


Cliques and cavities in the human connectome

Ann E. Sizemore^{1,2} · Chad Giusti¹ · Ari Kahn^{1,3} · Jean M. Vettel^{1,3,4} ·
Richard F. Betzel¹ · Danielle S. Bassett^{1,5} 

Received: 19 May 2017 / Revised: 30 September 2017 / Accepted: 27 October 2017 / Published online: 16 November 2017
© The Author(s) 2017. This article is an open access publication

Abstract Encoding brain regions and their connections as a network of nodes and edges captures many of the possible paths along which information can be transmitted as humans

Action Editor: Abraham Zvi Snyder

✉ Danielle S. Bassett
dsb@seas.upenn.edu

Ann E. Sizemore
annsize@seas.upenn.edu

Chad Giusti
cgiusti@seas.upenn.edu

Ari Kahn
arikahn@seas.upenn.edu

Jean M. Vettel
jvettel@gmail.com

Richard F. Betzel
rbetzel@seas.upenn.edu

¹ Department of Bioengineering, University of Pennsylvania, Philadelphia, PA, USA

² Broad Institute, Harvard University and the Massachusetts Institute of Technology, Cambridge, MA, USA

³ Human Research & Engineering Directorate, U.S. Army Research Laboratory, Aberdeen, MD, USA

⁴ Department of Psychological and Brain Sciences, University of California, Santa Barbara, CA, USA

⁵ Department of Electrical & Systems Engineering, University of Pennsylvania, Philadelphia, PA, USA

process and perform complex behaviors. Because cognitive processes involve large, distributed networks of brain areas, principled examinations of multi-node routes within larger connection patterns can offer fundamental insights into the complexities of brain function. Here, we investigate both densely connected groups of nodes that could perform local computations as well as larger patterns of interactions that would allow for parallel processing. Finding such structures necessitates that we move from considering exclusively pairwise interactions to capturing higher order relations, concepts naturally expressed in the language of algebraic topology. These tools can be used to study mesoscale network structures that arise from the arrangement of densely connected substructures called *cliques* in otherwise sparsely connected brain networks. We detect cliques (all-to-all connected sets of brain regions) in the average structural connectomes of 8 healthy adults scanned in triplicate and discover the presence of more large cliques than expected in null networks constructed via wiring minimization, providing architecture through which brain network can perform rapid, local processing. We then locate *topological cavities* of different dimensions, around which information may flow in either diverging or converging patterns. These cavities exist consistently across subjects, differ from those observed in null model networks, and – importantly – link regions of early and late evolutionary origin in long loops, underscoring their unique role in controlling brain function. These results offer a first demonstration that techniques from algebraic topology offer a novel perspective on structural connectomics, highlighting loop-like paths as crucial features in the human brain’s structural architecture.

Keywords Applied topology · Persistent homology · Network neuroscience

1 Introduction

Macroscopic computation and cognition in the human brain are affected by an intricately interconnected collection of neurophysical mechanisms (Bassett et al. 2010; Sporns et al. 2005). Unlike modern parallel computers, which operate through vast numbers of programs running in tandem and in isolation from one another, neural processes are supported on anatomically specialized brain regions that constantly share information among themselves through a network of white matter tracts (Hagmann et al. 2008). One approach for understanding the function of such a system begins with studying the organization of this white matter substrate using the language of networks (Sporns 2015; Bassett et al. 2011; Sporns 2013). Collections of regions that are pairwise tightly interconnected by large tracts, known as communities (Porter et al. 2009), modules (Meunier et al. 2009), and rich clubs (van den Heuvel and Sporns 2011; Senden et al. 2014), have been the subject of substantial prior study. Moreover, they have given critical insights into the large-scale structural units of the brain that give rise to many common cognitive functions (Chen et al. 2008; Medaglia et al. 2015). Such communities easily and rapidly transmit information among their members, facilitating local integration of information (Sporns and Betzel 2016).

Often left implicit in analyzes of structural networks, the *weakness* of connections to external regions is equally as important as the *strength* of internal connections within the community. This tendency to focus on strongly connected local regions arises naturally because standard network analyzes are based on local properties of the network at individual vertices, where local edge strength is the primary feature (Bassett and Bullmore 2006; Bullmore and Sporns 2009; Bullmore and Bassett 2011); the particular choice of quantitative language serves as a filter that diverts attention toward certain facets of the system. However, if one takes a more macro-scale view of the network, the small or absent white matter tracts intuitively serve to isolate processes carried on the strong white matter tracts from one another. Such structure facilitates more traditional conceptual models of parallel processing, wherein data is copied or divided into multiple pieces in order to rapidly perform distinct computations, and then recombined (Graham and Rockmore 2011). Together, the two notions of dense cliques and information-distributing cavities provide a picture of a system that performs complex computations by decomposing information into coherent pieces to be disseminated to local processing centers, and then aggregating the results.

To quantitatively characterize this macroscale structure, we must move from the language of graph theory to algebraic topology, which is sensitive to the interplay between weak and strong connections in systems (Ghrist 2008, 2014). In order to understand the interplay between strong and weak connections in the brain, we make use of two related lenses from algebraic topology. The first is an enumeration of the cliques, all-to-all connected subgraphs of the network, representing strongly-interconnected computational units. The number and size of such units gives a general sense for how intense local connections are across the brain. However, just as important is their context in the brain network: identical collections of processing units can be configured to perform very different tasks, depending on the way they pass information among themselves. Thus, we consider also how the cliques are arranged on a mesoscale level by examining the cycles they form. These structures, and the cavities they enclose, provide potential pathways along which data is disseminated and collected. Cycles enclosing voids correspond to extended paths of potential information transmission along which computations can be performed serially to effect cognition in either a divergent or convergent manner (i.e., distribution or integration of information), and we refer to these “enclosed spaces” as *topological cavities* in the network. We hypothesize that the spatial distributions of cliques and cavities will differ in their anatomical locations, corresponding to their differential putative roles in neural computations. Combined, these two perspectives provide a more complete view of the network’s capabilities than either does separately.

To test our predictions, we construct structural brain networks from diffusion spectrum imaging (DSI) data acquired from eight volunteers in triplicate. We measure node participation in cliques and compare these with a minimally wired null model (Betzel et al. 2016). To ensure this is an appropriate language for the structural connectome and to build intuition for later methods, we also demonstrate the correspondence between the anatomical location of cliques and the anatomical location of the brain’s hubs and structural rich club: a group of hubs that are densely connected to one another. Next, we study topological cavities using a recently developed method from algebraic topology which detects the presence and robustness, summarized by a quantity called *persistence*, of cavities in the network architecture. We recover all minimal length cycles corresponding to four highly persistent topological cavities in the consensus structure, and show that these features are robustly present across subjects through multiple scans. Our results demonstrate that while cliques are observed in the structural core, cycles enclosing topological cavities are observed to link regions of subcortex, frontal cortex, and parietal cortex in long loops, underscoring their unique role in controlling brain function (Gu et al. 2015a; Betzel et al. 2016; Muldoon et al. 2016b).

2 Materials and methods

2.1 Data acquisition, preprocessing, and network construction

Diffusion spectrum imaging (DSI) data and T1-weighted anatomical scans were acquired from eight healthy adult volunteers on 3 separate days (27 ± 5 years old, two female, and two left-handed) (Gu et al. 2015a). All participants provided informed consent in writing according to the Institutional Review Board at the University of California, Santa Barbara. Whole-brain images were parcellated into 83 regions (network nodes) using the Lausanne atlas (Hagmann et al. 2008), and connections between regions (network edges) were weighted by the number of streamlines identified using a deterministic fiber tracking algorithm. We represent this network as a graph $G(V, E)$ on V nodes and E edges, corresponding to a weighted symmetric adjacency matrix A . For clique calculations in the main text, the original network ($\rho = 0.9552$) was thresholded at $\rho = 0.25$ (corresponding to a weight = 261) to remove spurious connections (Zalesky et al. 2010; Zalesky et al. 2016; van den Heuvel et al. 2012) and for consistency with previous work (Sizemore et al. 2016). See Supporting Information and Refs (Cieslak and Grafton 2014; Gu et al. 2015a) for detailed descriptions of acquisition parameters, data preprocessing, and fiber tracking. In the supplement, we provide additional results for the case in which we correct edge weight definitions for the effect of region size Fig. 23.

2.2 Cliques versus cycles

In a graph $G(V, E)$ a k -clique is a set of k all-to-all connected nodes. It follows that any subset of a k -clique is a clique of smaller degree, called a face. Any clique that is not a face we call maximal. To assess how individual nodes contribute to these structures, we define node participation in maximal k -cliques as $P_k(v)$, and we record the total participation of a node as $P(v) = \sum_{k=1}^n P_k(v)$.

To detect cycles which enclose topological cavities, we computed the *persistent homology* using (Henselman and Ghrist 2016). We restrict our attention to dimensions 1–2 after finding no persistent features in dimension 3 (Sizemore et al. 2016).

Computing persistent homology involves first decomposing the weighted network into a sequence of binary graphs beginning with the empty graph and adding one edge at a time in order of decreasing edge weight (also called a Weight Rank Clique Filtration (Petri et al. 2013a, b). Formally, we translate edge weight information into a sequence of binary graphs called a filtration,

$$G_0 \subset G_1 \subset \dots \subset G_{|E|}$$

beginning with the empty graph G_0 and adding back one edge at a time following the decreasing edge weight ordering. To ensure all edge weights are unique we added random noise uniformly sampled from $[0, 0.0001]$. However, this has essentially no effect on the persistence diagrams, as stability theorems ensure that small perturbation of the filtration leads to small perturbation of the persistent homology (Chowdhury and Mémoli 2016; Cohen-Steiner et al. 2007). Noise can have a small effect on cycle representatives but in this study a great majority of edges within the thresholded networks are unique so the noise is not expected to largely alter cycle representatives – only to order those edges with tied edge weights.

Within each binary graph of this filtration, we extract the collection of all k -cycles, families of $(k + 1)$ -cliques which, when considered as a geometric object, form a closed shell with no boundary. Formally, as we are working with coefficients in \mathbb{Z}_2 , these are collections of $(k + 1)$ -cliques $\{\sigma_1, \dots, \sigma_n\}$ such that every k -subclique of some σ_i (called a *boundary*) appears as a subclique in the collection an even number of times. Two k -cycles are *equivalent* if they differ by a boundary of $k + 1$ -cliques. This relation forms equivalence classes of cycles with each non-trivial equivalence class representing a unique topological cavity. (In the mathematical literature, these are called *non-trivial homology classes*. However, due to the extensive and potentially confusing collision with the use of the word “homology” in the study of brain function, here we elect to use this new terminology outside of references and necessary mathematical discussion in the Methods and Supplementary Information. Throughout, the word “homology” refers to the mathematical, rather than the biological, notion.)

Constructing the sequence of binary graphs allows us to follow equivalence classes of cycles as a function of the edge density ρ . Important points of interest along this sequence are the edge density associated with the first G_i in which the equivalence class is found (called the birth density, ρ_{birth}) and the edge density associated with the first G_i in which the enclosed void is triangulated into higher dimensional cliques (called the death density, ρ_{death}). One potential marker of the relative importance of a persistent cavity to the weighted network architecture is its *lifetime* ($\rho_{death} - \rho_{birth}$). A large lifetime indicates topological cavities that persist over many edge additions, suggesting a greater importance of that cavity to the intrinsic structure of the complex. An alternative measure is the death to birth ratio $\pi = \rho_{death} / \rho_{birth}$ which highlights topological cavities that survive exceptionally long in spite of being born early, a feature that is interesting in geometric random graphs (see Bobrowski et al. 2015 and Supporting Information).

To study the role of each topological cavity in cognitive function, we extract the minimal representatives of each non-trivial equivalence class at the birth density. For

unfiltered complexes, the problem of finding a minimal generator for a given homology class is well known to be intractable (Chen and Freedman 2011; Dey et al. 2011). However, leveraging the filtration, we are able to answer the corresponding question in this context with relative ease. We used the persistent homology software Eirene (Henselman and Ghrist 2016) which returns the birth density and consequentially the starting edge of each persistent homology class. To recover the minimal cycle, we threshold the network at the density immediately preceding ρ_{birth} , then perform a breadth-first search (Rubinov and Sporns 2010) for a path from one vertex to the other, taking all minimum length paths as solutions. If for one persistent cavity we find multiple possible minimum-length paths arising from different equivalence classes, we still record and analyze each of the possible minimal generators, since any could be the homology class. For higher dimensional cycles we perform a similar process by hand, but we note that they could be algorithmically identified using appropriate generalizations of the graph search method and other approaches (Dey et al. 2011).

2.3 Standard graph statistics

In addition to the notions of cliques and cavities from algebraic topology, we also examined corresponding notions from traditional graph theory including communicability and rich-club architecture, computed using the Brain Connectivity Toolbox (Rubinov and Sporns 2010).

We first considered nodes that participated in many maximal cliques, and we assessed their putative role in brain communication using the notion of network communicability. The weighted communicability between nodes i and j is

$$C_{i,j} = (\exp(D^{-1/2}AD^{-1/2}))_{ij}$$

with $D := \text{diag}(s_i)$ for s_i the strength of node i in the adjacency matrix \mathbf{A} , providing a normalization step where each a_{ij} is divided by $\sqrt{d_i d_j}$ (Crofts and Higham 2009; Estrada and Hatano 2008). This statistic accounts for all walks between node pairs and scales the walk contribution according to the product of the component edge weights. The statistic also normalizes node strength to prevent high strength nodes from skewing the walk contributions. We refer to the sum of a node's communicability with all other nodes as node communicability, C_i .

Intuitively, nodes that participate in many maximal cliques may also play a critical role in the well-known rich club organization of the brain, in which highly connected nodes in the network are more connected to each other than expected in a random graph. For each degree k we compute the weighted rich club coefficient

$$\phi^w(k) = \frac{W_{>k}}{\sum_{l=1}^{E_{>k}} w_l^{\text{ranked}}}$$

where $W_{>k}$ is the summed weight of edges in the subgraph composed of nodes with degree greater than k , $E_{>k}$ is the number of edges in this subgraph, and w_l^{ranked} is the l -th greatest edge weight in \mathbf{A} . Rich club nodes are those that exist in this subgraph when $\phi^w(k)$ is significantly greater (one sided t -test) than $\phi_{\text{random}}^w(k)$, the rich club coefficient calculated from 1000 networks constructed by randomly rewiring the graph \mathbf{A} while preserving node strength (Rubinov and Sporns 2010).

Furthermore, highly participating nodes may also contribute to a hierarchical organization of the network. To evaluate this contribution, we compute the k -core and s -core decompositions of the graph (Hagmann et al. 2008; Chatterjee and Sinha 2007). The k -core is the maximally connected component of the subgraph with only nodes having degree greater than k . The s -core is similarly defined with summed edge weights in the subgraph required to be at least s .

2.4 Null model construction

We sought to compare the empirically observed network architecture to that expected in an appropriate null model. Due to the well-known spatial constraints on structural brain connectivity (Klimm et al. 2014; Lohse et al. 2014; Bullmore and Sporns 2012; Betzel et al. 2016) as well as the similarity in mesoscale homological features to the Random Geometric network (Sizemore et al. 2016) we considered a minimally wired network in which nodes are placed at the center of mass of anatomical brain regions. Each pair of nodes are then linked by an edge with weight $w_{i,j} = 1/d(i,j)$, where $d(i,j)$ is the Euclidean distance between nodes i and j . For consistency with the empirical data, we threshold this complete weighted network at an edge density of 0.25 for analyzes in which the DSI network is also thresholded. In each scan, the locations of region centers were collected. Thus, we considered a population of 24 model networks where differences between model networks arise from differences between scans. This null model allows us to assess what topological properties are driven by the precise spatial locations of brain regions combined with a stringent penalty on wiring length. Note that defining edge weights to be the inverse pairwise distance between points creates a filtered complex similar to that of either the Vietoris-Rips (Vietoris 1927; Hausmann et al. 1995) or Čech complex with an axis adjusted for edge rank instead of weight. We use the edge rank filtration for the null model here for consistency with the empirical data. Many ways of constructing simplicial complexes from graphs exist (Bergomi et al. 2017) but we have chosen the above methods because they are relatively well understood and do not require further assumptions about the data.

2.5 Cycles in individuals

Though we detected persistent cavities in the group-averaged DSI network using persistent homology, we also ask whether these patterns of connectivity and the corresponding cavities exist in multiple individuals and in multiple scans acquired from the same individual. To address this question, we asked whether a similar geometric loop is seen and whether a similar topological cavity is present in each scan. However, identifying similar topological cavities is not trivial and we next thoroughly discuss our method including our definition of “similar topological cavities”.

2.5.1 Considerations in per scan cycle validation

Persistent homology is a powerful tool with which to understand the mesoscale homological features of a weighted network. Determining all minimal generators for each of the long-lived topological cavities gives a finer resolution of such features, which can have biological implications as is the case with our DSI data. Isolating all minimal generators for each homology class additionally gives a geometric interpretation to these cavities. Then each cavity can be viewed from a biological, topological, and to a lesser extent geometric perspective.

This presents a challenge when looking for the “same” persistent homology classes in another clique complex. From the neuroscience perspective, two minimal cycles may be similar if the cycles include the same brain regions, or if the group of regions forming the second cycle performs the same function as those in the first. Geometrically we would perhaps require the same rigid shape of two cycle representatives to call them similar. Finally, through the lens of topology we might call two minimal cycles in two different complexes similar if we can find a map between the complexes which takes one cycle to the other. Less abstractly, we could instead ask if the minimal cycle of a homology class in the first clique complex exists in the second as a cycle in a nontrivial homology class but not necessarily as the minimal generator. The development of other definitions is an area of active research (Carlsson and De Silva 2010; Dey et al. 2014).

Because no universal method is available, we opt for a domain-specific heuristic to determine whether a persistent homology class found in an individual scan was the “same” as the persistent homology class in the average network. These requirements for similarity adequately capture some flexibility of topological similarity while being conservative enough to generally preserve the biological function of the cycle as well.

We consider each persistent homology class in turn. For a given persistent homology class found in the average DSI connectome, we denote the set of minimal generators of

the homology class at ρ_{birth} by L with elements ℓ_i for $i = 0, 1, 2, \dots, m$. Then for each ℓ_i there is a set of nodes N_i containing the nodes within this representative. We require both a non-trivial cycle formed by connections between at least one of N_0, N_1, \dots, N_m and a similar topological cavity to exist.

1. Nodes connected in a cycle. We first take the subgraph on N_i and ask if there is precisely one non-trivial homology class at any edge density. We then show the connection pattern at the edge density at which this class first appears. This first allows us to ask if these nodes ever form a non-trivial cycle throughout the filtration, which is possibly of interest from a geometric and neuroscience perspective. We also use this first test as a filter to see in which scans could these nodes surround a topological cavity. Then if we find a non-trivial cycle formed by any of N_0, N_1, \dots, N_m , this scan passes to the next stage.
2. Similar topological cavity. We then ask if a similar topological cavity exists. The algorithm from Henselman and Ghrist (2016) returns the birth density (and thus birth edge) of each persistent homology class. In order of increasing birth density, we ask if any of the nodes in N_0, N_1, \dots, N_m are in the birth edge. If this is true, we call this a similar cavity in an individual scan if any of the following hold:
 - (a) Let m_0, \dots, m_k be minimal generators of this homology class in the individual scan at ρ_{birth} . If any of m_0, \dots, m_k are the same as one of ℓ_0, \dots, ℓ_M or are in the same equivalence class, then we call this a similar topological cavity and we are done. This is the most straightforward and was most frequently observed within the unnormalized data.
 - (b) If there is some cycle within this non-trivial homology class at ρ_{birth} formed by at least all but one node of some N_i , along with no more than two additional nodes, and nodes from N_i are in the original order along the cycle, we call this similar.
 - (c) If either (a) or (b) hold for some ρ with $\rho_{birth} \leq \rho < \rho_{death}$, we call this a similar topological cavity. At ρ_{birth} , a minimal cycle contains seven nodes, four of which are the thalamus and caudate nucleus from both hemispheres. Following the minimal cycles throughout the lifetime of this persistent cavity we find at some edge density before ρ_{death} , a minimal representative consists of exclusively the thalamus and caudate nucleus regions from both hemispheres.

The first test covers the possibility of the same biological and geometric feature occurring in the individual scan. The second is perhaps the most important, however,

because it allows for matching the topological cavity itself. It is important to remember that the topological cavities are the features of interest, not the precise cycles themselves, though the two are clearly related. With the focus on the topological holes, the rationale for the three subrules 2a, 2b, and 2c, is more clear. Though labor intensive, this lets us keep the topological perspective when determining cycle similarity. Moreover, the rationale for focusing on cavities and not specific connections is similar to why large-scale organization such as communities (Betzel et al. 2016), cores (Hagmann et al. 2008), and rich-club organization (van den Heuvel and Sporns 2011) are studied with increased intensity. Composed of a plurality of interacting brain regions, these types of structures, and not the individual brain regions nor connections, form computational units that theoretically act to help segregate and integrate information flow across the brain.

One clear drawback of this method is the possibility of false negatives. For example, a persistent homology class may have been born which is similar to the cycle in the average data, yet the beginning edge did not include any of the cycle nodes and thus we would not detect this following the above procedure. This is a first attempt to identify similar topological cavities across subjects, and we expect more robust algorithms to be a topic of future research.

3 Results

To extract relevant architectural features of the human structural connectome, we first encoded diffusion spectrum imaging (DSI) data acquired from eight subjects in triplicate as undirected, weighted networks. In this network, nodes correspond to 83 brain regions defined by the Lausanne parcellation (Cammoun et al. 2012) and edges correspond to the density of white matter tracts between node pairs (Fig. 1a). We initially study a group-averaged network, and then demonstrate that our results are consistently observed across individuals in the group as well as across multiple scans from the same individual.

3.1 Cliques in the human structural connectome

Here, we use the group-averaged network thresholded at an edge density (ρ) of 0.25 to remove spurious edges (Zalesky et al. 2010, 2016; van den Heuvel et al. 2012) and for consistency with previous studies (Sizemore et al. 2016). Results at other densities are similar, and details can be found in the Appendix. As a null-model, we use minimally wired networks (Fig. 1d) created from assigning edge weights to the inverse Euclidean distance between brain region centers (see Methods) observed in each of 24 scans. This model mimics

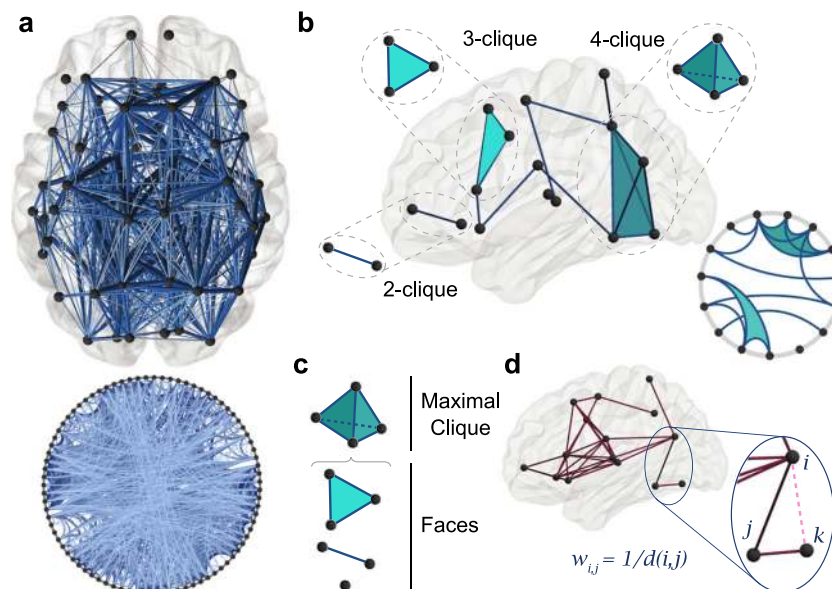


Fig. 1 Cliques are features of local neighborhoods in structural brain networks. **a** Diffusion spectrum imaging (DSI) data can be summarized as a network of nodes corresponding to brain regions, and weighted edges corresponding to the density of white matter streamlines reconstructed between them. Here we present a group-averaged network, where each edge corresponds to the mean density of white matter streamlines across eight subjects scanned in triplicate. We show the network at an edge density $\rho = 0.25$, and display its topology on the brain (top), and on a circle plot (bottom). This and all brain

networks are drawn with BrainNetViewer (Xia et al. 2013). **b** All-to-all connected subgraphs on k nodes are called k -cliques. For example, 2-, 3-, and 4-cliques are shown both as schematics and as features of a structural brain network. **c** A maximal 4-clique has 3-, 2-, and 1-cliques as faces. **d** For statistical validation, we construct a minimally wired null model by linking brain regions by edge weights equal to the inverse of the Euclidean distance between nodes corresponding to brain region centers. Here we show an example of this scheme on 15 randomly chosen brain regions

the tendency of the brain to conserve wiring cost by giving edges that connect physically close nodes higher weight than edges between distant nodes.

The first step in a topological analysis is an enumeration of all maximal k -cliques in the average structural network. Recall that a k -clique is a set of k nodes having all pairwise connections (see Fig. 1b for 2-, 3-, and 4-cliques representing edges, triangles, and tetrahedra, respectively.) By definition, a subgraph of a clique will itself be a clique of lower dimension, called a *face*. A *maximal* clique is one that is not a face of any other (see Fig. 1c for a maximal 4-clique, which contains 3-, 2-, and 1-cliques as faces).

To understand the anatomical distribution of maximal cliques in both real and null model networks, we count the number of maximal k -cliques in which a node is a member, and refer to this value as the node participation, $P_k(v)$ (see Methods). Summing over all k gives the total participation, $P(v)$. We observe that the distribution of maximal clique degrees is unimodal in the minimally wired null model and qualitatively bimodal in the empirical data (see Fig. 2a), though we report statistically that we cannot reject that it is unimodal ($p = 0.210$, dip test (Hartigan and Hartigan 1985)). Anatomically, we observe a general progression of maximal clique participation from anterior to posterior regions of cortex as we detect higher degrees (Fig. 2a, bottom and Fig. 8). Indeed, maximal cliques of 12–16 nodes contain nearly all of the visual cortex. This spatial distribution suggests that large interacting groups of brain regions are required for early information processing, while areas of frontal cortex driving higher-order cognition utilize smaller working clusters. We also observe that the human brain displays preferences for small (4–6 node), and large (12–16 node) processing units instead of medium-sized (approximately 8 node) units as in the minimally wired null model.

The anterior-posterior gradient of maximal clique size can be complemented by additionally analyzing regional variation in the cognitive computations being performed. Specifically, we ask whether node participation in maximal cliques differs in specific cognitive systems (Power et al. 2011) (Fig. 2b). We observe that the largest maximal cliques are formed by nodes located almost exclusively in the subcortical, dorsal attention, visual, and default mode systems, suggesting that these systems are tightly interconnected and might utilize robust topologically-local communication. This spatial distribution of the participation in maximal cliques differs significantly from the minimally wired null model, particularly in the cingulo-opercular and subcortical systems. We hypothesized that these differences may be driven by the excess of maximal 8-cliques in the minimally wired network (Fig. 2a). Expanding on the difference in node participation ($P_k^{DSI}(v) - P_k^{MW}(v)$), we see that the large discrepancies between empirical and null model networks in cingulo-opercular and subcortical systems are caused by a difference in maximal cliques of approximately eight nodes (Fig. 2b, bottom). Finally, we observe that the systems involved in the two peaks of the maximal clique distribution shown in Fig. 2a differ greatly from one another. The first peak composed of smaller cliques involves regions from nearly all systems, while the second peak is almost exclusively composed of regions in the default mode, subcortical, and visual systems. We observe the largest cliques in the subcortical, default mode, dorsal attention, and visual systems, though only the visual and dorsal attention systems have maximal clique distributions with significantly higher means than the rest of the brain regions ($p < 0.001$, $p < 0.05$, respectively). These data suggest that small, local processors may be a common feature across systems, while larger cliques may allow for rapid multi-system cross-talk.

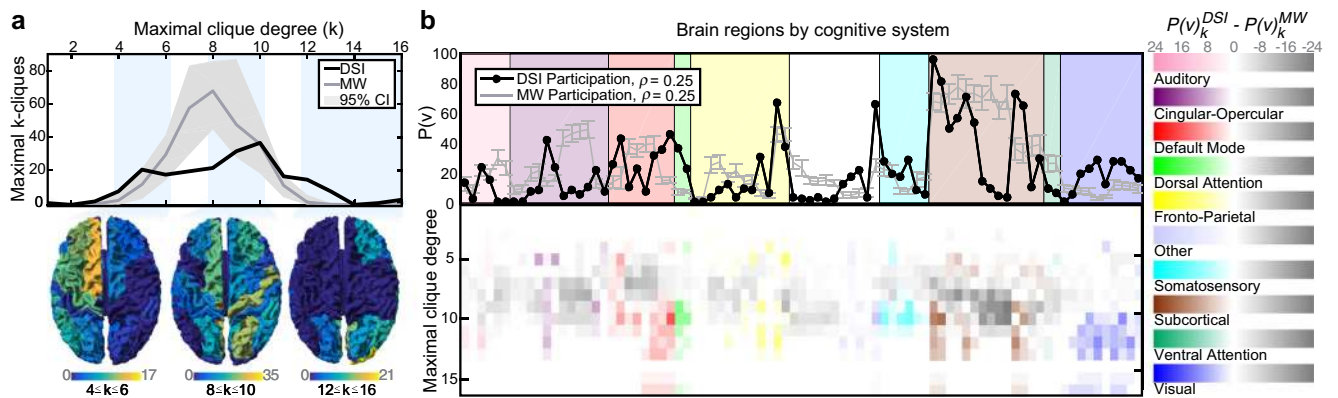


Fig. 2 Spatial distribution of maximal cliques varies between average DSI and minimally wired null model. **a** Distribution of maximal cliques in the average DSI (black) and individual minimally wired (gray) networks, thresholded at an edge density of $\rho = 0.25$. Heat maps of node participation on the brain for a range of clique degrees equal to 4–6 (left), 8–10 (middle), and 12–16 (right). **b** Node

participation in maximal cliques sorted by the putative cognitive system to which the node is affiliated in functional imaging studies (Power et al. 2011). We show individual node values (top) as well as the difference between real and null model ($P_k^{DSI} - P_k^{MW}$; bottom) according to the colormap (right). Individual node labels are listed in Fig. 9

We next check that the building blocks, here k -cliques, behave consistently with more common graph theoretic metrics. A node with high participation in maximal cliques must in turn be well connected locally (though the converse is not necessarily true – consider a node that only participates in one maximal 16-clique). Therefore we expect the participation of a node to act similarly to other measures of connectivity. To test this expectation, we examine the correlation of node participation with node strength, the summed edge weight of connections emanating from a node, as well as with node communicability, a measure of the strength of long distance walks emanating from a node (Fig. 3a). We find that both strength and communicability exhibit a strong linear correlation with the participation of a node in maximal cliques (Pearson correlation coefficient $r = 0.957$ and $r = 0.858$, respectively).

These results indicate that regions that are strongly connected to the rest of the brain by both direct paths and indirect walks also participate in many maximal cliques. Such an observation suggests the possibility that brain hubs – which are known to be strongly connected with one another in a so-called *rich-club* – play a key role in maximal cliques. To test this, we measure the association of brain regions to the rich-club using notions of coreness. A k -core of a graph G is a maximal connected subgraph of G in which all vertices have degree at least k , and an s -core is the equivalent notion for weighted graphs (see Methods). Using these notions, we consider how the k -core and s -core decompositions align with high participation (Fig. 3b). In both cases, nodes with higher participation often achieve higher levels in the k - and s -core decomposition. Moreover, we also observe the frequent existence of rich club connections between nodes with high participation (Fig. 3b, bottom). Together, these results suggest that

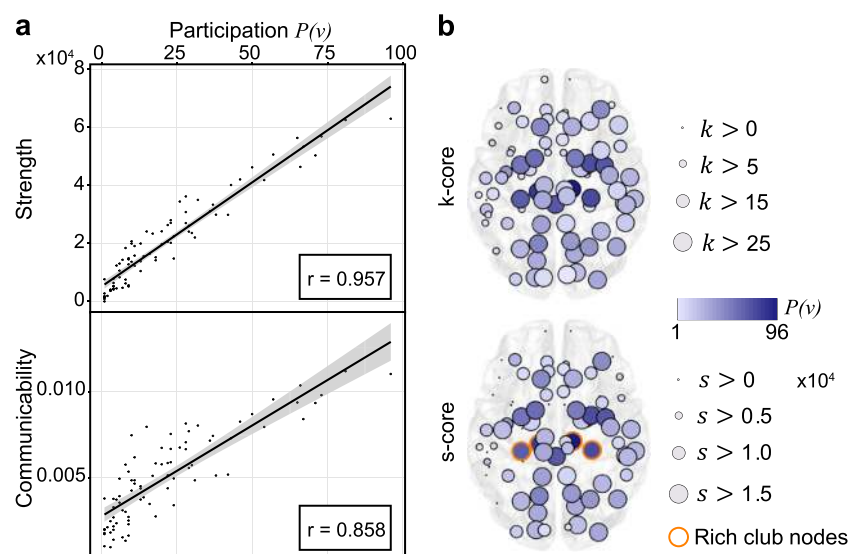
rich-club regions of the human brain tend to participate in local computational units in the form of cliques.

3.2 Cavities in the structural connectome

Whereas cliques in the DSI network act as neighborhood-scale building blocks for the computational structure of the brain, the relationships between these blocks can be investigated by studying the unexpected absence of strong connections, which can be detected as topological cavities in the structure of the brain network. Because connections are treated as communication channels along which brain regions can signal one another and participate in shared neural function, the absence of such connections implies a decreased capacity for communication which serves to enhance the segregation of different functions.

To identify topological cavities in a weighted network, we construct a sequence of binary graphs, each included in the next (Fig. 4a), known as a *filtration*. Beginning with the empty graph, we replace unweighted edges one at a time according to order of decreasing edge weight, and we index each graph by its *edge density* ρ , given by the number of edges in the graph divided by the number of possible edges. After each edge addition, we extract motifs of k -cliques called (non-trivial) $(k - 1)$ -cycles, each of which encloses a k -dimensional topological cavity in the structure. This shift in index is due to geometry: a 2-clique is a 1-dimensional line segment, a 3-clique is a 2-dimensional triangle, etc. When k is clear or not pertinent, we will suppress it from the notation, and refer simply to “cycles” and “cavities”. While any cavity is surrounded by at least one cycle, often multiple cycles surround the same cavity. However, any two $(k-1)$ -cycles that detect the same cavity will necessarily differ from one another by the boundaries of some collection of $(k + 1)$ -

Fig. 3 Maximal clique participation tracks with network measures. **a** Scatter plot of node participation and node strength (top) or communicability (bottom). **b** Calculated k -core (top) and s -core decomposition in relation to participation in maximal cliques with rich club nodes (shown $k_{RC} = 43$; see Methods and Fig. 10) indicated in orange (bottom). Size indicates maximum k -core or s -core level attained by the node, while color indicates the participation $P(v)$



cliques (see Supporting Information and Fig. 15). Any two such cycles are called *topologically equivalent*, so each topological cavity is detected by a non-trivial *equivalence class* of cycles. The equivalence class containing the cycle consisting of a single vertex is called *trivial* and bounds the “empty” cavity. We can represent a topological cavity using any of the cycles within the corresponding equivalence class, but for purposes of studying computational architectures it is reasonable to assume information will primarily travel along paths of minimal length; thus, in this analysis we will consider the collection of cycles in an equivalence class with the minimal number of nodes and call these the *minimal cycles* representing the cavity. Note in the absence of a filtration, there are serious computational issues involved in locating minimal-size representatives of equivalence classes. However, in this setting the

computation is easily performed using standard algorithms (see Methods).

As we move through the filtration by adding edges, the structure of the cycles, and thus of the cavities they represent, will evolve. We consider an example in Fig. 4a, showing a green minimal cycle surrounding a 2D cavity which first appears (is *born*) in the graph sequence at ρ_{birth} (cyan). As an edge completing a 3-clique is added, the minimal cycle representative shrinks to four nodes in size, then finally is tessellated by 3-cliques (*dies*) at ρ_{death} (orange). We record ρ_{birth} and ρ_{death} for all topological cavities (e.g., non-trivial equivalence classes of cycles) found within the filtration, and display them on a *persistence diagram* (Fig. 4b). Cavities that survive many edge additions have a long *lifetime*, defined as $\rho_{death} - \rho_{birth}$, or a large death-to-birth ratio, $\rho_{death}/\rho_{birth}$. Such cycles

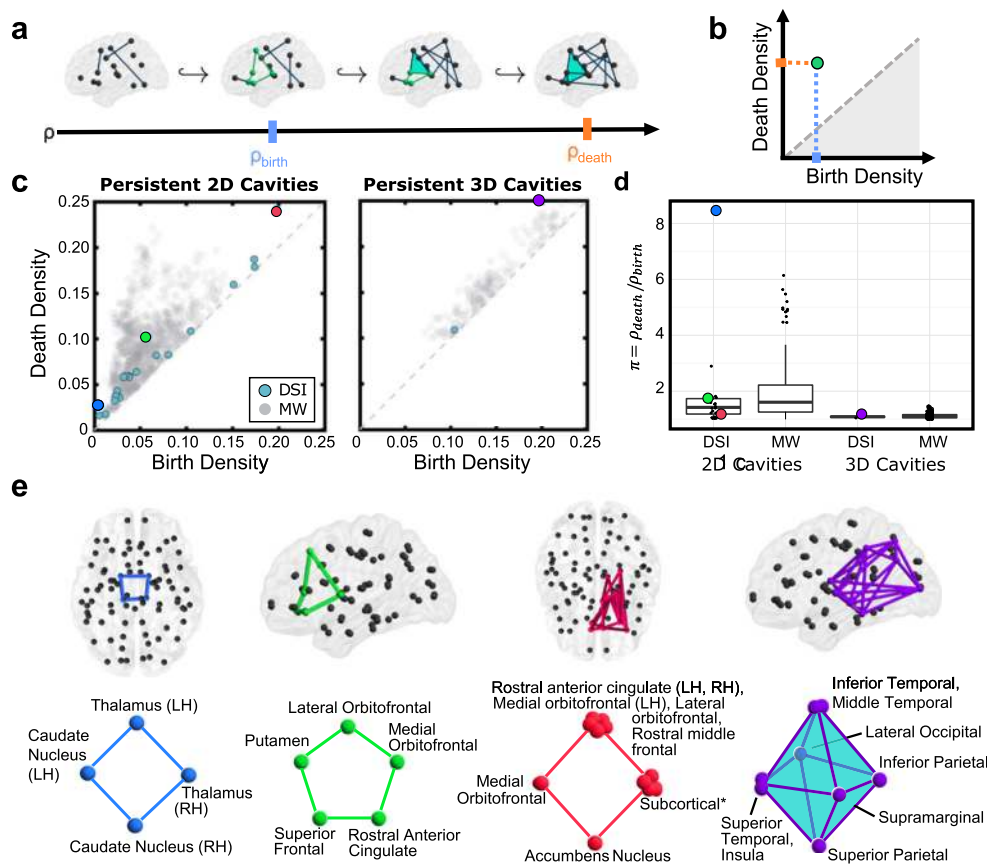


Fig. 4 Tracking clique patterns through a network filtration reveals key topological cavities in the structural brain network. **a** Example filtration of a network on 15 nodes shown in the brain across edge density (ρ). Blue line on the axis indicates the density of birth (ρ_{birth}) of the 2D cavity surrounded by the green minimal cycle. As edges are added, 3-cliques (cyan) form and shrink the cavity and consequentially the minimal green cycle is now four nodes in size. Finally, the orange line marks the time of death (ρ_{death}) when the cavity is now filled by 3-cliques. **b** Persistence diagram for the cavity surrounded by the green cycle from panel **a**. **c** Persistence diagrams for the group-averaged DSI (teal) and minimally wired null (gray) networks in dimensions one

(left) and two (right). Cavities in the group-averaged DSI network with long lifetime or high death-to-birth ratio are shown in unique colors and will be studied in more detail. **d** Box plots of the death-to-birth ratio π for cavities of two and three dimensions in the group-averaged DSI and minimally wired null networks. Colored dots correspond to those highlighted in panel **c**. The difference between π values for 3D topological cavities in the average DSI data versus the minimally wired null model was not found to be significant. **e** Minimal cycles representing each persistent cavity at ρ_{birth} noted in panels **c**, **d** shown in the brain (top) and as a schematic (bottom)

are commonly referred to as *persistent cavities* and in many applications are considered the “topological features” of the system.

We investigate the persistence of 2D and 3D cavities (respectively represented by equivalence classes of 1- and 2-cycles) in the group-average DSI network and minimally wired null networks (see Fig. 4c). There are substantially fewer persistent cavities in the group-average DSI network than in the null models. To illustrate the structure of these cavities, we select four representative cavities with exceedingly long lifetimes or a high ρ_{death} to ρ_{birth} ratio (Fig. 4c, d) in the empirical data, and for each we find the minimal-length representative cycles at ρ_{birth} (Fig. 4e). Such cycles for all of the persistent cavities found in the empirical data are illustrated in Figs. 20 and 21. The first persistent cavity appears as early as $\rho = 0.003$ and is minimally enclosed by the unique blue cycle composed of the thalamus and caudate nucleus of both hemispheres. The green cycle connecting the medial and lateral orbitofrontal, rostral anterior cingulate, putamen, and superior frontal cortex is the only minimal cycle surrounding a long-lived cavity in the left hemisphere. The final persistent 2D cavity in the average DSI data is found in the right hemisphere between the medial orbitofrontal, accumbens nucleus, any of the subcortical regions hippocampus, caudate nucleus, putamen, thalamus, and amygdala, and any of the rostral middle frontal, lateral orbitofrontal, medial orbitofrontal of the left hemisphere, and rostral anterior cingulate from both hemispheres (see Fig. 4e for all 12 minimal representatives). Finally, the purple octahedral cycle made from 3-cliques contains the inferior and middle temporal, lateral occipital, inferior parietal, supramarginal, superior parietal, and either of the superior temporal and insula of the left hemisphere, and encloses the longest-lived 3D cavity in the structural brain network. Though each minimal generator may have distinct biological implications, we observe a global pattern of subcortical–cortical connections within cycles. Indeed, 18 of the 20 recovered 1-cycles and both 2-cycles contain this motif. Additionally, the two persistent cycles that do not follow this motif comprise a third of persistent cycles robustly seen in the minimally wired network, suggesting that within-subcortical loops are more probable in this maximally efficient scheme.

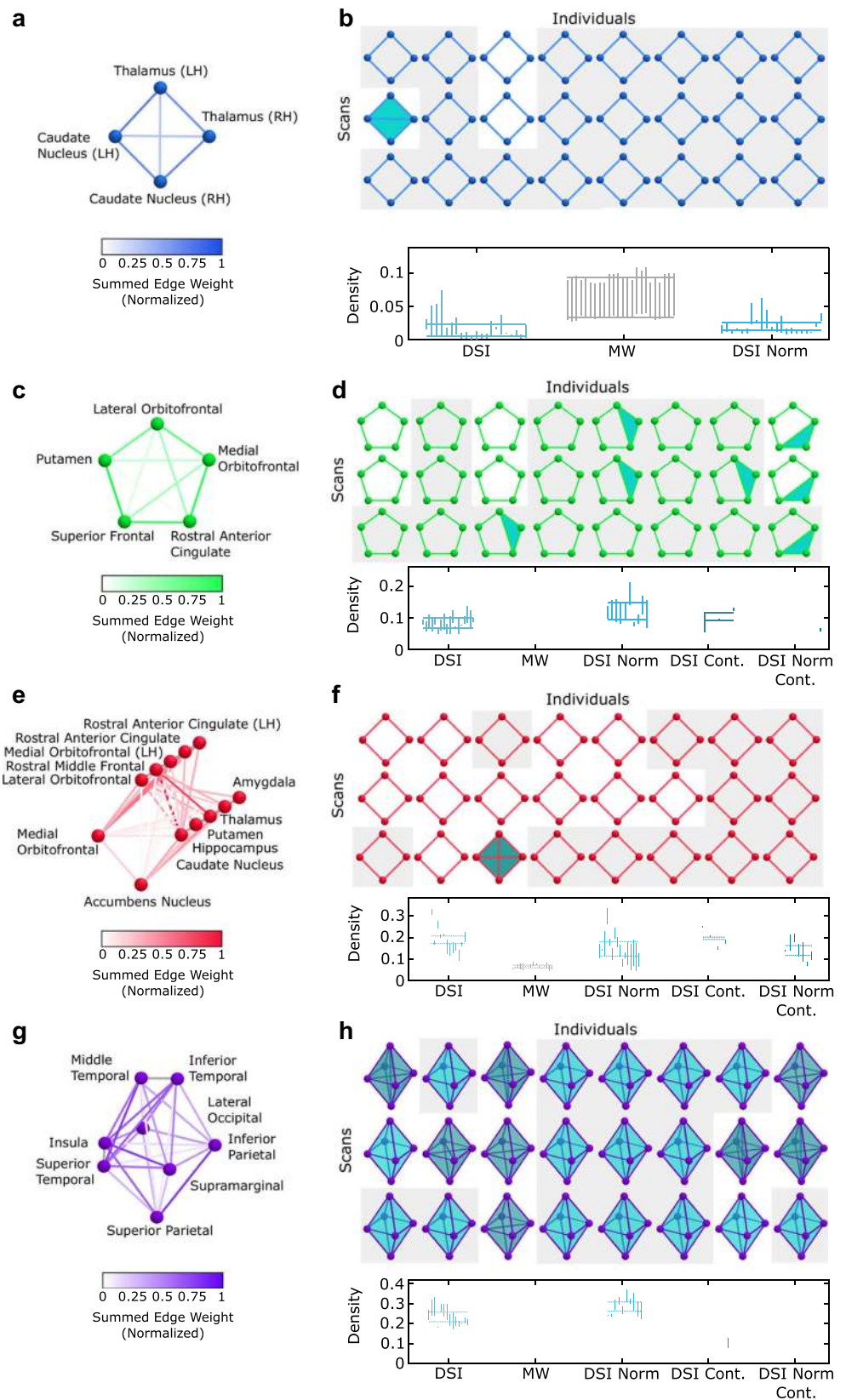
3.3 Test-ReTest reliability and other methodological considerations

It is important to ask whether the architectural features that we observe in the group-averaged DSI network can also be consistently observed across multiple individuals, and across multiple scans of the same individual to ensure these cavities are not artifacts driven by a few outliers. Comparison of persistent cavities arising from two different

networks is complicated by our notion of equivalence of cavities, and our desire to work with particular representative cycles. To capture the extent to which the cavities and their minimal representatives in the average DSI data are present in the individual scans, we record the collection of cliques that compose each minimal cycle representing the equivalence class (as seen in Fig. 4e), and check both for the existence of one of those collections of cliques, corresponding to the existence of the same strong fiber tracts, and, more stringently, for the presence of a topological cavity represented by that cycle in each individual’s DSI network (see Supporting Information for more details). We observed that the subcortical cycle (Fig. 4e, blue) exists and these nodes (thalamus and caudate nucleus of both hemispheres) surround an equivalent 2D cavity in at least one scan of all individuals and the late-developing subcortical-frontal cycle (Fig. 4e, red) surrounds a cavity found in seven of the eight individuals in at least one of three scans (Fig. 5b, f). The earlier arriving subcortical-frontal cycle (Fig. 4e, green) is present in all individuals and a similar cavity is seen at least once in all individuals (Fig. 5d). Finally, we observe that the octahedral connection pattern in posterior parietal and occipital cortex (Fig. 4e, purple) is present at least once in seven of eight individuals and these regions enclose a similar cavity at least once in six of these individuals (Fig. 5h). In the opposite hemisphere, the cyclic connection patterns and similar cavities appear though not as regularly (Fig. 5). Finally we check the existence of similar cavities within the minimally wired null models, and see cavities denoted by the green and purple cycles are never seen (Fig. 5). However, similar cavities to those represented by the red and blue minimal cycles appear frequently in the null model, though with different birth/death densities and lifetimes. In summary we find topological cavities observed in the group-averaged DSI network appear consistently across individuals, suggesting their potential role as conserved wiring motifs in the human brain.

In addition to consistency across subjects and scans, it is important to determine whether the known high connectivity from subcortical nodes to the rest of the brain may be artificially obscuring non-trivial cortico-cortical cavities important for brain function. To address this question, we examined the 66-node group-average DSI network composed only of cortical regions, after removing subcortical regions, insula, and brainstem. We recovered a long-lived topological cavity surrounded by four cycles of minimal length composed of nine nodes connecting temporal, parietal, and frontal regions (Fig. 6). Note in the schematic of Fig. 6a we see clearly two 2D cavities. The birth edge here was between the lateral orbitofrontal and superior temporal regions, which prevents us from determining whether the exact minimal cycle surrounding this cavity follows the superior frontal (LH)/posterior cingulate or the superior

Fig. 5 Cycles and similar cavities in the average DSI network are consistently seen across individuals. **a, c, e, g** Edge weights connecting nodes seen in the minimal cycle(s) recovered from the average DSI were summed then normalized for all individual scan data. **b, d, f, h** (Top) Within each scan, the network was thresholded at the minimal weight of any edge which would form the cycle seen in the average DSI data. At this threshold, any connection which exists between these cycle nodes is shown. A gray background indicates a similar cavity found in this scan. For those cycles seen which are not tessellated by higher cliques yet there is no gray background, there must exist some set of nodes which cone this cycle and thus make this loop equivalent to a point. (Bottom) Similar cycles found represented by vertical bars from birth to death density in the individual DSI networks, minimally wired networks, normalized data, and contralateral (cont.) hemispheres



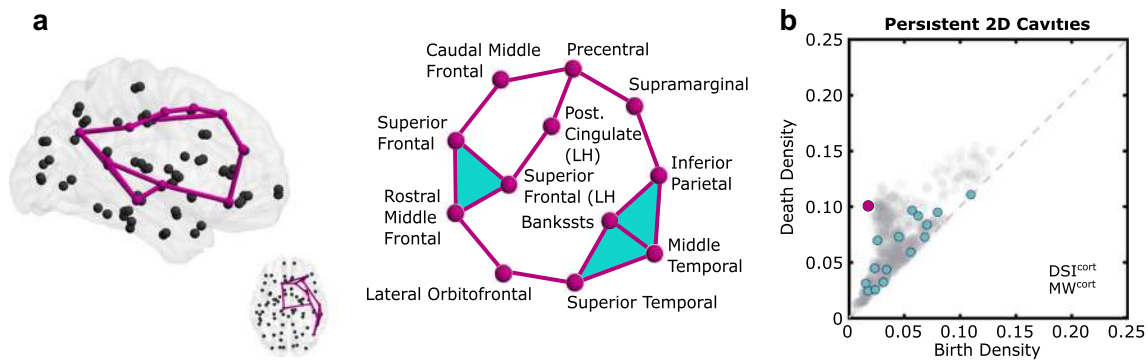


Fig. 6 Removal of subcortical nodes allows for detection of nine-node cortical cycle enclosing large 2D cavity. **a** Minimal cycles shown in the brain (left) and as a schematic (right). **b** Persistence diagram of DSI^{cort} and MW^{cort} . Persistent feature corresponding to minimal cycles in (a) indicated with maroon dot

frontal (RH)/caudal middle frontal branch of the top loop. Following either of these two branches (then either of the banks of the superior temporal sulcus or middle temporal route) gives four cycles in which two are equivalent to each other but not to either cycle in the other pair. We will accept all of these four as minimal maroon cycles since any of the four could be minimal representatives. Moreover, at least one of these minimal cycles and corresponding cavity was observed in each scan of every individual (Fig. 26c), and often in the opposite hemisphere as well (Fig. 26d). These results reveal that cortico-cortical cycles are indeed present and suggest their potential utility in segregating function across the brain.

4 Discussion

In this study, we describe a principled examination of multi-node routes within larger connection patterns that are not accessible to network analysis methods that exclusively consider pairwise interactions between nodes. Our approach draws on concepts from a discipline of mathematics known as *algebraic topology* to define sets of all-to-all connected nodes as structural units, called *cliques*, and then to use the clique architecture of the network to detect structural topological cavities, detected by the existence of non-trivial representative *cycles*. Using this approach, we show that node participation in maximal cliques varies spatially and by cognitive systems, suggesting a global organization of these neighborhood-scale features. These cliques form encapsulating patterns of connectivity in the human structural connectome, which separate relatively early-evolving regions of the subcortex with higher-order association areas in frontal, parietal, and temporal cortex that evolved on more recent time scales. We found the recovered topological cavities exist consistently across individuals and are not expected in a spatially embedded null model, emphasizing their importance in neural wiring and function. These results

offer a first demonstration that techniques from algebraic topology offer a novel perspective on structural connectomics, highlighting cavernous spaces as crucial features in the human brain's structural architecture.

4.1 Algebraic-topological tools for neural data analysis

Algebraic topology is a relatively young field of pure mathematics that has only recently been applied to the study of real-world data. However, the power of these techniques to measure structures that are inaccessible to common graph metrics has gained immediate traction in the neuroscience community. Here, we highlight a few notable examples from the growing literature; a more comprehensive recent account can be found in Giusti et al. (2016). At the neuron level, persistent has been used to detect intrinsic structure in correlations between neural spike trains (Giusti et al. 2015), expanding our understanding of the formation of spatial maps in the hippocampus (Dabaghian et al. 2012). Moreover, at the level of large-scale brain regions, these tools have been exercised to characterize the global architecture of fMRI data (Stolz 2014). Based on their unique sensitivity, we expect these algebraic-topological methods to provide novel contributions to our understanding of the structure and function of neural circuitry across all scales at which combinatorial components act together for a common goal: from firing patterns coding for memory (Rajan et al. 2016; Leen and Shea-Brown 2015) to brain regions interacting to enable cognition.

Our study uses algebraic topology in the classical form to obtain a global understanding of the structure, and in conjunction, it investigates particular topological features themselves and relates these features to cognitive function. Cycle representatives have previously been considered in biology (Chan et al. 2013; Petri et al. 2014; Lord et al. 2016; Kim et al. 2014; Emmett et al. 2016; Mamuye et al. 2016), but to our knowledge this is a first attempt to compare topological features in multiple brains.

4.2 Cliques and cavities for computations

Cliques and minimal cycles representing cavities are structurally positioned to play distinct roles in neural computations. Cliques represent sets of brain regions that may possess a similar function, operate in unison, or share information rapidly (Sizemore et al. 2016). Furthermore, the hierarchical organization of small cliques located more anteriorly and larger cliques connecting multiple systems allows for swift global sharing of information produced by local processing. Conversely, minimal cycles correspond to extended paths of potential information transmission along which computations can be performed serially to affect cognition in either a divergent or convergent manner. Indeed, the capsule-like or chain-like nature of cycles is a structural motif that has previously been – at least qualitatively – described in neuroanatomical studies of cellular circuitry. In this context, such motifs are known to play a key role in learning (Hermundstad et al. 2011), memory (Rajan et al. 2016), and behavioral control (Levy et al. 2001; Fiete et al. 2010). The presence of cycles suggests a possible role for polysynaptic connections and their importance to neural computations, consistent with evidence from the field of computational neuroscience highlighting the role of highly structured circuits in sequence generation and memory (Rajan et al. 2016; Hermundstad et al. 2011). Indeed, in computational models at the neuron level, architectures reminiscent of chains (Levy et al. 2001; Fiete et al. 2010) and rings are particularly conducive to the generation of sequential behavioral responses. It is interesting to speculate that the presence of these structures at the much larger scale of white matter tracts could support diverse neural dynamics and a broader repertoire of cognitive computations than possible in simpler and more integrated network architectures (Tang et al. 2016).

Another consideration concerns the apparent asymmetry of our results with respect to left and right cerebral hemispheres. While unanticipated, we note that in some cases they have intuitive mathematical underpinnings. For example, in Fig. 3, we explicitly count maximal cliques, so one edge difference between a region in the left and right hemisphere could result in a large difference in the number of observed maximal cliques. Interestingly, despite this fact we still observe a strong correlation between node strength and $P(v)$, instilling confidence in these results. From a neuroscience point of view, brain asymmetries are not wholly unexpected. There is a storied and ever-growing literature describing the lateralization (i.e., asymmetries) of brain function (Galaburda et al. 1978). While speech generation (Rasmussen and Milner 1977) and language processing (Desmond et al. 1995; Thulborn et al. 1999) are among the most commonly-cited functions to exhibit lateralization (Doron et al. 2012; Chai et al. 2016), such effects have also

been linked to a diverse group of other cognitive domains. These include emotion (Wager et al. 2003), processing of visual input (Sandi et al. 1993), and even working memory (Carpenter et al. 2000). In addition, a number of studies have also reported the emergence of pathological lateralization or the disruption of asymmetries with neurocognitive disorders including ADHD (Oades 1998). Our study does not offer a conclusive demonstration that the observed asymmetries arise from the lateralization of any specific brain function; we merely wish note that there is a precedent for such observations.

4.3 Evolutionary and developmental drivers

Network filtration revealed several persistent cavities in the macroscale human connectome. While each minimal cycle surrounding these cavities involved brain regions interacting in a distinct configuration, we also observed commonalities across these structures. One such commonality was these minimal cycles tended to link evolutionarily old structures with more recently-developed neo-cortical regions (Rakic 2009). For example, the green cycle depicted in Fig. 4e linked the putamen, an area involved in motor behavior (Middleton and Strick 2000), with the rostral anterior cingulate cortex, associated with higher-order cognitive functions such as error-monitoring (Braver et al. 2001) and reward processing (Kringelbach and Rolls 2004). This observation led us to speculate that the emergence of these cavities may reflect the disparate timescales over which brain regions and their circuitry have evolved (Gu et al. 2015b), through the relative paucity of direct connections between regions that evolved to perform different functions. This hypothesis can be investigated in future work comparing the clique and cavity structure of the human connectome with that of non-human connectomes from organisms with less developed neocortices.

4.4 Toward a global understanding of network organization

Though we highlighted minimal cycles in the brain, by nature persistence describes the global organization of the network. Often regions in the brain wire minimally to conserve wiring cost (Bassett et al. 2010; Bullmore and Sporns 2012; Klimm et al. 2014; Lohse et al. 2014), though there are exceptions that give the brain its topological properties such as its small-world architecture (Bassett and Bullmore 2006; Pessoa 2014; Hilgetag and Goulas 2016; Muldoon et al. 2016a; Bassett and Bullmore 2016). Following this idea, we could interpret the difference in the number of persistent cavities between the minimally wired and DSI networks as a consequence of the non-minimally wired edges, which tessellate cavities in the brain itself. Yet when

the subcortical regions are removed, the persistent cavities of the minimally wired and DSI networks are much more similar (Fig. 6b). This suggests that the wiring of cortical regions may be more heavily influenced by energy conservation than the wiring of subcortical regions. Additionally the drop in the number and lifetime of persistent cavities when subcortical regions are included indicates that these subcortical regions may prematurely collapse topological cavities. The often high participation of subcortical regions in maximal cliques suggests these well-connected nodes may have hub-like projections to regions involved in cortical cycles, thus tessellating the cortical cavity with higher dimensional cliques (topologically these subcortical nodes are *cone points*). Previous studies have found that networks with “star-like” configurations are optimally efficient in terms of shortest-path efficiency, but also efficient in terms of a random walk-based measure of efficiency (Goni et al. 2013). That is, networks optimized to have one or the other type of efficiency tend to have stars. Thus, stars appear to be useful configurations for fast communication, both along shortest paths and also in an unguided sense along random walks. The fact that we see star-like projections to cycles from subcortical regions may suggest that they are useful for efficient communication.

4.5 Methodological considerations

An important consideration relates to the data from which we construct the human structural connectome. DSI and tractography, non-invasive tools for mapping the brain’s white-matter connectivity, have some limitations. Tractography algorithms trade off specificity and sensitivity, making it challenging to simultaneously detect true connections while avoiding false connections (Thomas et al. 2014), fail to detect superficial connections (i.e. those that do not pass through deep white matter) (Reveley et al. 2015), and have challenges tracking “crossing fibers”, connections with different orientations that pass through the same voxel (Wedeen et al. 2008). Nonetheless, DSI and tractography represent the only techniques for non-invasive imaging and reconstruction of the human connectome. While such shortcomings limit the applicability of DSI and tractography, they may prove addressable with the development of improved tractography algorithms and imaging techniques (Pestilli et al. 2014).

4.6 Individual cavities in neuroscience applications

Though comparing persistent homology of weighted networks at the global level has been successful (for example Benzekry et al. 2015; Horak et al. 2009), scrutinizing individual persistent features may have more clinical relevance due to their size and understandability. Yet, multiple

questions remain to be answered before this goal can be achieved.

The first question pertains to the choice of representative cycle. As the current study presents an initial consideration of the persistent features of the structural connectome, we record *all* minimal generators, which reduces the number of choices made, and we define minimality using topological (hop) distance, which simplifies our analysis. However, a case could be made for using the representative with the minimal summed edge weight (Dey et al. 2011). Such a definition would further simplify the analysis by potentially giving a unique ‘minimal’ generator for each equivalence class. Additionally one might ask if a ‘minimal’ generator is even the appropriate representative cycle in the first place. Perhaps cycles of longer length have cognitive or clinical relevance beyond information distribution.

Second, it will be necessary to further develop the concept of similar persistent cavities. Here we used a region-matching process in order to incorporate perspectives from neuroscience and topology. An important open question is whether a more algorithmic matching could be devised that is better suited to the perspectives from both fields. Along the same lines, it is important to consider the birth, death time, and lifetime of a given persistent cycle (Stolz et al. 2017). We interpret longer-lived and earlier-born persistent cycles as more essential to the global architecture, and we hypothesize that this translates to healthy cognitive control and function as well. Then if two cavities are similar in terms of their regional composition, but are not similar in terms of birth or death times (for example, the blue cycle in the DSI *versus* MW networks in Fig. 5), it remains an open question whether the two cavities should be considered truly similar in a biological context.

Thirdly, with the development of algebraic-topological tools as described above, we speculate that comparing late-arriving persistent features could be important for clinical applications. Weaker connections have been shown to distinguish between health individuals and those with schizophrenia (Bassett et al. 2012), and have also been shown to predict individual differences in intelligence (Cole et al. 2012). Since late-born persistent cycles are a very particular arrangement of weak edges, we hypothesize that such cavities may be powerful biomarkers of individual brains, capable of distinguishing between diseased and normal connectomes.

5 Conclusion

In conclusion, we offer a unique perspective on the structural substrates of distinct types of neural computations. While traditional notions from graph theory and network science preferentially focus on local properties of the

network at individual vertices or edges (Bassett and Bullmore 2006, 2009; Bullmore and Sporns 2009; Bullmore and Bassett 2011), here we utilize an enriched network formalism that comes from the field of algebraic topology (Ghrist 2014). These tools are tuned to the interplay between weak and strong connections (Bassett et al. 2012), and therefore reveal architectural features that serve to isolate information transmission processes (Giusti et al. 2016). It will be interesting in the future to compare human and non-human connectomes across a range of spatial scales (Betzler and Bassett 2016) to further elucidate the evolutionary development of these features, and to link them to their functional (Hermundstad et al. 2013) and behavioral (Hermundstad et al. 2014) consequences.

Acknowledgments This work was supported from the John D. and Catherine T. MacArthur Foundation, the Alfred P. Sloan Foundation, the Army Research Laboratory and the Army Research Office through contract numbers W911NF-10-2-0022 and W911NF-14-1-0679, the National Institute of Mental Health (2-R01-DC-009209-11), the National Institute of Child Health and Human Development (1R01HD086888-01), the Office of Naval Research, and the National Science Foundation (CRCNS BCS-1441502 and CAREER PHY-1554488). We thank Scott T. Grafton for access to the DSI data.

Author Contributions DB and CG proposed the initial idea for the paper. AS performed research and prepared initial manuscript. All authors edited and approved the final manuscript.

Compliance with Ethical Standards

Conflict of interests The authors declare that they have no conflict of interest.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix: Data acquisition

All participants volunteered with informed consent in writing in accordance with the Institutional Review Board/Human Subjects Committee of the University of California, Santa Barbara. Diffusion spectrum imaging (DSI) scans were acquired from eight subjects (mean age 27 ± 5 years, two female, two left handed) on 3 separate days, for a total of 24 scans (Cieslak and Grafton 2014). DSI scans sampled 257 directions using a Q5 half-shell acquisition scheme with a maximum *b*-value of 5000 and an isotropic voxel size of 2.4 mm. We utilized an axial acquisition with the following parameters: repetition time (TR) = 11.4 s, echo time (TE) = 138 ms, 51 slices, field of view (FoV) (231,231,123 mm).

DSI data were reconstructed in DSI Studio (www.dsi-studio.labsolver.org) using *q*-space diffeomorphic reconstruction (QSDR) (Yeh and Tseng 2011). QSDR first reconstructs diffusion-weighted images in native space and computes the quantitative anisotropy (QA) in each voxel. These QA values are used to warp the brain to a template QA volume in Montreal Neurological Institute (MNI) space using the statistical parametric mapping (SPM) nonlinear registration algorithm. Once in MNI space, spin density functions were again reconstructed with a mean diffusion distance of 1.25 mm using three fiber orientations per voxel. Fiber tracking was performed in DSI studio with an angular cut-off of 55 degrees, step size of 1.0 mm, minimum length of 10 mm, spin density function smoothing of 0.0, maximum length of 400 mm and a QA threshold determined by DWI signal in the colony-stimulating factor. Deterministic fiber tracking using a modified FACT algorithm was performed until 100,000 streamlines were reconstructed for each individual.

In addition to diffusion scans, a three-dimensional high-resolution T1-weighted sagittal sequence image of the whole brain was obtained at each scanning session by a magnetization-prepared rapid acquisition gradient-echo sequence with the following parameters: TR = 15.0 ms; TE = 4.2 ms; flip angle = 9 degrees, 3D acquisition, FOV = 256 mm; slice thickness = 0.89 mm, matrix = 256×256 . Anatomical scans were segmented using FreeSurfer (Dale et al. 1999) and parcellated according to the Lausanne 2008 atlas included in the connectome mapping toolkit (Hagmann et al. 2008). A parcellation scheme including 83 regions was registered to the B0 volume from each subject's DSI data. The B0 to MNI voxel mapping produced via QSDR was used to map region labels from native space to MNI coordinates. To extend region labels through the gray–white matter interface, the atlas was dilated by 4 mm. Dilation was accomplished by filling non-labeled voxels with the statistical mode of their neighbors' labels. In the event of a tie, one of the modes was arbitrarily selected. Each streamline was labeled according to its terminal region pair.

Additional neighborhood-scale computations

In the main text we count maximal cliques at an edge density of 0.25 (Fig. 2). To ensure our interpretation would not fluctuate based on this choice of ρ , we also show the maximal clique distribution for $0 \leq \rho \leq 0.25$ for the average DSI network (Fig. 7a). For comparison, we include the average maximal clique distribution for $0 \leq \rho \leq 0.25$ of the minimally wired null models (Fig. 7b).

To address the extent to which an anterior-posterior gradient of maximal cliques exists, we calculated the correlation coefficient of $P_k(v)$ with the position of the node along

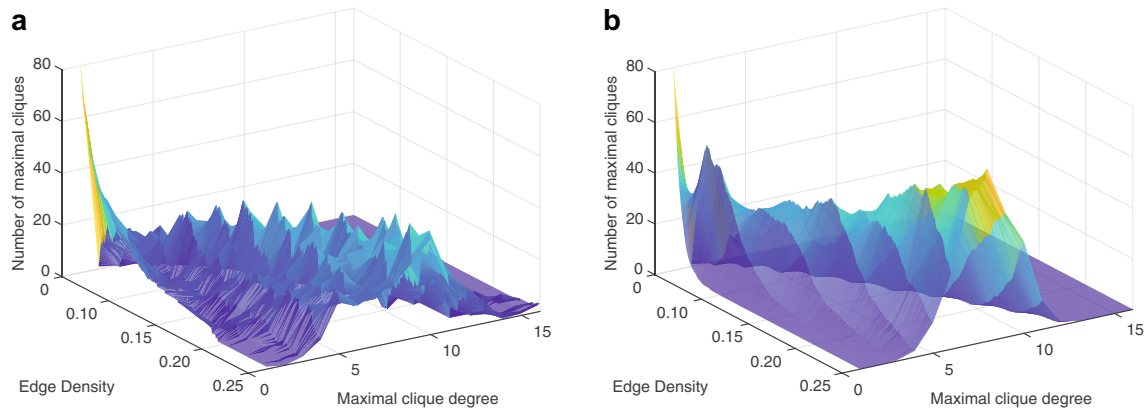


Fig. 7 Maximal clique distribution along the filtration. **a** Distribution of maximal cliques at each edge density for the average DSI network, and **b** average maximal clique distribution for the minimally wired networks

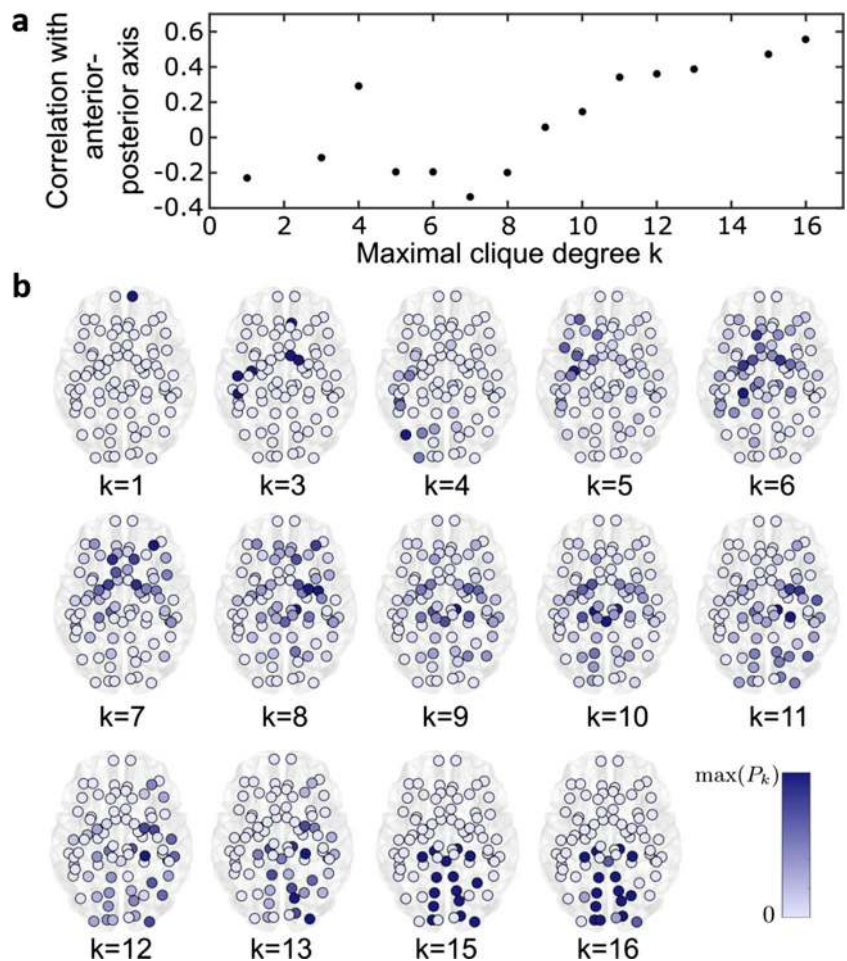
this axis. Fig. 8 shows generally the maximal participation of a node is more highly correlated with anterior-posterior position for higher degree cliques. To complement this calculation, Fig. 8b shows the normalized $P_k(v)$ of each node for all maximal clique degree k .

We then asked if node participation varies by cognitive system, perhaps reflecting each system’s unique function. Results are shown in Fig. 2. The specific ordering of nodes for this figure are shown below (Fig. 9b). For each (right,

left) hemisphere pair, the brain region in the right hemisphere was listed first, immediately followed by that in the left hemisphere.

Additionally we are interested in comparing node participation to other measures of connectedness, as we expect they should generally agree. One such measure is the rich club. Following the work of van den Heuvel and Sporns (van den Heuvel and Sporns 2011), we calculated ϕ , ϕ_{rand} , and ϕ_{norm} for each value of k (Fig. 10).

Fig. 8 Maximal clique correlation with anterior-posterior gradient. **a** Pearson correlation coefficient of $P_k(v)$ with the coordinate along the anterior-posterior axis. **b** Spatial distribution of $P_k(v)$ for each k . Color of node corresponds to the value of $P_k(v)$ between zero and the maximum participation of any node for the given degree k



Auditory	Banks of superior temporal sulcus (RH, LH)
	Superior temporal (RH, LH)
	Transverse temporal (RH, LH)
Cingulo-Opercular	Pars orbitalis (RH, LH)
	Pars opercularis (RH, LH)
	Rostral middle frontal (RH, LH)
	Rostral anterior cingulate (RH, LH)
	Caudal anterior cingulate (RH, LH)
	Supramarginal (RH, LH)
Default Mode	Superior frontal (RH, LH)
	Posterior cingulate (RH, LH)
	Isthmus cingulate (RH, LH)
	Precuneus (RH, LH)
Dorsal Attention	Superior parietal (RH, LH)
Fronto-Parietal	Frontal pole (RH, LH)
	Medial orbitofrontal (RH, LH)
	Pars triangularis (RH, LH)
	Caudal middle frontal (RH, LH)
	Inferior parietal (RH, LH)
	Insula (RH, LH)
Other	Parahippocampal (RH, LH)
	Entorhinal (RH, LH)
	Temporal pole (RH, LH)
	Inferior temporal (RH, LH)
	Middle temporal (RH, LH)
	Brainstem
Somatosensory	Precentral (RH, LH)
	Paracentral (RH, LH)
	Postcentral (RH, LH)
Subcortical	Thalamus proper (RH, LH)
	Caudate (RH, LH)
	Putamen (RH, LH)
	Pallidum (RH, LH)
	Accumbens area (RH, LH)
	Hippocampus (RH, LH)
	Amygdala (RH, LH)
Ventral Attention	Lateral orbitofrontal (RH, LH)
Visual	Cuneus (RH, LH)
	Pericalcarine (RH, LH)
	Lateral occipital (RH, LH)
	Lingual (RH, LH)
	Fusiform (RH, LH)

Fig. 9 Order of brain regions for Fig. 2b

Persistent homology

We are interested in finding mesoscale structural features, specifically *non-trivial minimal cycles* within our weighted network. Though these minimal cycles may *geometrically*

be quite large and span a large portion of the brain, we emphasize that these are mesoscale features from a topological perspective. Persistent homology strings together these features across network snapshots in a *filtration*, offering a global picture of network architecture. We include a brief description of the method here, and we advise the interested reader to consult (Carlsson 2009; Ghrist 2014; Zomorodian and Carlsson 2005) for additional details.

Complexes

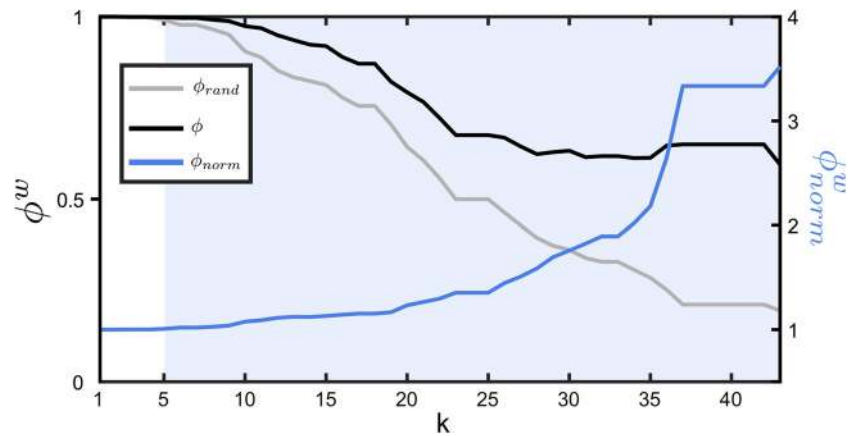
Cliques First, we will transform our network (equivalently, graph) of interest into an algebraic object so that we can use powerful computational tools from linear algebra to compute intuitive topological features. We begin by selecting building blocks from which to assemble larger, mesoscale structures. Drawing on classical graph theory and our intuition about the type of structures we are looking for, we are led to a natural (and well studied) choice of such blocks: sets of all-to-all connected nodes called *cliques*. In the context of brain networks, cliques are groups of brain regions that are able to rapidly and effectively share information. Formally, a $(k + 1)$ -clique of a graph G is a set of $(k + 1)$ nodes for which all pairwise edges are in G . Thus, a single node is a 1-clique, an edge a 2-clique, a triangle a 3-clique, and so on. Any subgraph of a clique must itself be a clique of lower degree, called a *face*. A *maximal* clique is thus any clique that is not a face. Intuitively, we will think of cliques as “filled in” regions, rather than hollow collections of edges (Fig. 11a).

Clique complex We study the structure formed by all cliques induced by the graph G , a combinatorial object called the *clique complex* (Fig. 11b). More specifically, we build the abstract simplicial complex formed from the correspondence of k -simplices and $(k + 1)$ -cliques. See Carlsson (2009), Hatcher (2002), and Ghrist (2014) for more details.

The clique complex of a graph G is the collection of all the cliques in G , formally denoted $X(G) = \{X_0(G), X_1(G), \dots, X_N(G)\}$ where $X_k(G)$ is the set of all $(k + 1)$ -cliques in G . Historically, the index is chosen to correspond to the dimension of the enclosed region, and we adopt this index shift here for consistency. The clique complex is an object which allows us to formally manipulate certain important geometric properties (as we explore in more detail in the following sections), and, through these computations, discover mesoscale features of interest.

Chain group In order to perform computations, we move from sets of cliques to vector spaces. We define the *chain group* $C_k(X(G))$ (abbreviated to C_k when the underlying clique complex is understood) as the vector space with basis $X_k(G)$. We denote by $\sigma_{i_1, i_2, \dots, i_k} \in C_k(X(G))$ the

Fig. 10 Defining the rich club of the DSI network. Rich club coefficient of the DSI network ($\phi(k)$) is shown in black, the average rich club coefficient of randomized networks ($\phi_{rand}(k)$) in gray, and the normalized rich club coefficient $\phi_{norm}(k)$ in blue. Shaded regions indicate values of k for which $\phi(k)$ significantly exceeded $\phi_{rand}(k)$



basis element corresponding to a $(k + 1)$ -clique on nodes $\{i_0, i_1, \dots, i_k\}$. Though this definition can be made for any scalar field, we use vector spaces over the field with two elements, $\mathbb{F}_2 = \{0, 1\}$, as is standard in topological data analysis. Elements of $C_k(X(G))$ are linear combinations of k -chains which correspond to collections of $(k + 1)$ -cliques.

For example, consider the clique complex $X(G)$ shown in Fig. 12. Elements of C_1 are linear combinations of edges, or 2-cliques. One such element is $b = \sigma_{5,6} + \sigma_{6,7} + \sigma_{7,8}$, shown in blue in Fig. 12. This is intuitively an undirected path from v_5 to v_8 that passes through v_6 and v_7 . We could also take the purple path $a \in C_1$. This path begins at v_0 and follows $\sigma_{0,1}, \sigma_{1,2}, \sigma_{2,5}$, then $\sigma_{0,5}$ which returns us to v_0 . Because we work over \mathbb{F}_2 , this algebraic encoding is not sensitive to clique direction, only the parity of the number of times a clique appears in a chain. In C_2 , an element is a linear combination of 3-cliques. Highlighted in Fig. 12 (right) is one such example: the element $c \in C_2$ with $c = \sigma_{2,3,4} + \sigma_{2,4,5}$. Because we are working in \mathbb{F}_2 , if we took this path twice, we would have the chain $c + c = \sigma_{2,3,5} + \sigma_{2,4,5} + \sigma_{2,3,5} + \sigma_{2,4,5} = 2\sigma_{2,3,5} + 2\sigma_{2,4,5} = 0$.

Boundary operator Recall that our goal is to detect topological cavities in our algebraic object. Note the structure of cycles is subtle and not necessarily indicative of physical cavities in a general sense. However, in the case of these relatively sparse 3D graphs this is usually the case. Cavities exist when cliques are arranged in a loop or capsule, but there are no higher dimensional cliques that “fill in”

the enclosed space – that is, the capsule is not the “boundary” of some collection of higher dimensional cliques. To detect this computationally, we use the *boundary operator* $\partial_k : C_k \rightarrow C_{k-1}$, which takes a collection of $(k + 1)$ -cliques (an element of C_k) and sends them to their boundary (an element of C_{k-1}).

Geometrically, the boundary of a k -clique is the family of $(k - 1)$ -cliques obtained by removing each vertex in succession. The boundary of a contiguous collection of (one or more) k -cliques is a “capsule” of $(k - 1)$ -cliques surrounding the original collection, inside of which the boundaries of neighboring $(k - 1)$ -cliques overlap. We can detect this pattern computationally when chains corresponding to the shared faces cancel. In Fig. 13 the boundary of $c \in C_2$ is the chain corresponding to the surrounding four edges (2-cliques), as the interior edge ($\sigma_{2,4}$) cancels. Formalizing this intuition, we define the boundary operator (with coefficients in \mathbb{F}_2) on the basis $X_k(G)$ to be

$$\partial_k(\sigma_{0,1,\dots,k}) = \sum_{i=0}^k \sigma_{0,1,\dots,\hat{i},\dots,k}$$

where \hat{i} indicates that vertex i is not included in the set of vertices that form the clique, and we extend this map linearly to all of $C_k(X(G))$. Then, for example, in Fig. 13,

$$\begin{aligned} \partial_2(c_3) &= \partial(\sigma_{2,3,4} + \sigma_{2,4,5}) = \partial_2(\sigma_{2,3,4}) + \partial_2(\sigma_{2,4,5}) \\ &= (\sigma_{3,4} + \sigma_{2,4} + \sigma_{2,3}) + (\sigma_{4,5} + \sigma_{2,5} + \sigma_{2,4}) \\ &= \sigma_{3,4} + \sigma_{2,3} + \sigma_{4,5} + \sigma_{2,5}. \end{aligned}$$

Fig. 11 From cliques to a clique complex. **a** Cliques are all-to-all connected sets of nodes which we use as “filled in” building blocks. **b** The clique complex is created by inserting these building blocks into the completely connected subgraphs of G

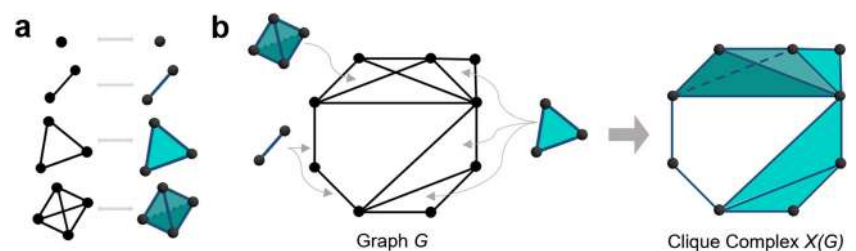
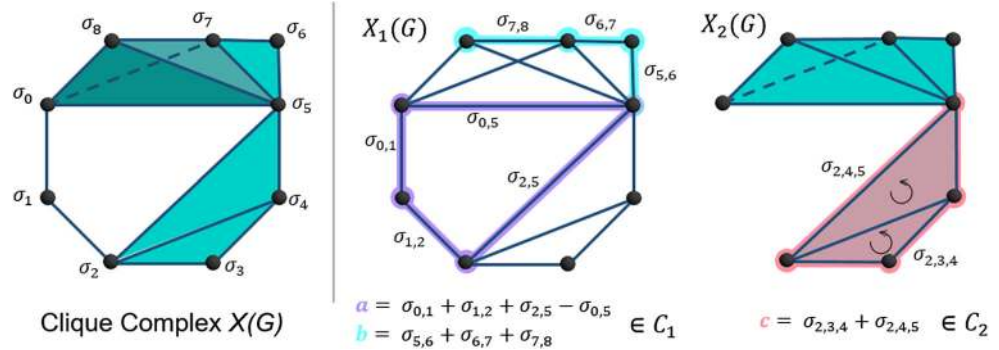


Fig. 12 Chain group elements are linear combinations of cliques. See Appendix text for a complete description of these graphs



Because the boundary of $c_3 \in C_2$ is itself an element of C_1 , we can apply ∂_1 to it in turn. As illustrated in Fig. 13,

$$\begin{aligned} \partial_1(\partial_2(c_3)) &= \partial_1(\sigma_{2,3} + \sigma_{3,4} + \sigma_{4,5} + \sigma_{2,5}) \\ &= \sigma_3 + \sigma_2 + \sigma_4 + \sigma_3 + \sigma_5 + \sigma_4 + \sigma_5 + \sigma_2 \\ &= 2\sigma_2 + 2\sigma_3 + 2\sigma_4 + 2\sigma_5 \\ &= 0. \end{aligned}$$

This example illustrates a crucial property of the boundary operator: $\partial_{k-1} \circ \partial_k = 0$, which will be more thoroughly discussed in the Homology section below.

Chain complex We now have a boundary operator that lets us move from k -chains to $(k - 1)$ -chains for every k . Note the boundary of a 0-chain is defined to be 0, since a node is a single point with no geometric boundary. These operators link together the chain groups into a sequence

$$\xrightarrow{\partial_{k+1}=0} C_k \xrightarrow{\partial_k} C_{k-1} \xrightarrow{\partial_{k-1}} \dots \xrightarrow{\partial_2} C_1 \xrightarrow{\partial_1} C_0 \xrightarrow{\partial_0=0} 0$$

called the *chain complex* for $X(G)$. This is our fundamental algebraic tool for studying the structure of the clique complex.

In summary, we have taken an unweighted, undirected graph G and, from an enumeration of its cliques, formed the clique complex $X(G)$ (Fig. 14, left). We then used the cliques of each dimension as basis elements in the chain groups $C_0(X(G)), C_1(X(G)), \dots, C_N(X(G))$ (Fig. 14, middle). Finally, we defined the boundary operator ∂ that finds the boundary of a chain (which represents a collection of $(k + 1)$ -cliques), itself a (possibly empty)

chain representing a collection of k -cliques, and we used this function to string together the chain groups into the chain complex (Fig. 14, right).

Homology

We turn now to the definitions and concepts needed to compute homology. Homology discovers features of interest in the clique complex by separating *cycles*, mesoscale patterns constructed from cliques, which surround a cavity from those that are the boundary of a collection of cliques.

Cycles Though we have seen examples of cliques strung together as paths, we are particularly interested in paths that form closed structures called *cycles*, the 1-dimensional analog of which are graph-theoretic circuits. Consider the three closed circuits in Fig. 15, each can be thought of as a linear combination of elements in $C(X_1(G))$. If we begin at any 1-clique (node) on the cycle, for example σ_2 in ℓ_1 , and traverse each 2-clique in the cycle in order, we will end at our starting 1-clique. Since the boundary of any path $\in C(X_1(G))$ is $\sigma_{end} + \sigma_{begin}$, the boundary of any cycle $\ell \in C(X_1(G))$ must be

$$\partial_1(\ell) = \sigma_{end} + \sigma_{begin} = 2\sigma_{begin} = 0.$$

Though we have thus far focused on the familiar notion of cycles built of 2-cliques, the notion that boundaries should cancel allows us to construct cycles in any dimension. We define a k -cycle to be any element $\ell \in C_k$ with $\partial_k(\ell) = 0$. Since the cycles are exactly the elements that are sent to 0 by

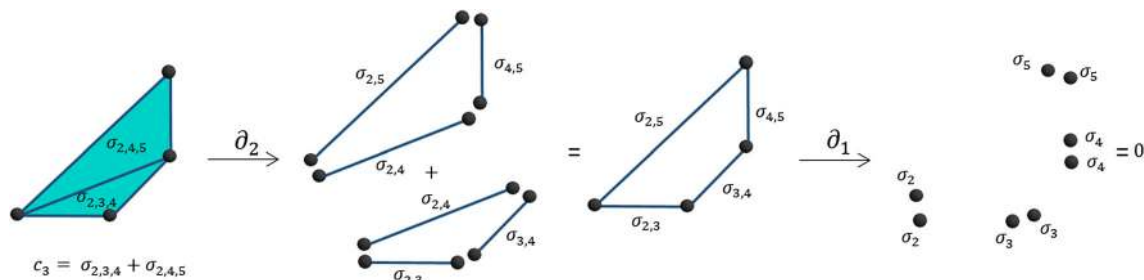
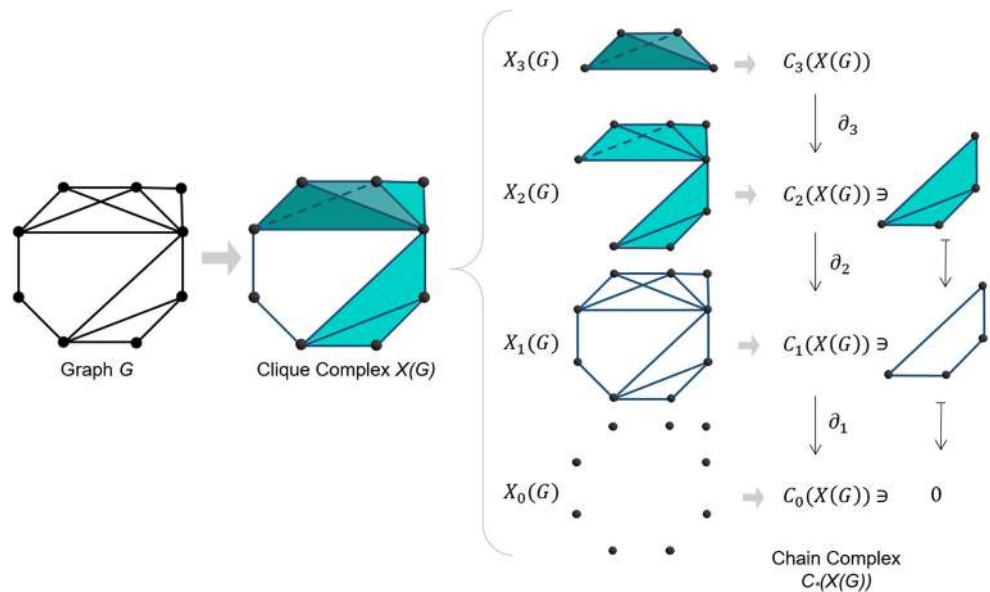


Fig. 13 Example of the boundary operator in C_2 . See Appendix text for a complete description of these graphs

Fig. 14 Illustration of creating from G the clique complex $X(G)$. Also shown are the induced chain complex $C_*(X(G))$ and an example of boundary calculations on an element in $C_2(X(G))$



the boundary operator, the subspace of k -cycles is precisely the kernel (or nullspace), denoted $\ker(\partial_k) \subset C(X_k(G))$.

As noted above, cycles can surround either cavities or a collection of cliques, and since we are strictly interested in cycles of the first type, we must determine how to differentiate between these two options. Figure 15 depicts three 1-cycles found in the clique complex shown on the left. Looking strictly at $X_1(G)$, we cannot distinguish which of these three cycles belong to which category.

However if we include information about 3-cliques, the separation becomes apparent, in the same way looking at the full depiction of the clique complex in Fig. 15 (left) makes it apparent that this object surrounds one cavity. We need consider only the image of the boundary map from $\partial_2 : C_2(X(G)) \rightarrow C_1(X(G))$: if a 1-cycle ℓ surrounds a collection of higher dimensional cliques, it must in particular surround a collection of 2-cliques (2-faces of these larger cliques). In our example in Fig. 15, this means ℓ_1 is the boundary of some element in $C_2(X(G))$ (this element is $\sigma_{2,3,4} + \sigma_{2,4,5}$).

We can repeat such an argument for any k -cycle that surrounds a collection of higher dimensional cliques, which allows us to define k -boundaries as elements in $\text{im}(\partial_{k+1}) \subseteq C_K(X(G))$. Furthermore it must be true that $\text{im}(\partial_{k+1}) \subseteq \ker(\partial_{k+1})$ per our previous observation that $\partial_k \circ \partial_{k+1} = 0$.

However, not all cycles are necessarily boundaries: ℓ_2 and ℓ_3 are in $\ker(\partial_1)$ but neither are elements of $\text{im}(\partial_1)$. The k -cycles that surround cavities are thus those that are in $\ker(\partial_k) - \text{im}(\partial_k)$ is not enough to produce a proper list of cavities in our clique complex, because we will suffer from redundancy. For example, knowing either ℓ_2 or ℓ_3 tells us the cavity they both enclose exists. Certainly $\ell_2 \neq \ell_3$, but we should consider them equivalent since they both reveal the same feature of our complex. So we need a way to count more carefully.

Equivalence The solution to our enumeration problem will depend on what we regard as “the same”. Above we mentioned we should consider ℓ_2 to be equivalent to ℓ_3 because

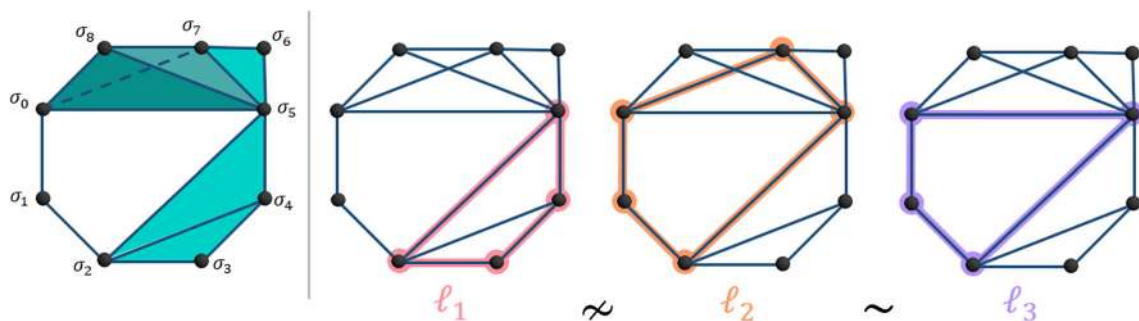


Fig. 15 Cycles. Examples of a cycle that is also a boundary (ℓ_1) and two equivalent, non-boundary cycles (ℓ_2 and ℓ_3)

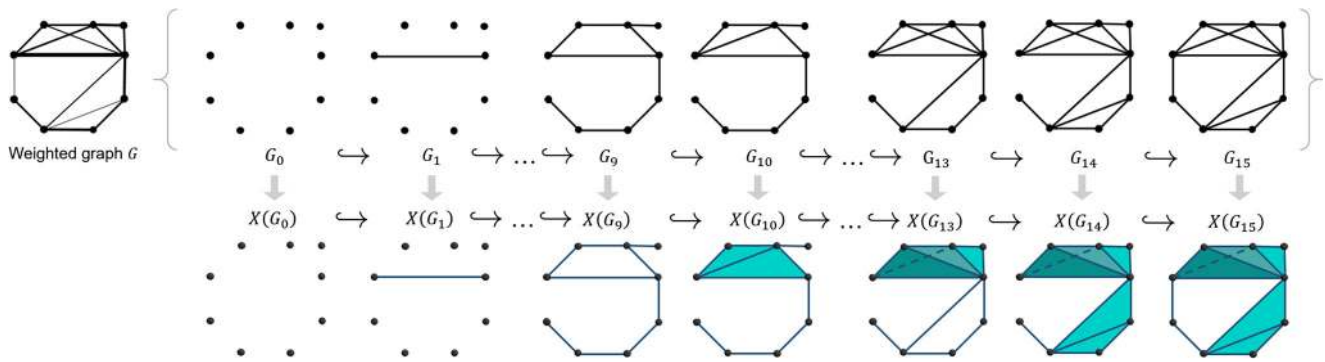


Fig. 16 Filtrations and inclusion maps. Edge weights indicated by line thickness induce a filtration on the weighted graph G . The inclusion maps $G_i \hookrightarrow G_{i+1}$ induce inclusion maps on the corresponding clique complexes $X(G_i) \hookrightarrow X(G_{i+1})$

they surround the same cavity. How is it that we understood this? We see they both enclose this cavity, while ℓ_2 also surrounds one 3-clique. But this 3-clique (specifically $\sigma_{0,5,7}$) does not change the cavity or add a new one, so we decided this difference of a higher dimensional clique should be insubstantial, and thus the two cycles are equivalent. Generalizing this example provides a method for correctly enumerating the cavities in the complex.

Two k -cycles, ℓ_i and ℓ_j , are called *equivalent* if their sum, (working over \mathbb{F}_2) $\ell_i + \ell_j$ is the boundary of a $(k + 1)$ -chain, e.g. $\ell_i \sim \ell_j$ if $\ell_i + \ell_j \in \text{im}(\partial_{k+1})$. In Fig. 15, we have

$$\begin{aligned} \ell_2 + \ell_3 &= (\sigma_{2,5} + \sigma_{5,7} + \sigma_{7,0} + \sigma_{0,1} + \sigma_{1,2}) \\ &\quad + (\sigma_{2,5} + \sigma_{5,0} + \sigma_{0,1} + \sigma_{1,2}) \\ &= \sigma_{5,7} + \sigma_{7,0} + \sigma_{0,5} \\ &= \partial_2(\sigma_{0,5,7}) \in \text{im}(\partial_2) \end{aligned}$$

so indeed we see $\ell_2 \sim \ell_3$.

This, finally, provides us with a proper count: if we only count one cycle from each set of (non-trivial) equivalent cycles, then we will have precisely the number of topological cavities of a given dimension within the clique complex. The clique complex in Fig. 14 by eye has only one cavity surrounded by 1-cycles, and our computations agree. Any closed loop of 2-cliques either is equivalent to ℓ_2 or it is strictly a boundary of higher dimensional cliques and thus is trivial. So, as desired, we have a sole 2-dimensional cavity.

The *equivalence class* of a k -cycle ℓ is $[\ell] = \{v \in Z_k \mid v \sim \ell\}$. Note the equivalence class of boundary loops $b \in \text{im}(\partial_k)$ contain the empty set, since $b - \emptyset = b \in \text{im}(\partial_k)$. This means for any $\ell \in \ker(\partial_k)$ and $b \in \text{im}(\partial_k)$, we have $\ell + b \sim \ell + \emptyset \sim \ell$, confirming our requirement that cycles differing by boundaries are equivalent. By abuse, it is common to refer to an equivalence class of k -cycles as a k -cycle, and we will continue with this convention.

Homology groups The heavy lifting is now complete and we are left with only the formal definition of homology to conclude the section. Recalling the equivalence classes we

have discussed above, we define the homology group of dimension n as

$$H_n := \ker(\partial_n) / \text{im}(\partial_{n+1})$$

which is simply the vector space spanned by equivalence classes of n -cycles. The dimension of H_n is the number of nontrivial n -cycles and thus the number of $(n + 1)$ -dimensional topological cavities of our clique complex. In summary we can now take a graph of nodes and edges, convert it to an algebraic object called the clique complex, then use the boundary operator to find equivalence classes of cycles that describe essential mesoscale architecture of our network in the form of topological cavities.

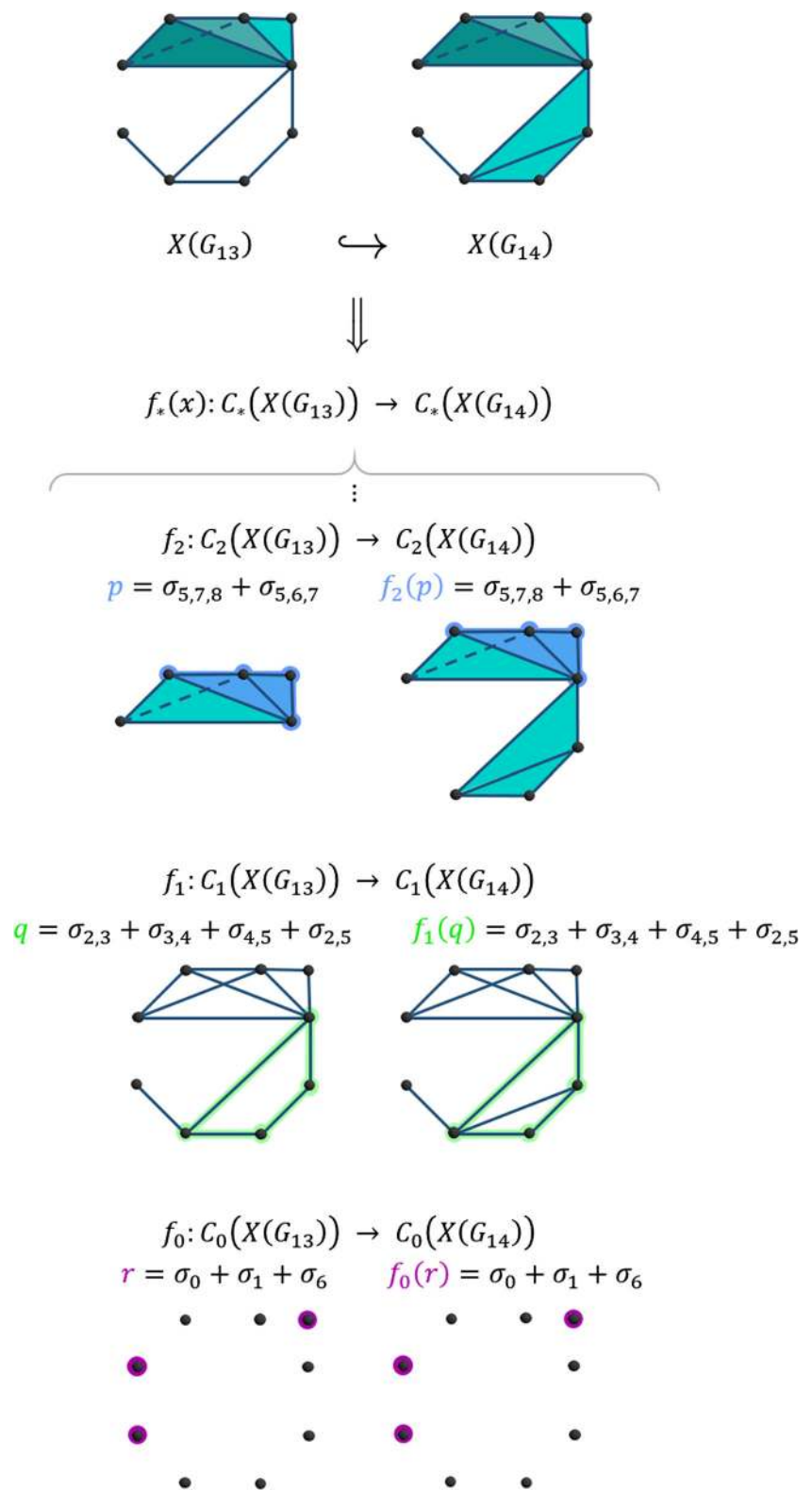
Homology for weighted networks: persistent homology

While homology detects cavities in binary graphs, the DSI data (and many other sources in biology) create a weighted network. *Persistent homology* was originally developed (Carlsson 2009; Zomorodian and Carlsson 2005) to describe topological features of high-dimensional point clouds, but has since been adapted to address the current problem of finding topological cavities within weighted networks. This method uses the edge weights to unravel the weighted network into a sequence of binary networks on which we can then compute homology, in a manner related to but more principled than standard thresholding techniques. Overall persistent homology perceives how the features seen with homology evolve with the weighted network.

Filtrations Given G a weighted network, we first construct a sequence of binary graphs that will allow us to use homology on each graph in the sequence. The edge weights induce a natural ordering on the edges from highest to lowest weight. Then, beginning with the empty graph, we replace edges following this ordering. This process creates a *filtration*

$$G_0 \subset G_1 \subset \dots \subset G_{|E|} = G$$

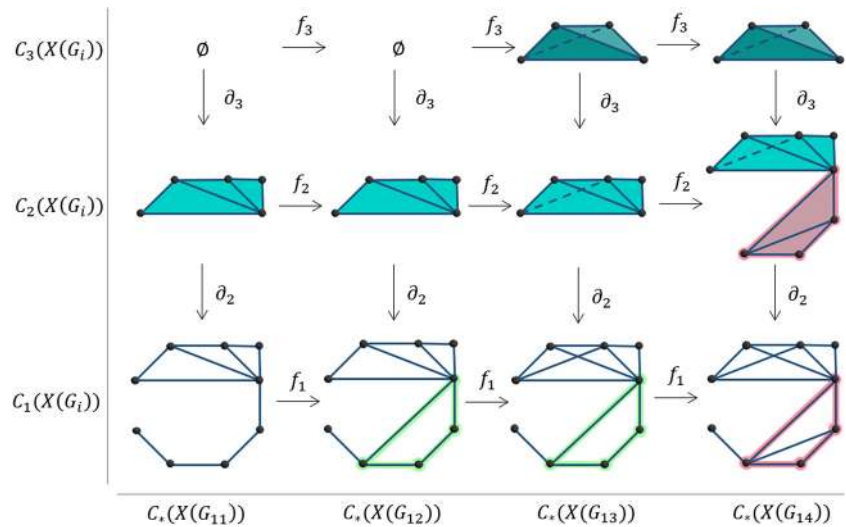
Fig. 17 Inclusion maps between clique complexes induce maps between the corresponding chain complexes. See Appendix text for a complete description of these graphs



where each G_{i+1} contains one more edge than G_i . Since G_{i+1} contains G_i (and one more edge), we obtain an inclusion map $i : G_i \hookrightarrow G_{i+1}$ which describes how G_i maps into G_{i+1} . In our case this is quite natural, G_i is sent to

itself, now a subgraph of G_{i+1} (Fig. 16, top row). This process to create a filtration from a weighted graph has been used previously in Petri et al. (2013a, b) and Giusti et al. (2015).

Fig. 18 Illustration of the persistence complex of the weighted graph G . The green 1-cycle is first seen in $X(G_{12})$, is mapped through filtrations, and finally becomes the boundary of a collection of 3-cliques (pink) in $X(G_{14})$



Having an inclusion of G_i into G_{i+1} means we can also get an inclusion of $X(G_i)$ into $X(G_{i+1})$ in a similar fashion, where cliques in $X(G_i)$ map to their corresponding selves in $X(G_{i+1})$ (Fig. 16, bottom row).

But now knowing how one clique complex maps into the next clique complex means we get maps between the chain groups as well. For example, in Fig. 17 we look only at the inclusion of $X(G_{13})$ into $X(G_{14})$. This inclusion map tells us how to take cliques from $X(G_{13})$ and fit them into $X(G_{14})$, which means we can figure out how to take some combination of cliques and fit them into $X(G_{14})$ as well. The functions that perform this task are defined

$$f_* : C_*(X(G_{13})) \rightarrow C_*(X(G_{14}))$$

where the $*$ refers to the set of functions indexed by dimension. We show the first three with examples in Fig. 17. If we have a 0-chain $r = \sigma_0 + \sigma_1 + \sigma_6 \in C_0(X(G_{13}))$, it gets mapped by f_0 to a chain in $C_0(X(G_{14}))$, explicitly $f_0(r) = \sigma_0 + \sigma_1 + \sigma_6$.

We can do this in the higher dimensions as well. Figure 17 also shows the green 1-chain $q = \sigma_{2,3} + \sigma_{3,4} + \sigma_{4,5} + \sigma_{2,5} \in C_1(X(G_{13}))$ and how it maps into $C_1(X(G_{14}))$ as well. It is interesting here to note that in $C_1(X(G_{13}))$, the 1-chain q is also a 1-cycle, but is equivalent to the trivial cycle in $C_1(X(G_{14}))$. Again we can move to the 2-chains and observe how $p = \sigma_{5,7,8} + \sigma_{5,6,7}$ is sent to $f_2(p) = \sigma_{5,7,8} + \sigma_{5,6,7} \in C_2(X(G_{14}))$.

Generally filtrations are a powerful way to understand weighted networks. Here, we will use these chain maps f_* to track particular chains throughout the filtration to see how they may change as new edges (and thus cliques) are added.

Persistent homology As we are interested in cycles, we now turn to tracking specifically cycles throughout the filtration. A k -loop is a k -chain, so it can be tracked horizontally from clique complex to clique complex in the

filtration. Additionally, we have vertical boundary maps that tell us if the k -loop in question is a cycle or a boundary loop within the particular clique complex. More generally we are combining the information from the filtration and its between-complex induced maps (Figs. 16, 17) with the boundary loop information from the within-complex boundary operators (Fig. 14) to observe how cycles change as we add edges of decreasing weight.

Formally these maps and complexes form the *persistence complex* of our weighted graph G (Fig. 18). Armed with inclusion and boundary maps between chain groups, we can compute the homology of each graph in the filtration and therefore obtain maps $H_*(X(G_i)) \rightarrow H_*(X(G_{i+1}))$ that describe how cycles (equivalence classes of cycles) in $X(G_i)$ change (map directly, shrink in length, become a boundary loop) in $X(G_{i+1})$.

For example, in Fig. 18 we see the green 1-cycle first appears in G_{12} . We say the cycle is *born* at this edge density $\rho_{birth} = (\# \text{ edges present}) / (\# \text{ edges possible}) = 12/36$. The green cycle continues to exist until it maps to a cycle that is the boundary of the pink 2-chain in $C_2(X(G_{14}))$. Since this cycle is now a boundary, it is equivalent to the trivial cycle in $H_1(X(G_{14}))$. We say the cycle *dies* at this edge density $\rho_{death} = 14/36$.

Cycles that exist over many edge additions must evade becoming triangulated by cliques, thus becoming a boundary. Therefore we consider such cycles more essential if they *persist* for many edge additions. We measure cycle persistence in two ways. First we record cycle *lifetime* $l = \rho_{death} - \rho_{birth}$, which is commonly used in persistent homology calculations (Carlsson 2009) and displayed on a persistence diagram. For our cycle which is born at $\rho = 12/36 = 1/3$ and dies at $\rho = 14/36 = 7/18$, we see an example persistence diagram in Fig. 19. However, recent work (Bobrowski et al. 2015) suggests alternatively considering $\pi = \rho_{death} / \rho_{birth}$ which allows for cycle

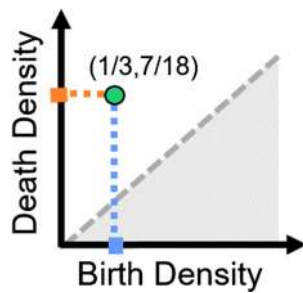


Fig. 19 Example persistence diagram for green cycle shown in Fig. 18. See Appendix text for a complete description of these graphs

persistence comparison at difference scales and underscores the importance of cycles forming at low edge densities.

To summarize, persistent homology tracks interesting connection patterns (cycles) through network frames induced by edge weights, recovering a parameter-free perspective on essential structural features in a weighted network.

Comparison with alternative loop-finding algorithms

One may ask how our method compares with other loop-finding algorithms. While such programs can be powerful, two fundamental differences exist. The first is in the definition of cycles identified. Recall that we extract equivalence classes of cycles, so we will find only cycles that enclose a structural cavity, while loop-finding algorithms will extract all loops that are boundaries of higher cliques (Tucker 2006). Additionally, persistent homology detects cycles in multiple dimensions with much less computational effort than loop algorithms (Johnson 1975).

Additionally one might ask how small changes in edge weights or edge ordering may affect these findings. Cohen-Steiner et al. showed generally small changes in the edge ordering will result in small changes in the persistence diagram (Cohen-Steiner et al. 2007). This makes persistent homology relatively robust to noise and consequentially a powerful tool in neuroscience (Giusti et al. 2016).

Cycles in the average DSI data

To understand the function non-boundary cycles may have in the structural brain network, we recover all minimal generators at ρ_{birth} for each persistent homology class found in the averaged DSI data (Fig. 4c). These cycles for all 20 of the 2D cavities and the two 3D cavities are shown below in Figs. 20, 21, respectively. To summarize this information we plot all minimal representatives with edges weighted by their participation in minimal representatives. This summarization is similar to the frequency scaffold (Lord et al.

2016; Petri et al. 2014) in Fig. 22, though here we are unable to assign one minimal representative to each persistent equivalence class so if an edge is part of any of the minimal representatives of one equivalence class it gets an added weight of one. Cycles reach most areas of the brain, and as seen in Fig. 20, many follow the cortical to subcortical theme. The edge involved in the highest number of dimension-one minimal generators in the average DSI data links the left and right thalamus. For dimension 2 we see each edge only exists within one minimal generator.

Confirming topological cavities in contralateral hemisphere

In the main text we show validation of the four highlighted cycles in individual scans. Following the procedure above, we next ask if these cycles are seen in the contralateral hemisphere to assess symmetry of these features. Figure 23 shows these features are seen in the contralateral hemisphere, though with less frequency than in the original.

Cavities in the normalized dataset

When studying the network formed from DSI, it is important to consider any potential bias created by the different sizes of the 83 brain regions. To account for this potential bias, we normalized the original network of streamline counts by the geometric mean of the end point region sizes and checked to see which cycles were still present (Hagmann et al. 2008). More precisely, the normalized edge weight $A_{i,j}$ between nodes i and j is streamline count $_{ij}/(\text{volume}_i \text{volume}_j)^{1/2}$ (Bassett et al. 2011).

After this normalization, we asked if the cycles found in the streamline counts data are present in the normalized networks. Figs. 23 (DSI Norm, DSI Norm cont) show the cycles are found to a similar extent across scans in the original and contralateral hemispheres.

Locating all cavities from the group-averaged DSI in the minimally wired networks

Noting many persistent cycles seem likely sampled from the minimally wired distribution of persistent cycles, we asked if we detect the 20 cycles observed in the average data in the null model. Figure 24 show the lifespan of each of these persistent cycles within the individual scans (black) and the minimally wired null model (gray). Each vertical bar represents a persistent cavity within a scan, and scans where the cavity was not validated are removed. Average birth and death densities are indicated with horizontal dashed lines. We surprisingly see very few of the persistent homology classes of the DSI data have counterparts in the

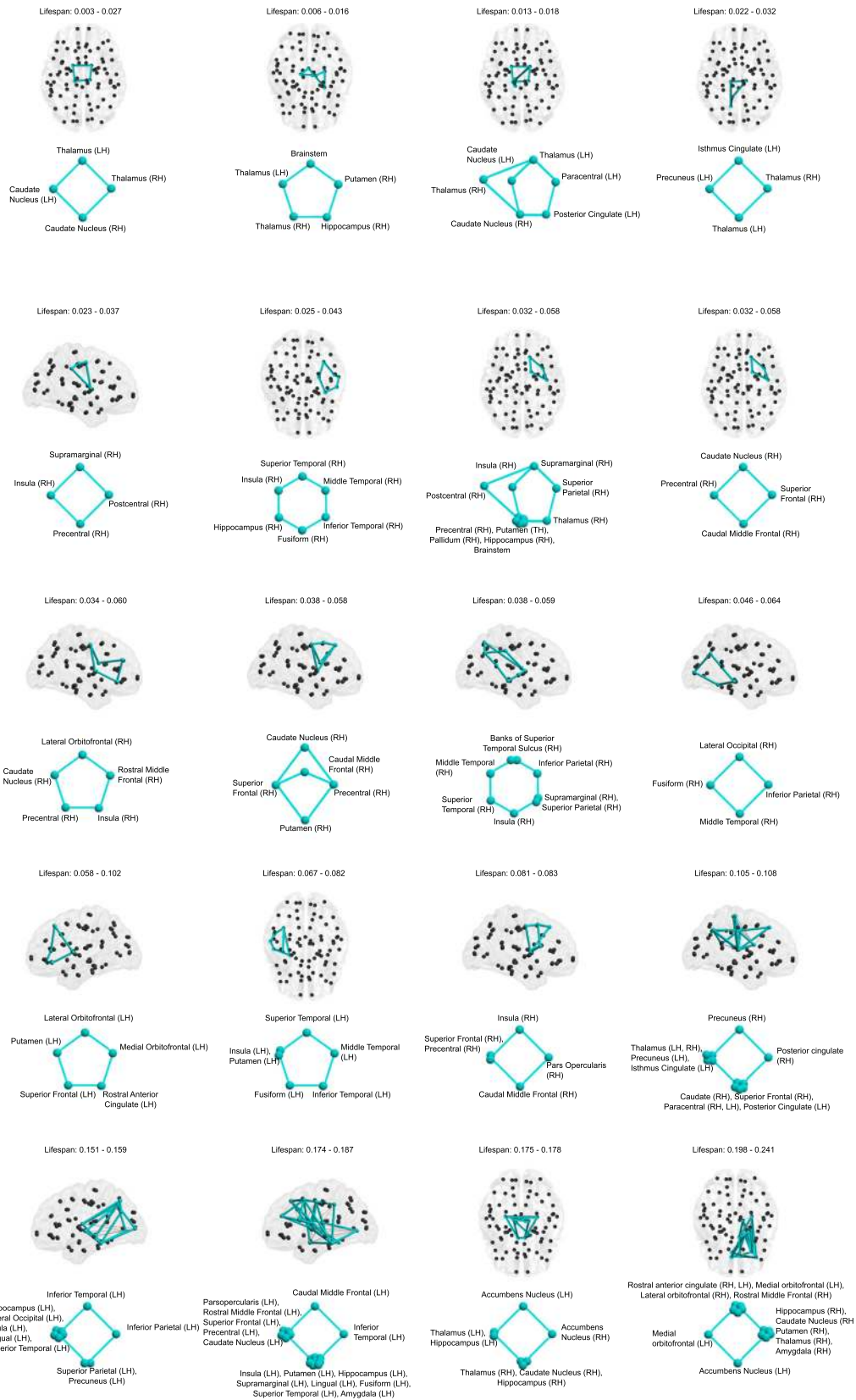
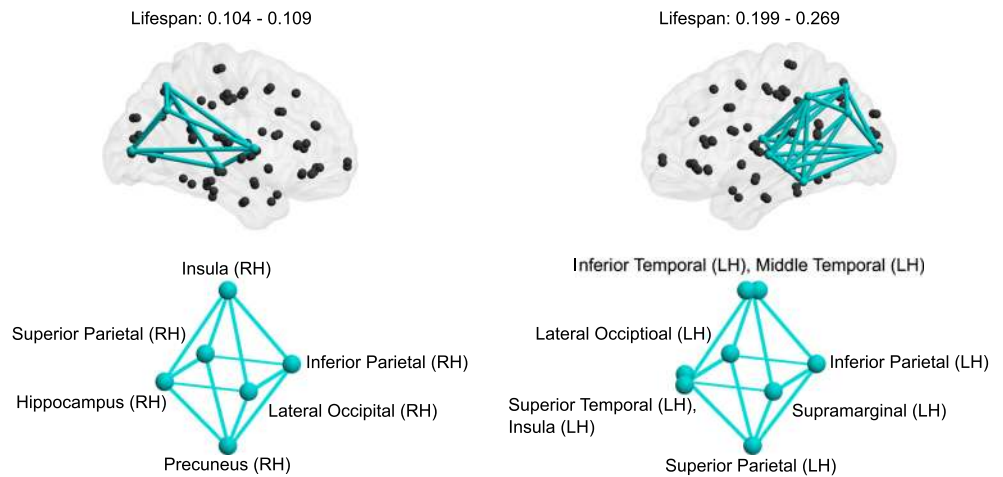


Fig. 20 Minimal representatives at ρ_{birth} of all 2D cavities found in the average DSI data, listed in order of increasing birth density. For each topological cavity, the lifespan ($\rho_{birth} - \rho_{death}$), location in

the brain, and schematic is shown. For the third, seventh, and tenth appearing cavities, we could not isolate exactly one unique equivalence class

Fig. 21 Minimal representatives at ρ_{birth} of all 3D cavities found in the average DSI data, listed in order of increasing birth density. For each, the lifespan ($\rho_{birth} - \rho_{death}$), location in the brain, and schematic is shown



minimally wired null model. Of those that do, often the average birth and death times are quite different, underscoring the importance of the filtration in this method (Figs. 24 and 25).

Cortical cavities

Densely connected subcortical nodes may prevent the longevity of nonzero homology classes by forming cross-cycle edges or cliques which tessellate the cycle completely. We asked what cavities could be found when removing these subcortical nodes, forming DSI^{cort} as described in the main text. Here, Fig. 26 shows a 1-cycle on nine nodes recovered from DSI^{cort} within the brain and as a schematic (panel (a)). The persistence diagram for 2D cavities within DSI^{cort} in Fig. 26b shows the four minimal cycles marked

in maroon. Importantly, because of the connection patterns between nodes at the density of cycle formation, we will refer to any of these four cycles as the minimal cycle. Two of these cycles are equivalent loops which involve the superior frontal (RH) and the caudal middle frontal regions. The other two are equivalent to each other but not to the first two loops, and involve the superior frontal (LH) and posterior cingulate (LH). The edge added at ρ_{birth} connects the lateral orbitofrontal to the superior temporal. The cycle formed by the superior frontal (RH, LH), caudal middle frontal, precentral, and posterior cingulate (LH) is itself a minimal cycle surrounding a separate topological cavity. This information along with the connection patterns at ρ_{birth} mean we cannot claim either pair are the two minimal generators, instead it is either one pair or the other. The smaller, five

Fig. 22 Spatial distribution of minimal generators at ρ_{birth} of 2D (top) and 3D (bottom) persistent cavities. Edge thickness reflects the number of minimal generators in which they participate

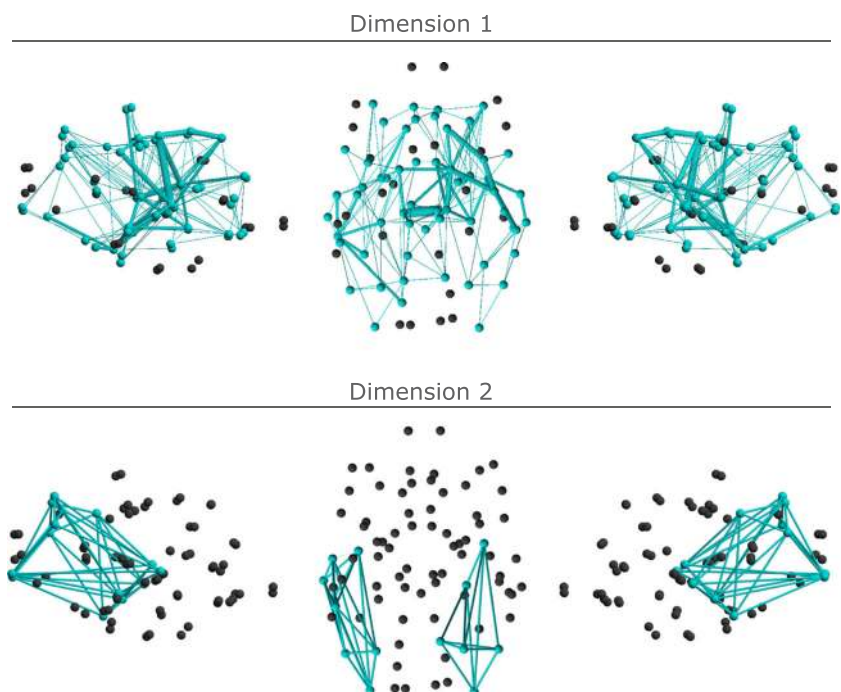


Fig. 23 Validation of similar topological cavities in additional data. For each of the four minimal generators highlighted in the main text, bars indicate cavity lifetime for all collected data. Dotted lines indicate average birth or death edge density

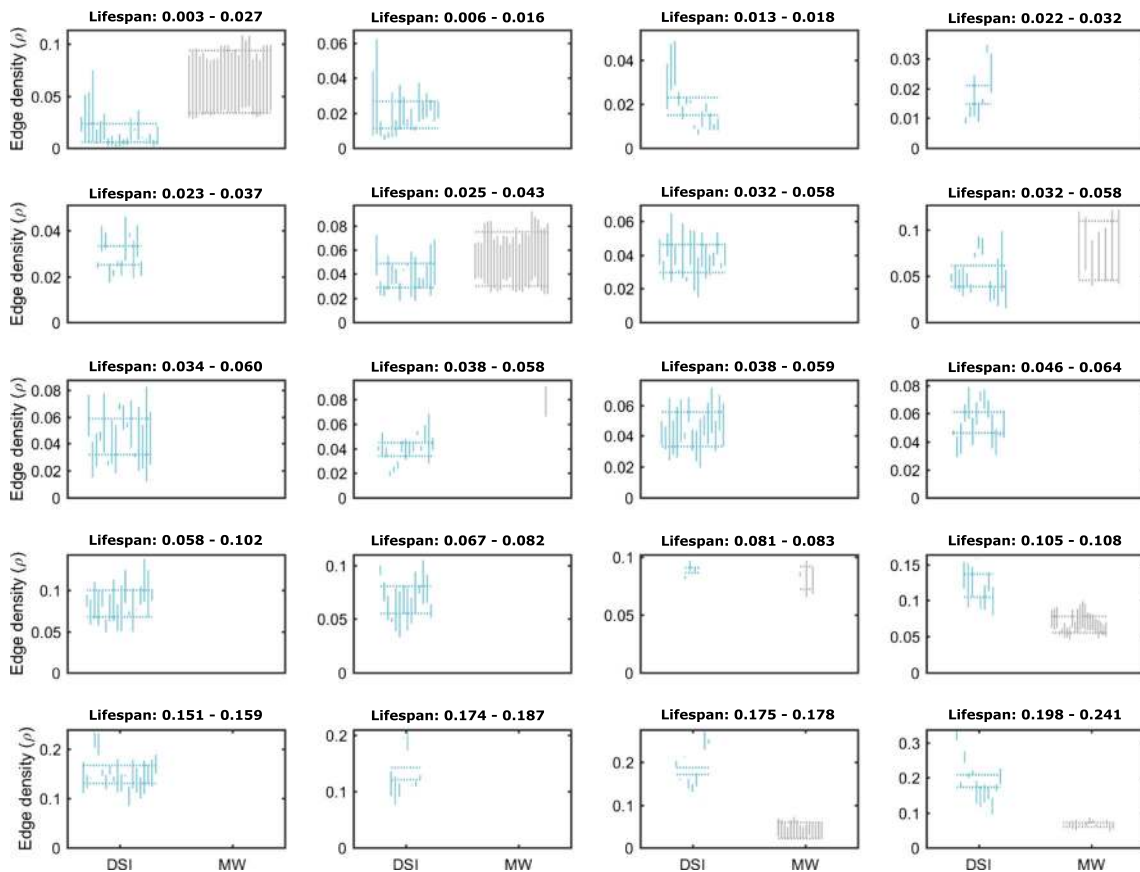
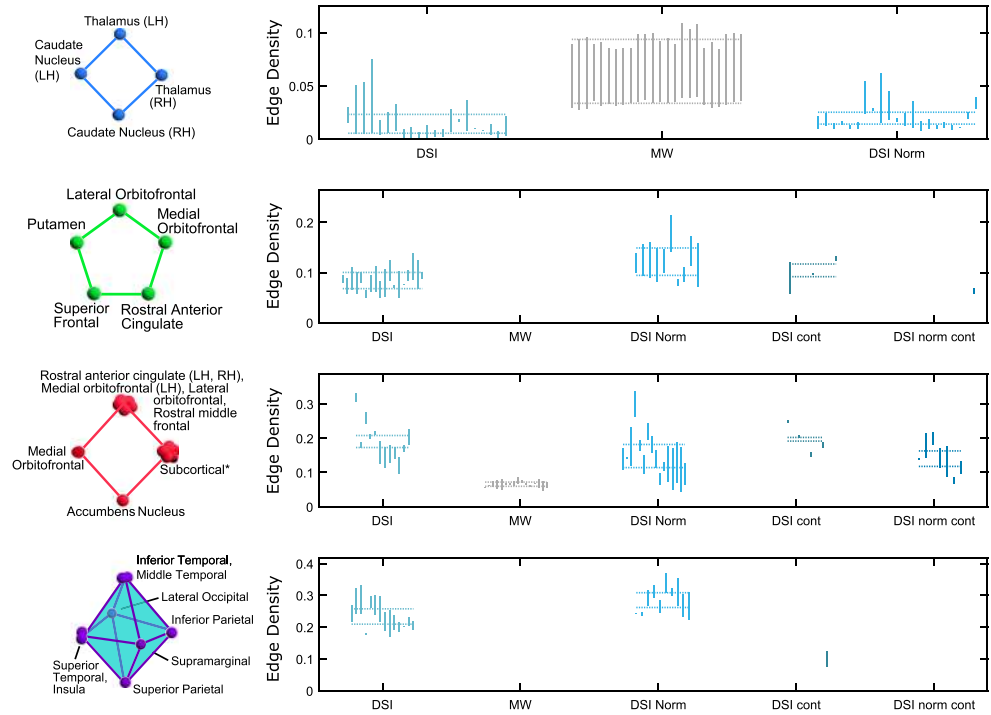
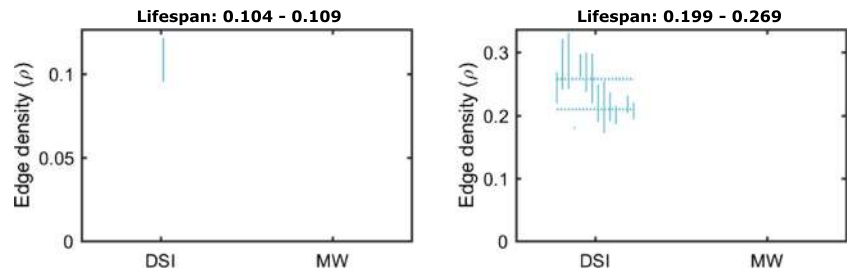


Fig. 24 Lifetimes of all 20 persistent 2D cavities in the individuals (black bars) and minimally wired models (gray bars). Dashed lines indicate the average birth and death densities of each class

Fig. 25 Lifetimes of both persistent 3D cavities in the individuals (black bars) and minimally wired models (gray bars). Dashed lines indicate the average birth and death densities of each class



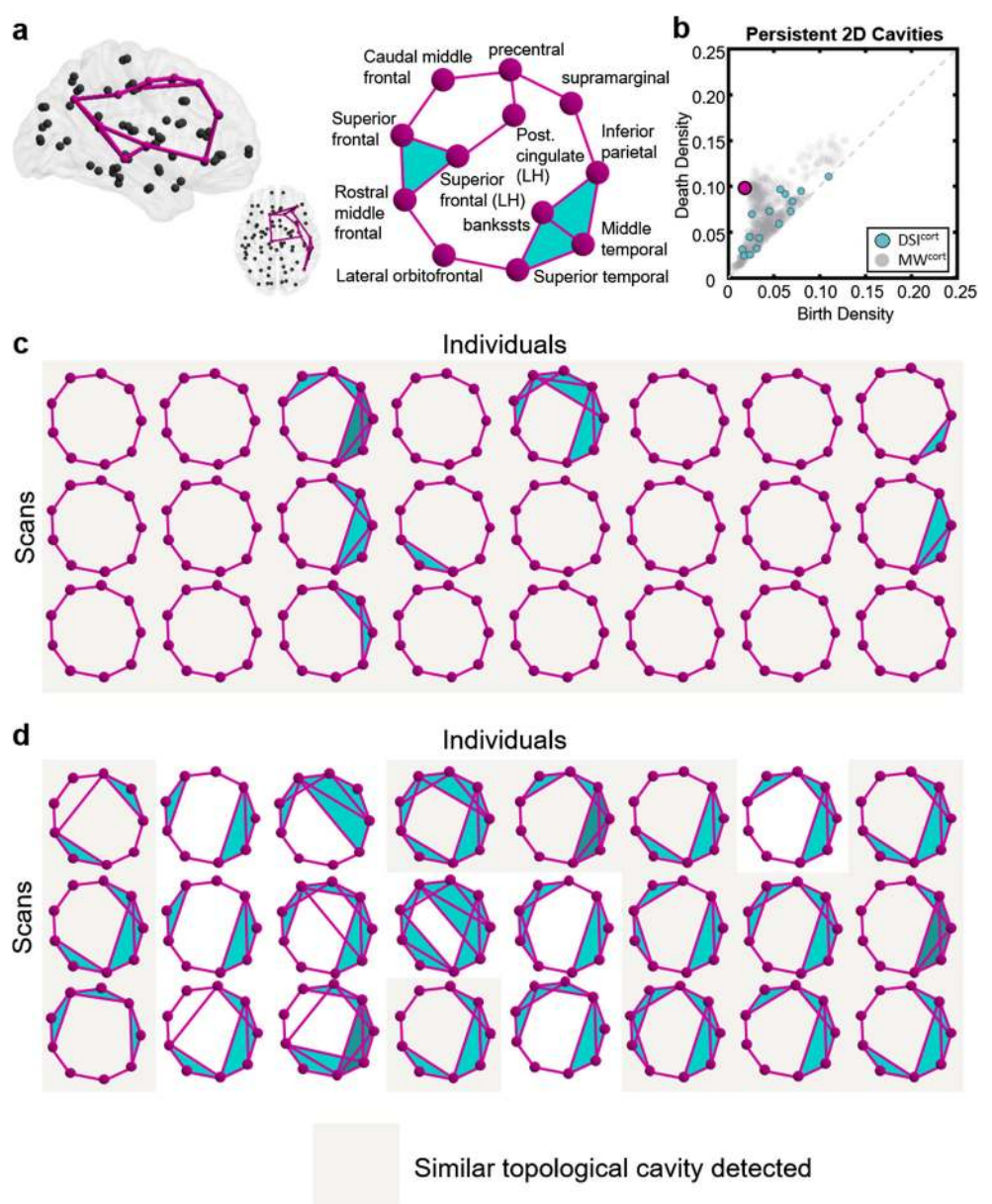
node cycle was already in existence, so either of these possible paths (but not both simultaneously) completes the larger maroon cycle.

We see the pattern of connectivity is not often exactly seen in all individuals, yet the large 2-dimensional cavity

enclosed is present in every scan (Fig. 26c) in the original hemisphere, and often in the opposite hemisphere (Fig. 26d), suggesting its importance in neural structure.

The number and pattern of persistent cycles in Fig. 26b matches that of the minimally wired null model much more

Fig. 26 Recovered 1-cycle on nine nodes. **a** Minimal representatives at ρ_{birth} shown in the brain (left) and as a schematic (right). **b** Persistence diagram of DSI^{cort} and MW^{cort} . Topological cavity in (a) circled in maroon. **c** Patterns of connectivity between maroon loop nodes found for the original (c) and contralateral **d** hemispheres in each scan. If the exact pattern is not found, the pattern at the edge density when all cycle edges first exist is shown. For each scan, the connection pattern of nodes in the minimal generator with the fewest cross-edges is shown



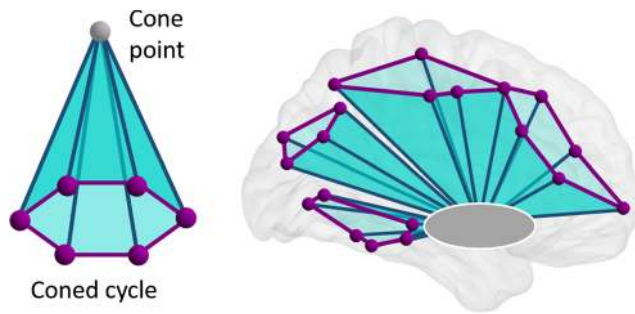


Fig. 27 Subcortical regions as cone points in the brain network. A loop (maroon, left) may be the base of a cone, where the cone point (gray) triangulates the loop interior thus making the loop a boundary loop. In the brain, the greater number and longevity of topological cavities seen after removing subcortical nodes indicates these subcortical regions (gray, right) may act as cone points for many cortical cycles

closely than the full DSI network. This suggests first that the cortical wiring of the brain is globally arranged as if it was wired minimally. Yet the difference in the cortical only and full DSI persistence diagrams also implies the subcortical regions drive the reduction of homology. Knowing the subcortical regions are highly connected and participate in many high-dimensional cliques (Fig. 2), we conclude the subcortical regions are acting as cone points in the brain network (Fig. 27, left). Finally, this adds more detail to our understanding of the global wiring of the brain, as we imagine many cortical loops that are coned by sets of subcortical regions (Fig. 27, right).

References

- Bassett, D.S., & Bullmore, E. (2006). Small-world brain networks. *The Neuroscientist*, *12*(6), 512–523.
- Bassett, D.S., & Bullmore, E.T. (2009). Human brain networks in health and disease. *Current Opinion in Neurology*, *22*(4), 340–347.
- Bassett, D.S., & Bullmore, E.T. (2016). Small-world brain networks revisited. *Neuroscientist* Epub ahead of print 1073858416667720.
- Bassett, D.S., Greenfield, D.L., Meyer-Lindenberg, A., Weinberger, D.R., Moore, S.W., Bullmore, E.T. (2010). Efficient physical embedding of topologically complex information processing networks in brains and computer circuits. *PLoS Computational Biology*, *6*(4), 1000748.
- Bassett, D.S., Brown, J.A., Deshpande, V., Carlson, J.M., Grafton, S.T. (2011). Conserved and variable architecture of human white matter connectivity. *NeuroImage*, *54*(2), 1262–1279.
- Bassett, D.S., Nelson, B.G., Mueller, B.A., Camchong, J., Lim, K.O. (2012). Altered resting state complexity in schizophrenia. *NeuroImage*, *59*(3), 2196–2207.
- Benzekry, S., Tuszynski, J.A., Rietman, E.A., Klement, G.L. (2015). Design principles for cancer therapy guided by changes in complexity of protein-protein interaction networks. *Biology Direct*, *10*(1), 32.
- Bergomi, M.G., Ferri, M., Zuffi, L. (2017). Graph persistence. arXiv:1707.09670.
- Betzell, R.F., & Bassett, D.S. (2016). Multi-scale brain networks. *NeuroImage*, *S1053-8119*(16), 30615–2.
- Betzell, R.F., Gu, S., Medaglia, J.D., Pasqualetti, F., Bassett, D.S. (2016). Optimally controlling the human connectome: the role of network topology. *Science Reports*, *6*, 30770.
- Betzell, R.F., Medaglia, J.D., Papadopoulos, L., Baum, G., Gur, R.E., Gur, R.C., Roalf, D., Satterthwaite, T.D., Bassett, D.S. (2016). The modular organization of human anatomical brain networks: accounting for the cost of wiring. *Network Neuroscience* In Press.
- Bobrowski, O., Kahle, M., Skraba, P. (2015). Maximally persistent cycles in random geometric complexes. arXiv:1509.04347.
- Braver, T.S., Barch, D.M., Gray, J.R., Molfese, D.L., Snyder, A. (2001). Anterior cingulate cortex and response conflict: effects of frequency, inhibition and errors. *Cerebral Cortex*, *11*(9), 825–836.
- Bullmore, E., & Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, *10*(3), 186–198.
- Bullmore, E., & Sporns, O. (2012). The economy of brain network organization. *Nature Reviews Neuroscience*, *13*(5), 336–349.
- Bullmore, E.T., & Bassett, D.S. (2011). Brain graphs: graphical models of the human brain connectome. *Annual Review of Clinical Psychology*, *7*, 113–140.
- Cammoun, L., Gigandet, X., Meskaldji, D., Thiran, J.P., Sporns, O., Do, K.Q., Maeder, P., Meuli, R., Hagmann, P. (2012). Mapping the human connectome at multiple scales with diffusion spectrum MRI. *Journal of Neuroscience Methods*, *203*(2), 386–397.
- Carlsson, G. (2009). Topology and data. *Bulletin of the American Mathematical Society*, *46*(2), 255–308.
- Carlsson, G., & De Silva, V. (2010). Zigzag persistence. *Foundations of Computational Mathematics*, *10*(4), 367–405.
- Carpenter, P.A., Just, M.A., Reichle, E.D. (2000). Working memory and executive function: evidence from neuroimaging. *Current Opinion in Neurobiology*, *10*(2), 195–199.
- Chai, L.R., Mattar, M.G., Blank, I.A., Fedorenko, E., Bassett, D.S. (2016). Functional network dynamics of the language system. *Cereb Cortex* Epub ahead of print.
- Chan, J.M., Carlsson, G., Rabadan, R. (2013). Topology of viral evolution. *Proceedings of the National Academy of Sciences*, *110*(46), 18566–18571.
- Chatterjee, N., & Sinha, S. (2007). Understanding the mind of a worm: hierarchical network structure underlying nervous system function in *C. Elegans*. *Progress in Brain Research*, *168*, 145–153.
- Chen, C., & Freedman, D. (2011). Hardness results for homology localization. *Discrete & Computational Geometry*, *45*(3), 425–448.
- Chen, Z.J., He, Y., Rosa-Neto, P., Germann, J., Evans, A.C. (2008). Revealing modular architecture of human brain structural networks by using cortical thickness from MRI. *Cerebral Cortex*, *18*(10), 2374–2381.
- Chowdhury, S., & Mémoli, F. (2016). Persistent homology of asymmetric networks: an approach based on dower filtrations. arXiv:1608.05432.
- Cieslak, M., & Grafton, S. (2014). Local termination pattern analysis: a tool for comparing white matter morphology. *Brain Imaging and Behavior*, *8*(2), 292–299.
- Cohen-Steiner, D., Edelsbrunner, H., Harer, J. (2007). Stability of persistence diagrams. *DCG*, *37*(1), 103–120.
- Cole, M.W., Yarkoni, T., Repovš, G., Anticevic, A., Braver, T.S. (2012). Global connectivity of prefrontal cortex predicts cognitive control and intelligence. *Journal of Neuroscience*, *32*(26), 8988–8999.
- Crofts, J.J., & Higham, D.J. (2009). A weighted communicability measure applied to complex brain networks. *Journal of the Royal Society Interface*, rsif-2008.
- Dabaghian, Y., Mémoli, F., Frank, L., Carlsson, G. (2012). A topological paradigm for hippocampal spatial map formation using persistent homology. *PLoS Computational Biology*, *8*(8), 1002581.

- Dale, A.M., Fischl, B., Sereno, M.I. (1999). Cortical surface-based analysis: i. Segmentation and surface reconstruction. *NeuroImage*, 9(2), 179–194.
- Desmond, J.E., Sum, J., Wagner, A., Demb, J., Shear, P., Glover, G., Gabrieli, J., Morrell, M. (1995). Functional mri measurement of language lateralization in wada-tested patients. *Brain: A Journal of Neurology*, 118(6), 1411–1419.
- Dey, T.K., Hirani, A.N., Krishnamoorthy, B. (2011). Optimal homologous cycles, total unimodularity, and linear programming. *SIAM Journal on Computing*, 40(4), 1026–1044.
- Dey, T.K., Fan, F., Wang, Y. (2014). Computing topological persistence for simplicial maps. In *Proceedings of the thirtieth annual symposium on computational geometry* (p. 345): ACM.
- Doron, K.W., Bassett, D.S., Gazzaniga, M.S. (2012). Dynamic network structure of interhemispheric coordination. *Proceedings of the National Academy of Sciences of the United States of America*, 109(46), 18661–18668.
- Emmett, K., Schweinhart, B., Rabadan, R. (2016). Multiscale topology of chromatin folding. In *Proceedings of the 9th EAI international conference on bio-inspired information and communications technologies (formerly BIONETICS)* (pp. 177–180): ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering).
- Estrada, E., & Hatano, N. (2008). Communicability in complex networks. *Physical Review E*, 77(3), 036111.
- Fiete, I.R., Senn, W., Wang, C.Z., Hahnloser, R.H.R. (2010). Spike-time-dependent plasticity and heterosynaptic competition organize networks to produce long scale-free sequences of neural activity. *Neuron*, 65, 563–576.
- Galaburda, A.M., LeMay, M., Kemper, T.L., Geschwind, N. (1978). Right-left asymmetries in the brain. *Science*, 199(4331), 852–856.
- Ghrist, R. (2008). Barcodes: the persistent topology of data. *Bulletin of the American Mathematical Society*, 45(1), 61–75.
- Ghrist, R. (2014). Elementary Applied Topology. CreateSpace Independent Publishing Platform. <http://researchbooks.org/1502880857>.
- Giusti, C., Pastalkova, E., Curto, C., Itskov, V. (2015). Clique topology reveals intrinsic geometric structure in neural correlations. *Proceedings of the National Academy of Sciences*, 112(44), 13455–13460.
- Giusti, C., Ghrist, R., Bassett, D.S. (2016). Two's company, three (or more) is a simplex: algebraic-topological tools for understanding higher-order structure in neural data. *Journal of Complex Networks* In Press.
- Goni, J., Avena-Koenigsberger, A., Velez de Mendizabal, N., van den Heuvel, M.P., Betzel, R.F., Sporns, O. (2013). Exploring the morphospace of communication efficiency in complex networks. *PLoS One*, 8(3), 58070.
- Graham, D., & Rockmore, D. (2011). The packet switching brain. *Journal of Cognitive Neuroscience*, 23(2), 267–276.
- Gu, S., Pasqualetti, F., Cieslak, M., Telesford, Q.K., Alfred, B.Y., Kahn, A.E., Medaglia, J.D., Vettel, J.M., Miller, M.B., Grafton, S.T., et al. (2015a). Controllability of structural brain networks. *Nature Communications*, 6.
- Gu, S., Satterthwaite, T.D., Medaglia, J.D., Yang, M., Gur, R.E., Gur, R.C., Bassett, D.S. (2015b). Emergence of system roles in normative neurodevelopment. *Proceedings of the National Academy of Sciences*, 112(44), 13681–13686.
- Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C.J., Wedeen, V.J., Sporns, O. (2008). Mapping the structural core of human cerebral cortex. *PLoS Biology*, 6(7), 159.
- Hartigan, J.A., & Hartigan, P.M. (1985). The dip test of unimodality. *The Annals of Statistics*, 13, 70–84.
- Hatcher, A. (2002). *Algebraic topology*. Cambridge: Cambridge University Press.
- Hausmann, J.-C. et al. (1995). On the Vietoris-Rips complexes and a cohomology theory for metric spaces. *Annals of Mathematics Studies*, 138, 175–188.
- Henselman, G., & Ghrist, R. (2016). Matroid filtrations and computational persistent homology. arXiv preprint arXiv:1606.00199.
- Hermundstad, A.M., Brown, K.S., Bassett, D.S., Carlson, J.M. (2011). Learning, memory, and the role of neural network architecture. *PLoS Computational Biology*, 7, 1002063.
- Hermundstad, A.M., Bassett, D.S., Brown, K.S., Aminoff, E.M., Clewett, D., Freeman, S., Frithsen, A., Johnson, A., Tipper, C.M., Miller, M.B., Grafton, S.T., Carlson, J.M. (2013). Structural foundations of resting-state and task-based functional connectivity in the human brain. *Proceedings of the National Academy of Sciences of the United States of America*, 110(15), 6169–6174.
- Hermundstad, A.M., Brown, K.S., Bassett, D.S., Aminoff, E.M., Frithsen, A., Johnson, A., Tipper, C.M., Miller, M.B., Grafton, S.T., Carlson, J.M. (2014). Structurally-constrained relationships between cognitive states in the human brain. *PLoS Computational Biology*, 10(5), 1003591.
- Hilgetag, C.C., & Goulas, A. (2016). Is the brain really a small-world network? *Brain Structure and Function*, 221(4), 2361–2366.
- Horak, D., Maletić, S., Rajković, M. (2009). Persistent homology of complex networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2009(03), 03034.
- Johnson, D.B. (1975). Finding all the elementary circuits of a directed graph. *SIAM Journal on Computing*, 4(1), 77–84.
- Kim, E., Kang, H., Lee, H., Lee, H.-J., Suh, M.-W., Song, J.-J., Oh, S.-H., Lee, D.S. (2014). Morphological brain network assessed using graph theory and network filtration in deaf adults. *Hearing Research*, 315, 88–98.
- Klimm, F., Bassett, D.S., Carlson, J.M., Mucha, P.J. (2014). Resolving structural variability in network models and the brain. *PLOS Comput. Biol.*, 10(3), 1003491.
- Kringelbach, M.L., & Rolls, E.T. (2004). The functional neuroanatomy of the human orbitofrontal cortex: evidence from neuroimaging and neuropsychology. *Progress in Neurobiology*, 72(5), 341–372.
- Leen, D.A., & Shea-Brown, E. (2015). A simple mechanism for beyond-pairwise correlations in integrate-and-fire neurons. *The Journal of Mathematical Neuroscience (JMN)*, 5(1), 1–13.
- Levy, N., Horn, D., Meilijson, I., Ruppin, E. (2001). Distributed synchrony in a cell assembly of spiking neurons. *Neural Networks*, 14, 815–824.
- Lohse, C., Bassett, D.S., Lim, K.O., Carlson, J.M. (2014). Resolving anatomical and functional structure in human brain organization: identifying mesoscale organization in weighted network representations. *PLoS Computational Biology*, 10(10), 1003712.
- Lord, L.-D., Expert, P., Fernandes, H., Petri, G., Van Hartevelt, T., Vaccarino, F., Deco, G., Turkheimer, F., Kringelbach, M. (2016). Insights into brain architectures from the homological scaffolds of functional connectivity networks. *Frontiers in Systems Neuroscience*, 10, 85. <https://doi.org/10.3389/fnsys.2016.00085>.
- Mamuye, A.L., Rucco, M., Tesei, L., Merelli, E. (2016). Persistent homology analysis of RNA. *Molecular Based Mathematical Biology* 4(1).
- Medaglia, J.D., Lynall, M.-E., Bassett, D.S. (2015). Cognitive network neuroscience. *Journal of Cognitive Neuroscience*.
- Meunier, D., Achard, S., Morcom, A., Bullmore, E. (2009). Age-related changes in modular organization of human brain functional networks. *NeuroImage*, 44(3), 715–723.

- Middleton, F.A., & Strick, P.L. (2000). Basal ganglia and cerebellar loops: motor and cognitive circuits. *Brain Research Brain Research Reviews*, 31(2–3), 236–250.
- Muldoon, S.F., Bridgeford, E.W., Bassett, D.S. (2016a). Small-world propensity and weighted brain networks. *Science Reports*, 6, 22057.
- Muldoon, S.F., Pasqualetti, F., Gu, S., Cieslak, M., Grafton, S.T., Vettel, J.M., Bassett, D.S. (2016b). Stimulation-based control of dynamic brain networks. *PLoS Computational Biology*, 12(9), 1005076.
- Oades, R.D. (1998). Frontal, temporal and lateralized brain function in children with attention-deficit hyperactivity disorder: a psychophysiological and neuropsychological viewpoint on development. *Behavioural Brain Research*, 94(1), 83–95.
- Pessoa, L. (2014). Understanding brain networks and brain organization. *Physics of Life Reviews*, 11(3), 400–435.
- Pestilli, F., Yeatman, J.D., Rokem, A., Kay, K.N., Wandell, B.A. (2014). Evaluation and statistical inference for human connectomes. *Nature Methods*, 11(10), 1058–1063.
- Petri, G., Scolamiero, M., Donato, I., Vaccarino, F. (2013a). Topological strata of weighted complex networks. *PLoS one*, 8(6), 66506.
- Petri, G., Scolamiero, M., Donato, I., Vaccarino, F. (2013b). Networks and cycles: a persistent homology approach to complex networks. In *Proceedings of the european conference on complex systems 2012* (pp. 93–99): Springer.
- Petri, G., Expert, P., Turkheimer, F., Carhart-Harris, R., Nutt, D., Hellyer, P., Vaccarino, F. (2014). Homological scaffolds of brain functional networks. *Journal of The Royal Society Interface*, 11(101), 20140873.
- Porter, M.A., Onnela, J.-P., Mucha, P.J. (2009). Communities in networks. *Notices of the American Mathematical Society*, 56(9), 1082–109711641166.
- Power, J.D., Cohen, A.L., Nelson, S.M., Wig, G.S., Barnes, K.A., Church, J.A., Vogel, A.C., Laumann, T.O., Miezin, F.M., Schlaggar, B.L., et al. (2011). Functional network organization of the human brain. *Neuron*, 72(4), 665–678.
- Rajan, K., Harvey, C.D., Tank, D.W. (2016). Recurrent network models of sequence generation and memory. *Neuron*, 90(1), 128–142.
- Rakic, P. (2009). Evolution of the neocortex: a perspective from developmental biology. *Nature Reviews Neuroscience*, 10(10), 724–735.
- Rasmussen, T., & Milner, B. (1977). The role of early left-brain injury in determining lateralization of cerebral speech functions. *Annals of the New York Academy of Sciences*, 299(1), 355–369.
- Reveley, C., Seth, A.K., Pierpaoli, C., Silva, A.C., Yu, D., Saunders, R.C., Leopold, D.A., Ye, F.Q. (2015). Superficial white matter fiber systems impede detection of long-range cortical connections in diffusion mr tractography. *Proceedings of the National Academy of Sciences of the United States of America*, 112(21), 2820–2828.
- Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. *NeuroImage*, 52(3), 1059–1069.
- Sandi, C., Patterson, T.A., Rose, S. (1993). Visual input and lateralization of brain function in learning in the chick. *Neuroscience*, 52(2), 393–401.
- Senden, M., Deco, G., de Reus, M.A., Goebel, R., van den Heuvel, M.P. (2014). Rich club organization supports a diverse set of functional network configurations. *NeuroImage*, 96, 174–182.
- Sizemore, A., Giusti, C., Bassett, D.S. (2016). Classification of weighted networks through mesoscale homological features. *Journal of Complex Networks* In Press.
- Sporns, O. (2013). The human connectome: origins and challenges. *NeuroImage*, 80, 53–61.
- Sporns, O. (2015). Cerebral cartography and connectomics. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences*, 370, 1668.
- Sporns, O., & Betzel, R.F. (2016). Modular brain networks. *Annual Review of Psychology*, 67, 613–640.
- Sporns, O., Tononi, G., Kötter, R. (2005). The human connectome: a structural description of the human brain. *PLoS Computational Biology*, 1(4), 42.
- Stolz, B. (2014). Computational topology in neuroscience. Master's Thesis, University of Oxford.
- Stolz, B.J., Harrington, H.A., Porter, M.A. (2017). Persistent homology of time-dependent functional networks constructed from coupled time series. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 27(4), 047410.
- Tang, E., Giusti, C., Baum, G., Gu, S., Kahn, A.E., Roalf, D., Moore, T.M., Ruparel, K., Gur, R.C., Gur, R.E., et al. (2016). Structural drivers of diverse neural dynamics and their evolution across development. arXiv:1607.01010.
- Thomas, C., Ye, F.Q., Irfanoglu, M.O., Modi, P., Saleem, K.S., Leopold, D.A., Pierpaoli, C. (2014). Anatomical accuracy of brain connections derived from diffusion MRI tractography is inherently limited. *Proceedings of the National Academy of Sciences of the United States of America*, 111(46), 16574–16579.
- Thulborn, K.R., Carpenter, P.A., Just, M.A. (1999). Plasticity of language-related brain function during recovery from stroke. *Stroke*, 30(4), 749–754.
- Tucker, A. (2006). Chapter 2: covering circuits and graph colorings. *Applied Combinatorics*, 49.
- van den Heuvel, M.P., & Sporns, O. (2011). Rich-club organization of the human connectome. *The Journal of Neuroscience*, 31(44), 15775–15786.
- van den Heuvel, M.P., Kahn, R.S., Goñi, J., Sporns, O. (2012). High-cost, high-capacity backbone for global brain communication. *Proceedings of the National Academy of Sciences*, 109(28), 11372–11377.
- Vietoris, L. (1927). Über den höheren zusammenhang kompakter räume und eine klasse von zusammenhangstreuen abbildungen. *Mathematische Annalen*, 97(1), 454–472.
- Wager, T.D., Phan, K.L., Liberzon, I., Taylor, S.F. (2003). Valence, gender, and lateralization of functional brain anatomy in emotion: a meta-analysis of findings from neuroimaging. *NeuroImage*, 19(3), 513–531.
- Wedeen, V.J., Wang, R.P., Schmahmann, J.D., Benner, T., Tseng, W.Y., Dai, G., Pandya, D.N., Hagmann, P., D'Arceuil, H., de Crespigny, A.J. (2008). Diffusion spectrum magnetic resonance imaging (DSI) tractography of crossing fibers. *NeuroImage*, 41(4), 1267–1277.
- Xia, M., Wang, J., He, Y. (2013). Brainnet viewer: a network visualization tool for human brain connectomics. *PLoS one*, 8(7), 68910.
- Yeh, F.-C., & Tseng, W.-Y.I. (2011). Ntu-90: a high angular resolution brain atlas constructed by q-space diffeomorphic reconstruction. *NeuroImage*, 58(1), 91–99.
- Zalesky, A., Fornito, A., Harding, I.H., Cocchi, L., Yücel, M., Pantelis, C., Bullmore, E.T. (2010). Whole-brain anatomical networks: does the choice of nodes matter? *NeuroImage*, 50(3), 970–983.
- Zalesky, A., Fornito, A., Cocchi, L., Gollo, L.L., van den Heuvel, M.P., Breakspear, M. (2016). Connectome sensitivity or specificity: which is more important? *NeuroImage*, 142, 407–420.
- Zomorodian, A., & Carlsson, G. (2005). Computing persistent homology. *DCG*, 33(2), 249–274.