Hodge Laplacian of Brain Networks

D. V. Anand, Sixtus Dakurah and Moo K. Chung

***Abstract*— The closed loops or cycles in a brain network embeds higher order signal transmission paths, which provide fundamental insights into the functioning of the brain. In this work, we propose an efficient algorithm for systematic identification and modeling of cycles using persistent homology and the Hodge Laplacian. Various statistical inference procedures on cycles are developed. We validate the our methods on simulations and apply to brain networks obtained through the resting state functional magnetic resonance imaging. The computer codes for the Hodge Laplacian are given in `https://github.com/laplcebeltrami/hodge`.**

***Index Terms*— Hodge Laplacian, Wasserstein distance, brain networks, Cycle basis, Heat kernel smoothing**

I. Introduction

Understanding the collective dynamics of brain networks has been a long standing question and continues to remain elusive. Many symptoms of the brain diseases such as schizophrenia, epilepsy, autism, and Alzheimer's disease (AD) have shown possible connections with abnormally high levels of synchrony in neural activity [1]. The mechanisms underlying the emergence of this synchronous behaviour, is often attributed to the higher order interactions that occur at multiple topological scales [2], [3]. The higher order interactions are evidenced across multiple spatial scales in neuroscience such as collective firing of neurons [1], simultaneous activation of multiple brain regions during cognitive tasks [4]. The consideration of higher-order interactions can be highly informative for understanding neuronal synchronisation and co-activation of brain areas at different scales of the network [5].

Over the past several decades, significant progress has been made in understanding the structural and functional behavior of the human brain using functional magnetic resonance images (fMRI). In typical fMRI network studies, the brain is usually modelled as a graph whose nodes are specific brain regions and their connectivity is determined by the strength of dependency between brain regions. Often graph theory based methods have been applied to analyze the brain networks using quantitative measures such as centrality, modularity and small-worldness [6]–[10], which allows to interpret and understand the spatial and functional organization of the brain. Besides, graph measures also provide reliable and quantifiable biomarkers that can discriminate normal and clinical populations [11]. Hence, the graph measures are used to identify and quantify the differences in the functional networks at both the individual and group level [6]. The graph comparisons are often performed in the form of either distance-based comparisons or statistics applied theory features [6], [12], [13].

Although graph-based methods can be used to identify graph attributes at disparate scales ranging from local scales at the node level up to global scales at the community level, their power is limited to mostly pairwise dyadic relations [8]. The inherent *dyadic* assumption limits the types of neural structure and function that the graphs can model [14], [15]. Therefore, brain network models built on top of graphs cannot encode higher order interactions, i.e., three- and four-way interactions, beyond pairwise connectivity *without* additional analysis [16]. Despite these limitations the graph-based approaches were often used in brain network analysis [17]. To overcome these limitations, we propose to use topological data analysis (TDA). The TDA has gained a lot of traction in recent years due to its simplistic construct in systematically extracting information from hierarchical layers of abstraction [18]. The algebraic topology in TDA has mathematical ingredients that can effectively manipulate structures with higher order relations. One such tool is the *simplicial complex* which captures many body interactions in complex networks using basic building blocks called simplices [14]. The simplicial complex representation easily encode higher order interactions by the inclusion of 2-simplices (faces) and 3-simplices (volumes) to graphs. We can further adaptively increases the complexity of connectivity hierarchically from simple node-to-node interaction to more complex higher order connectivity patterns easily. Simplicial complexes have been used to represent and analyse the brain data [14], [19], [20]. The modular structure of network can easily be recognized by means of connected components, which is the first topological invariant that characterizes the shape of the network [18]. The cycle on the other hand is a second topological invariant which are loops in the network [21]–[23].

The persistent homology (PH), main TDA technique deeply rooted in simplicial complexes, enables network representation at different spatial resolution and provides a coherent framework for obtaining higher order topological features [18], [24]. The PH based approaches are becoming increasingly popular to understand the brain imaging data [13], [21], [25]. The main approach of PH applied to brain networks is to generate a series of nested networks over every possible parameter through a filtration [26]. In particular, the *graph filtration* is the most often used filtration specifically designed to uncover the hierarchical structure of the brain networks in a sequential manner [23].

This paragraph of the first footnote will contain the date on which you submitted your paper for review. This study is funded by NIHR01 EB022856, EB02875, NSF MDS-2010778. Correspondence: Moo K. Chung.

D. V. Anand. Sixtus Dakurah and Moo K. Chung are with the Department of Biostatistics and Medical Informatics, University of Wisconsin–Madison, Madison, WI 53706 USA (e-mail: mkchung@wisc.edu).

Topology-based comparison methods infer the similarity and dissimilarity of networks based on PH feature summaries such as persistent diagrams and persistent landscapes [21], [27]–[29]. Typically, a topological discriminating function acts on these PH summaries to discern their topological similarity or dissimilarity [21], [24], [27], [28]. The common network distances in the literature for comparing brain networks are the Gromov-Hausdorff (GH) distance and bottle-neck (BN) distances [13]. The GH and BN distances are regarded as PH-based distances since they can naturally act on PH feature summaries [13], [24].

In the last two decades, the persistent homology techniques have made significant inroads in neuroimaging analysis particularly for uncovering global topological features beyond pairwise interactions [25]. These global features are the topological invariants such as number of connected components, number of cycles or holes in a network [30], [31]. Traditional persistent homology based methodologies in neuroimaging have mostly focussed on using these topological invariants as biomarkers for identifying and characterising the topological disparities between the control and diseased populations [22], [32]. While the connected structures of the brain network has been extensively investigated, the studies on the cycles in modeling brain networks is very limited [2], [8], [22], [23], [28]. The presence of more cycles in a network signifies a dense connection with stronger connectivity. The cycles in the brain network not only determines the propagation of information but also controls the feedback [33]. Since the information transfer through cycles can occur in two different paths, it is sometimes interpreted as redundant connections. Further, it is also associated with the information diffusion, dissemination, redundancy and information bottleneck problems [34]–[36].

While cycles appear naturally in networks, it is not easy to extract or enumerate them. The cycles are often computed using brute-force depth-first search algorithms [37]. Recently, a scalable algorithm for computing the number of cycles in the network was proposed [28]. The cycle or holes is usually identified by manipulating the boundary matrix in the persistent homology [24], [38]. A better approach to determine cycles is by computing the eigenvectors corresponding to zero eigenvalues of Hodge Laplacian [22]. This approach generalizes graph Laplacian (0-Laplacian) applied to nodes (0-simplices) to higher order simplexes [39]. Although these algorithms are useful to extract cycles in small networks, it is computationally not feasible to construct and manipulate higher order simplices and extract cycles for large networks. Ideally, we need algorithms that can capture the essence of higher order interactions and yet retain the simplicity of graph-based approaches.

We propose a new spectral method using the Hodge Laplacian that can explicitly identify the connections associated with the cycles. The method is further capable of localizing the connections contributing to the difference and extract the most discriminative cycles in a network. This is made possible by computing the independent cycle basis and then subsequently building a new statistical inference framework that identifies the most discriminating cycles. To the best of our knowledge there is no efficient algorithm in literature to extract and quantify cycles from brain networks. For the numerical implementation, we propose an efficient new algorithm based on the birth death decomposition of graphs [27].

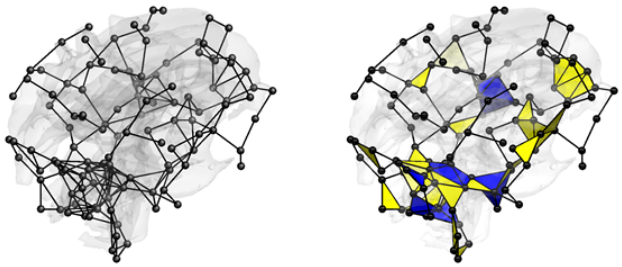

Fig. 1: (a) Illustration of brain network representation using graph and simplicial complex. The graph (left) has only nodes and edges. The simplicial complex (right) shows higher dimensional objects such as triangles (yellow) and tetrahedrons(blue) in addition to nodes and edges.

II. Method

In this section, a detailed explanation on topological data analysis tools such as simplicial complexes, birth death decomposition, Hodge Laplacian over simplicial complexes and the algebraic representation of cycles is presented.

A. Graphs as a simplicial complex

1) Simplicial complex: Consider an undirected complete graph $G = (V, w)$ with vertex set V and edge weight matrix $w = (w_{ij})$ [6], [13]. We assume there are p number of nodes. A binary graph $G_\epsilon = (V, w_\epsilon)$ is a graph consisting of the node set V and the binary edge weights $w_\epsilon = (w_{\epsilon,ij})$ given by

$$w_{\epsilon,ij} = \begin{cases} 1 \text{ if } w_{ij} > \epsilon, \\ 0 \text{ otherwise.} \end{cases} \tag{1}$$

Denote E_ϵ the edge set consisting of all the edges with nonzero weights. Then we may also represent the binary graph G_ϵ as $G_\epsilon = (V, E_\epsilon)$ if there is no confusion.

A p-simplex $\sigma_p = [v_0, v_1, \cdots, v_p]$ is the convex hull of $p+1$ algebraically independent points $v_0, v_1, \cdots, v_p$. A simplicial complex is a collection of simplices such as nodes ($0-$simplices), edges (1-simplices), triangles (2-simplices), a tetrahedron (3-simplices) and higher dimensional counterparts. A simplicial complex can be viewed as the higher dimensional generalization of a graph [24]. Figure 1 illustrates the difference between graphs and simplical complexes in representing a brain network.

2) Chain complex: A p-chain is a sum of p-simplices in K denoted as $c = \sum_i \alpha_i \sigma_i$, where σ_i are the p-simplices and the α_i are either 0 or 1 [40]. The collection of p-chains forms a group and the sequence of these groups is called a chain complex. To relate chain groups, we denote a boundary operator $\partial_p : C_p \to C_{p-1}$, where C_p denotes the p-th chain group. For an oriented p-simplex σ_p with the ordered vertex

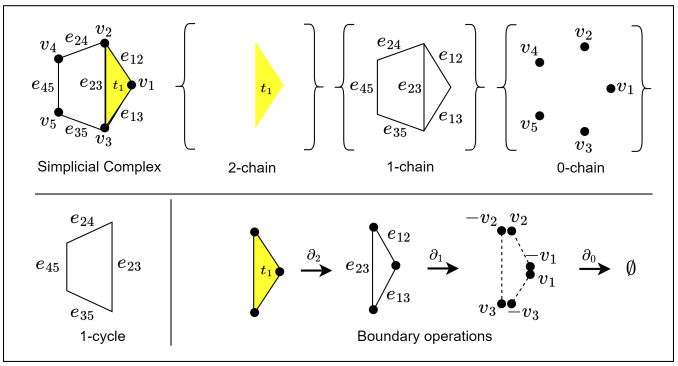

Fig. 2: Top left: A simplicial complex with five vertices (0-simplex), six edges (1-simplex) and a triangle (2-simplex). The triangle is represented by $t_1 = [v_1, v_2, v_3]$ with a filled-in face (colored yellow). Top left: Illustration of chain complex showing 2-chain (set of triangles), 1-chain (set of edges) and 0-chain (set of nodes). Bottom left: The 1-cycle which is present in the simplicial complex. Bottom right: A sequence of boundary operations applied to t_1. After boundary operation ∂_2, we get the 1-simplices $[v_1, v_2] + [v_2, v_3] - [v_1, v_3] = e_{12} + e_{23} - e_{13}$ which is the boundary of the triangle [22], [25].

set, the boundary operator is defined as

$$\partial_p \sigma_p = \sum_{i=0}^{p} (-1)^i [v_0, v_1, \cdots, \widehat{v}_i, \cdots, v_p],$$

where $[v_0, v_1, \cdots, \widehat{v}_i, \cdots, v_p]$ is a $(p-1)$-simplex generated from $\sigma_p = [v_0, v_1, \cdots, v_p]$ excluding $\widehat{v}_i$. The boundary operator maps a simplex to its boundaries. Thus, $\partial_2 \sigma_2$ maps a triangle to its three edges. We can algebraically show that [24]

$$\partial_{p-1} \partial_p \sigma_p = 0.$$

Figure 2 displays a toy example of a simplicial complex with five vertices (0-simplex), six edges (1-simplex) and a triangle (2-simplex). The triangle is represented by $t_1 = [v_1, v_2, v_3]$ with a filled-in face (colored yellow). A schematic representation of chain complex showing 2-chain (set of triangles), 1-chain (set of edges) and 0-chain (set of nodes) is shown on the top right. On the bottom left is the 1-cycle present in the simplicial complex and on the bottom right, a sequence of boundary operations is applied to t_1. After boundary operation ∂_2, we get the 1-simplices $[v_1, v_2] + [v_2, v_3] - [v_1, v_3] = e_{12} + e_{23} - e_{13}$ which is the boundary of the triangle [22], [25].

3) Cycles: A p-cycle is a p-chain whose boundary is zero. In a graph (1-skeleton), 1-cycles are loops and 0-cycles are the number of nodes. To compute p-cycles, we use the kernel and image for the boundary operator and establish their relation to the p-cycle [24], [41]. Let Z_p be the collection of all the p-cycles given by

$$Z_p = ker \partial_p = \{\sigma_p \in C_p | \partial_p \sigma_p = 0\}.$$

Let B_p be the boundaries obtained as

$$B_p = img \partial_{p+1} = \{\sigma_p \in C_p | \sigma_p = \partial_{p+1} \sigma_{p+1}, \sigma_{p+1} \in C_{p+1}\}.$$

Since any boundary $\partial_{p+1} \sigma_{p+1} \in B_p$ satisfies $\partial_p \partial_{p+1} \sigma_{p+1} = 0$, it is a p-cycle and $B_p \subset Z_p$. Thus, we can partition Z_p into cycles that differ from each other by boundaries through the quotient space

$$H_p = Z_p / B_p,$$

which is called the p-th homology group. The p-th Betti number β_p counts the number of algebraically independent p-cycles, i.e.,

$$\beta_p = rank H_p = rank Z_p - rank B_p.$$

In graph G, which is 1-skeleton, Betti numbers $\beta_0(G)$ and $\beta_1(G)$ counts the number of connected components (0-cycles) and number of loops (1-cycles) respectively. Betti numbers other than β_0 and β_1 are all zero in graphs.

4) Birth-death decomposition: The graph filtration of G is defined as a sequence of nested binary networks [13], [23]:

$$G_{\epsilon_0} \supset G_{\epsilon_1} \supset \cdots \supset G_{\epsilon_k}$$

where $\epsilon_0 < \epsilon_1 < \cdots < \epsilon_k$ are the sorted edge weights [13], [23]. The birth and death of k-cycles during the process of filtration is quantified using *persistence*, which is the duration of filtration values from birth to death. The persistence is usually represented as 1D intervals as persistent barcode (PB) or 2D scatter points as a persistent diagram (PD) [24].

During the filtration, once a component is born, it does not die. All the death values of connected components are ∞ and can be ignored. Then the total number $\mathcal{P}$ of birth values of connected components (0-cyles) is

$$\mathcal{P} = \beta_0(G_\infty) - 1 = p - 1. \quad (2)$$

The 0D barcode corresponding to 0-cycles consists of a set of increasing birth values

$$B(G) = b_1 < b_2 < \cdots < b_{\mathcal{P}}.$$

During the filtration, cycle is assumed to be born at $-\infty$. All the birth values of 1-cycles can be ignored. Thus, we have $\mathcal{Q} = (p-1)(p-2)/2$ number of death values of 1-cycles. The 1D barcode corresponding to 1-cycles consists of a set of increasing death values

$$D(G) = d_1 < d_2 < \cdots < d_{\mathcal{Q}}.$$

During the filtration, the birth of a component and the death of a cycle cannot occur at the same instant and this can more formally stated as [42]:

Theorem 1 (Birth-death decomposition [42]): The set of 0D birth values $B(G)$ and 1D death values $D(G)$ partition the edge weight set W such that $W = B(G) \cup D(G)$ with $B(G) \cap D(G) = \emptyset$. The cardinalities of $B(G)$ and $D(G)$ are $p-1$ and $(p-1)(p-2)/2$ respectively.

5) Wasserstein distance on 1-cycles: Since the barcodes embed the topological information about the network, the topological similarity or dissimilarity between the networks can be inferred from the differences between barcodes [43]. The Wasserstein distance is a metric that is often used to quantify the underlying differences in the barcodes [42], [44], [45]. Let $\Omega = (V^\Omega, w^\Omega)$ and $\Psi = (V^\Psi, w^\Psi)$ be two given networks with p nodes. Their persistent diagrams denoted as

P_Ω and P_Ψ are expressed in terms of scatter points as $x_1 = (b_1^\Omega, d_1^\Omega), \cdots, x_q = (b_\mathcal{Q}^\Omega, d_\mathcal{Q}^\Omega)$ and $y_1 = (b_1^\Psi, d_1^\Psi), \cdots, y_q = (b_\mathcal{Q}^\Psi, d_\mathcal{Q}^\Psi)$ respectively. We can show that the 2-*Wasserstein distance* on persistent diagrams is given by

$$\mathcal{D}(P_\Omega, P_\Psi) = \inf_{\tau: P_\Omega \to P_\Psi} \Big(\sum_{x \in P_\Omega} \|x - \tau(x)\|^2 \Big)^{1/2}$$

over every possible bijection τ between P_Ω and P_Ψ [42]. For graph filtrations, since persistent diagrams are 1D scatter points, the bijection τ is simply given by matching sorted scatter points [42]:

Theorem 2: The 2-Wasserstein distance between the 1D persistent diagrams (1-cycles) for graph filtration is given by

$$\mathcal{D}_1(P_\Omega, P_\Psi) = \Big[\sum_{i=1}^{q} (d_{(i)}^\Omega - d_{(i)}^\Psi)^2 \Big]^{1/2},$$

where $d_{(i)}^\Omega$ and $d_{(i)}^\Psi$ are the i-th smallest death values associated with 1-cycles (loops).

B. Hodge Laplacian over simplicial complexes

The Hodge Laplacian generalizes the usual graph Laplacian for nodes (0-simplices) to p-simplices. The Laplacian matrix $\mathcal{L}_0$ for a graph is given by $\mathcal{L}_0 = D - A$. The D is the degree matrix and A is the adjacency matrix. In general, a higher-dimensional Laplacian can be defined for each dimension p using two matrices that perform the role of upper and lower adjacency matrices:

$$\mathcal{L}_p = \mathcal{L}_p^U + \mathcal{L}_p^L$$

where $\mathcal{L}_p^U$ and $\mathcal{L}_p^L$ are called the upper and lower adjacency Laplacians [15].

1) Hodge Laplacian: The higher dimensional Laplacian $\mathcal{L}_p$ is usually referred to as the Hodge Laplacian or the p-Laplacian that connects the p-simplices with their adjacent $(p+1)$-(upper adjacency) and $(p-1)$-simplices (lower adjacency). To enable efficient computation of Hodge Laplacian, we represent the boundary operator ∂_p using the boundary matrix $\mathcal{B}_p$ defined as [46]

$$(\mathcal{B}_p)_{ij} = \begin{cases} 1, & \text{if } \sigma_{p-1}^i \subset \sigma_p^j \text{ and } \sigma_{p-1}^i \sim \sigma_k^j \\ -1, & \text{if } \sigma_{p-1}^i \subset \sigma_p^j \text{ and } \sigma_{p-1}^i \nsim \sigma_k^j \\ 0, & \text{if } \sigma_{p-1}^i \not\subset \sigma_p^j \end{cases}, \tag{3}$$

where σ_{p-1}^i is the i-th $(p$-1)-simplex and σ_p^j is the j-th p-simplex. Notations $\sim$ and $\nsim$ denote similar (positive) and dissimilar (negative) orientations respectively.

Then the p-th Hodge Laplacian matrix $\mathcal{L}_p$ of K is defined using the boundary matrices, which is the matrix form of the boundary operators:

$$\mathcal{L}_p = \mathcal{B}_p^T \mathcal{B}_p + \mathcal{B}_{p+1} \mathcal{B}_{p+1}^T. \tag{4}$$

More specifically, $\mathcal{L}_p$ is viewed as the sum of the Laplacians composed of boundary matrices from the lower dimensional simplices [47]–[50]: $\mathcal{L}_p^L = \mathcal{B}_p^T \mathcal{B}_p$ and upper dimensional simplices $\mathcal{L}_p^U = \mathcal{B}_{p+1} \mathcal{B}_{p+1}^T$. Since $\mathcal{B}_0 = 0$, the Hodge Laplacian for a 1-skeleton is $\mathcal{L}_0 = \mathcal{B}_1 \mathcal{B}_1^T$, which is popularly referred as the graph Laplacian [39]. The boundary matrix $\mathcal{B}_1$ which relates nodes to edges is commonly referred as incidence matrix in graph theory. Further, we also have $\mathcal{L}_1 = \mathcal{B}_1^T \mathcal{B}_1 + \mathcal{B}_2 \mathcal{B}_2^T$. In case of a 1-skeleton, Since there is only 0-simplex and 1-simplex, the boundary matrix $\mathcal{B}_2 = 0$, thus the second term in the Hodge Laplacian $\mathcal{L}_1$ vanishes and we have

$$\mathcal{L}_1 = \mathcal{L}_1^L = \mathcal{B}_1^T \mathcal{B}_1.$$

C. Algebraic representation of 1-cycles

The spectral decomposition of Hodge Laplacian is performed to identify p-cycles of the underlying network [39], [47]. The p-th homology group H_p is a kernel of Hodge Laplacian $\mathcal{L}_p$ [22], [46], [47], [51], i.e.,

$$H_p = ker \mathcal{L}_p.$$

The eigenvectors with zero eigenvalues of $\mathcal{L}_p$ span the kernel space of $\mathcal{L}_p$. Thus, numerically we find the eigenvectors corresponding to the zero eigenvalues of $\mathcal{L}_p$. We first solve

$$\mathcal{L}_p \mathbf{U}_p = \Lambda_p \mathbf{U}_p,$$

where Λ_p is a diagonal matrix of eigenvalues and $\mathbf{U}_p$ is a matrix of eigenvectors. The multiplicity of the zero eigenvalue of Hodge Laplacian $\mathcal{L}_p$ is the Betti number β_p, the rank of the kernel space of $\mathcal{L}_p$. This is related to the algebraic connectivity and generalizes from the well known fact that the number of zero eigenvalues of the graph Laplacian is the number of connected components [39]. Similarly, the number of zero eigenvalues of the $\mathcal{L}_0$, $\mathcal{L}_1$ and $\mathcal{L}_2$ matrix corresponds to the number of 0-cycles (connected components), 1-cycles (closed loops) and 2-cycles (voids or cavities) respectively. Since the eigenvectors corresponding to the zero eigenvalues are related to the homology generators, we represent a 1-cycle using the coefficients of the eigenvectors. Let $A = (a_{l(i,j),m})$ be the collection the columns of $\mathbf{U}_1$ that corresponds to the zero eigenvalues, where $a_{l(i,j),m}$ correspondents to edge e_{ij}. The size of A_1 is $q \times \beta_1$ with Betti number β_1. Each column of A corresponds to 1-cycles. The m-th 1-cycle $\mathcal{C}^m$ can be represented as

$$\mathcal{C}^m = \sum_{e_{ij} \in E} a_{l(i,j),m} e_{ij}. \tag{5}$$

$\mathcal{C}^m$ can be represented as a vector by putting coefficient $a_{l(i,j),m}$ into the corresponding position in the lexicographically ordered edge set $[e_{12}, e_{13}, \cdots, e_{23}, e_{24}, \cdots, e_{q-1,q}]^T$.

The eigen decompostion is performed on the Hodge Laplacian $\mathcal{L}_1$ results in the eigenvalues $[\mathbf{0.00}, 0.83, 2.00, 2.69, 4.48]$ and the eigenvector corresponding to the zero eigenvalue is obtained as $[0.00, 0.50, -0.50, 0.50, -0.50]^T$. The 1-cycle is then represented as

$$\mathcal{C}^1 = 0.5e_{23} - 0.5e_{24} + 0.5e_{35} - 0.5e_{45}.$$

Figure 3-bottom left is a 1-skeleton representation for this example. It is a graph since there are no higher order simplices beyond nodes and edges. In such a case, we obtain 1-cycle representation as seen in Figure 3-bottom right with only the edges that constitute the 1-cycle. This example illustrates that in order to identify and extract the 1-cycles, we need to break down the graph into series of subgraph such that each subgraph contains only one 1-cycle.

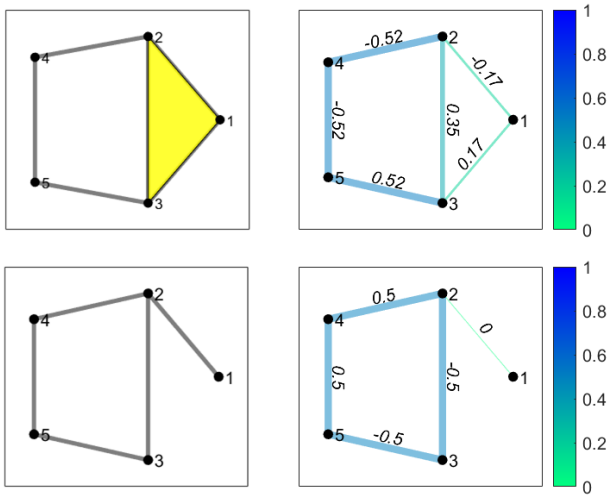

Fig. 3: Top-left: A 2-skeleton representation network in Example 1 made of five vertices connected by six edges. Top-Right: The 1-cycle is formed by the vertices v_2, v_4, v_5, v_3 is identified by the eigen decomposition of the Hodge Laplacian ($\mathcal{L}_1$). The edge colors indicate the absolute value of coefficients of the cycle representation C^1. Bottom-left: 1-skeleton representation of the network in Example 2 made of five vertices connected by five edges. Bottom-right: The 1-cycle identified along with the edges that constitute the cycle.

1) Computation of 1-cycle basis: The representation (5) uses all the edges representing a 1-cycle. Even the edges that are not a part of a cycle are used in the representation. This has been the main limitation of using Hodge Laplacian in identifying 1-cycles in the past [22]. In the proposed method, we split the graph into a series of subgraphs such that each subgraph has only one 1-cycle.

Recall that the graph filtration partitions the edges in a given network uniquely into the birth and death sets. While the edges in the birth set are responsible for creating components, the edges in the death set accounts for destroying cycles. The edges in the birth set forms the maximums spanning tree (MST) with no cycles. Upon add an edge from the death set to MST a 1-cycle is formed. The process is repeated sequentially till we use up all the edges in the death set. We claim the resulting 1-cycles form a basis.

Theorem 3: Let $M(G) = (V, T)$ be the MST of graph G. When an edge d_k from the death set $D(G)$ is added to the MST, 1-cycle $\mathcal{C}^k$ is born. The collection of cycles $\mathcal{C}^1, \cdots, \mathcal{C}^Q$ spans $ker\mathcal{L}_p$.

Proof. Let E_k be the edge set of the cycle $\mathcal{C}^k$. Since E_k and E_l differ at least by an edge d_k and d_l, they are algebraically independent. Hence, all the cycles $E_1, \cdots, E_Q$ are independent from each other. Since there should be Q number of independent cycles in the p-th Homology group $H_p = ker\mathcal{L}_p$, they form a basis. □

The 1-cycles can now be sequentially extracted by using the Hodge Laplacian of the subgraph $G_k = (V, T \cup \{d_k\})$, which contains a cycle $\mathcal{C}^k$. We get exactly one eigenvector corresponding to the zero eigenvalue. The entries of eigenvector will be all zero on the edges that are not part of cycle. Thus, we can represent 1-cycle $\mathcal{C}^k$ only using edges that contribute to the cycle:

$$\mathcal{C}^k = \sum_{e_{ij} \in E_k} a_{l(i,j),k} e_{ij}. \tag{6}$$

Here $a_{l(i,j),k}$ is the entries of eigenvector of the Hodge Laplacian corresponding to edge e_{ij}. The representation (6) contains only the edges that form the cycle. All other terms are zero. Thus, $\mathcal{C}^k$ can be represented as vectors by putting $a_{l(i,j),k}$ into the corresponding position in the vectorized edge set $[e_{12}, e_{13}, \cdots, e_{23}, e_{24}, \cdots, e_{q-1,q}]^T$. Subsequently, all the 1-cycle basis $\mathcal{C}^1, \cdots, \mathcal{C}^Q$ can be systematically extracted and efficiently stored as a sparse matrix. Since $\mathcal{C}^1, \cdots, \mathcal{C}^Q$ forms a basis, any cycle in the graph can be represented as a linear combination $\sum_{j=1}^{Q} \alpha_j \mathcal{C}^j$.

The extraction of 1-cycle basis of a network can be summarized into three parts as illustrated in Figure 4. Firstly, a given network is represented as a complete graph G. The birth-death decomposition is used in extracting birth and death values. The edges $[e_{15}, e_{25}, e_{35}, e_{45}]$ form the birth set $B(G)$ and the remaining edges $[e_{12}, e_{13}, e_{14}, e_{23}, e_{24}, e_{34}]$ become the death set $D(G)$. The edges in the birth set correspond to the maximum spanning tree (MST). Secondly, the subgraphs having only one cycle each are then created by adding an edge from the death set to the MST. Lastly, the subgraphs are subjected to Theorem 3 to obtain the algebraically independent cycle basis of G.

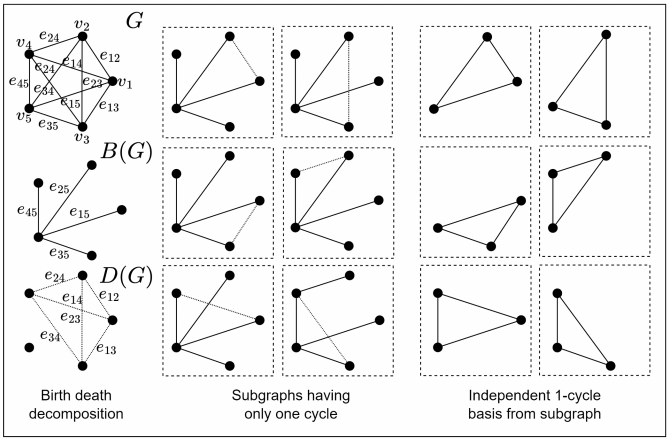

Fig. 4: Left: A graph (G) is decomposed into birth set $B(G)$ with edges $[e_{15}, e_{25}, e_{35}, e_{45}]$ and death set $D(G)$ with edges $[e_{12}, e_{13}, e_{14}, e_{23}, e_{24}, e_{34}]$. Middle: The subgraphs constructed using the edges from birth set and adding an edge from the death set. Right: The independent 1-cycles obtained by constructing Hodge Laplacian on the subgraphs and identifying the cycles in the kernel of the Hodge Laplacian.

D. Statistical analysis on 1-cycles

We present how to use topologeal features such as death values and length of cycles in analyzing collection of brain networks. Let $\Omega = \{\Omega_1, \cdots, \Omega_m\}$ and $\Psi = \{\Psi_1, \cdots, \Psi_n\}$ be a collection of m and n complete networks each consisting

of p number of nodes. There are exactly $\mathcal{Q} = (p-1)(p-2)/2$ number of cycles in each network. We are interested in developing new statistical inference procedures testing the topological equivalence of two groups of networks Ω and Ψ.

1) Inference on death values: We test the topological equivalence of two groups of networks Ω and Ψ using the Wasserstein distances within groups $\mathfrak{L}_W$ and between groups $\mathfrak{L}_B$ [42]:

$$\mathfrak{L}_W = \frac{\sum_{i<j} \mathcal{D}_1(\Omega_i, \Omega_j) + \sum_{i<j} \mathcal{D}_1(\Psi_i, \Psi_j)}{\binom{m}{2} + \binom{n}{2}} \quad \text{and}$$

$$\mathfrak{L}_B = \frac{\sum_{i=1}^{m} \sum_{i=1}^{n} \mathcal{D}_1(\Omega_i, \Psi_j)}{mn}.$$

Note we are only using the Wasserstein distance between cycles, which are computed as the squared sum of sorted death values. Then we use the ratio $\mathfrak{L}_{B/W} = \mathfrak{L}_B / \mathfrak{L}_W$ as the test statistic. If the two groups are close, $\mathfrak{L}_B$ becomes small while $\mathfrak{L}_W$ becomes large. Thus the ratio $\mathfrak{L}_{B/W}$ can be used to as test statistic.

Since the probability distribution of $\mathfrak{L}_{B/W}$ is unknown, we used the permutation test [52]–[56]. For large sample sizes m and n as in our study, the permutation test will be computationally costly. Thus, we adapted for the scalable *transposition test* that sequentially update the test statistic over transpositions [42], [52].

Unlike the permutation test that shuffles all the networks, the permutation test only shuffles one network per group. Computing the statistic $\mathfrak{L}_{B/W}$ over each permutation requires the recomputation of the Wasserstein distance. Instead, we perform the transposition of swapping only one network per group and setting up iteration of how the test statistic change over the transposition. In this study, we generate the test statistics with sufficiently large number of 500000 random transpositions while injecting a random permutation for every 500 transpositions. The intermix of transpositions and permutations has the effect of speeding up the convergence [52].

E. Spectral approach of using 1-cycle basis

So far we investigated how to use topological features of cycles in statistical inference. In this section, we present the spectral geometry approach of using 1-cycle basis in modeling brain connectivity. Brain images are often denoised to increase the signal-to-noise ratio (SNR) and enhance statistical sensitivity. Denoising induces many nice statistical properties such as variance reduction and improves sensitivity [58]. Brain connectivity matrices are noisy so it is necessary to denoise the matrices as well. We extend the concept of heat diffusion defined on brain surfaces [58] to simplicial complexes. We propose to perform heat diffusion over 1-simplices (edges) using the eigenvectors of Hodge Laplacian as follows.

1) Heat diffusion on connectivity matrices: Let $\mathcal{K}$ be a collection of 1-simplices with cardinality $|\mathcal{K}|$. Let f be the initial observed data defined over $\mathcal{K}$. For instance, f can be the vectorization of upper triangle of edge weight matrix $w = (w_{ij})$. We will perform heat diffusion smoothing over $\mathcal{K}$:

Theorem 4: The unique solution to diffusion equation over $\mathcal{K}$

$$\frac{\partial g(t)}{\partial t} = \mathcal{L}_1 g(t) \tag{7}$$

with initial condition $g(t=0) = f$ is given by

$$g(t) = \sum_{j=0}^{|\mathcal{K}|-1} e^{-\lambda_j t} f_j \psi_j \tag{8}$$

where λ_j and ψ_j are the j-th eigenvalue and eigenvector of the Laplacian matrix $\mathcal{L}_1$, and $f_j = f^T \psi_j$.
The diffusion time t serves as smoothing bandwidth. Unlike existing heat diffusion methods which defines the smoothing over nodes [58], we are defining along edges. Our approach avoids an ad-hoc procedure of converting nodes to edges and edges to nodes in performing smoothing data defined on edges [59]. The smoothing process is illustrated in Figure 5. Starting with the original network data with bandwidth of 0, as the bandwidth increases, the network gradient decreases reducing high frequency noise.

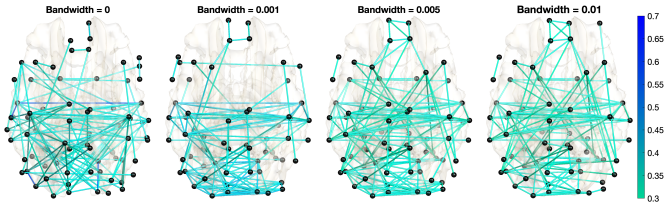

Fig. 5: The smoothed connectivity maps, thresholded for better visualization. The top left represents the initial edge data with some potential high frequency noise influencing the blue-gradient edges. After smoothing with a bandwidth of 0.001, the noise component appears to be removed for edges with high frequency noise, whiles the other edge signals are not significantly altered.

We can show that the variance across subjects at each edge is reduced after smoothing. If we denote $\mathbb{V}_i f$ to be the variance of functional measurement f across subject at the i-th edge, we can algebraically show that $\mathbb{V}_i g(t) \leq \mathbb{V}_i f$ for all t. After heat diffusion smoothing, the variance will be reduced and statistical sensitivity of group discrimination will increase. This variance reduction property is demonstrated in Figure 6. We smoothed the brain networks of 400 subjects with three different smoothing bandwidths 0.001, 0.005, and 0.01. We observe that the variation of the smoothed data across all edges is consistently less compared to the initial non-smoothed data. This reduced variation is as a result of the heat diffusion smoothing removing high frequency noise, and hence has the effect of increasing statistical sensitivity.

2) Common 1-cycle basis across subjects: If 1-cycle basis change from one subject to next, it is difficult to use the basis expansion itself as a feature statistical analysis. Thus, we propose to extract common 1-cycle basis in the network template obtained by averaging correlation matrices of all the subjects. Then we encode subject-level variability in the expansion coefficients with the fixed 1-cycle basis across subjects. To achieve this we used the average correlation

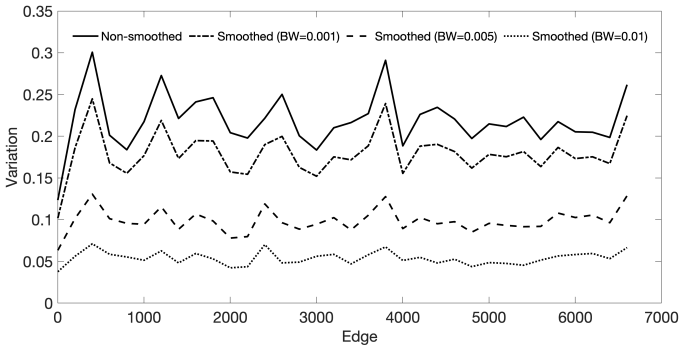

Fig. 6: Illustration of the level of variation with three diffusion bandwidths for 400 subjects. A bandwidth (BW) of $0.001, 0.005$, and 0.01 were applied to the connectivity maps.

matrix of all the networks as a functional network template. We then obtain $\mathcal{Q} = (p-1)(p-2)/2$ number of 1-cycle basis $\boldsymbol{\phi} = \left[\mathcal{C}^1, \mathcal{C}^2, \cdots, \mathcal{C}^{\mathcal{Q}}\right]$ for the functional template. Here $\boldsymbol{\phi}$ denotes the common 1-cycle basis obtained from (6). Subsequently, the vectorized upper triangle entries of correlation matrix of each subject $\mathcal{W}$ is expanded as

$$\mathcal{W} = \alpha_1 \mathcal{C}^1 + \alpha_2 \mathcal{C}^2 + \cdots + \alpha_{\mathcal{Q}} \mathcal{C}^{\mathcal{Q}}.$$

The coefficients $\boldsymbol{\alpha} = [\alpha_1, \cdots, \alpha_{\mathcal{Q}}]^T$ are then estimated in the least squares fashion

$$\boldsymbol{\alpha} = (\boldsymbol{\phi}^T \boldsymbol{\phi})^{-1} \boldsymbol{\phi}^T \mathcal{W}. \tag{9}$$

The estimated coefficients $\boldsymbol{\alpha}$ for all the subjects are then used in discriminating two groups of networks Ω and Ψ. Let $\bar{\alpha}_j^{\Omega}$ and Let $\bar{\alpha}_j^{\Psi}$ be the mean of j-th 1-cycle basis in group Ω and Ψ respectively. Then we used the maxim difference

$$\mathfrak{L}(\Omega, \Psi) = \max_{1 \le j \le \mathcal{Q}} |\bar{\alpha}_j^{\Omega} - \bar{\alpha}_j^{\Psi}| \tag{10}$$

as the test statistic in discriminating between two groups of networks. The statistical significance is determined using the permutation test. Unlike previous analysis that cannot localize specific cycles, the test statistic gives a way to localize most discriminating cycles by identifying the j-th cycle that gives the maximum.

F. Validation

Since cycles can be modelled to embed complex interactions, it can potentially uncover hidden topological patterns which are hitherto impossible in conventional graph models. We validate the proposed method in a simulation study with the ground truth. We generate three types of networks with different number of loops. Some well known curved shapes such as a circle, leminiscate, quadrifolium are chosen as the ground truth signal and then Gaussian noise is added $\mathcal{N}(0, 0.02^2)$ to the coordinates (Figure 7). The circle has a single loop, the leminiscate has two loops and the quadriform has four loops. The number of nodes to construct the network are chosen as $p = 64$ for all the types. This ensures we have the same number of cycles ($Q = 1953$ independent 1-cycles) in each type of simulated network.

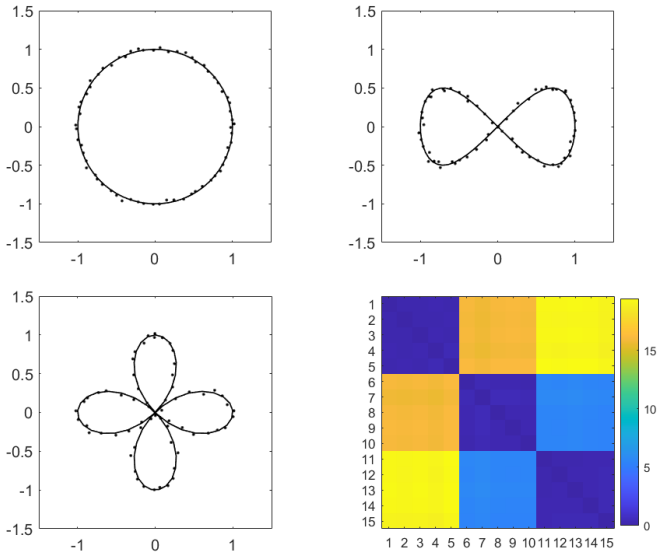

Fig. 7: The three types of cycles with different topology: circle (1-loop), lemniscate (2-loops), quadrifolium (4-loops) used in the simulation study. The Gaussian noise $\mathcal{N}(0, 0.02^2)$ is added to the coordinates of curves. Bottom right: the pairwise Wasserstein distance matrix computed using the death values of the 1-cycles on 5 networks in each group.

1) Death values: The topological distances between the simulated networks were measured by computing the 2-Wasserstein distance between 1D persistent diagram of 1-cycles. To compare between the different simulated loop structures, we generated five networks in each type (1-loop, 2-loops and 4-loops) such that they are clustered into three distinct groups. We then computed the pairwise Wasserstein distance between networks. Figure 7 shows the Wasserstein distance matrix between three groups. The clear clustering pattern demonstrates the Wasserstein distance applied to 1D topological feature works as expected. Networks with similar topology have smaller distances while networks with different topology have relatively large distances.

Using the proposed ratio statistic, we computed p-values comparing different groups. Table I shows the average p-values obtained after 50 independent simulations. Each simulation is done with 100000 permutations. Networks of the same topology have large p-values indicating they are shown to be statistically not different. Networks of different topology have small p-values indicating they are shown to be statistically different. The results indicate the proposed method perform well in discriminating networks of different topology and not produces any false positives when there is no topological differences. Thus, the method perform well as expected. As the number of networks increase in each group, the p-values get smaller showing increased statistical power over increased sample size.

2) Common 1-cycles basis across subjects: We used the maximum gap between coefficients of 1-cycle basis as the test statistic on the same simulation study. The tests are repeated

TABLE I: The performance results of Wasserstein distance on 1-cycles are summarized as average p-values for testing various combinations of cycles. Here 1 vs 2 means we compare the circle (1-loop) against a leminiscate (2-loops) and the columns 6 networks, 8 networks, 10 networks and 12 networks indicate the number of networks that we consider for each type. The smaller p-values indicate that our method can discriminate network differences.

loop-type	6 networks	8 networks	10 networks	12 networks
1 vs. 2	1.4×10^{-3}	6.8×10^{-5}	6.1×10^{-6}	4.4×10^{-7}
1 vs. 4	1.1×10^{-3}	5.4×10^{-5}	4.9×10^{-6}	5.2×10^{-7}
2 vs. 4	1.2×10^{-3}	6.6×10^{-5}	3.2×10^{-6}	2.0×10^{-7}
1 vs. 1	0.3954	0.5336	0.9790	0.7834
2 vs. 2	0.6516	0.8404	0.3458	0.5376
4 vs. 4	0.5943	0.8294	0.7561	0.5403

for 10 times and the average p-values are reported. Each simulation is done with 100000 permutations. Table II shows the p-values obtained for this study. The p-values are low for networks with differences while the values are large when the network has no difference. The method performed better than the cycle length based analysis.

TABLE II: The performance results of common 1-cycle basis are summarized as average p-values for testing various combinations of cycles. Here 1 vs 2 means we compare the circle (1-loop) against a leminiscate (2-loops) and the columns 6 networks, 8 networks, 10 networks and 12 networks indicate the number of networks that we consider for each type. The smaller p-values indicate that our method can discriminate network differences.

loop-type	6 networks	8 networks	10 networks	12 networks
1 vs. 2	2.1×10^{-3}	2.0×10^{-4}	1.0×10^{-5}	0.0000
1 vs. 4	1.9×10^{-3}	1.2×10^{-4}	2.0×10^{-5}	0.0000
2 vs. 4	1.8×10^{-3}	1.4×10^{-4}	1.0×10^{-5}	0.0000
1 vs. 1	0.4263	0.6606	0.8736	0.6735
2 vs. 2	0.3962	0.8919	0.9620	0.5590
4 vs. 4	0.7988	0.7365	0.4598	0.9815

In literature, there is no baseline method for explicitly modeling cycles in a network. Also, there is no ground truth in real brain data; so even if we apply the baseline methods to real data, it is unclear which method provides the best answer. Thus we first compared our method in a simulation study with the ground truth. To perform a simulations, we construct topologically different shapes by combining circular arcs with and without a gap. The simulation networks are generated by sampling points from different topological shapes as shown in Figure 8. We consider three topologically different networks with the difference in their number of loops in each group. Group 1 has three loops, Group 2 has two loops and Group 3 has one loop. An individual network in each group is generated by first sampling the coordinates y_i along the ground truth patterns. The coordinates of y_i are perturbed with Gaussian noise $\mathcal{N}(0, 0.05^2)$. The weight w_{ij} between any two nodes is

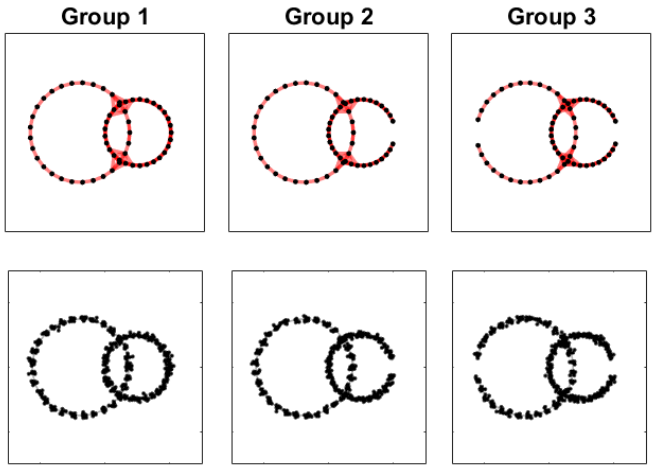

Fig. 8: Top: Three topologically different network shapes with different number of loops in each group. Group 1 has three loops, Group 2 has two loops and Group 3 has one loop. Bottom: The sample points for the simulation networks generated using the Gaussian noise $\mathcal{N}(0, 0.05^2)$ on the base network.

given by the Euclidean distance between the coordinates y_i and y_j. To retain only the dominant loops in the network, we applied the following thresholding scheme

$$w'_{ij} = w_{ij}(1 - I_{ij}) + 10^{-3} I_{ij} \cdot U(0, 1),$$

where $U(0, 1)$ is the uniform distribution on the interval $(0, 1)$ and the indicator $I_{ij} = 1$ if $w_{ij} > 0.5$ and 0 otherwise. The edge weights w'_{ij} are constructed such that the connections larger than the threshold 0.5 are replaced with random noise to retain only the dominant loops in the networks. In the simulation, we generated $N = 60$ random networks per group.

We compared our model to graph theory features (Q-modularity, Betweenness) and persistent homology methods (Gromov-Hausdorff distance and bottle neck distances). Table III shows the performance results with the average p-values with the standard deviations. The false negative rates and false positive rates are also indicated within brackets. The $\mathfrak{L}_Q$, $\mathfrak{L}_{bet}$ are based on the Q-modularity and betweenness [11]. The $\mathfrak{L}_{GH}$ and $\mathfrak{L}_{BN}$ are based on the Gromov-Hausdorff [60] and bottleneck distances [61]. The $\mathfrak{L}_{\mathfrak{w}}$ is Wasserstein distance based on death values. $\mathfrak{L}_{\mathfrak{c}}$ is the statistical inference based on the cycle basis. We use the test statistic (10) and the statistical significance is determined using the permutation test.

In testing topological differences (first three rows of Table III), the existing methods did not performing well failing to identify the topological differences. The proposed methods $\mathfrak{L}_{\mathfrak{w}}$ and $\mathfrak{L}_{\mathfrak{c}}$ performed very well and were able to differentiate topological differences. In testing *no* topological difference (last three rows of Table III), all the methods performed reasonably well and did not report any false positives. In summary, if there are subtle topological differences that are difficult to differentiate, existing methods will likely to fail while the topological method will likely to detect signals.

TABLE III: The performance results showing average p-values with the standard deviations. The false positive and false negative rates are shown in the brackets. Smaller error rates are preferred. The graph theory features Q-modularity $\mathfrak{L}_Q$ and betweenness $\mathfrak{L}_{bet}$ are used. The Gromov-Hausdorff $\mathfrak{L}_{GH}$ and bottleneck $\mathfrak{L}_{BN}$ in persistent homology are used. The $\mathfrak{L}_{\mathfrak{w}}$ is the proposed Wasserstein distance on death values. $\mathfrak{L}_{\mathfrak{c}}$ is the proposed test statistic on the cycle basis.

Groups	$\mathfrak{L}_Q$	$\mathfrak{L}_{bet}$	$\mathfrak{L}_{GH}$	$\mathfrak{L}_{BN}$	$\mathfrak{L}_{\mathfrak{w}}$	$\mathfrak{L}_{\mathfrak{c}}$
1 vs. 2	0.4192 ±0.28 (0.94)	0.5167 ±0.28 (0.98)	0.4880 ±0.30 (0.88)	0.4495 ±0.30 (0.92)	0.0023 ±0.00 (0.00)	0.0000 ±0.00 (0.00)
1 vs. 3	0.3521 ±0.31 (0.76)	0.4673 ±0.27 (0.98)	0.4903 ±0.29 (0.94)	0.5419 ±0.30 (0.98)	0.0000 ±0.00 (0.00)	0.0000 ±0.00 (0.00)
2 vs. 3	0.5671 ±0.28 (0.96)	0.4672 ±0.30 (0.96)	0.5503 ±0.29 (0.98)	0.4362 ±0.29 (1.00)	0.0144 ±0.03 (0.06)	0.0000 ±0.00 (0.00)
1 vs. 1	0.5399 ±0.26 (0.04)	0.5170 ±0.28 (0.04)	0.4251 ±0.26 (0.08)	0.4793 ±0.27 (0.04)	0.5069 ±0.29 (0.04)	0.4917 ±0.25 (0.06)
2 vs. 2	0.5487 ±0.30 (0.02)	0.5291 ±0.26 (0.02)	0.5153 ±0.29 (0.04)	0.5031 ±0.30 (0.04)	0.4524 ±0.26 (0.04)	0.5164 ±0.32 (0.14)
3 vs. 3	0.4836 ±0.30 (0.08)	0.4608 ±0.26 (0.06)	0.5322 ±0.25 (0.04)	0.5464 ±0.33 (0.10)	0.5069 ±0.32 (0.04)	0.5086 ±0.31 (0.04)

III. Application

A. Dataset and preprocessing

In this study, we used the the subset of the resting-state fMRI data collected in the Human Connectome Project (HCP) [62], [63]. The fMRI data were acquired for approximately 15 minutes for each scan. The participants are at rest with eyes open with relaxed fixation on a projected bright cross-hair on a dark back-ground [62]. The fMRI data were collected on a customized Siemens 3T Connectome Skyra scanner using a gradient-echoplanar imaging (EPI) sequence with multiband factor 8, repetition time (TR) $720ms$, time echo (TE) $33.1ms$, flip angle $52°$, 104×90 (RO $\times$ PE) matrix size, 72 slices, $2mm$ isotropic voxels, and 1200 time points is used.

The standard minimal preprocessing pipelines [63] such as spatial distortion removal [64], motion correction [65], bias field reduction [66], registration to the structural MNI template, and data masking using the brain mask obtained from FreeSurfer [63] is performed on the fMRI scans. This resulted in the resting-state functional time series with $91 \times 109 \times 91$, $2mm$ isotropic voxels at 1200 time points. The subjects were in the age group ranging from 22 to 36 years with average age 29.24 ± 3.39 years for 172 males and 240 females. Subsequently, the Automated Anatomical Labeling (AAL) template was applied to parcellate the brain volume into 116 non-overlapping anatomical regions [67]. The fMRI across voxels within each brain parcellation is averaged (spatial denoising), resulting in 116 average fMRI time series with 1200 time points for each subject.

The scrubbing is done to remove fMRI volumes with spatial artifacts in functional connectivity [68] due to significant head motion [68], [69]. The framewise displacement (FD) from the three translational displacements and three rotational displacements at each time point to measure the head movement from one volume to the next is calculated. The volumes with FD larger than $0.5mm$ and their neighbors were scrubbed [68], [69]. About 12 subjects having excessive head movement are excluded from the dataset, resulting in a refined fMRI dataset of 400 subjects (168 males and 232 females). Additional details on the dataset can be found here [42], [69].

B. Cycle computation

For each subject, we measured the whole-brain functional connectivity by computing the Pearson correlation matrix $\rho = (\rho_{ij})$ over while time points across 116 brain regions resulting in 400 correlation matrices of size 116×116. Since the dataset contains $p = 116$ nodes, the total number of edges in the brain network is computed as $q = p(p-1)/2 = 6670$. The edges in the transformed correlation matrix is now decomposed into birth and death sets following Theorem 1. The number of edges in the birth set is $\mathcal{P} = p - 1 = 116 - 1 = 115$. The number of edges in the death set is $\mathcal{Q} = q - \mathcal{P} = 6555$. The edges from the death set are then sequentially added to the birth set to generate a sequence of 6555 subnetworks. Each subnetwork has only one cycle which is identified using the Hodge Laplacian.

Figure 9 shows how the number of the topological invariants β_0 (number of connected components) β_1 (number of cycles) changes over graph filtration on edge weights $w = (w_{ij})$. β_0 remains at one for a long duration and begins to increase towards the end and eventually reaches 116 which is the number of independent components or nodes. On the other hand, β_1 begins with $\mathcal{Q} = 6555$ cycles for a complete network and then gradually keeps decreasing as the edges are removed sequentially and goes to zero when all the cycles are dead. Once all the cycles are identified and extracted we primarily consider the death values of cycles. These topological quantities are used as test statistics for discriminating males from females.

C. The topological features of 1-cycles

The topological similarity between the networks can be measured by computing the 2-Wasserstein distance between persistent diagrams [42]. A distance matrix is constructed considering the pairwise Wasserstein distance between the subjects. Once we have the distance matrix, the group statistics can be carried out by calculating the within and between group statistics. Since the permutation test is computationally more demanding we adapt a scalable computation strategy using transpositions [52]. The transposition test is applied to determine the statistical significance in discriminating the 232 females and 168 male subjects. The observed test statistic is 1.0232 and corresponding p-value is 0.049 based for 500000 random transpositions.

We also accessed the topological disparity between the groups using the length of the cycle. The test statistics were formulated for the length of the cycles following the proposed procedure. The observed statistic was found to be $\mathfrak{L} = 0.303$ which corresponds to p-value of 0.64 based on 500000 random permutations. Based on the simulation study and real data, we

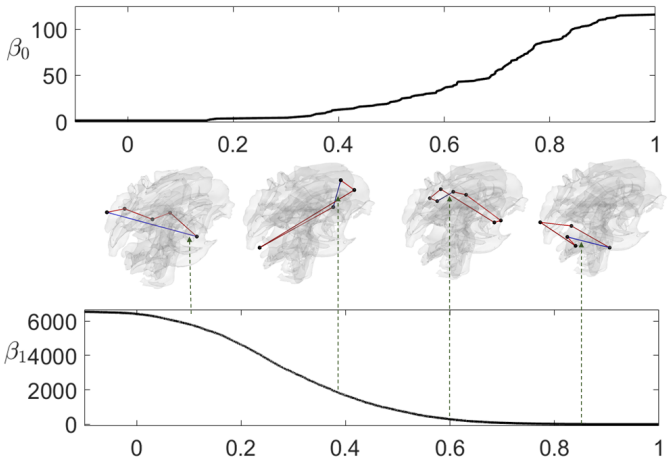

Fig. 9: Graph filtration of the average brain network of 400 subjects. β_0 is monotonically increasing while β_1 is monotonically decreasing. over the graph filtration. We have total 6555 cycles in the brain network. Middle: Four 1-cycles chosen at specific death values are shown. The edge weights are set to a constant value to visualize the cycles. The edges that destroy the cycles are shown in blue color.

conclude that the death values of cycles seem to be useful feature but the length of cycles are not useful discriminating features of cycles.

We also tested the standard geometric and PH measures on the brain data using the test statistic designed to evaluate most discriminative cycles. We compared the discriminating power of our method against the PH methods Gromov-Hausdorff (GH) and bottleneck (BN) distances in comparing male and female brain networks. The computed p-values are 0.540 and 0.277 respectively and not able to discriminate the networks. Both the GH and BN distances do not perform well in the real data. We also used graph theory features Q-modularity and betweenness measures [11] using the Brain Connectivity Toolbox [11] and obtained p-values of 0.035 and 0.6202 respectively. Among all 4 baseline methods, Even though Q-modularity performed well, it cannot be used to identify connections that are responsible for the differences and explicitly localize regions that cause significant topological disparity.

D. Common 1-cycle basis

The common 1-cycle basis is first obtained using the average correlation matrices of 400 subjects. We then solved for the coefficients for each network using equation (9) and computed the mean coefficients for females and males separately for each cycle. We then used the maximum difference between mean coefficients as the test statistic. The observed statistic was found to be 0.408, which corresponds to p-value of 0.03. The statistical inference is based on 500000 permutations. The five most discriminating cycles are identified based on the maximum values in the test statistics namely., 0.408, 0.407, 0.405, 0.396 and 0.393 which correspond to the cycle IDs 2446, 1140, 4090, 3683 and 831 respectively. Figure 10 shows five most discriminating cycles corresponding to the maximum observed statistics. It can be seen that some brain connections consistently appear in all the five cycles. The five most discriminating 1-cycles include the following brain regions: superior parietal gyrus (Parietal-Sup-L), inferior parietal lobule (Parietal-Inf-L), Precentral gyrus (Precentral-L), Postcentral gyrus (Postcentral-L), the rolandic operculum (Rolandic-Oper-L, Rolandic-Oper-R), the median cingulate and para cingulate gyri (Cingulum-Mid-R, Cingulum-Mid-L) and the Insula. The connectivities between these regions highlight their importance in discriminating males and females. The symmetric connection between the left and right rolandic operculum, superior parietal lobule and the middle cingulate appear in at least 3 most dominating cycles. We can further localize these regions using the frequency of occurrence f_e of the a particular connection (edge) in each cycle given as

$$f_e = \frac{N_e}{N_c},$$

where is the frequency of occurrence, N_e number of cycles in which a particular edge is present and N_c is the total number of most discriminating cycles chosen for analysis. Figure 10 (Bottom right) specifically shows the edges that have $f_e > 0.5$. The edges connecting the regions Parietal-Sup-L, Precentral-L, Postcentral-L and Rolandic-Oper-L appear in all the 5 cycles and has $f_e = 1.0$. There is known sex difference in the parietal region involved in spatial ability, and particularly involved in mental rotation [70]. [71] reported sex differences in the left parietal, precentral and postcentral regions in a rs-fMRI study, where Kendall's coefficient of concordance (KCC) was used to measure the similarity of the ranked time series of a given voxel to its nearest 26 neighbor voxels [71], [72]. The sex difference is reported in the left rolandic operculum in rs-fMRI study [73]. While all these pervious studies are reporting the sex differences at the node level, we are consistently identifying them at the cycle-level within 5 most dominant cycles. The edges connecting Rolandic-Oper-L, Rolandic-Oper-R and Insula appear in 4 cycles and has $f_e = 0.8$. The edges connecting Parietal-Sup-L and Parietal-Inf-L and the edges connecting Cingulum-Mid-R, Cingulum-Mid-L and Insula-R occur in 3 cycles and has a $f_e = 0.6$. We believe these brain regions can act as discriminating biomarkers for sexual dimorphism studies including Alzheimer's disease which affects disproportionately more women than men [74].

IV. Conclusion

A cycle in the brain network is one of the most fundamental topological features that one has to identify, extract and quantify in order to understand and model higher order interactions. In this work, an efficient scalable algorithm to identify and extract the 1-cycles in a network is proposed. We combine the ideas from persistent homology and the Hodge Laplacian to facilitate an easy detection of 1-cycles. The method is demonstrated with an illustration and applied to the resting state brain networks from Human Connectome Project (HCP). The proposed algorithm is efficient for typical brain network data which has few hundred nodes ($p \sim 100$). Even for larger networks ($p \sim 1000$), computation can be done

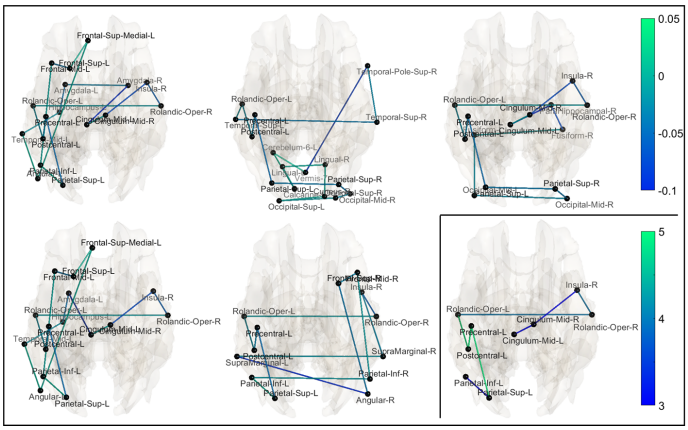

Fig. 10: The five most discriminating cycles having maximum values in the test statistics namely., 0.408, 0.407, 0.405, 0.396 and 0.393 are shown. The colorbar in the range -0.10 to 0.05 shows the difference between the average correlation matrices of the female and male subjects for the five cycles. Bottom Right: The edges that frequently occur in all the five cycles are shown. The blue edges occur in atleast 3 cycles whereas the green edges appear in all the 5 cycles.

quickly in $\mathcal{O}(p \log p)$ run time through the maximum spanning trees (MST).

One of our major goals in the study is to discriminate networks having different loops. To capture this topological characteristic, we used the 1-cycle basis to precisely encode this information without redundancy. Although the information about the number of loops is present in the cycle basis, sometimes it can get hidden or lost in the large number of cycles. It is not even clear how to represent all the cycles without overlaps. Through the combination of MST and the Hodge Laplacian, we were able to extract and represent 1-cycle basis as a sparse matrix.

We designed a new topological inference procedure based on the 1-cycle attributes such as length and death values of cycles. These statistical frameworks are used to examine in discriminating the brain networks of males and females. Our studies emphasize that it is meaningful to study and model the higher order interactions using the 1-cycle basis for brain network analysis.

Based on the proposed 1-cycle basis, any cycle in the graph can be represented as linear combination of basis: $\sum_{j=1}^{Q} \alpha_j \mathcal{C}^j$. Such vectorization enables us to build more complex models such as sparse network models or joint identification of common cycles across subjects [22]. This is left as a future study. The Wasserstein distance between cycles $\mathcal{C}^i$ and $\mathcal{C}^j$ is simply the squared difference of death values $(d_i - d_j)^2$. Such squared norm makes computation involving cycles straightforward.

REFERENCES

[1] P. J. Uhlhaas and W. Singer, "Neural synchrony in brain disorders: relevance for cognitive dysfunctions and pathophysiology," *Neuron*, vol. 52, no. 1, pp. 155–168, 2006.

[2] H. J. Park and K. Friston, "Structural and functional brain networks: from connections to cognition," *Science*, vol. 342, no. 6158, 2013.

[3] R. F. Betzel and D. S. Bassett, "Multi-scale brain networks," *Neuroimage*, vol. 160, pp. 73–83, 2017.

[4] L. Pessoa, "Understanding brain networks and brain organization," *Physics of Life Reviews*, vol. 11, no. 3, pp. 400–435, 2014.

[5] R. Ghorbanchian, J. G. Restrepo, J. J. Torres, and G. Bianconi, "Higher-order simplicial synchronization of coupled topological signals," *Communications Physics*, vol. 4, no. 1, pp. 1–13, 2021.

[6] E. T. Bullmore and O. Sporns, "Complex brain networks: graph theoretical analysis of structural and functional systems," *Nature Reviews Neuroscience*, vol. 10, no. 3, pp. 186–198, 2009.

[7] M. E. Newman, "Finding community structure in networks using the eigenvectors of matrices," *Physical Review E*, vol. 74, no. 3, p. 036104, 2006.

[8] O. Sporns, "Graph theory methods: applications in brain networks," *Dialogues in Clinical Neuroscience*, vol. 20, no. 2, p. 111, 2018.

[9] M. Rubinov and O. Sporns, "Weight-conserving characterization of complex functional brain networks," *Neuroimage*, vol. 56, no. 4, pp. 2068–2079, 2011.

[10] D. S. Bassett and E. T. Bullmore, "Small-world brain networks revisited," *The Neuroscientist*, vol. 23, no. 5, pp. 499–516, 2017.

[11] M. Rubinov and O. Sporns, "Complex network measures of brain connectivity: uses and interpretations," *Neuroimage*, vol. 52, no. 3, pp. 1059–1069, 2010.

[12] A. Mheich, F. Wendling, and M. Hassan, "Brain network similarity: methods and applications," *Network Neuroscience*, vol. 4, no. 3, pp. 507–527, 2020.

[13] M. K. Chung, H. Lee, V. Solo, R. J. Davidson, and S. D. Pollak, "Topological distances between brain networks," in *International Workshop on Connectomics in Neuroimaging*, pp. 161–170, Springer, 2017.

[14] 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.

[15] F. Battiston, G. Cencetti, I. Iacopini, V. Latora, M. Lucas, A. Patania, J.-G. Young, and G. Petri, "Networks beyond pairwise interactions: structure and dynamics," *Physics Reports*, vol. 874, pp. 1–92, 2020.

[16] P. Skardal and A. Arenas, "Higher order interactions in complex networks of phase oscillators promote abrupt synchronization switching," *Communications Physics*, vol. 3, pp. 1–6, 2020.

[17] F. V. Farahani, W. Karwowski, and N. R. Lighthall, "Application of graph theory for identifying connectivity patterns in human brain networks: a systematic review," *frontiers in Neuroscience*, vol. 13, p. 585, 2019.

[18] G. Carlsson, "Topology and data," *Bulletin of the American Mathematical Society*, vol. 46, no. 2, pp. 255–308, 2009.

[19] M. W. Reimann, M. Nolte, M. Scolamiero, K. Turner, R. Perin, G. Chindemi, P. Dłotko, R. Levi, K. Hess, and H. Markram, "Cliques of neurons bound into cavities provide a missing link between structure and function," *Frontiers in Computational Neuroscience*, vol. 11, p. 48, 2017.

[20] 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, p. 20140873, 2014.

[21] 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, no. 3, pp. 674–694, 2019.

[22] H. Lee, M. K. Chung, H. Kang, and D. S. Lee, "Hole detection in metabolic connectivity of alzheimer's disease using k- laplacian," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 297–304, Springer, 2014.

[23] H. Lee, M. K. Chung, H. Kang, B.-N. 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*, pp. 302–309, Springer, 2011.

[24] H. Edelsbrunner, J. Harer, *et al.*, "Persistent homology-a survey," *Contemporary Mathematics*, vol. 453, pp. 257–282, 2008.

[25] A. E. Sizemore, J. E. Phillips Cremins, 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.

[26] G. Petri, M. Scolamiero, I. Donato, and F. Vaccarino, "Topological strata of weighted complex networks," *PloS One*, vol. 8, no. 6, p. e66506, 2013.

[27] T. Songdechakraiwut and M. K. Chung, "Topological learning for brain networks," *arXiv preprint arXiv:2012.00675*, 2020.

[28] M. K. Chung, S. G. Huang, A. Gritsenko, L. Shen, and H. Lee, "Statistical inference on the number of cycles in brain networks," in *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, pp. 113–116, IEEE, 2019.

[29] L. Wasserman, "Topological data analysis," *arXiv preprint arXiv:1609.08227*, 2016.

[30] 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, no. 12, pp. 2267–2277, 2012.

[31] H. Lee, M. K. Chung, H. Kang, H. Choi, Y. K. Kim, and D. S. Lee, "Abnormal hole detection in brain connectivity by kernel density of persistence diagram and Hodge Laplacian," in *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pp. 20–23, IEEE, 2018.

[32] M. Chung, J. Hanson, J. Ye, R. Davidson, and S. Pollak, "Persistent homology in sparse regression and its application to brain morphometry," *IEEE Transactions on Medical Imaging*, vol. 34, pp. 1928–1939, 2015.

[33] P. G. Lind, M. C. Gonzalez, and H. J. Herrmann, "Cycles and clustering in bipartite networks," *Physical Review E*, vol. 72, no. 5, p. 056127, 2005.

[34] H. Lee, K. H. Chung, M.K., and D. Lee, "Hole detection in metabolic connectivity of Alzheimer's disease using k-Laplacian," in *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), Lecture Notes in Computer Science*, pp. 297–304, 2014.

[35] A. Sizemore, C. Giusti, A. Kahn, J. Vettel, R. Betzel, and D. Bassett, "Cliques and cavities in the human connectome," *Journal of computational neuroscience*, vol. 44, pp. 115–145, 2018.

[36] M. Chung, S.-G. Huang, A. Gritsenko, L. Shen, and H. Lee, "Statistical inference on the number of cycles in brain networks," in *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, pp. 113–116, IEEE, 2019.

[37] R. Tarjan, "Depth-first search and linear graph algorithms," *SIAM Journal on Computing*, vol. 1, no. 2, pp. 146–160, 1972.

[38] C. Chen and D. Freedman, "Measuring and computing natural generators for homology groups," *Computational Geometry*, vol. 43, no. 2, pp. 169–181, 2010.

[39] F. R. Chung and F. C. Graham, *Spectral graph theory*. No. 92, American Mathematical Society., 1997.

[40] H. Edelsbrunner and J. Harer, *Computational Topology: an introduction*. American Mathematical Society., 2010.

[41] K. Xia and G. W. Wei, "Persistent homology analysis of protein structure, flexibility, and folding," *International Journal for Numerical Methods in Biomedical Engineering*, vol. 30, no. 8, pp. 814–844, 2014.

[42] T. Songdechakraiwut, L. Shen, and M. Chung, "Topological learning and its application to multimodal brain network integration," *Medical Image Computing and Computer Assisted Intervention (MICCAI)*, pp. 166–176, 2021.

[43] Y. Mileyko, S. Mukherjee, and J. Harer, "Probability measures on the space of persistence diagrams," *Inverse Problems*, vol. 27, no. 12, p. 124007, 2011.

[44] L. Mi, W. Zhang, X. Gu, and Y. Wang, "Variational wasserstein clustering," in *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 322–337, 2018.

[45] L. Mi, W. Zhang, and Y. Wang, "Regularized wasserstein means for aligning distributional data," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, pp. 5166–5173, 2020.

[46] Z. Meng and K. Xia, "Persistent spectral–based machine learning (PerSpect ML) for protein-ligand binding affinity prediction," *Science Advances*, vol. 7, no. 19, p. eabc5329, 2021.

[47] D. Horak and J. Jost, "Spectra of combinatorial Laplace operators on simplicial complexes," *Advances in Mathematics*, vol. 244, pp. 303–336, 2013.

[48] S. Barbarossa and S. Sardellitti, "Topological signal processing over simplicial complexes," *IEEE Transactions on Signal Processing*, vol. 68, pp. 2992–3007, 2020.

[49] M. T. Schaub, A. R. Benson, P. Horn, G. Lippner, and A. Jadbabaie, "Random walks on simplicial complexes and the normalized hodge 1-laplacian," *SIAM Review*, vol. 62, no. 2, pp. 353–391, 2020.

[50] S. Mukherjee and J. Steenbergen, "Random walks on simplicial complexes and harmonics," *Random structures & Algorithms*, vol. 49, no. 2, pp. 379–405, 2016.

[51] L. H. Lim, "Hodge Laplacians on graphs," *SIAM Review*, vol. 62, no. 3, pp. 685–715, 2020.

[52] M. K. Chung, L. Xie, S. G. Huang, Y. Wang, J. Yan, and L. Shen, "Rapid acceleration of the permutation test via transpositions," in *International Workshop on Connectomics in Neuroimaging*, pp. 42–53, Springer, 2019.

[53] P. M. Thompson, T. D. Cannon, K. L. Narr, T. Van Erp, V. P. Poutanen, M. Huttunen, J. Lönnqvist, C. G. Standertskjöld Nordenstam, J. Kaprio, M. Khaledy, *et al.*, "Genetic influences on brain structure," *Nature Neuroscience*, vol. 4, no. 12, pp. 1253–1258, 2001.

[54] A. Zalesky, A. Fornito, I. H. Harding, L. Cocchi, M. Yücel, C. Pantelis, and E. T. Bullmore, "Whole-brain anatomical networks: does the choice of nodes matter?," *Neuroimage*, vol. 50, no. 3, pp. 970–983, 2010.

[55] T. E. Nichols and A. P. Holmes, "Nonparametric permutation tests for functional neuroimaging: a primer with examples," *Human Brain Mapping*, vol. 15, no. 1, pp. 1–25, 2002.

[56] A. M. Winkler, G. R. Ridgway, G. Douaud, T. E. Nichols, and S. M. Smith, "Faster permutation inference in brain imaging," *Neuroimage*, vol. 141, pp. 502–516, 2016.

[57] K. Turner, Y. Mileyko, S. Mukherjee, and J. Harer, "Fréchet means for distributions of persistence diagrams," *Discrete & Computational Geometry*, vol. 52, pp. 44–70, 2014.

[58] M. K. Chung, A. Qiu, S. Seo, and H. K. Vorperian, "Unified heat kernel regression for diffusion, kernel smoothing and wavelets on manifolds and its application to mandible growth modeling in ct images," *Medical image analysis*, vol. 22, no. 1, pp. 63–76, 2015.

[59] W. Kim, N. Adluru, M. Chung, O. Okonkwo, S. Johnson, B. Bendlin, and V. Singh, "Multi-resolution statistical analysis of brain connectivity graphs in preclinical Alzheimer's disease," *NeuroImage*, vol. 118, pp. 103–117, 2015.

[60] H. Lee, M. Chung, H. Kang, B.-N. Kim, and D. Lee, "Computing the shape of brain networks using graph filtration and Gromov-Hausdorff metric," *MICCAI, Lecture Notes in Computer Science*, vol. 6892, pp. 302–309, 2011.

[61] M. Chung, H. Lee, A. Gritsenko, A. DiChristofano, D. Pluta, H. Ombao, and V. Solo, "Topological brain network distances," *arXiv preprint arXiv:1809.03878*, 2018.

[62] D. C. Van Essen, K. Ugurbil, E. Auerbach, D. Barch, T. E. Behrens, R. Bucholz, A. Chang, L. Chen, M. Corbetta, S. W. Curtiss, *et al.*, "The Human Connectome Project: a data acquisition perspective," *Neuroimage*, vol. 62, no. 4, pp. 2222–2231, 2012.

[63] M. F. Glasser, S. N. Sotiropoulos, J. A. Wilson, T. S. Coalson, B. Fischl, J. L. Andersson, J. Xu, S. Jbabdi, M. Webster, J. R. Polimeni, *et al.*, "The minimal preprocessing pipelines for the Human Connectome Project," *Neuroimage*, vol. 80, pp. 105–124, 2013.

[64] J. L. Andersson, S. Skare, and J. Ashburner, "How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging," *Neuroimage*, vol. 20, no. 2, pp. 870–888, 2003.

[65] M. Jenkinson and S. Smith, "A global optimisation method for robust affine registration of brain images," *Medical Image Analysis*, vol. 5, no. 2, pp. 143–156, 2001.

[66] M. F. Glasser and D. C. Van Essen, "Mapping human cortical areas in vivo based on myelin content as revealed by T1-and T2-weighted MRI," *Journal of Neuroscience*, vol. 31, no. 32, pp. 11597–11616, 2011.

[67] N. Tzourio Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, and M. Joliot, "Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain," *Neuroimage*, vol. 15, no. 1, pp. 273–289, 2002.

[68] J. D. Power, K. A. Barnes, A. Z. Snyder, B. L. Schlaggar, and S. E. Petersen, "Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion," *Neuroimage*, vol. 59, no. 3, pp. 2142–2154, 2012.

[69] S. G. Huang, S. B. 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, p. 108480, 2020.

[70] T. Koscik, D. O'Leary, D. Moser, N. Andreasen, and P. Nopoulos, "Sex differences in parietal lobe morphology: relationship to mental rotation performance," *Brain and cognition*, vol. 69, pp. 451–459, 2009.

[71] Y. Xu and M. Lindquist, "Dynamic connectivity detection: an algorithm for determining functional connectivity change points in fMRI data," *Frontiers in Neuroscience*, vol. 9, p. 285, 2015.

[72] Y. Zang, T. Jiang, Y. Lu, Y. He, and L. Tian, "Regional homogeneity approach to fmri data analysis," *Neuroimage*, vol. 22, pp. 394–400, 2004.

[73] L. Rubin, L. Yao, S. Keedy, J. Reilly, J. Bishop, C. Carter, H. Pournajafi-Nazarloo, L. Drogos, C. Tamminga, and G. Pearlson, "Sex differences in associations of arginine vasopressin and oxytocin with resting-state functional brain connectivity," *Journal of neuroscience research*, vol. 95, pp. 576–586, 2017.

[74] D. Fisher, D. Bennett, and H. Dong, "Sexual dimorphism in predisposition to Alzheimer's disease," *Neurobiology of aging*, vol. 70, pp. 308–324, 2018.