e., there are many other possible subsets of edges for representing a hole, the harmonic hole method finds the unique representation of a hole. In experiments, the proposed method was applied to FDG PET data of Alzheimer's disease neuroimaging initiative (ADNI) and showed that the proposed method can find the abnormal shape of network as AD progressed.

2. METHODS

2.1. Network construction

The ADNI FDG-PET dataset consists of three groups, 181 normal controls (NC), 168 mild cognitive impairment (MCI), and 135 Alzheimer's disease (AD) subjects (Age: 73.7 ± 5.9, range $56.1 - 90.1$). The MCI group was divided into two groups, 91 stable MCI (sMCI) and 77 progressive MCI (pMCI), depending on whether a subject remained stable or progressed to AD after three years. Details of data sets and preprocessing are in [10]. The whole brain image was parcellated into 94 regions of interest (ROIs) based on automated anatomical labeling (AAL2) except for cerebellum [11]. Each ROI serves as a node and its measurement is obtained by averaging FDG uptakes in the ROI. The distance between two nodes was estimated by the diffusion distance on positive correlations between measurements of two ROIs. The diffusion distance considers an average distance of all direct and indirect paths between two nodes via random walk.

2.2. Hole

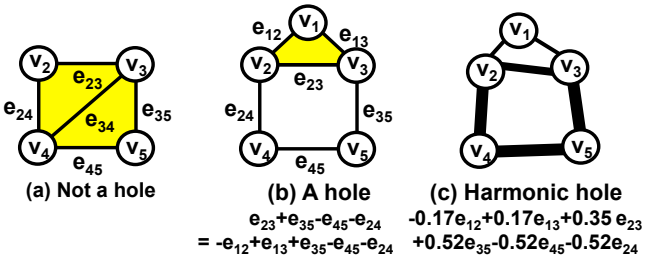

Fig. 1: Example of Rips complex (a) without a hole and (b) with a hole, represented in two different binary representations. (c) Harmonic hole of (b), represented by the weighted sum of edges.

Suppose that a weighted network $\mathcal{N}(V, E, \boldsymbol{L})$ consists of p nodes in a set V and q edges in E, and an edge distance matrix $\boldsymbol{L} \in \mathbb{R}^{p\times p}$. The entry of $\boldsymbol{L} = [l(e_{ij})]$ is the distance of the edge $e_{ij} \in E$ connecting two nodes v_i and v_j ($v_i, v_j \in V$). Given a threshold $\epsilon > 0$, Rips complex, denoted by $\mathcal{R}_N(\epsilon)$, is the collection of nodes V, edges satisfying $l(e) < \epsilon, e \in E$ and triangles satisfying $l(e_{ij}), l(e_{jk}), l(e_{ki}) < \epsilon$, $e_{ij}, e_{jk}, e_{ki} \in E$, $v_i, v_j, v_k \in V$. When the sequence of thresholds is given by $\epsilon_0 = 0 \leqslant \epsilon_1 \leqslant \epsilon_2 \leqslant \cdots$, the sequence of Rips complexes is estimated by $\mathcal{R}_N(\epsilon_0) \subseteq \mathcal{R}_N(\epsilon_1) \subseteq \mathcal{R}_N(\epsilon_2) \subseteq \cdots$. This procedure is called the Rips filtration [4].

Definition 1. Holes *in $\mathcal{R}_N(\epsilon)$ is a subset $H \subseteq E$ where consecutive edges form a cycle, but are not a boundary of any consecutive triangles (See Fig. 1) [9]. The number of holes in $\mathcal{R}_N(\epsilon)$ is* the first Betti number*, denoted as β_1.*

Definition 2. *If a hole appears at the threshold ξ and disappears at τ ($0 \leqslant \xi \leqslant \tau < \infty$), it is encoded into a 2-dimensional point $t = (\xi, \tau) \in \mathbb{R}^2$. Given weighted network $\mathcal{N}(V, E, \boldsymbol{L})$, if m holes appear and disappear during the filtration of $\mathcal{N}$, they are represented by a set of m points $P = \{t_1, \ldots, t_m\}$. The scatter plot of P is called a* persistence diagram (PD) *of $\mathcal{N}$ [9]. Because $\xi_i \geqslant \tau_i$ for $\forall i$, the points are always in the upper regions of the diagonal line $y = x$ in $\mathbb{R}^2$.*

2.3. Kernel density estimation of persistence diagram

Let $u_1, \ldots, u_m \in \mathbb{R}^2$ be an independent, identically distributed random sample from an unknown density p. Kernel density estimation can be expressed as

$$\hat{p}_m(u) = \frac{1}{mh^2} \sum_{i=1}^{m} K\left(\frac{u - u_i}{h}\right),$$

where $K : \mathbb{R}^2 \to \mathbb{R}$ is a smooth kernel function and $h > 0$ is the smoothing bandwidth that controls the amount of smoothing. The kernel function should satisfy two requirements, normalization $\int_{-\infty}^{\infty} K(u)du = 1$ and symmetry $K(-u) = K(u)$ for all u.

Since a point $t = (\xi, \tau)$ in a persistence diagram is bounded by $\xi, \tau > 0$ and $\xi < \tau$, we cannot directly apply the traditional kernel function. Thus, we transform $t = (\xi, \tau)$ to $u(t) = (v(t), w(t)) = (-\log(\tau + \xi), -\log(\tau - \xi))$ *on the entire real plane*. The function u is *one-to-one transformation*. The Jacobian of the transformation function u is $\left|\frac{\partial(v,w)}{\partial(\xi,\tau)}\right| = \begin{vmatrix} \frac{\partial v}{\partial \xi} & \frac{\partial v}{\partial \tau} \\ \frac{\partial w}{\partial \xi} & \frac{\partial w}{\partial \tau} \end{vmatrix} = \begin{vmatrix} -\frac{1}{\tau+\xi} & -\frac{1}{\tau+\xi} \\ \frac{1}{\tau-\xi} & -\frac{1}{\tau-\xi} \end{vmatrix} = \frac{2}{(\tau+\xi)(\tau-\xi)}$. Then, we can write the kernel density of persistence diagram

$$\begin{aligned} \hat{p}(t) &= \hat{p}_m(u) \left|\frac{\partial(v,w)}{\partial(\xi,\tau)}\right| \qquad (1) \\ &= \left(\frac{1}{mh^2} \sum_{i=1}^{m} K\left(\frac{u(t) - u_i(t)}{h}\right)\right) \frac{2}{(\tau+\xi)(\tau-\xi)}. \end{aligned}$$

We use the student's t-distribution for the kernel K because it is more robust to errors with heavier trails than normal distributions

Definition 3. *Given N random weighted networks, we can obtain N persistence diagrams, $P_1, \ldots, P_N$, and the corresponding kernel density $\hat{p}_1, \ldots, \hat{p}_N$ in (1). Then, the probability density of persistence diagram is estimated by the average kernel density map $\bar{p} = \frac{1}{N}\sum_{i=1}^{N} \hat{p}_i$. If the birth and death of a hole is in the region of significance level $< .05$ in $\bar{p}$, we call it an* abnormal hole *of which birth and death are rarely found in random networks.*

2.4. Hodge Laplacian for hole localization

Since the boundary of an edge is two nodes, we denote an edge by $e^i = v_1^i - v_2^i$. If two edges $e^i = v_1^i - v_2^i$ and $e^j = v_1^j - v_2^j$ have common node with the same orientation, i.e., $v_1^i = v_1^j$ or $v_2^i = v_2^j$, we say that e^i and e^j are lower adjacent with similar orientation, denoted as $e^i \frown_+ e^j$. If they have common node with different orientation, i.e., $v_1^i = v_2^j$ or $v_2^i = v_1^j$, we say that they are lower adjacent with dissimilar orientation, denoted as $e^i \frown_- e^j$. If two edges e_i and e_j belong to the same triangle, we say that they are upper adjacent, denoted as $e_i \smile e_j$. The number of triangles to which e_i belongs is denoted as $d_u(e_i)$.

Definition 4. *Suppose that Rips complex $\mathcal{R}_N(\epsilon)$ has the ordered edges, $e^1, \ldots, e^q$. The first Hodge Laplacian $\boldsymbol{H}_1(\epsilon) \in \mathbb{R}^{q\times q}$ is defined by*

$$[\boldsymbol{H}_1(\epsilon)]_{ij} = \begin{cases} d_u(e_i) + 2 & i = j, \\ 1 & i \neq j, \sim (e_i \smile e_j), e_i \frown_+ e_j, \\ -1 & i \neq j, \sim (e_i \smile e_j), e_i \frown_- e_j, \\ 0 & \textit{otherwise}, \end{cases} \quad (2)$$

where $\sim(\cdot)$ is 1 if $(\cdot)$ is 0, and 0 if $(\cdot)$ is 1 [8, 5]. The number of zero eigenvalues of $\boldsymbol{H}_1$ is equal to the number of holes β_1 in $\mathcal{R}_N(\epsilon)$. The corresponding null eigenvectors is called the harmonic holes.

Let a hole $t = (\xi, \tau)$ is abnormal. If we choose ϵ in $[\xi, \tau]$, the null eigenvectors of $\boldsymbol{H}_1(\epsilon)$ include the harmonic hole of t. While t is a subset of edges, the harmonic hole is the weighted sum of edges (See Fig. 1). If $\boldsymbol{H}_1(\epsilon)\boldsymbol{z} = 0$, the harmonic hole is represented by $\sum_{i=1}^{q} z_i e_i$, where $\boldsymbol{z} = [z_i]$. The absolute value $|z_i|$ is proportional to the contribution of e_i to the hole. If it is large, the edge is an unique path on the hole. Otherwise, alternative paths exist on the hole. To find the harmonic hole of t, we estimate $\boldsymbol{H}_1(\epsilon_1), \boldsymbol{H}_1(\epsilon_2)$, and $\boldsymbol{H}_1(\epsilon_3)$ for $\xi = \epsilon_1 < \epsilon_2 < \epsilon_3 = \tau$ and find their common null eigenvector.

3. RESULTS

We generated 5000 random networks in a group by permutations and estimated the kernel density of persistence diagram in Definition 3. The kernel density of persistence diagram of NC, sMCI, pMCI, and AD were shown in Fig. 2. 27, 26, 25, and 35 holes were found in NC, sMCI, pMCI, and AD, respectively. The blue and red dots represented holes and abnormal holes, respectively. Among them, only pMCI and AD had an abnormal hole ($p < .05$). t_1 was an abnormal hole of pMCI and t_2 and t_3 were that of AD. The duration of the hole t_3 was not persistent enough, however, it was selected as an abnormal hole by the proposed method. To see the hole structure of sMCI, we chose t_4 in (b) which has the smallest p value in sMCI ($p = 0.122$). $t_1, \ldots, t_4$ are shown in Fig. 3.

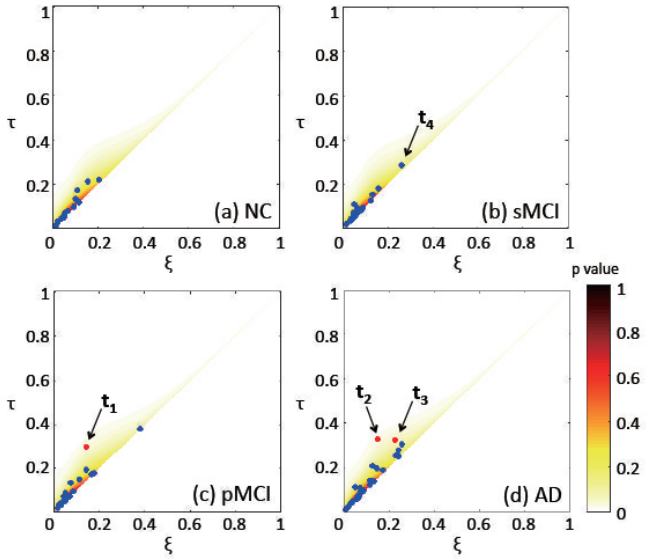

Fig. 2: Kernel density of persistence diagram of (a) NC, (b) sMCI, (c) pMCI, and (d) AD. 27, 26, 25, and 35 holes are plotted by blue dots in (a) NC, (b) sMCI, (c) pMCI, and (d) AD, respectively. Three holes t_1, t_2 and t_3 in (c) pMCI and (d) AD were significant ($p < .05$). t_4 is the hole with the smallest p value in sMCI. t_4 and three abnormal holes are shown in Fig. 3.

4. DISCUSSION AND CONCLUSIONS

The brain regions of NC and sMCI may be well-connected to each other because there was no abnormal hole in NC and sMCI. pMCI and AD had similar abnormal hole where the connections between two large modules, the fronto-subcortical and parieto-occipital regions, were disturbed by right interior, middle, superior temporal gyrus (ITG, MTG, STG), right supplementary motor area (SMA), posterior cingulate cortex (PCC) and paracentral lobule (PCL). The results show that the abnormal hole was found in the functional connectivity when AD severely progressed. By the localization of abnormal holes, we can see how the shape of network after the deterioration of brain. However, further discussion is needed on the biological meaning of abnormal holes. We

also need to investigate how the hole structure will vary with age, gender, apoE, and so on in a group in the future.

5. REFERENCES

[1] E. Bullmore and O. Sporns, "Complex brain networks: graph theoretical analysis of structural and functional systems," *Nature Rev. Neurosci.*, vol. 10, pp. 186–198, 2009.

[2] M. Daianu, N. Jahanshad, T. M. Nir, C. R. Jack Jr., M. W. Weiner, M. A. Bernstein, P. M. Thompson, and the Alzheimer's Disease Neuroimaging Initiative, "Rich club analysis in the Alzheimer's disease connectome reveals a relatively undisturbed structural core network," *Hum Brain Mapp*, vol. 36, no. 8, pp. 3087–3103, 2015.

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

[4] H. Lee, M. K. Chung, H. Kang, B. N. Kim, and D. S. Lee, "Persistent brain network homology from the perspective of dendrogram," *IEEE T. Med. Imaging*, vol. 31, pp. 2267–2277, 2012.

[5] H. Lee, M. K. Chung, H. Kang, and D. S. Lee, "Hole detection in metabolic connectivity of Alzheimer's disease using k-Laplacian," in *MICCAI*, Tokyo, Japan, 2014, vol. 8675, pp. 297–304.

[6] G. Petri, P. Expert, F. Turkheimer, R. Carhart-Harris, D. Nutt, P. J. Hellyer, and F. Vaccarino, "Homological scaffolds of brain functional networks," *Journal of The Royal Society Interface*, vol. 11, no. 101, 2014.

[7] A. Sizemore, C. Giusti, A. Kahn, R. F. Betzel, and D. S. Bassett, "Cliques and cavities in the human connectome," *arXiv:1608.03520*, 2016.

[8] V. de Silva, "Point-cloud topology via harmonic forms," In Workshop on Algorithms for Modern Massive Data Sets, 2006.

[9] A. Zomorodian and G. Carlsson, "Computing persistent homology," *Discrete Comput. Geom.*, vol. 33, pp. 249–274, 2005.

[10] H. Choi and K. H. Jin, "Predicting cognitive decline with deep learning of brain metabolism and amyloid imaging," *arXiv:1704.06033*, 2017.

[11] E. T. Rolls, M. Joliot, and N. Tzourio-Mazoyer, "Implementation of a new parcellation of the orbitofrontal cortex in the automated anatomical labeling atlas," *NeuroImage*, vol. 122, pp. 1–5, 2015.

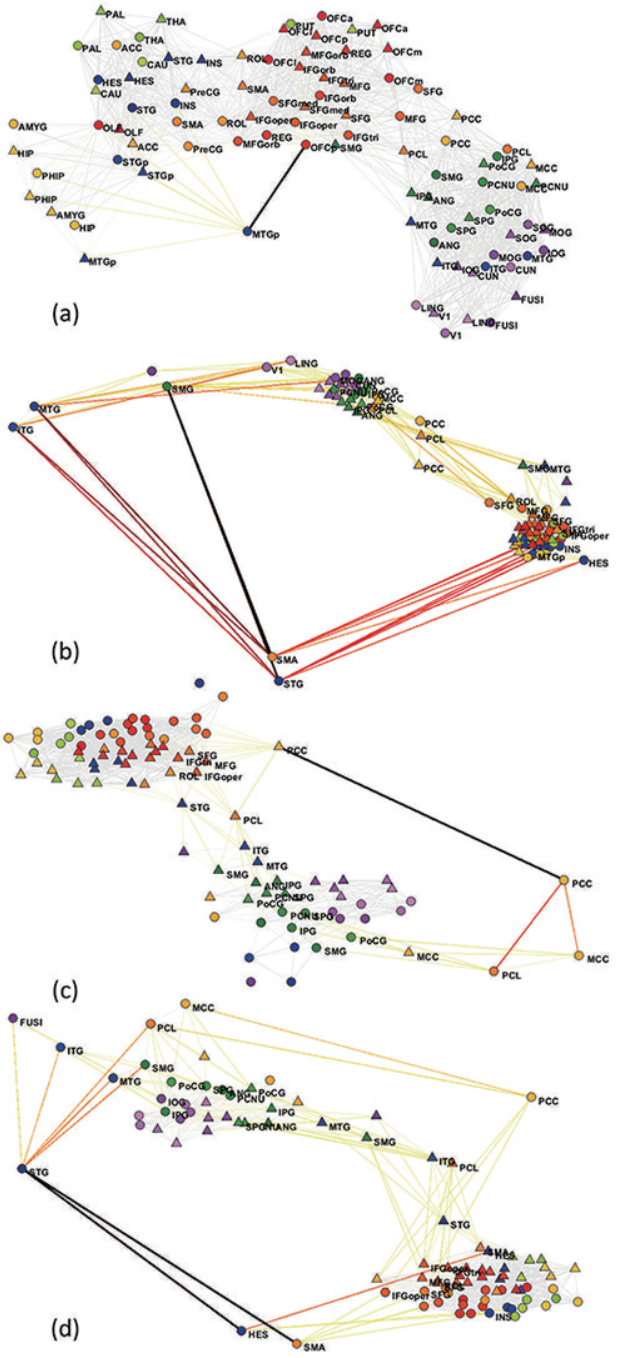

Fig. 3: (a) Hole t_4 in Fig. 2 (b) sMCI, (b) abnormal hole t_1 in Fig. 2 (c) pMCI, and (c,d) abnormal holes t_2 and t_3 in Fig. 2 (d) AD. The node color represents the lobe location (red: frontal, green: parietal, blue: temporal, purple: occipital, yellow: limbic, green: basal ganglia, circle: right, triangle: left hemisphere). The color of edge is determined by the corresponding $|z_i|$ in Sec. 2.4. As $|z_i|$ increases, the edge color is changed from gray through yellow to dark red.