

Learning Higher-Order Interactions in Brain Networks via Topological Signal Processing

Breno C. Bispo¹, Stefania Sardellitti², Fernando A. N. Santos³, Juliano B. Lima¹

¹Dept. of Electronics and Systems, Federal University of Pernambuco, Recife, Brazil

²Engineering and Sciences Dept., Universitas Mercatorum, Rome, Italy

³Dutch Institute for Emergent Phenomena, KdVI, University of Amsterdam, Amsterdam, The Netherlands

e-mail: breno.bispo@ufpe.br, stefania.sardellitti@unimercatorum.it, f.a.nobregasantos@uva.nl, juliano.lima@ufpe.br

Abstract—Our goal in this paper is to leverage the potential of the topological signal processing (TSP) framework for analyzing brain networks. Representing brain data as signals over simplicial complexes allows us to capture higher-order relationships within brain regions of interest (ROIs). Here, we focus on learning the underlying brain topology from observed neural signals using two distinct inference strategies. The first method relies on higher-order statistical metrics to infer multiway relationships among ROIs. The second method jointly learns the brain topology and sparse signal representations, of both the solenoidal and harmonic components of the signals, by minimizing the total variation along triangles and the data-fitting errors. Leveraging the properties of solenoidal and irrotational signals, and their physical interpretations, we extract functional connectivity features from brain topologies and uncover new insights into functional organization patterns. This allows us to associate brain functional connectivity patterns of conservative signals with well-known functional segregation and integration properties. Our findings align with recent neuroscience research, suggesting that our approach may offer a promising pathway for characterizing the higher-order brain functional connectivity.

Index Terms—Brain networks, topology learning, topological signal processing, simplicial complexes.

I. INTRODUCTION

Advances in neuroimaging technologies, particularly functional magnetic resonance imaging (fMRI), have enabled non-invasive, high-resolution mapping of brain activity. This has allowed for indirect explorations of anatomical and functional interactions between distinct brain regions of interest (ROIs), improving our understanding of neural activity and facilitating the development of predictive models of brain function. Modeling brain networks as graphs has become a natural approach for representing these complex systems, where nodes correspond to ROIs and links capture functional and structural connectivity [1]. Although graph-based techniques have been widely used to analyze neural signals at the nodes of brain networks [2], they are inherently limited to modeling only pairwise (dyadic) relationships, which constrains their capacity to capture and characterize higher-order structures and functional interactions.

This work was funded by the FIN-RIC Project TSP-ARK, supported by Universitas Mercatorum under grant No. 20-FIN/RIC. Additional support was provided by CAPES (88881.311848/2018-01, 88887.899136/2023-00), CNPq (442238/2023-1, 312935/2023-4, 405903/2023-5), and FACEPE (APQ-1226-3.04/22).

On the other hand, recent studies [3], [4] have leveraged multivariate interdependence metrics to infer higher-order interactions (HOIs), i.e., simultaneous inter-dependence among three or more nodes, aiming to uncover novel functional connectivity (FC) patterns that remain elusive to solely pairwise-based approaches. To efficiently characterize HOIs, brain networks have been represented on more structured topological spaces than graphs, such as hypergraphs [5].

More recently, hypergraph signal processing (HGSP) has shown potential applications in characterizing higher-order FC [6]. However, the interpretation of its algebraic operators remains a challenge in extracting meaningful insights into the dynamics of HOIs. Nevertheless, a particular class of hypergraphs, such as simplicial complexes, offers a rich algebraic structure enhancing data interpretability and features extraction. Built upon this algebraic structure, a recently developed framework Topological Signal Processing (TSP) [7] has emerged as a powerful tool for analyzing and processing signals defined over simplicial complexes. This framework promises to enhance brain data analysis and interpretability by enabling the identification of critical cycles and detection of brain regions with significant activity.

One of the main challenges in processing information over simplicial complexes is learning their underlying structures from data in a way that accurately reflects their features. In this paper, we address this challenge by developing strategies to learn the brain simplicial complex (SCX) topology from real brain datasets using two different approaches. The first approach leverages (non-)dyadic statistical metrics to infer the presence of edges and filled triangles, capturing pairwise and triple-wise interdependencies between ROIs of the brain SCX. The second approach aims to jointly determine the brain SCX topology and sparse signal representations by minimizing the data-fitting error and the total variation of the signals. Analyzing HOIs in brain networks through the lens of TSP, we leverage key operators that provide physical interpretations of signals over a SCX, such as the divergence and curl operators. We analyze resting-state fMRI (rs-fMRI) signals to identify the primary sources and sinks of information in the brain, as well as significant circulations among ROIs within our learned topologies. Our findings reveal that the mean divergence patterns in the brain align with established neuroscience literature. Additionally, the observed conservative circulations remind the

idea of functional segregation and integration, reflecting key organizational principles of brain functioning.

II. OVERVIEW OF TOPOLOGICAL SIGNAL PROCESSING

In this section we briefly recall the basic TSP principles [7]. Let us consider a finite set of N nodes $\mathcal{V} = \{v_0, v_1, \dots, v_{N-1}\}$. A k -simplex $\tau^k = \{v_{j_0}, \dots, v_{j_k}\}$ is an unordered subset of \mathcal{V} with $k + 1$ distinct nodes. A face of the k -simplex $\tau^k = \{v_{j_0}, \dots, v_{j_k}\}$ is a $(k - 1)$ -simplex $\tau^{k-1} = \{v_{j_0}, \dots, v_{j_{n-1}}, v_{j_{n+1}}, \dots, v_{j_k}\}$ for some $0 \leq n \leq k$. Therefore, a node is a 0-simplex, an edge a 1-simplex, a triangle a 2-simplex, and so on. A simplicial complex (SCX) \mathcal{X}^K of order K is a finite set of k -simplices τ^k for $k \in \{0, 1, \dots, K\}$, that is closed under the inclusion property, i.e., if $\tau_i^k \in \mathcal{X}^K$, all faces of τ_i^k are in \mathcal{X}^K . The order of a simplicial complex is the maximum dimension of its simplices. The topological structure of a simplicial complex is captured by the incidence relations between its simplices. Specifically, two simplices of order k , τ_i^k and τ_j^k are upper adjacent in \mathcal{X} if both are faces of a simplex of order $k + 1$ or lower adjacent if they have a common face of order $k - 1$. A $(k - 1)$ -face τ_j^{k-1} of a k -order simplex τ_i^k is called a boundary element of τ_i^k , denoted as $\tau_j^{k-1} \prec_b \tau_i^k$. Given an orientation of the simplices [7], we denote as $\tau_i^{k-1} \sim \tau_j^k$ two coherently oriented simplices. Then, the structures of the simplicial complex is described by the incidence matrices \mathbf{B}_k , whose entries establish which $(k - 1)$ -simplices are incident to which k -simplices.

Specifically, we have $B_k(i, j) = 1$ (or $B_k(i, j) = -1$) if $\tau_i^{k-1} \prec_b \tau_j^k$ and $\tau_i^{k-1} \sim \tau_j^k$ (or $\tau_i^{k-1} \not\sim \tau_j^k$), while $B_k(i, j) = 0$ if $\tau_i^{k-1} \not\prec_b \tau_j^k$. An important property of the incidence matrices is that it always holds $\mathbf{B}_k \mathbf{B}_{k+1} = \mathbf{0}$. In this work, we restrict our attention w.l.o.g. to simplicial complexes \mathcal{X}^2 of order $K = 2$ due to practical application and computational feasibility. Let us denote with \mathcal{V} , \mathcal{E} and \mathcal{T} , the set of vertices, edges and triangles, respectively, with cardinalities $|\mathcal{V}| = N$, $|\mathcal{E}| = E$, $|\mathcal{T}| = T$. Given the node-edge incidence matrix $\mathbf{B}_1 \in \mathbb{R}^{N \times E}$ and the edge-triangles incidence matrix $\mathbf{B}_2 \in \mathbb{R}^{E \times T}$, the topological structure of \mathcal{X}^2 is described by the graph Laplacian $\mathbf{L}_0 = \mathbf{B}_1 \mathbf{B}_1^\top \in \mathbb{R}^{N \times N}$, which encodes the upper adjacencies of 0-simplices, and by the first-order Hodge Laplacian matrix $\mathbf{L}_1 = \mathbf{B}_1^\top \mathbf{B}_1 + \mathbf{B}_2 \mathbf{B}_2^\top \in \mathbb{R}^{E \times E}$, where $\mathbf{L}_{1,\ell} = \mathbf{B}_1^\top \mathbf{B}_1$ and $\mathbf{L}_{1,u} = \mathbf{B}_2 \mathbf{B}_2^\top$ are the lower and upper Laplacian matrices, respectively. They encode the lower and upper adjacencies of all 1-simplices, respectively [7].

Topological signals representation. Denoting by N_k the number of k -order simplices, a k -simplicial signal $\mathbf{s}^k = [s^k(1), \dots, s^k(N_k)]^\top \in \mathbb{R}^{N_k}$ is a map $s^k : \mathcal{S}^k \rightarrow \mathbb{R}^{N_k}$, where the entry $s^k(i)$ corresponds to the i -th k -simplex [7] and \mathcal{S}^k is the subset of simplices of order k . We denote with $\mathbf{s}^0 \in \mathbb{R}^N$, $\mathbf{s}^1 \in \mathbb{R}^E$, $\mathbf{s}^2 \in \mathbb{R}^T$ the signals over the N nodes, E edges and T triangles of \mathcal{X}^2 , respectively. Suitable bases to represent k -simplicial signals are the eigenvectors of the k -th order Hodge Laplacian matrix [7]. Let us consider the eigendecomposition $\mathbf{L}_k = \mathbf{U}_k \mathbf{\Lambda}_k \mathbf{U}_k^\top$ with $\mathbf{U}_k \in \mathbb{R}^{N_k \times N_k}$ the eigenvectors matrix and $\mathbf{\Lambda}_k$ the diagonal matrix with entries the associated eigenvalues $\lambda_{k,i}$, $i = 1, \dots, N_k$. The

Simplicial Fourier Transform (SFT) [7] of \mathbf{s}^k can be defined as $\hat{\mathbf{s}}^k = \mathbf{U}_k^\top \mathbf{s}^k$ and, then, the signal can be represented in terms of its SFT coefficients as $\mathbf{s}^k = \mathbf{U}_k \hat{\mathbf{s}}^k$. A signal \mathbf{s}^k is bandlimited if it admits a sparse representation on the eigenvectors bases.

One of the key aspects in TSP that bridges mathematical concepts to physical interpretations is the possibility to decompose the space of a k -simplicial signal \mathbb{R}^{N_k} into the direct sum of three orthogonal subspaces, according to the so called Hodge decomposition:

$$\mathbb{R}^{N_k} \equiv \text{img}(\mathbf{B}_k^\top) \oplus \ker(\mathbf{L}_k) \oplus \text{img}(\mathbf{B}_{k+1}). \quad (1)$$

In the domain of the edge flows (1-simplices), which can be considered as a discrete counterpart of vector fields [7], $\text{img}(\mathbf{B}_1^\top)$, $\text{img}(\mathbf{B}_2)$ and $\ker(\mathbf{L}_1)$ are the gradient, curl and harmonic subspaces with dimensions N_G , N_C and N_H , respectively. Therefore, we can partition the eigenvectors of \mathbf{U}_1 into three groups of eigenvectors: i) the gradient eigenvectors, $\mathbf{U}_G \in \mathbb{R}^{E \times N_G}$, consisting of eigenvectors of $\mathbf{L}_{1,\ell}$ (associated with non-zeros eigenvalues) that generates the gradient subspace $\text{img}(\mathbf{B}_1^\top)$; ii) the curl (or solenoidal) eigenvectors $\mathbf{U}_C \in \mathbb{R}^{E \times N_C}$, i.e., the eigenvectors of $\mathbf{L}_{1,u}$, associated with non-zeros eigenvalues and spanning the curl subspace $\text{img}(\mathbf{B}_2)$; and iii) the harmonic eigenvectors $\mathbf{U}_H \in \mathbb{R}^{E \times N_H}$ spanning the kernel of \mathbf{L}_1 . The dimension of $\ker(\mathbf{L}_1)$, known as the Betti number of order 1, represents a key topological invariant of the SCXs, counting the number of holes (unfilled triangles) within the complex.

From (1), each edge signal observed over the SCX admits the following Hodge decomposition

$$\mathbf{s}^1 = \mathbf{B}_1^\top \mathbf{s}^0 + \mathbf{B}_2 \mathbf{s}^2 + \mathbf{s}_H, \quad (2)$$

with $\mathbf{s}_{irr}^1 = \mathbf{B}_1^\top \mathbf{s}^0$ being the irrotational edge signal, i.e., the gradient of the node signal observed at the vertices of each edge, $\mathbf{s}_s^1 = \mathbf{B}_2 \mathbf{s}^2$ the solenoidal signal reflecting the circulation (curl) along triangles, and, finally, \mathbf{s}_H the harmonic signal component. Two key operators that provide meaningful physical interpretations related to the signals observed over a SCX are the divergence and curl operators [7]. The *divergence operator*, defined as

$$\text{div}(\mathbf{s}^1) = \mathbf{B}_1 \mathbf{s}^1, \quad (3)$$

is a node signal, revealing which nodes act as sources or sinks of information within a SCX. The *curl operator*, given by

$$\text{curl}(\mathbf{s}^1) = \mathbf{B}_2^\top \mathbf{s}^1, \quad (4)$$

quantifies the circulation of flows along the triangles. Note that solenoidal signals are divergence-free, since it holds $\mathbf{B}_1 \mathbf{s}_s^1 = \mathbf{0}$, while irrotational signals are curl-free since we have $\mathbf{B}_2^\top \mathbf{s}_{irr}^1 = \mathbf{0}$.

III. STATISTICAL LEARNING OF THE BRAIN TOPOLOGY

In this section, we propose a method based on statistical metrics to infer a 2-order brain SCX from real resting-state fMRI (rs-fMRI) signals, where each rs-fMRI time series, in a brain dataset with N ROIs, is treated as a node signal. To construct the skeleton of a brain SCX (i.e., the graph)

from node signals, we leverage a conventional pairwise-based metric commonly used in neuroscience, the Pearson correlation. Among the various multivariate interaction metrics used in neuroscience to infer HOIs, including multivariate cumulants [8], time series co-fluctuations [9], and multivariate information-theoretic measures [4], [10], we draw inspiration from the latter. These information-theoretic approaches have been shown to reveal unique higher-order organizational patterns in the brain that may remain elusive by solely pairwise-based methods [11], [12]. Specifically, within our context, the most suitable metric for determining the weights of 2-simplices (triangles), representing the strength of three-way simultaneous interdependence, seems to be the *total correlation* (TC), defined as [10]

$$TC_{\mathbf{x}_i \mathbf{x}_j \mathbf{x}_k} = H(\mathbf{x}_i) + H(\mathbf{x}_j) + H(\mathbf{x}_k) - H(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k) \quad (5)$$

where \mathbf{x}_i , \mathbf{x}_j and \mathbf{x}_k are random variables vectors. In our context, these random vectors represent rs-fMRI time series of ROIs observed over the distinct graph vertices v_i, v_j, v_k for $i, j, k \in \{1, \dots, N\}$, $H(\cdot)$ is the Shannon entropy, while $H(\cdot, \cdot)$ and $H(\cdot, \cdot, \cdot)$ are the joint-entropies of the arguments. This approach is justified by the complementary property shared between (non-)pairwise correlations that characterize redundant-dominated structures [10]. Applying this metric to a set of N rs-fMRI time series, we can compute all HOI weights.

Using the Human Connectome Project (HCP) public dataset (<http://www.humanconnectome.org/>), we select 226 sets of $N = 116$ rs-fMRI time series obtained from a cohort of 126 unrelated healthy individuals (young adults aged 21-35 consisting of 60 males and 66 females). Following the Schaefer's brain atlas reference used in Fig. 1(a) of [6], each ROI is mapped to its corresponding functional subnetwork (FS) and color-coded accordingly.

Our goal is to investigate the second-order brain SCX that best represents the mean individual in the HCP dataset. To accomplish this, we compute the absolute values of the entries in the correlation matrices derived from the participants' rs-fMRI time series (available at <https://doi.org/10.5281/zenodo.6770120>) and then calculate their average. Since some studies suggest that anticorrelations may carry meaningful biological information [13], we include negative correlation values by taking their absolute values to preserve these potential insights. Moreover, given the absence of a standardized thresholding procedure [13], we retain only the top 5% of the strongest connections (resulting in a network with 333 edges) for practical purposes.

Subsequently, we compute all 253460 triangle weights for each set of rs-fMRI time series using (5). To enable population-level comparisons, we compute the *z-score* of all triangle weights at the individual level and then take their average. From the mean individual triangle weights, we select the 200 strongest HOIs as filled triangles in the mean brain SCX. This number is chosen to ensure a fair comparison with the learned brain topology introduced in the following section. To maintain the inclusion property of our brain SCX, we incorporate edges that are not subsets of the filled triangles.

This results in a learned 2-order brain SCX with $N = 116$ nodes, $E = 491$ edges, and $T = 200$ filled triangles. Furthermore, after selecting orientations for simplices, we build the incidence matrices \mathbf{B}_1 and \mathbf{B}_2 . It is important to emphasize that the orientation of edges and triangles for defining the incidence matrices is merely a formalism and does not imply causality in information exchange between brain ROIs, as we are considering undirected SCXs. In order to provide an overall analysis of the brain signals over the learned mean brain SCX, we first compute the mean individual node signals by z-scoring the rs-fMRI time series at the individual level and then averaging them across all volunteers' time series, resulting in a set of mean node signals given by the matrix $\mathbf{S}^0 = [\mathbf{s}^0(1), \dots, \mathbf{s}^0(M)] \in \mathbb{R}^{N \times M}$, where $M = 2400$ is the number of time samples of the entire rs-fMRI recording. Given the node signals, we derive the edge signals by computing the 1-order instantaneous co-fluctuation magnitude. This is obtained by taking the absolute value of element-wise multiplication of pairs of z-scored node signals, a common approach in neuroscience for assessing topological organizations from fMRI signals [9]. As a result, the edge signals are represented by the matrix $\mathbf{S}^1 = [\mathbf{s}^1(0), \dots, \mathbf{s}^1(M)] \in \mathbb{R}^{E \times M}$.

Given the brain SCX and edge signals, we perform a comprehensive analysis of averaged edge signal divergence to identify the primary sources and sinks of information throughout the mean individual rs-fMRI recording. For each edge signal $\mathbf{s}^1(m)$ we compute its divergence using (3) and compute their average across all M time samples given by $\overline{\text{div}}(\mathbf{S}^1) = \sum_{m=1}^M \text{div}(\mathbf{s}^1(m))/M$. In Fig. 1(a) we show a histogram illustrating the distribution of $\overline{\text{div}}(\mathbf{S}^1)$ across nodes. In Fig. 1(b), the values of $\overline{\text{div}}(\mathbf{S}^1)$ are visualized as node signals over the ROIs, where colors indicate their respective values, and node sizes are scaled according to their magnitudes. Analyzing the most negative values of $\overline{\text{div}}(\mathbf{S}^1)$, we can infer that the overall sink of information originates from the left hemisphere of the brain. Specifically, the primary sinks are located in ROIs from the left side of the visual (VIS) subnet, followed by ROIs from default mode network (DMN) and dorsal attention (DA) subnets. On the other hand, exploring the most positive values of $\overline{\text{div}}(\mathbf{S}^1)$, we can infer that the right hemisphere serves as a source of information, with the primary sources located in ROIs from VIS and ventral attention (VA) subnets, followed by ROIs from DA and DMN subnets. These results exhibit patterns similar to recent findings on the characterization of brain wave dynamics [14].

Furthermore, we conduct an analysis on the mean circulations of flows along the triangles of the inferred brain SCX. Similarly to divergence analysis of the edge signals $\mathbf{s}^1(m)$, we compute their curl counterparts using (4) and compute their average across time as $\overline{\text{curl}}(\mathbf{S}^1) = \sum_{m=1}^M \text{curl}(\mathbf{s}^1(m))/M$. Fig. 1(c) depicts a histogram illustrating the component distribution of $\overline{\text{curl}}(\mathbf{S}^1)$. In this study, we focus on conservative circulations, i.e., the smallest triangle circulation values in magnitude, highlighting its interpretation from a neuroscience point of view. Fig. 1(d) illustrates the 20 weakest circulation values (in magnitude) over the filled triangles, where the cones

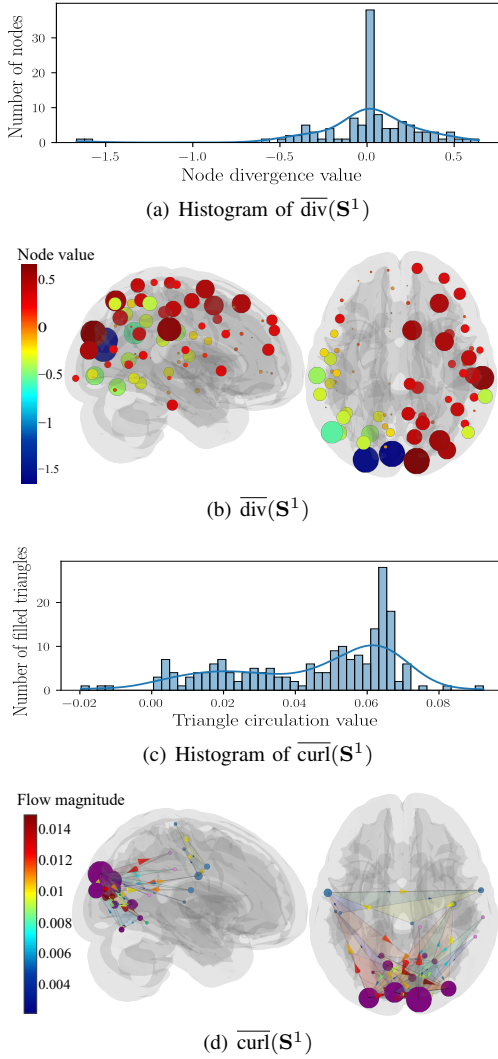


Fig. 1: Histograms and 3D brain representations of the averaged divergence and circulation of the edge signals.

along the edges indicate their (counter-)clockwise orientation. The cone size and color are proportional to the flow magnitude over each triangle. The ROI size is scaled according to the number of participation among the triangles, while its color corresponds to its FS. We can easily infer that ROIs from the VIS subnet (in purple) play a crucial role in the overall context, interacting with ROIs from somatomotor (SM) subnet (in steelblue). From a neuroscience perspective, the circulation patterns involving the VIS and SM subnets simultaneously can be attributed to functional integration, which characterizes global communication between distinct brain modules [15]. Specifically, this kind of integration can be explained by the underlying redundant-dominated structure present in these regions, as highlighted in a recent study [3]. On the other hand, circulations that involve solely ROIs from either the SM or VIS subnet reflect the concept of functional segregation, which refers to specialized neuronal processing within distinct modules [15].

IV. JOINT LEARNING OF TOPOLOGY AND SPARSE SIGNALS

Simplicial complex learning strategies based on edge signal smoothness have been proposed in [7], [16]. In this section, we design a strategy to jointly learn the simplicial complex topology and sparse spectral representations of both the solenoidal and harmonic signals from brain data. Assuming the graph topology given, our goal is to infer a topological structure where solenoidal and harmonic signals admit sparse spectral representations, minimizing jointly the total signal circulations and the data-fitting error. Denoting with \hat{s}_s^1, \hat{s}_H^1 , the SFT coefficients of, respectively, the solenoidal and harmonic signals, their spectral representations can be written as $s_s^1 = U_C \hat{s}_s^1$, $s_H^1 = U_H \hat{s}_H^1$. Given a set of M edge signals $\mathbf{Y}^1 = [\mathbf{y}^1(1), \dots, \mathbf{y}^1(M)] \in \mathbb{R}^{E \times M}$, we need to learn from data the incidence matrix \mathbf{B}_2 , or equivalently, the upper Laplacian matrix $\mathbf{L}_{1,u}$. Denoting with T the number of possible triangles in the complex, this matrix can be written in the equivalent form $\mathbf{L}_{1,u} = \sum_{n=1}^T q_n \mathbf{b}_n \mathbf{b}_n^T$ where \mathbf{b}_n are the $E \times 1$ -dimensional columns of the matrix \mathbf{B}_2 , while $q_n \in \{0, 1\}$ are binary coefficients selecting the filled triangles according to the chosen optimization rule. Then, similarly to the learning method in [7], we first check if data have components on the solenoidal subspace by projecting the observed edge flow signals onto the space orthogonal to the one spanned by the irrotational component as $\mathbf{y}_{sH}^1(m) = (\mathbf{I} - \mathbf{U}_G \mathbf{U}_G^T) \mathbf{y}^1(m)$, $\forall m$. Then, we calculate the Frobenius norm of the matrix $\mathbf{Y}_{sH}^1 = [\mathbf{y}_{sH}^1(1), \dots, \mathbf{y}_{sH}^1(M)]$ and if this norm is higher than a given threshold, we proceed to learn the upper Laplacian matrix. Therefore, defining the matrices $\hat{\mathbf{S}}_s^1 = [\hat{s}_s^1(1), \dots, \hat{s}_s^1(M)]$, $\hat{\mathbf{S}}_H^1 = [\hat{s}_H^1(1), \dots, \hat{s}_H^1(M)]$, we formulate the following non-convex optimization problem

$$\begin{aligned}
 & \min_{\hat{\mathbf{S}}_s^1, \hat{\mathbf{S}}_H^1, \mathbf{q}} \sum_{n=1}^T q_n \|\mathbf{Y}_{sH}^1 \mathbf{b}_n\|_F^2 + \beta \|\mathbf{Y}_{sH}^1 - \mathbf{U}_C \hat{\mathbf{S}}_s^1 - \mathbf{U}_H \hat{\mathbf{S}}_H^1\|_F^2 \\
 & \text{s.t.} \quad \mathbf{U}_C, \mathbf{U}_H, \Lambda_s \succeq \mathbf{0} \\
 & \quad a) \quad \|\hat{s}_s^1(m)\|_1 \leq \alpha_1, \|\hat{s}_H^1(m)\|_1 \leq \alpha_2, \forall m, \\
 & \quad b) \quad \sum_{n=1}^T q_n \mathbf{b}_n \mathbf{b}_n^T \mathbf{U}_C = \mathbf{U}_C \Lambda_s, \Lambda_s(i, j) = 0, \forall i \neq j, \\
 & \quad c) \quad \left(\sum_{n=1}^T q_n \mathbf{b}_n \mathbf{b}_n^T + \mathbf{B}_1^T \mathbf{B}_1 \right) \mathbf{U}_H = \mathbf{0}, \\
 & \quad d) \quad \|\mathbf{q}\|_0 \leq q^*, \quad \mathbf{q} \in \{0, 1\}^T, \quad (\mathcal{P}_t)
 \end{aligned} \tag{6}$$

where the first term in the objective function is the energy of the signals circulation along all triangles, while the second term is the data-fitting error controlled by the positive coefficient $\beta > 0$. The l_1 -norm constraints *a)* enforce spectral sparsity of the signals through the positive coefficients α_1, α_2 , while *b)* and *c)* constrain the matrices \mathbf{U}_C and \mathbf{U}_H to be the eigenvectors associated with the non-zeros eigenvalues of $\mathbf{L}_{1,u}$, and the eigenvectors spanning the kernel of \mathbf{L}_1 , respectively. Finally, constraint *d)* imposes that the number of filled triangles doesn't exceed a maximum number q^* . Problem in (6) is non-convex due to the constraints *b)*, *c)* and *d)*, therefore we propose an alternating iterative suboptimal

strategy solving it as a sequence of convex-subproblems. Specifically, by varying the value of the required number of filled cells q^* from 1 to the maximum value T , we alternate between the solutions of two convex optimization problems and then select the number q^* minimizing the data-fitting error. Specifically, we iterate for $i = 1, \dots, T$ between the following two steps:

S.1 Solve problem \mathcal{P}_t with respect to \mathbf{q} with $\beta = 0$, $q^* = i$ and considering only constraints b), c) and d). This problem admits the closed form solution obtained by computing $a_n = \|\mathbf{Y}_{sH}^{1T} \mathbf{b}_n\|_F^2$, sorting these coefficients in increasing order and selecting the indices $\mathcal{T}_t = \{i_1, i_2, \dots, i_{q^*}\}$ of the q^* lowest coefficients a_n . Hence, we get $\hat{\mathbf{L}}_{1,u}^i = \sum_{n \in \mathcal{T}_t} \mathbf{b}_n \mathbf{b}_n^\top$;

S.2 Given $\hat{\mathbf{L}}_{1,u}^i$ find the eigenvectors matrices \mathbf{U}_C^i and, using the estimated 1-order Laplacian $\hat{\mathbf{L}}_1^i = \mathbf{L}_{1,\ell} + \hat{\mathbf{L}}_{1,u}^i$, derive the eigenvectors matrix \mathbf{U}_H^i spanning its kernel. Then, solve the following convex optimization problem

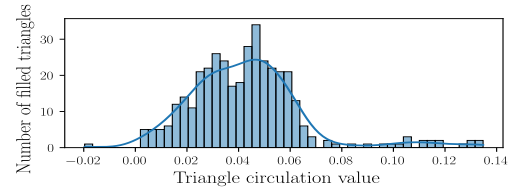
$$\begin{aligned} \min_{\hat{\mathbf{S}}_s^1, \hat{\mathbf{S}}_H^1} \quad & \|\mathbf{Y}_{sH}^1 - \mathbf{U}_C^i \hat{\mathbf{S}}_s^1 - \mathbf{U}_H^i \hat{\mathbf{S}}_H^1\|_F^2 \quad (\mathcal{P}_i) \\ \text{s.t.} \quad & a) \|\hat{\mathbf{s}}_s^1(m)\|_1 \leq \alpha_1, \|\hat{\mathbf{s}}_H^1(m)\|_1 \leq \alpha_2, \forall m. \end{aligned}$$

At each iteration i , we find the optimal data fit error $g(i) = \|\mathbf{Y}_{sH}^1 - \mathbf{U}_C^i \hat{\mathbf{S}}_s^1 - \mathbf{U}_H^i \hat{\mathbf{S}}_H^1\|_F^2$ by alternating between steps S.1 and S.2. Finally, we derive the filled triangles as $q^* = \arg \min_{i \in \{1, \dots, T\}} g(i)$ and the learned Laplacian $\hat{\mathbf{L}}_1 = \mathbf{L}_{1,\ell} + \hat{\mathbf{L}}_{1,u}^{q^*}$.

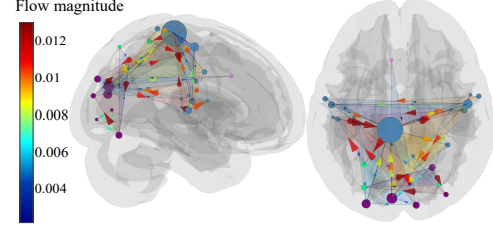
As numerical test, we apply the proposed iterative algorithm to learn the brain topology using the same brain graph and edge signals \mathbf{S}^1 from Sec. III. Hence, we get a mean brain SCX with $T = 381$ filled triangles. Subsequently, we analyze the flow circulations along the filled triangles while the divergence patterns persist, as shown in Fig. 1(b). Fig. 2(a) shows the component distribution of $\overline{\text{curl}}(\mathbf{S}^1)$, while Fig. 2(b) highlights its primary conservative circulations. Similar patterns to Fig. 1(d) are observed, including circulations involving both intra- and inter-module HOIs among ROIs from SM and VIS subnets. This reinforces the concepts of functional segregation [15] and integration, driven by the redundant-dominated structure shared between these subnets [3].

V. CONCLUSION

In this paper we characterized HOIs in the brain for the first time through the lens of the TSP framework. We developed two approaches for learning the mean brain topology from real brain datasets using higher-order statistical measures and TSP tools. By analyzing the divergence and circulation properties of brain edge signals, our study reveals that brain conservative signals, i.e., circulation signals with smallest magnitudes, reflect concepts of functional segregation and integration. For instance, both learning approaches reveal functional integration between the SM and VIS subnets, consistent with recent neuroscience findings on redundant-dominated structures in the brain.



(a) Histogram of $\overline{\text{curl}}(\mathbf{S}^1)$



(b) $\overline{\text{curl}}(\mathbf{S}^1)$

Fig. 2: Histograms and 3D brain representations of the averaged divergence and circulation of the signals estimated using the joint learning strategy.

REFERENCES

- [1] D. S. Bassett and O. Sporns, "Network neuroscience," *Nat. Neurosci.*, vol. 20, no. 3, pp. 353–364, 2017.
- [2] R. Li et al., "Graph signal processing, graph neural network and graph learning on biological data: A systematic review," *IEEE Rev. Biomed. Eng.*, vol. 16, pp. 109–135, 2023.
- [3] T. Varley, M. Pope, J. Faskowitz, and O. Sporns, "Multivariate information theory uncovers synergistic subsystems of the human cerebral cortex," *Commun. Biol.*, vol. 6, 04 2023.
- [4] F. E. Rosas, P. A. M. Mediano, M. Gastpar, and H. J. Jensen, "Quantifying high-order interdependencies via multivariate extensions of the mutual information," *Phys. Rev. E*, vol. 100, pp. 032305, Sep 2019.
- [5] F. Battiston et al., "Networks beyond pairwise interactions: Structure and dynamics," *Phys. Rep.*, vol. 874, pp. 1–92, 2020.
- [6] B. C. Bispo, F. A. N. Santos, J. R. O. Neto, and J. B. Lima, "Emergence of higher-order functional brain connectivity with hypergraph signal processing," in *2024 32nd EUSIPCO*, 2024, pp. 1332–1336.
- [7] S. Barbarossa and S. Sardellitti, "Topological signal processing over simplicial complexes," *IEEE Trans. Signal Process.*, vol. 68, pp. 2992–3007, March 2020.
- [8] R. Hindriks et al., "Higher-order functional connectivity analysis of resting-state functional magnetic resonance imaging data using multivariate cumulants," *Hum. Brain Mapp.*, vol. 45, no. 5, pp. e26663–e26663, 2024.
- [9] A. Santoro, F. Battiston, G. Petri, and E. Amico, "Higher-order organization of multivariate time series," *Nat. Phys.*, vol. 19, pp. 221–229, 2023.
- [10] N. Timme, W. Alford, B. Flecker, and J. Beggs, "Synergy, redundancy, and multivariate information measures: An experimentalist's perspective," *J. Comput. Neurosci.*, vol. 36, 07 2013.
- [11] Q. Li, "Functional connectivity inference from fMRI data using multivariate information measures," *Neural Net.*, vol. 146, pp. 85–97, 2022.
- [12] F. A. N. Santos et al., "Emergence of high-order functional hubs in the human brain," *bioRxiv*, 2023.
- [13] E. Centeno et al., "A hands-on tutorial on network and topological neuroscience," *Brain Structure and Function*, vol. 227, pp. 741–762, 2022.
- [14] J. A. Roberts et al., "Metastable brain waves," *Nat. Commun.*, vol. 10, no. 1, pp. 1056, 2019.
- [15] O. Sporns, "Network attributes for segregation and integration in the human brain," *Curr. Opin. Neurobiol.*, vol. 23, no. 2, pp. 162–171, 2013.
- [16] S. Gurugubelli and S. P. Chepuri, "Simplicial complex learning from edge flows via sparse clique sampling," in *2024 32nd EUSIPCO*. IEEE, 2024, pp. 2332–2336.