# Network dissimilarity based on harmonic holes

Hyekyoung Lee<sup>1</sup> Moo K. Chung<sup>2</sup> Hongyoon Choi<sup>1</sup> Hyejin Kang<sup>3</sup> Seunggyun Ha<sup>1</sup> Yu Kyeong Kim<sup>4</sup> and Dong Soo Lee<sup>1,3</sup>

- 1) Seoul National University Hospital, Seoul, 03080, KOREA
- 2) University of Wisconsin-Madison, WI, 53706, USA
- 3) Seoul National University College of Medicine, Seoul, 03080, KOREA
- 4) SMG-SNU Borame Medical Center, Seoul, 07061, KOREA

Corresponding Author : Hyekyoung Lee, hklee.brain@gmail.com

### ABSTRACT

Persistent homology has been applied to brain network analysis for finding the shape of brain networks across multiple thresholds. In the persistent homology, the shape of networks is often quantified by the sequence of k-dimensional holes and Betti numbers. The Betti numbers are more widely used than holes themselves in topological brain network analysis. However, the holes show the local connectivity of networks, and they can be very informative features in analysis. In this study, we proposed a new method of measuring network differences based on the dissimilarity measure of harmonic holes (HHs). The HHs, which represents the substructure of brain networks, are extracted by the Hodge Laplacian of brain networks. We also found the most contributed HHs to the network difference based on the HH dissimilarity. In clinical application, the proposed method was applied to clustering the networks of 4 groups, normal control (NC), stable and progressive mild cognitive impairment (MCI), and Alzheimer's disease (AD). The results showed that the clustering performance of the proposed method was better than that of network distances based on only the global change of topology.

#### **NETWORK CONSTRUCTION**

We used FDG PET images in ADNI data set (http://adni.loni.usc.edu). The ADNI FDG-PET dataset consists of 4 groups: 181 NC, 91 sMCI, 77 pMCI, and 135 AD (Age:  $73.7\pm5.9$ , range 56.1–90.1). FDG PET images were preprocessed by using SPM8. The whole brain image was parcellated into 94 regions of interest (ROIs) based on AAL2 excluding cerebellum [3]. The distance between two nodes was estimated by the diffusion distance on positive correlation between the measurements. We constructed 600 bootstrapped networks from 600 bootstrap samples in each group by diffusion distance. The total number of generated brain networks was 2400.

## DISSIMILARITY BETWEEN BRAIN NETWORKS BASED ON HHS

We estimated persistent HHs from each network by the first Hodge Laplacian [1,2]. Each HH is an eigenvector with zero eigenvalue of the Hodge Laplacian. It is denoted by x of which dimension is the number of edges in a network.

Suppose that two networks  $K_a$  and  $K_b$  have m and n persistent HHs, denoted by  $H_a = [x_1^a, \dots, x_m^a]$  and  $H_b = [x_1^b, \dots, x_m^b]$ , respectively. We assume that the difference between networks is determined by the difference of local substructures of networks, and the persistent HHs represent the substructure of a network. Then, the network dissimilarity based on persistent HHs is defined by

$$D_H(K_a, K_b) = \inf_{\psi: H_a \to H_b} \frac{1}{\min(m, n)} \sum_{x \in H_a} d_h(\boldsymbol{x}, \psi(\boldsymbol{x})),$$

where  $d_h(\boldsymbol{x}, \psi(\boldsymbol{x}))$  is the difference between two HHs,  $\boldsymbol{x} \in H_a$  and  $\psi(\boldsymbol{x}) \in H_b$ , estimated by  $= 1 - |\boldsymbol{x}^\top \psi(\boldsymbol{x})|$ . It is the smallest singular value of the matrix  $[\boldsymbol{x} \ \psi(\boldsymbol{x})]$ .

#### **CITATION OF HH**

We can also find which substructures, i.e., HHs, make the difference between two networks or two groups of networks. The most discriminative pair of HHs between two networks are

$$(\boldsymbol{x},\psi(\boldsymbol{x})) = \operatorname{argmax}_{x\in H_a,\psi(x)\in H_b} \inf_{\psi:H_a \to H_b} d_h(\boldsymbol{x},\psi(\boldsymbol{x})).$$

If we compare two groups of networks, denoted by A and B, the citation of a HH  $x \in H$  in a network  $K \in A$  is defined by the ratio of the difference between groups to the difference within a group such that

$$\xi(\boldsymbol{x}) = \frac{\sum_{K_b \in B} \inf_{\psi: H \to H_b} d_h(\boldsymbol{x}, \psi(\boldsymbol{x}))}{\sum_{K_a \in A} \inf_{\psi: H \to H_a} d_h(\boldsymbol{x}, \psi(\boldsymbol{x}))}.$$

The most cited HH is the substructure of a network maximizing the difference between groups and minimizing the difference within a group.

#### RESULTS

The network distance between 2400 brain networks was estimated by (a) L2-norm, (b) Gromov-Hausdorf (GH) distance, (c) Kolmogorov-Smirnov distance (KS) of connected components (KS0), (d) KS of cycles (KS1), (e) bottleneck distance of holes, and (f) HH dissimilarity in Fig. 1 [4–6]. We clustered them into 4 groups and the clustering accuracy was shown in Table 1. The 4 most cited HHs between sMCI and pMCI were shown in Fig. 2.

	Distance	4 groups	2 groups
		(NC, sMCI, pMCI, and AD)	(sMCI and pMCI)
(a)	L2	66.09 %	98.50 %
(b)	GH	45.96 %	87.58 %
(c)	$KS_0$	52.54 %	74.00 %
(d)	$KS_1$	77.38 %	79.83 %
(e)	Bottleneck	45.71 %	76.58 %
(f)	HH	100 %	100 %

Table 1 Clustering accuracy



Figure 1. Distance of 2400 networks. (a) L2, (b) GH, (c)  $KS_0$ , (d)  $KS_1$ , (e) Bottleneck, and (f) HH. The 2400 networks were sorted in the order of NC, sMCI, pMCI, and AD. Each group had 600 networks. The clustering accuracy is shown in Table .



Figure 2. (a) Clustering of the 600 most cited HHs when sMCI and pMCI were compared. (b) Representative HHs in cluster 1, 2, 3 and 4. The left two columns showed HHs in sMCI and the right two columns showed the corresponding HHs in pMCI. Each HH was visualized in a brain and in a 2-dimensional plane. The shape of the HH was more clearly shown in the plane, and the location of the HH could be checked in the brain. The color of nodes was determined by the location of nodes in a brain: frontal (red), parietal (blue), temporal (green), occipital (purple), subcortical (yellow), and limbic (orange) regions. If the edge weight was larger in a HH, the color of edge was darker and the width of edge was larger.

### CONCLUSIONS

We proposed a new network dissimilarity, called HH dissimilarity, and extracted network submodules that discriminated between groups.

### ACKNOWLEDGEMENTS

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.

### REFERENCES

- 1. Kim, Y. -J. and Kook, W., "Harmonic cycles for graphs", *Linear and Multilinear Algebra*, Vol. 2, 2018, pp. 1-11.
- 2. Lee, H., Chung, M. K., Choi, H., Kang, H., Ha, S., Kim, Y. K. and Lee, D. S., "Harmonic holes as the submodules of brain network and network dissimilarity", *Workshop on Computational Topology in Image Context, Lecture Notes in Computer Science*, Vol. 11382, pp 110-122, 2019.
- 3. Rolls, E. T., Joliot, M. and Tzourio-Mazoyer, N., "Implementation of a new parcellation of the orbitofrontal cortex in the automated anatomical labeling atlas", *NeuroImage*, Vol. 122, 2015, pp. 1-5.
- 4. Carlsson, G. and Mémoli, F., "Persistent Clustering and a Theorem of J. Kleinberg", *ArXiv e-prints*, 2008.
- 5. Chung, M. K., Lee, H., Gritsenko, A., DiChristofano, A., Pluta, D., Ombao, H. and Solo, V., "Topological Brain Network Distances", *arXiv:1809.03878*, 2018.
- 6. Cohen-Steiner, D., Edelsbrunner, H. and Harer, J., "Stability of Persistence Diagrams", *Discrete Comput. Geom.*, Vol. 37, 2007, pp. 103-120.