

HODGE-DECOMPOSITION OF BRAIN NETWORKS

D Vijay Anand¹, Moo K. Chung²

¹ University College London, London, UK

² University of Wisconsin, Madison, USA

ucahvad@ucl.ac.uk.edu, mkchung@wisc.edu

ABSTRACT

We analyze brain networks by decomposing them into three orthogonal components through the Hodge decomposition, a technique advantageous for capturing complex topological features. We quantify the magnitude and relative strength of each component through simulation studies with known ground truths. The method is then applied to human brain networks obtained from resting-state functional magnetic resonance imaging studies. Our results indicate statistically significant differences in the topological features between male and female brain networks.

1. INTRODUCTION

Traditional graph models employed in functional magnetic resonance imaging (fMRI) studies have predominantly focused on capturing pairwise interactions between brain regions. This focus often results in the neglect of higher-order interactions, which are crucial for a comprehensive understanding of brain network topology [1]. While the significance of higher-order interactions has been increasingly acknowledged, research in this direction has been constrained by challenges such as computational complexity and the lack of effective analytical tools. To overcome these challenges, topological data analysis (TDA) has emerged as a promising technique. Persistent homology (PH), a specialized technique within TDA, leverages mathematical constructs, particularly simplicial complexes, to encode higher-order interactions in a systematic and computationally efficient manner [2, 3]. Simplicial complexes consist of simplices, each corresponding to a different level of interaction within the network. Specifically, nodes represent 0-way interactions, edges correspond to 1-way interactions, triangles denote 2-way interactions, and tetrahedra signify 3-way interactions. PH allows for a nuanced representation of networks across different spatial resolutions, thereby enriching brain imaging data [1, 4].

Persistent Homology (PH) quantifies multiscale topological features of data using topological invariants through a filtration process [5]. Hodge theory provides a unified framework combining simplicial homology and spectral geometry, offering insights into network topology [6–8]. While the

Hodge Laplacian, a generalization of the graph Laplacian, offers insights into the topological features of higher order simplices, the Hodge decomposition allows to establish relationships between simplices of different dimensions [7]. Hodge decomposition breaks data defined on edges (edge flow) into three orthogonal components: gradient, curl, and harmonic flows, each providing unique topological insights. The gradient flow, driven by node gradients, represents the network's gradient-like behavior. The curl flow, arising from triangle-induced flows, captures rotational patterns, while the harmonic flow exposes loop structures and topological signatures [7]. Using a Wasserstein distance-based statistical approach on each component, this study assesses the topological similarities and differences between loop and non-loop flows. Further, leveraging on the properties of the decomposed networks, the study seeks to elucidate the most discriminating topological disparities in female and male functional brain networks.

2. METHOD

2.1. Boundary and coboundary operators

A simplicial complex is a collection of simplices that includes nodes (0-simplices), edges (1-simplices), triangles (2-simplices) and their higher-dimensional counterparts. The 0-skeleton of a simplicial complex consists only of nodes, while a 1-skeleton comprises both nodes and edges. Graphs are examples of 1-skeletons. A k -chain is a formal linear combination of k -simplices. The set of all such k -chains constitutes a group, denoted as \mathcal{K}_k . A sequence of these groups forms a chain complex. To relate different chain groups, boundary operators are used.

For two successive chain groups, \mathcal{K}_k and \mathcal{K}_{k-1} , the boundary operator $\partial_k : \mathcal{K}_k \rightarrow \mathcal{K}_{k-1}$ for a given k -simplex σ_k is defined as

$$\partial_k(\sigma_k) = \sum_{i=0}^k (-1)^i(v_0, \dots, \hat{v}_i, \dots, v_k), \quad (1)$$

where $(v_0, \dots, \hat{v}_i, \dots, v_k)$ represents the $(k-1)$ -faces of σ_k , obtained by omitting the vertex \hat{v}_i . The boundary operator's

matrix representation, $\mathbb{B}_k = (\mathbb{B}_k^{ij})$, is given by

$$\mathbb{B}_k^{ij} = \begin{cases} 1, & \text{if } \sigma_{k-1}^i \subset \sigma_k^j \text{ and } \sigma_{k-1}^i \sim \sigma_k^j, \\ -1, & \text{if } \sigma_{k-1}^i \subset \sigma_k^j \text{ and } \sigma_{k-1}^i \not\sim \sigma_k^j, \\ 0, & \text{otherwise,} \end{cases} \quad (2)$$

where \sim and $\not\sim$ indicate similar and dissimilar orientations, respectively [9, 10].

Coboundary operators δ_k are duals of the boundary operators, mapping k -cochains to $(k+1)$ -cochains: $\delta_k : \mathcal{K}^k \rightarrow \mathcal{K}^{k+1}$. The coboundary operator is the adjoint of the boundary operator ∂_{k+1} . The matrix representation of δ_k is the transpose of the matrix representation of ∂_{k+1} : $\delta_k = \mathbb{B}_{k+1}^\top$.

2.2. Hodge Decomposition

Let \mathcal{C}^k be the space of functions over k -simplices. The Hodge decomposition separates an edge flow $X \in \mathcal{C}^1$ into three orthogonal components: gradient X_G , curl X_C and harmonic flows X_H :

$$X = X_G + X_C + X_H = \delta_0 s + \delta_1^T \phi + X_H \quad (3)$$

with potential functions $s \in \mathcal{C}^0$ and $\phi \in \mathcal{C}^2$. The components are determined through orthogonal projections onto their respective subspaces. X_G represents the function on the k -chain that can be described as the gradient of a scalar field. Similarly, X_C captures the component that behaves like the curl of a vector field, and X_H denotes the harmonic component, which is divergence-free and curl-free. X_G and X_C are obtained by minimizing the residual in the projection as

$$X_G = \min_{s \in \mathcal{C}^0} \|X - \delta_0 s\|, \quad X_C = \min_{\phi \in \mathcal{C}^2} \|X - \delta_1^T \phi\|. \quad (4)$$

The harmonic component as $X_H = X - (X_G + X_C)$ is obtained as the residual. The sum of the curl and harmonic component forms the loop flow while the gradient component is referred as the non-loop flow. Figure 1 illustrates the Hodge decomposition from a complete graph.

2.3. Topological Inference on the Hodge Decomposition

To measure topological distance between graphs, we employ the *birth-death decomposition* (BDD), which partition graphs into topologically distinct subgraphs [11, 12]. We first apply graph filtration, a technique involving the sequential removal of edges from a graph G , starting with the smallest edge weight and progressing to the largest [3, 5]. We identify the birth set $B(G)$, associated with the emergence of connected components, by computing the maximum spanning tree (MST) of G using Kruskal's or Prim's algorithms [3]. The death set $D(G)$ then consists of the edges not present in $B(G)$ (Figure 1), which consists of death values of cycles (loops) during the filtration. We perform BDD independently on both non-loop and loop flows, allowing us to characterize

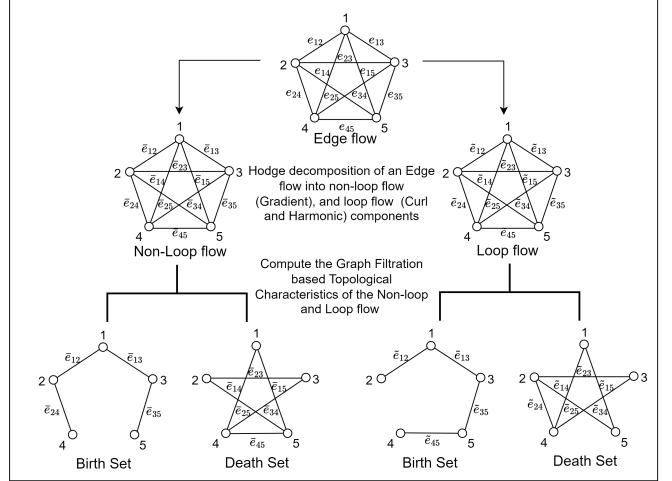


Fig. 1: Illustration of the Hodge decomposition, which decomposes the edge flow into non-loop and loop flows. These networks are then separately subjected to birth-death decomposition to obtain the topological features.

the topology of each component of the Hodge decomposition. To measure the topological disparities between components, we use the Wasserstein distance applied to their respective BDD.

We evaluate the difference between two groups of networks $\Omega = \{\Omega_1, \Omega_2, \dots, \Omega_m\}$ and $\Psi = \{\Psi_1, \Psi_2, \dots, \Psi_n\}$. The Wasserstein distance is used as a test statistic [13, 14]:

$$\begin{aligned} \mathfrak{L}_\infty(\Omega, \Psi) &= \mathfrak{L}_\infty^b(\Omega, \Psi) + \mathfrak{L}_\infty^d(\Omega, \Psi) \\ &= \max_{1 \leq j \leq q_0} |\bar{b}_j^\Omega - \bar{b}_j^\Psi| + \max_{1 \leq j \leq q_1} |\bar{d}_j^\Omega - \bar{d}_j^\Psi|. \end{aligned} \quad (5)$$

$\mathfrak{L}_\infty^b(\Omega, \Psi)$ computes the ∞ -Wasserstein distance, where \bar{b}_j^Ω and \bar{b}_j^Ψ are the means of the j -th smallest birth values of connected components in Ω and Ψ . $\mathfrak{L}_\infty^d(\Omega, \Psi)$ computes the ∞ -Wasserstein distance, where \bar{d}_j^Ω and \bar{d}_j^Ψ are the means of the j -th smallest death values of cycles in Ω and Ψ [15]. q_0 and q_1 are the number of birth and death values respectively. Under the null hypothesis of topological equivalence between the two groups, we expect $\mathfrak{L}_\infty(\Omega, \Psi)$ to be close to zero. Deviations from this value would suggest a topological discrepancy between Ω and Ψ . Given that the null distribution of $\mathfrak{L}_\infty(\Omega, \Psi)$ is not pre-established, we propose to approximate it through a permutation test, from which we subsequently derive the p -value.

3. VALIDATIONS

In our simulations, we generate random modular networks, with edge weights drawn from a Beta distribution [15]. The Beta distribution $Beta(\alpha, \beta)$ is defined on the interval $[0, 1]$ and is parameterized by positive shape parameters α and β . These parameters allows us to create networks with varying

Table 1: The performance results of the bottleneck and Gromov-Hausdorff (GH) distances compared against the Wasserstein distance on the non-loop and loop components. Smaller p -values are better when there are network differences (top rows) and larger p -values are better when there are no network differences (bottom rows).

Nodes p	Modules c	Graph distances			
		Bottleneck	GH	nonloop	loop
12 vs. 12	2 vs. 3	0.7056	0.7786	0.0000	0.0081
	3 vs. 6	0.5589	0.8816	0.0077	0.0000
18 vs. 18	2 vs. 3	0.3488	0.7822	0.0087	0.0078
	3 vs. 6	0.5596	0.7269	0.0000	0.0000
24 vs. 24	2 vs. 3	0.1519	0.1575	0.0075	0.0080
	3 vs. 6	0.8153	0.2625	0.0085	0.0000
	2 vs. 2	0.8144	0.6885	0.7322	0.3500
24 vs. 24	3 vs. 3	0.9397	0.2398	0.9923	0.2228
	6 vs. 6	0.1337	0.3578	0.9906	0.7718

strengths of connectivity for comparison of their topological similarity and dissimilarity. We used $(\alpha, \beta) = (2, 4), (4, 2)$ to generate random networks (Figure 2). To construct a modular network, we assigned edge weights within the same module from $\text{Beta}(\alpha, \beta)$ where $\alpha > \beta$, and edge weights between different modules from $\text{Beta}(\beta, \alpha)$. We set the number of nodes $p = 12, 18, 24$, and the number of modules $c = 2, 3, 6$, ensuring an even distribution of nodes among the modules. We employed the Wasserstein distance proposed in Section 2.3 to evaluate the topological similarity and dissimilarity in a two-group comparison setting, where each group consists of 10 networks each. We compared our method against the bottleneck [5] and Gromov-Hausdorff (GH) [3, 16] distances, which have been adapted in brain network analysis. Statistical inferences were done using a permutation test with 100,000 permutations, where the average network of each group in each permutation served as the input for computing the bottleneck and GH distances. We repeated the simulations independently 10 times, reporting the average p -value in Table 1, where top rows test false negatives while the bottom rows test for false positives.

The proposed ∞ -Wasserstein distance-based test statistic exhibits robust performance on both the loop and non-loop flows. The bottleneck and GH distances demonstrated suboptimal performance for dissimilar networks, whereas \mathcal{L}_∞ effectively discriminated networks in both the non-loop (gradient) and loop (curl) components when network differences were present (as shown in the top rows of the table). In scenarios with no network differences (bottom rows of the table), all distances yielded satisfactory results. This underscores that the modularity in the network is aptly captured by both the non-loop and loop components of the Hodge decomposition, and that our ∞ -Wasserstein distance is capable of discerning variations in modularity.

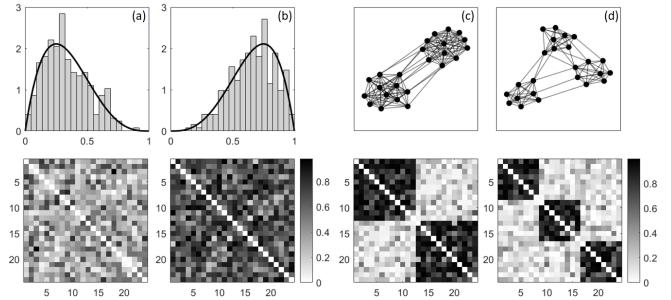


Fig. 2: Edge weights following the Beta distributions with parameters (a) $(\alpha = 2, \beta = 4)$ (b) $(\alpha = 4, \beta = 2)$ with their corresponding connectivity matrices. The modular graphs obtained using the Beta distributions with (c) two and (d) three modules. The networks are thresholded at 0.4 to enable better display of the modules.

4. APPLICATION

4.1. Functional brain imaging data and preprocessing

We used the resting-state fMRI (rs-fMRI) in the Human Connectome Project [17]. rs-fMRI are collected at 2 mm isotropic voxels and 1200 time points. Data that was subjected to the standard minimal preprocessing pipelines [18] was used. Volumes with framewise displacement larger than 0.5mm and their neighbors were scrubbed [17]. Excessive head movement were excluded from the study. Subsequently, the Automated Anatomical Labeling (AAL) template is used to parcelate and average rs-fMRI spatially into 116 non-overlapping anatomical regions. The details on image processing is given in [19]. The final data is comprised of the fMRI of 400 subjects of which 168 are males and 232 are females.

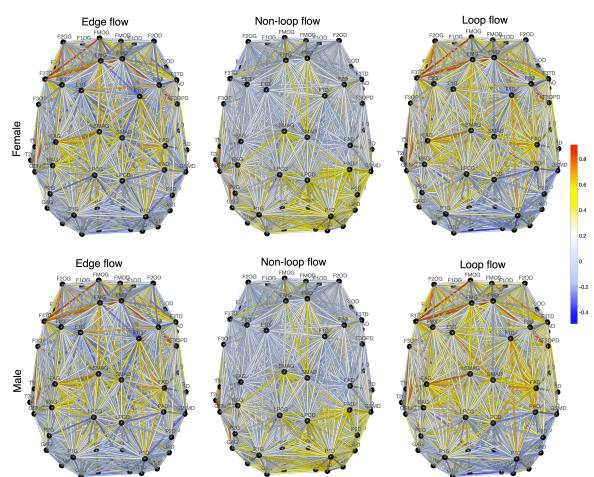


Fig. 3: The average connectivity (edge flow), non-loop flow (middle) and the loop flow (right) of the female (top) and male networks (bottom).

4.2. Hodge decomposition of the brain network

Following the proposed method, we decomposed individual brain networks using the Hodge decomposition. Figure 3 displays the average decomposition of 400 subjects. We first assessed if there are topological difference between females and males in the original connectivity (edge flow). Following the test procedure outlined in Section 2.3, the Wasserstein distances \mathcal{L}_∞^b on birth values for testing 0D topology difference and \mathcal{L}_∞^d on death values for testing 1D topology difference are separately used. We performed the permutation test for the birth set (first term) and death set (second term) separately and obtained the p -values 0.0177 and 0.0110 respectively showing there are topological difference present in the original connectivity matrices.

We further determined if we can detect topological differences in the decomposed components. There are several interesting properties of the Hodge decomposition that motivate us to study each component. The gradient component sum to zero along any cycles. The curl components are zero for edges that are not a 2-simplex boundary and the entries sum to zero around each node. The harmonic component sums to zero around each node, and it also sums to zero along each 2-simplex. We tested the topological equivalence of female brain networks and male brain networks using the Wasserstein distance based test statistic (5). We first considered the network constructed from the gradient component and performed the birth death decomposition. We then carried out the permutation test (p -value = 0.008). Also we performed the permutation test on the curl component (p -value = 0.0296). We conclude that the non-loop and loop components were able to detect the group difference. We further tested if we can detect 0D and 1D topological signals separately in each component. The Wasserstein distances \mathcal{L}_∞^b and \mathcal{L}_∞^d were separately used in each component. For non-loop component, the permutation test gave the p -values of 0.0088 and 0.0080 for the birth (0D topology) and and death (1D topology) values respectively. For loop component, the permutation test gave the p -values of 0.0019 and 0.1582 for the birth (0D topology) and death (1D topology) values respectively. We conclude that each composition can be used for detecting 0D and 1D topological signals.

5. CONCLUSION

This study detailed the use of Hodge decomposition—gradient, curl, and harmonic components—in topological data analysis of brain networks. Applying this technique to rs-fMRI correlation networks, we subsequently categorized these components into loop and non-loop flows. By incorporating Wasserstein distance with birth and death values of topological features, we differentiated network types, showcasing the components' ability to detect topological differences. The maximum gap statistic based on the Wasserstein distance offers a

simplified summary p -value , avoiding multiple testing issues and quantifying network differences. An advantage of the Wasserstein distance-based maximum gap statistic is the circumvention of multiple comparison corrections and provides a singular summary p -value, serving as a quantifiable measure of network differentiation. Tested on static functional brain networks, our method discriminated between male and female brain networks, with potential for extension to dynamic network analysis and exploration of time-dependent topological changes as future work.

6. ACKNOWLEDGEMENT

This study is funded by NIH EB028753, MH133614 and NSF DMS-2010778.

7. REFERENCES

- [1] A.E. Sizemore, J.E. Phillips, R. Ghrist, and D. S. Bassett, “The importance of the whole: topological data analysis for the network neuroscientist,” *Network Neuroscience*, vol. 3, no. 3, pp. 656–673, 2019.
- [2] C. Giusti, R. Ghrist, and D.S. Bassett, “Two’s company, three (or more) is a simplex,” *Journal of Computational Neuroscience*, vol. 41, no. 1, pp. 1–14, 2016.
- [3] H. Lee, M.K. Chung, H. Kang, B. Kim, and D.S. Lee, “Computing the shape of brain networks using graph filtration and Gromov-Hausdorff metric,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2011, pp. 302–309.
- [4] M.K. Chung, H. Lee, A. DiChristofano, H. Ombao, and V. Solo, “Exact topological inference of the resting-state brain networks in twins,” *Network Neuroscience*, vol. 3, pp. 674–694, 2019.
- [5] M.K. Chung, H. Lee, V. Solo, R.J. Davidson, and S.D. Pollak, “Topological distances between brain networks,” *International Workshop on Connectomics in Neuroimaging*, vol. 10511, pp. 161–170, 2017.
- [6] L.H. Lim, “Hodge Laplacians on graphs,” *SIAM Review*, vol. 62, no. 3, pp. 685–715, 2020.
- [7] S. Barbarossa and S. Sardellitti, “Topological signal processing over simplicial complexes,” *IEEE Transactions on Signal Processing*, vol. 68, pp. 2992–3007, 2020.
- [8] D.V. Anand and M.K. Chung, “Hodge-Laplacian of brain networks,” *IEEE Transactions on Medical Imaging*, vol. 42, pp. 1563–1473, 2023.

- [9] H. Edelsbrunner and J. Harer, *Computational topology: An introduction*, American Mathematical Society, 2010.
- [10] J. Huang, M.K. Chung, and A. Qiu, “Heterogeneous graph convolutional neural network via hodge-laplacian for brain functional data,” in *International Conference on Information Processing in Medical Imaging*. Springer, 2023, pp. 278–290.
- [11] T. Songdechakraiut, L. Shen, and M.K. Chung, “Topological learning and its application to multimodal brain network integration,” *Medical Image Computing and Computer Assisted Intervention (MICCAI)*, vol. 12902, pp. 166–176, 2021.
- [12] T. Songdechakraiut and M.K. Chung, “Topological learning for brain networks,” *Annals of Applied Statistics*, vol. 17, pp. 403–433, 2023.
- [13] Z. Su, W. Zeng, Y. Wang, Z.-L. Lu, and X. Gu, “Shape classification using wasserstein distance for brain morphometry analysis,” in *International Conference on Information Processing in Medical Imaging*. Springer, 2015, pp. 411–423.
- [14] J. Shi and Y. Wang, “Hyperbolic wasserstein distance for shape indexing,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, pp. 1362–1376, 2019.
- [15] S. Das, D.V. Anand, and M.K. Chung, “Topological data analysis of human brain networks through order statistics,” *PLOS One*, vol. 18, no. 3, pp. e0276419.
- [16] H. Lee, H. Kang, M.K. Chung, B.-N. Kim, and D.S. Lee, “Persistent brain network homology from the perspective of dendrogram,” *IEEE Transactions on Medical Imaging*, vol. 31, pp. 2267–2277, 2012.
- [17] D.C. Van Essen, K. Ugurbil, E. Auerbach, D. Barch, T.E.J. Behrens, R. Bucholz, A. Chang, L. Chen, M. Corbetta, and S.W. Curtiss, “The human connectome project: a data acquisition perspective,” *NeuroImage*, vol. 62, pp. 2222–2231, 2012.
- [18] M.F. Glasser, S.M. Smith, D.S. Marcus, J.L.R. Andersson, E.J. Auerbach, T.E.J. Behrens, T.S. Coalson, M.P. Harms, M. Jenkinson, and S. Moeller, “The human connectome project’s neuroimaging approach,” *Nature Neuroscience*, vol. 19, pp. 1175, 2016.
- [19] S.-G. Huang, S.-T. Samdin, C.M. Ting, H. Ombao, and M.K. Chung, “Statistical model for dynamically-changing correlation matrices with application to brain connectivity,” *Journal of Neuroscience Methods*, vol. 331, pp. 108480, 2020.