Received: 25 August 2023 | Revised: 6 March 2024 | Accepted: 12 March 2024

DOI: 10.1002/hbm.26669

RESEARCH ARTICLE

WILEY

Community detection in the human connectome: Method types, differences and their impact on inference

Skylar J. Brooks[1,2] | Victoria O. Jones[3] | Haotian Wang[4] | Chengyuan Deng[4] | Staunton G. H. Golding[1] | Jethro Lim[1] | Jie Gao[4] | Prodromos Daoutidis[3] | Catherine Stamoulis[1,5]

[1]Boston Children's Hospital, Department of Pediatrics, Boston, Massachusetts, USA

[2]University of California Berkeley, Helen Wills Neuroscience Institute, Berkeley, California, USA

[3]University of Minnesota, Department of Chemical Engineering and Material Science, Minneapolis, Minnesota, USA

[4]Rutgers University, Department of Computer Science, Piscataway, New Jersey, USA

[5]Harvard Medical School, Department of Pediatrics, Boston, Massachusetts, USA

Correspondence
Catherine Stamoulis, Boston Children's Hospital, Department of Pediatrics, Boston, MA 02115 USA.
Email: caterina.stamoulis@childrens.harvard.edu

Funding information
National Science Foundation, Grant/Award Numbers: 1649865, 1938914, 1940094, 2116707, 2126582, 2207733

Abstract

Community structure is a fundamental topological characteristic of optimally organized brain networks. Currently, there is no clear standard or systematic approach for selecting the most appropriate community detection method. Furthermore, the impact of method choice on the accuracy and robustness of estimated communities (and network modularity), as well as method-dependent relationships between network communities and cognitive and other individual measures, are not well understood. This study analyzed large datasets of real brain networks (estimated from resting-state fMRI from $n = 5251$ pre/early adolescents in the adolescent brain cognitive development [ABCD] study), and $n = 5338$ synthetic networks with heterogeneous, data-inspired topologies, with the goal to investigate and compare three classes of community detection methods: (i) modularity maximization-based (Newman and Louvain), (ii) probabilistic (Bayesian inference within the framework of stochastic block modeling (SBM)), and (iii) geometric (based on graph Ricci flow). Extensive comparisons between methods and their individual accuracy (relative to the ground truth in synthetic networks), and reliability (when applied to multiple fMRI runs from the same brains) suggest that the underlying brain network topology plays a critical role in the accuracy, reliability and agreement of community detection methods. Consistent method (dis)similarities, and their correlations with topological properties, were estimated across fMRI runs. Based on synthetic graphs, most methods performed similarly and had comparable high accuracy only in some topological regimes, specifically those corresponding to developed connectomes with at least quasi-optimal community organization. In contrast, in densely and/or weakly

connected networks with difficult to detect communities, the methods yielded highly dissimilar results, with Bayesian inference within SBM having significantly higher accuracy compared to all others. Associations between method-specific modularity and demographic, anthropometric, physiological and cognitive parameters showed mostly method invariance but some method dependence as well. Although method sensitivity to different levels of community structure may in part explain method-dependent associations between modularity estimates and parameters of interest, method dependence also highlights potential issues of reliability and reproducibility. These findings suggest that a probabilistic approach, such as Bayesian inference in the framework of SBM, may provide consistently reliable estimates of community structure across network topologies. In addition, to maximize robustness of biological inferences, identified network communities and their cognitive, behavioral and other correlates should be confirmed with multiple reliable detection methods.

KEYWORDS
community detection, data-driven synthetic graphs, fMRI, graph Ricci flow, human brain networks, stochastic block modeling

Practitioner Points

- Community detection in human brain networks significantly depends on both the type of method used and underlying network topology.
- Different types of community detection methods can yield highly dissimilar modularity estimates, and may significantly impact correlations of modularity estimates with physiological, cognitive, and other individual characteristics.
- A probabilistic approach, such as Bayesian inference within the framework of stochastic block modeling, may be more accurate and robust for community detection in brain networks.

1 | INTRODUCTION

The optimal organization of structural and functional circuits in the adult human brain is modular, with a small-world topology that facilitates efficient, domain-specific localized computations and information transmission to highly connected regions (hubs) for synthesis (Bassett & Bullmore, 2006; Bullmore & Sporns, 2009; Cohen & D'Esposito, 2016; Fransson et al., 2018; van den Heuvel & Sporns, 2013). These characteristics, which are shared across optimally organized complex systems, may be necessary to maximize the brain's flexibility, adaptability to dynamically changing environments, learning, but also resilience (Simon, 1962). Across spatial scales, locally connected neuronal ensembles and brain regions form functionally specialized communities that support local information processing and interact with each other through robust but sparse long-range connections (Bertolero et al., 2015; Crossley et al., 2013). This topological organization is optimized over a period of two decades of human development (Fair et al., 2009; Venon, 2013).

Whether structural or functional, brain network communities have partly distinct neurodevelopmental trajectories. Highly connected structural modules (hubs) may be present even at birth, primarily in association brain areas (parietal and superior frontal) and subcortical regions (Ball et al., 2014). These hubs remain topologically consistent throughout life. In contrast, functional modules and hubs undergo significant changes as a function of neural maturation. They begin to emerge in the first year of life (Wen et al., 2019), and are detectable in primary sensory and motor areas, but by adulthood are present in frontal, visual, temporal, and subcortical areas (Ball et al., 2014; Huang et al., 2015). Furthermore, overall functional network modularity increases from childhood to young adulthood, (Chen & Deem, 2015), likely peaking in young adulthood and decreasing thereafter (Cao et al., 2014; Chong et al., 2019; Geerligs et al., 2015; Iordan et al., 2018; Onoda & Yamaguchi, 2013; Song et al., 2014). Age-related differences in modularity have been reported across brain regions, including frontoparietal control, attention, and visual networks, as a function of differences in the rate of neural maturation and protracted development of some of these regions (Betzel et al., 2014).

In addition to its role in the brain's overall flexibility, adaptability, and robustness, modularity has been linked to cognitive processing across domains (Lorenz et al., 2011; Pradhan et al., 2011). Findings based on diffusion MRI suggest that the development of structural brain modules, particularly within the frontoparietal control network, may be associated with age-related improvements in executive

function in youth (Baum et al., 2017). Prior work based on functional MRI (fMRI) has shown associations between modularity and working memory, relational reasoning, language processing, and social cognition (Bertolero et al., 2018; Stevens et al., 2012). Furthermore, inter- and intra-individual variability of functional network modularity has been mapped onto variability in cognitive performance (Stevens et al., 2012) and response to interventions for cognitive enhancement (Arnemann et al., 2015; Baniqued et al., 2018; Chaddock-Heyman et al., 2020; Gallen & D'Esposito, 2019).

Prior studies have also reported disease/disorder-related changes in the brain's modular organization. For example, epilepsy patients may have higher structural and functional network modularity compared to healthy individuals (Pedersen et al., 2015; Takeda et al., 2017; Vaessen et al., 2013). In contrast, lower modularity has been reported in patients with schizophrenia and depression (Alexander-Bloch et al., 2010, 2012; Peng et al., 2014). Children with attention-deficit hyperactivity disorder (ADHD) or autism spectrum disorder may also have reduced modularity (Belmonte et al., 2004; Qian et al., 2019; Rudie et al., 2013; Wass, 2011). Finally, lower modularity has also been reported in patients with neurodegenerative diseases, such as Alzheimer's and Parkinson's disease (Brier et al., 2014; Göttlich et al., 2013; Ng et al., 2021).

A wide range of community detection methods are used by the Neuroscience community to estimate the modular organization of brain networks. However, there is currently no true standard or systematic approach for method selection. In other fields, method dependence of modularity estimates has been documented, and has been shown to limit the interpretability, reproducibility, and/or relevance of findings, as well as scientific inference (Laender et al., 2020; Leskovec et al., 2010). In the context of the brain, to date, there are limited systematic assessments and comparisons of these methods as a function of network topology, particularly in the developing brain, which is both highly heterogeneous and undergoes profound age-related topological changes. There are a few detailed studies focusing on a single type, for example, based on modularity maximization (Garcia et al., 2018), but very limited cross-type investigations. It is also unclear how method choice impacts the accuracy of modularity estimates in the developing brain and their associations with cognitive, physiological, and other individual measures. Thus, there is a significant unmet need to systematically evaluate different types of methods and identify robust approaches for estimating modularity in brain networks. It is also important to identify advantages and shortcomings of widely used and/or promising community detection methods, particularly in settings where inter-network variability is high.

To address this significant need, and contribute to ongoing efforts to increase reproducibility and robustness of brain graph theoretic analyses in Neuroscience, this study aimed to elucidate differences between types of community detection methods applied to incompletely developed brain networks of early adolescents, and the impact of specific choices on correlational analyses using modularity as the topological measure of interest. The study focused on classes of methods that use distinct approaches for community detection, based on edge-removal, modularity optimization, statistical inference, and geometric analysis of the graph topology. It also focused specifically on maturating, inherently noisy, and heterogeneous developing brains, where community detection methods with different sensitivity to network noise and weak/redundant connections can lead to disparate results and method-dependent inferences. Investigated methods were applied to both a large dataset of heterogeneous graphs estimated from resting-state (rs) fMRI data from a sample of n>5000 early adolescents, and a large dataset of data-inspired synthetic graphs with controlled and systematically varied parameters. The latter were used to investigate the impact of topological changes associated with extensive reorganization of brain networks during development and neural maturation on method performance and differences between modularity estimates. Finally, in real-data analyses, modularity estimates and their similarity were correlated with cognitive measures and other individual youth characteristics. These mappings aimed to further highlight the impact of method choice on biological associations between brain and behavior/cognition (and related inferences), and the need to confirm research findings with multiple methods in order to maximize their biological relevance.

2 | MATERIALS AND METHODS

Coordinated functional activity between brain areas can be described mathematically by a graph $\mathcal{G}(E,V)$, in which discretized brain regions represent nodes V and functional interactions between their edges E. Brain graphs are typically estimated from regional measurements of electrical/electromagnetic activity (EEG and MEG) or blood oxygenation (fMRI) using various correlation techniques, from simple pairwise correlation in the time or frequency domains (e.g., peak cross-correlation and coherence) to probabilistic (information theoretic) methods (Bastos & Schoffelen, 2016; Li et al., 2009; Rossini et al., 2019), and directional techniques for effective connectivity (Harush & Baruch, 2017; Lopez-Madrona et al., 2019; Stephan & Friston, 2010). Connectivity matrices are further processed (e.g., via thresholding or model-related approaches (Bielczyk et al., 2018)) to eliminate edges that represent weak and/or artifact-related regional interactions and obtain binary or weighted adjacency matrices, based on which communities and other topological graph characteristics can be estimated.

2.1 | Community detection algorithms

Three types of community detection methods were investigated and compared: (1) modularity maximization-based, including Newman and Louvain methods, (2) a model-based probabilistic approach based on stochastic block modeling (SBM) (Holland et al., 1983) and Bayesian inference, and (3) a geometric approach (Ni et al., 2019), based on Ricci flow (Hamilton, 1982; Ollivier, 2007). Their respective mechanisms, computational cost, and examples of previous applications to real and synthetic brain networks are summarized in Table 1. The first

TABLE 1 Community detection methods analyzed in the study, and corresponding approach, computational cost, and previous application to brain networks.

Method	Time complexity	Mechanism	Examples of prior brain applications
Girvan-Newman	$O(m^2n)$ or $O(n^3)$	Edge removal	Meunier et al. (2009) Bordier et al. (2017)
Louvain	$O(n\log n)$	Modularity maximization	Bassett et al. (2013) Rudie et al. (2013) Ji et al. (2019)
SBM	$O\left(n\log^2 n\right)$	Bayesian inference	Betzel et al. (2018) Faskowitz and Sporns (2020)
Ricci flow	$O(mn\log n)$	Geometric approach	Weber et al. (2017)
Infomap	$O(m)$	Minimization of description length	Sporns and Betzel (2016) Seitzman et al. (2020)

Note: n corresponds to number of nodes, and m the number of edges.

two are state-of-the-art methods that have been used extensively in Neuroscience for community detection in structural and functional brain networks, whereas SBM and Ricci flow are much less frequently used in the field. Each method's advantages and shortcomings are discussed in terms of accuracy, computational cost, and interpretability. Their methodological differences are also highlighted below. Finally, in a secondary set of comparisons, the *Infomap* method (Rossval & Bergstrom, 2008), which is gaining popularity in brain network studies (Sporns & Betzel, 2016), was also applied to the analyzed data. Although it uses a different optimization metric for community detection, this method is not fundamentally different than the Louvain method. Both are greedy algorithms but Louvain uses modularity maximization whereas Infomap minimizes the description length of a random walk in the network for partitioning.

2.1.1 | Girvan–Newman method

The Newman method is one of the most widely used community detection algorithms (Newman, 2004), and is based on progressive removal of network edges based on betweenness centrality. The latter quantifies a node's importance in a network based on how many shortest paths between pairs of nodes pass through it. Mathematically, betweenness centrality C_b is defined as:

$$C_b(v) = \sum_{s,t \in V} \frac{\sigma(s,t|v)}{\sigma(s,t)}, \tag{1}$$

where $\sigma(s,t)$ is the total number of shortest paths between nodes s and t and $\sigma(s,t|v)$ is the number of those paths that cross node v (Brandes, 2008). As edges are progressively removed, betweenness centrality is recalculated for nodes affected by the previous removal, and these steps are repeated until no edges are left in the graph. During this process, groups of nodes become disconnected and community structure is unveiled. This structure is inherently hierarchical.

In this study, the following definition of modularity Q was used (Arenas et al., 2008; Newman, 2016):

$$Q(\gamma) = \frac{1}{2m}\sum_{ij}\left(A_{ij} - \gamma\frac{k_i k_j}{2m}\right)\delta_{g_i g_j}, \tag{2}$$

where m is the total number of edges within the network, A_{ij} the adjacency matrix, $k_i = \sum_j A_{ij}$ the degree of a node i, and similarly for k_j, δ_{ij} the Kronecker delta, g_i, g_j community assignments of nodes i, j, and γ a resolution parameter. When $\gamma = 1$, Equation (2) corresponds to the traditional definition of modularity. For $\gamma > 1$, the network partitioning favors the detection of a larger number of smaller communities, whereas $\gamma < 1$ leads to fewer but larger communities. In this study, the modularity estimation was performed using a range of gamma values (0.8–1.3, with increment of 0.02). The results are based on the γ value at the first inflection point (Khambati et al., 2015). Though simple and widely used, the Newman method is computationally expensive, particularly for large networks. The optimization of γ also increases computational cost.

2.1.2 | Louvain method

The Louvain method aims to optimize modularity as an objective function. It begins by assigning each node to its own community and then aggregates communities together based on change in modularity (Blondel et al., 2008). Similarly to several other methods, it has the advantage of being aparametric, thus not requiring a priori assumptions on the number of communities. It uncovers primarily assortative community structure via modularity maximization, comparing the number of edges connecting nodes within a community to the number of edges connecting nodes between two different communities. Recently, a major shortcoming of this approach has been discovered, in that it may yield arbitrarily disconnected or badly connected communities (i.e., disjoint, island-like groups of nodes). This flaw has been

addressed via a new formulation of Louvain called the Leiden method (Traag et al., 2019), which guarantees that communities remain connected. The algorithm attempts to maximize difference between the "true" number of intracommunity edges and the expected number of these edges. Similarly, to the Newman method, the resolution parameter γ was varied for the Louvain algorithm, too. Because the algorithm is non-deterministic, we performed 10 repetitions at each gamma value, using the median modularity value to select community affiliations, a process that also increased computational cost.

2.1.3 | Bayesian inference and stochastic block modeling

SBM is often used to generate synthetic graphs. The number of nodes and blocks (communities), and a probability matrix are provided as inputs, and graphs are generated with a topological structure dictated by the probability matrix (Faust & Wasserman, 1992; Holland et al., 1983; Lee & Wilkinson, 2019). This matrix effectively governs the group membership of the nodes, which is specified by a partition b. Because of the probabilistic generation of b, SBM does not limit the resulting graph's community structure to a single type. Thus, it may be assortative, as is detected by Girvan–Newman and Louvain algorithms, core-periphery, disassortative, or a combination of these (Guimera & Sales-Pardo, 2009). For community detection in real networks, SBM is also used to identify optimal partitions (communities). Assuming that the real graph has been generated by a process similar to that assumed by the SBM, Bayesian inference is used to find an optimal graph partitioning and assign each node to a community. Thus, Bayesian inference is applied within the SBM framework, and the two together become a community detection tool. This probabilistic framework ensures that detected communities are rooted in the statistics of the real network's structure.

Degree correction and nested SBM

In the most basic SBM, edge placement within a community is random. Thus, communities tend to contain nodes of similar degrees, which are unlikely to occur in real-world networks. This bias is accounted for in the degree-corrected SBM, wherein additional parameters are included to permit degree heterogeneity within groups (Karrer & Newman, 2011; Yan et al., 2014). In this study, the degree-corrected variant of SBM was used. Another variant used in this study is the degree-corrected nested SBM (nSBM), a hierarchical extension of SBM that utilizes a hierarchy of priors and hyperpriors and improves the SBM's resolution (Amini et al., 2019). First, a set of communities is detected for the network via SBM. Then, the detected set is treated as its own graph, so that each community is a node. This recursive process continues until a single community is obtained. Thus, the information gleaned from coarser levels is considered prior information in more granular levels (Peixoto, 2014b). nSBM allows for the detection of smaller communities by circumventing the maximum group scaling associated with the regular SBM and can detect multilevel hierarchy within the network.

Bayesian inference

This approach is used to identify node partitioning into communities (Peixoto, 2016), based on maximization of the posterior distribution of possible partitions:

$$P(b|A) = \frac{P(A|\theta,b)P(\theta,b)}{P(A)}, \tag{3}$$

where A is an adjacency matrix, b a partition, and θ a model parameter controlling the partition. An equivalent information-theoretic perspective is the minimization of description length, Σ:

$$\Sigma = -\ln P(A|\theta,b) - \ln P(\theta,b). \tag{4}$$

Thus, the optimal partition is one that describes a given network with as little information as possible. In this study, the minimum description length (MDL) was used both for model selection and for choosing representative partitions for each brain. A Markov chain Monte Carlo (MCMC) agglomerative algorithm (Peixoto, 2014a) was implemented for this analysis. The MCMC sampling is asymptotically exact and is thus more accurate than other approaches, such as variational inference, which use appropriate distributions (Lange et al., 2022). The algorithm's time complexity is $O\left(n\log^2 n\right)$, making this approach computationally feasible for large-scale (in number of graphs and number of nodes per graph) network analyses. Because this method is stochastic, an optimal partition cannot be guaranteed. Thus, as best practice, the algorithm was called 10 times for each graph, and the partition yielding the lowest MDL (i.e., the best fit) was selected.

2.1.4 | Ricci flow

The Ricci flow approach is based on the geometric notion of curvature, which quantifies how spaces are bent at each point. Ricci flow (Hamilton, 1982) deforms the metric of a Riemannian manifold in a way formally analogous to the diffusion of heat, smoothing out irregularities in the metric. Under the Ricci flow, regions in a space of large positive curvature shrink to points, whereas regions of very negative curvature spread out. Discretized curvature (Ollivier, 2007) and Ricci flow have been developed to study graphs, for example, to identify bottleneck edges and discover community structures in social networks (Ni et al., 2015, 2019). The discrete Ollivier–Ricci curvature on a network edge $(x,y) \in E$ is defined as:

$$\kappa_{xy} = 1 - \frac{W(m_x,m_y)}{d(x,y)}, \tag{5}$$

where $W(m_x,m_y)$ is the minimum total weighted travel distance (optimal transport distance, or Wasserstein distance) to move a distribution m_x on the neighbors of vertex x to a distribution m_y on the neighbors of vertex y, and $d(x,y)$ is the distance between x and y in the graph. The curvature of an edge (x,y) is positive if the two vertices have a well-connected/overlapping neighborhood, and negative if (x,y)'s neighborhood are largely disjoint.

Discrete Ricci flow is a process that deforms the metric (i.e., edge length) by its Ricci curvature until edge curvature evolves to be uniform everywhere. For any pair of adjacent nodes x and y on a graph $G=(V,E)$, weight $w(x,y)$ of edge (x,y) is adjusted by the curvature $\kappa(x,y)$:

$$w_{i+1}(x,y) = w_i(x,y) - \epsilon \cdot \kappa_i(x,y) \cdot w_i(x,y), \forall (x,y) \in E, \quad (6)$$

where $\kappa_i(x,y)$ is computed using the current edge weight $w_i(x,y)$. The step size is controlled by $\epsilon > 0$. After each iteration edge weights are rescaled so the total edge weight in the graph remains the same. The Ricci flow process expands negatively curved edges and shrinks positively curved edges. Eventually, nodes connected by intra-community edges are condensed, and inter-community edges are stretched. Network "surgery" is performed to remove edges with large weights, in order to separate the network into different communities. The threshold used to remove edges can be chosen by either the evaluating threshold insensitivity or maximizing graph modularity (Ni et al., 2019). For networks with hierarchical community structures, multiple rounds of network surgery and Ricci flow can be performed to further separate communities at different scales. Running time of the discrete Ricci flow is dominated by two factors: (a) the search of shortest paths distances between nodes that are three hops away in the network, which is at most $O(mn\log n)$ for a graph of n vertices and m edges; (b) the number iterations and surgeries. Typically the number of iterations and surgery operations are both small constants (Ni et al., 2019). Given the geometric nature of the method, community detection by Ricci flow favors dense graphs. It also assumes that edges across communities are less dense than edges within communities. In graphs when this assumption does not hold, the method does not detect multiple communities, that is, considers the entire graph as a single community (Sia et al., 2019, 2022). This assumption is not necessary in other methods, for example, SBM.

It is important to note that like most, if not all, community detection algorithms, there are both detectability and resolution limits that impact the investigated methods. Regardless of the algorithm, the detectability limit depends of the graph characteristics, including number of intercommunity connections, density, and degree distribution (Richardt & Leone, 2008). Furthermore, the resolution limit is $\sqrt{m}$ for Newman and Louvain, $\sqrt{n}$ for SBM, and $\frac{n}{\log n}$ for nSBM. Based on the real and synthetic data analyzed in this study, these translate to hundreds of communities and thus resolution limits that are much higher than the typical number of communities in (even high-resolution) brain networks. In the Ricci flow method, the number of communities is controlled by the threshold used in the surgery step for edge removal. When graphs are sufficiently sparse, Ricci flow may remove edges and reach multiple singleton communities. Finally, in theory, all investigated methods are scalable to higher-resolution networks. However, in practice, scalability will depend on each method's time complexity.

2.2 | Datasets

2.2.1 | Networks estimated from real fMRI data

The study analyzed minimally preprocessed rs fMRI data from the adolescent brain cognitive development (ABCD) study, a longitudinal investigation of almost 12,000 children (Casey et al., 2018), measured at 21 sites across the United States. A cohort of 5251 participants in pre/early adolescence (from the ABCD baseline dataset, median age = 120.0 months, inter-quartile range (IQR) = 13.0 months) was selected for analysis, following exclusions based on poor-quality imaging data, homogeneously high connectivity across the brain (likely associated with motion and other artifacts), clinical findings in the structural MRI, or history of bipolar disorder or attention deficit hyperactivity disorder (ADHD). Both disorders have been associated with aberrant functional connectivity that could impact the estimation of community organization (Chase & Phillips, 2016; Konrad & Eickhoff, 2010). All imaging data were from Release 2.0.1. Additional details on inclusion criteria for connectivity analysis, and imaging systems and protocols used by the ABCD are provided in (Brooks et al., 2021) and (Hagler et al., 2019), respectively. All neuroimaging data from this release were acquired with 3T Siemens or GE scanners. T1w acquisition (1 mm isotropic) included scanner-based motion correction. Repetition time (TR) for fMRI (2.4 mm isotropic) was 0.8 s, and thus sampling rate was 1.25 samples/s. Preprocessing included correction for B0 distortion. Further custom processing is outlined below.

The Next-Generation Neural Data Analysis (NGNDA) platform (https://github.com/cstamoulis1/Next-Generation-Neural-Data-Analysis-NGNDA) was used to further process the fMRI data. Processing included structural MRI (T1w) segmentation, coregistration of each participant's fMRI to their structural MRI, normalization to MNI space, initial frame removal, and slice-time correction. These steps were performed using a combination of the SPM12 software package (Friston et al., 1994, 2007) and the NGNDA. They were followed by motion and artifact suppression and denoising of fMRI voxel time series. Also, nuisance signals were regressed out. Breathing and other artifacts in the frequency range 0.28–0.46 Hz were removed from these signals using a third-order elliptical bandstop filter. Signals were filtered in both directions to eliminate potential distortions associated with the nonlinear phase of the IIR filter. Frames with displacement >0.3 mm were censored. The statistical modeling approach of Power et al. was used to regress motion-related contributions out of voxel time series (Power et al., 2014). To reduce the very high voxel resolution for further analysis, images were parcellated (Siegel et al., 2014), using the high-resolution Schaefer-1000 cortical atlas (1000 parcels), Melbourne subcortical atlas (54 parcels), and a probabilistic MR cerebellar atlas (34 parcels). This resulted in a reduction of the spatial dimension from ~300,000 voxels to 1088 voxel-averaged parcels (Diedrichsen et al., 2009; Schaefer et al., 2018; Tian et al., 2020). A third-order elliptical bandpass filter with cutoffs 0.01 and 0.25 Hz

(typical physiological frequency range of the BOLD signals), respectively, was then applied to each parcel time series (Yuen et al., 2019), to suppress high-frequency artifacts. Time series were further denoised through time-domain signal decomposition (using a variation of the ensemble empirical mode decomposition) (Torres et al., 2011; Wu & Huang, 2009) and component exclusion based on their amplitudes and characteristic frequencies. Scanner effects on signal amplitude were finally assessed, and all signals were harmonized through amplitude normalization (by the median of the absolute signals).

Each participant included in this study had up to 4 five-minute rs fMRI runs. The cohort of $n = 5251$ included participants with at least one run in which ≤10% of frames had been censored for motion (with a displacement cutoff of 0.3 mm). A subcohort had more than one run that met this criterion. To test reliability and reproducibility of method performance and similarity across datasets, a subcohort of $n = 3820$ participants with a second run that met the frame censoring cutoff was also analyzed. Best and second best runs are referred to as first and second run hereafter. Note that each run represents a snapshot of rs activity during a 5-min period. The two selected (quality-based) runs were not necessarily consecutive. In addition, the ABCD study protocol is such that, to avoid fatigue, falling asleep, and compliance issues, a movie clip (~20 s long) is shown between runs.

Connectivity was calculated as the peak cross-correlation between signal pairs, resulting in a 1088×1088 connectivity matrix. Although a number of measures can be used to obtain edge weights, this estimator explicitly accounts for time-dependent similarity between signals, and the peak corresponds to maximal similarity. Statistical and percolation thresholds were estimated, and a cohort-level threshold (estimated via bootstrapping) corresponding to the moderate outlier of peak cross-correlation (defined as median + 1.5× IQR) was selected. It was then applied to all connectivity matrices to obtain corresponding weighted and binary adjacency matrices. Correlation values below this threshold were set to 0. For each participant, in addition to imposing a threshold for frame censoring, the run with the lowest median nonzero connectivity was selected at the best-quality run, under the assumption that the brain at rest is weakly coordinated (with the exception of select networks such as the default mode network (Greicius et al., 2003)). In most brains, this run also coincided with that with the lowest number of frames censored for motion (median = 1.6% of censored frames, interquartile range (IQR) = 4.53%). Additional details on threshold estimation and selection of fMRI runs for analysis are provided in (Brooks et al., 2021). The same criteria were used to select the second best-quality run (median = 1.6% of censored frames, IQR = 4.0%). Statistics of network characteristics were: median (IQR) degree = 40 (52) and 56 (78) for the two runs, clustering was 0.33(0.06) and 0.35 (0.06) for the two runs, and median weight was 0.65(0.08) for both runs. Community structure of each brain, from each run and each method was estimated using the parallel processing resources of the high-performance computing (HPC) clusters in the collaborating institutions.

2.2.2 | Synthetic networks

Network generation algorithm

Given that the true community structure (the ground truth) of networks estimated from real brain data is unknown, synthetic networks with controllable properties, including community structure, were generated. The Lancichinetti–Fortunato–Radicchi (LFR) algorithm (Lancichinetti et al., 2008) was used for this purpose. A prior study on functional brain network modularity also used this model (Bordier et al., 2017). The algorithm allows the variation of multiple graph parameters. Here, these were varied based on the topological properties of the real rs networks. The Python library Networkx (https://networkx.org/documentation/stable/reference/generated/networkx.generators.community.LFR_benchmark_graph.html) was used to generate the synthetic graph dataset. This algorithm was chosen because its parameters are biologically interpretable and can be mapped to topological characteristics of brain networks that vary as a function of age, for example, during development.

The LFR takes several inputs that define the characteristics of the graph. The number of nodes is controlled by a parameter n. The node degree and community size are assumed to follow a power law distribution. Power law exponents τ_1 and τ_2 influence degree and community size, respectively. The parameter $\mu \in [0, 1]$, determines the ratio of inter-community edges for each node. A value of $\mu = 1$ results in a graph where all edges are between nodes belonging to different communities, and $\mu = 0$ in a graph where only nodes within the same community are connected. The average node degree can also be specified. To generate a graph, each node is assigned a degree by drawing from the power law distribution with exponent τ_1. Each community's size is assigned based on the power law distribution with exponent τ_2. Median node degree must be equal to the average degree that was given as input, and the sum of the community sizes must equal the total number of nodes n. Each node is added to a community randomly. In communities that become too large, nodes are randomly removed and reassigned. After community assignments are complete, each node u is given $(1-\mu) \times \text{degree}(u)$ edges within its community and $\mu \times \text{degree}(u)$ edges outside of its community.

3 | DATA-DRIVEN NETWORKS

Synthetic networks were generated using model parameters that were statistically derived from the real data. Based on modularity estimated using the Newman method, the number of communities and median community size were calculated for each participant. Bootstrapping with replacement was used to obtain cohort-wide community statistics (25th and 75th percentiles) for number of communities, median community size, median node degree, and median ratio of inter-community connections. Based on these estimates, the number of communities in synthetic graphs was varied in the range [2, 22]. As shown in Table 2, when applied to real data, both Newman and Louvain methods yielded a relatively small number of functional communities, in the range of those reported in prior work (Tooley et al., 2022).

TABLE 2 Summary statistics (minimum, maximum, median, and inter-quartile range (IQR)) for number of communities detected with each method, applied to brain networks estimated from two fMRI runs.

Number of communities—first run				Modularity
Method	**Minimum**	**Maximum**	**Median (IQR)**	**Median (IQR)**
Newman	2	22	5 (2)	0.522 (0.166)
Louvain	2	17	6 (3)	0.534 (0.167)
SBM	9	87	48 (17)	0.197 (0.164)
nSBM	10	124	89 (13)	0.126 (0.097)
Ricci flow	2	103	42 (38)	0.298 (0.198)
Infomap	2	47	18 (13)	0.467 (0.213)
Number of communities—second run				**Modularity**
Method	**Minimum**	**Maximum**	**Median (IQR)**	**Median (IQR)**
Newman	2	20	4 (3)	0.471 (0.188)
Louvain	2	16	6 (3)	0.483 (0.191)
SBM	13	87	52 (19)	0.149 (0.146)
nSBM	21	132	89 (11)	0.097 (0.086)
Ricci flow	2	96	31 (35)	0.249 (0.267)
Infomap	1	50	16 (11)	0.422 (0.292)

Note: Corresponding median and IQR for modularity and also included.

To represent the variability of the real connectomes in the synthetic graphs, as well as vary their topological properties so that they reflect developmental changes in functional circuits (Venon, 2013), a range of average degree, μ, τ_1, and τ_2 inputs were used based on the estimated real network statistics. First, average degree and μ were varied within the range of real estimates of these parameters' 25th and 75th percentiles. Since appropriate ranges for τ_1 and τ_2 were not a priori known, initially a wide range of values were used. However, only ranges of τ_1 and τ_2 that consistently produced biologically plausible numbers of communities were used to select a final set of synthetic graphs. Once all input ranges were set, each parameter was individually varied (holding others constant) so that every combination of parameters would generate a distinct graph. Some combinations failed to produce a graph, likely because one or more distribution assumptions had been violated. Furthermore, some graphs were discarded because they contained an unrealistic large number of communities. Following these exclusions, a total of 2669 valid binary graphs were used in further analyses.

To create weighted graphs, a range of correlation values (the edge weights) based on statistical thresholds estimated from the real rs networks were used. Brain circuits in pre/early adolescence are differentially maturated. Some are fairly well-developed (e.g., those supporting sensory processing), but others are underdeveloped (e.g., frontoparietal control and DMN). In this study, the large-scale networks identified by Yeo et al. (2011), were considered. For simplicity, they were classified as developed, partially/moderately developed, and underdeveloped, so that ranges of correlation values could be established for these categories. For example, visual networks were assumed to represent fairly well-developed circuitry, the somatomotor network moderately developed circuitry, and the frontoparietal control and limbic networks under-developed circuitry. For each participant, median edge weights were calculated in each of these three categories of networks. Then, in each synthetic graph, communities were classified as fully developed, moderately developed, or underdeveloped, and within-community edge weights were randomly sampled from the corresponding median connectivity distributions of three categories of real networks. Weights between communities were also assigned based on a classification of low/high connectivity estimated from the real data. This approach simplified method comparisons as a function of graph parameters, given an already complex space of multiple parameter variations. It also maximized the representation of differentially developed communities in the synthetic graphs, similar to real brain networks. For each graph with a specific intra-community connectivity, two graphs were generated, for high and low inter-community connectivity. Median values for the two were 0.639 and 0.730, respectively. Based on this approach, a total of 5338 weighted synthetic graphs (twice the original dataset size) were generated and further analyzed. The synthetic data were generated using the parallel processing resources of the HPC cluster at Harvard Medical School.

3.1 | Statistical analysis

Modularity was compared across methods using real and synthetic networks and two commonly used measures of similarity, the adjusted Rand index (ARI)—adjusted for similarity by chance, and normalized mutual information (NMI), which was unadjusted. Both statistics were estimated at the graph level. In real data, only inter-method comparisons were possible, as the ground truth was unknown. In synthetic networks, absolute comparisons were possible, given each network's known community structure (the ground truth). Inter-method differences in community structure and modularity, and differences in associations between method-specific modularity and individual demographic, physiological, and cognitive measures were also investigated. Simple (unadjusted) statistical comparisons of ARI and NMI in between groups of participants used the nonparametric Wilcoxon rank sum test for unpaired samples and ordinary linear regression models with ARI and NMI as the dependent variables, and participant characteristics as independent variables. Additional multivariate regression models were developed with appropriate adjustments for age, race, ethnicity, family income, and experimental site, to assess correlations between method-specific modularity and cognitive task measures. Correlograms were generated as visual representations of correlations between methods and topological network properties. Box plots were generated to compare ARI and NMI in individual topological properties' quartiles. Given the non-normal distribution of most variables, median and IQR were used as the relevant summary statistics. Across models and comparisons, all p-values were adjusted for the false discovery rate (Benjamini & Hochberg, 1995).

4 | RESULTS

Summary statistics for the number of communities estimated by each method for connectomes derived from the first and second fMRI runs are provided in Table 2. Overall, both Newman and Louvain estimated on average a relatively small number of communities (<10), whereas SBM, nSBM, and Ricci, estimated over 10 times more communities. Finally, infomap estimated a higher number of communities than those obtained with the Louvain method but a lower number than SBM.

4.1 | Comparisons between methods applied to real data

Pairwise ARI and NMI statistics for modularity estimated from both sets of fMRI runs with each of the investigated methods are summarized in Table 3. At least modest similarity (based on ARI and NMI >0.5) was estimated between the Newman and Louvain methods, both modularity optimization approaches, and SBM and nSBM. Otherwise, inter-method similarity was overall low to moderate (ARI ≤0.25, NMI ≤0.68). For Infomap, its highest similarly was to Louvain in both runs (0.55 (0.25) and 0.54 (0.32) respectively), and lowest similarity to nSBM (0.08 (0.08) and 0.07 (0.08), respectively for the two runs).

Intra-method modularity similarity was also estimated in the $n = 3820$ brains with two rs fMRI runs. Corresponding statistics are summarized in Table 4. Similarity measured with either ARI or NMI was low for Newman, Louvain, and Ricci (ARI = 0.15–0.34; NMI = 0.32–0.48). For SBM and nSBM, ARI was low (0.17–0.20), but NMI was moderate (0.54–0.61). Similar results were estimated for Infomap (see Table S1), with median (IQR) ARI = 0.27 (0.22), and median (IQR) NMI = 0.41 (0.19). Note that modularity of rs networks may vary substantially between runs.

To assess the impact of overall graph topology on the similarity of modularity estimates, eight global network properties were calculated from each adjacency matrix: degree assortativity, mean connectivity, median connectivity, natural connectivity (a measure of network robustness), smallworldness, efficiency, global clustering coefficient, and topological stability. Correlograms were generated for both ARI and NMI and showed variable correlations between method agreement and topological characteristics. Correlations were estimated using Spearman's ρ given the non-normal distribution of some of these parameters. Overall, for each property, the direction of correlation was similar across methods, except for SBM–nSBM and Newman–Louvain pairs, which showed either no correlation or weak correlation in the opposite direction for all properties. Similarly, correlations between Ricci-Newman (and Ricci–Louvain) ARI and topological properties were also low ($\rho = -0.17$ to 0.26). For all other pairs, negative correlations were estimated between ARI (and similarly for NMI) and topological robustness ($\rho = -0.60$ to -0.86) global efficiency ($\rho = -0.59$ to -0.88), and global clustering ($\rho = -0.55$ to -0.83). Correlations between ARI (and

TABLE 3 Summary statistics (median and inter-quartile range (IQR)) for adjusted Rand index (ARI) and normalized mutual information (NMI) estimated between each pair of community detection methods applied to brain networks estimated from 2 fMRI runs.

Method agreement based on real functional networks		Ricci and SBM	Ricci and nSBM	Ricci and Newman	Ricci and Louvain	SBM and nSBM	SBM and Newman	SBM and Louvain	nSBM and Newman	nSBM and Louvain	Newman and Louvain
First run ($N = 5251$)											
Adjusted Rand index	Median	0.254	0.207	0.160	0.175	0.513	0.109	0.119	0.063	0.063	0.611
	IQR	0.329	0.327	0.104	0.112	0.077	0.071	0.069	0.039	0.030	0.166
Normalized mutual information	Median	0.661	0.682	0.435	0.501	0.813	0.441	0.495	0.409	0.466	0.632
	IQR	0.266	0.292	0.143	0.120	0.027	0.109	0.092	0.105	0.083	0.109
Second run ($N = 3820$)											
Adjusted Rand index	Median	0.101	0.078	0.153	0.175	0.516	0.094	0.103	0.057	0.058	0.628
	IQR	0.301	0.266	0.139	0.134	0.077	0.062	0.062	0.031	0.025	0.174
Normalized mutual information	Median	0.576	0.585	0.402	0.470	0.817	0.421	0.469	0.395	0.444	0.635
	IQR	0.383	0.407	0.209	0.173	0.026	0.105	0.096	0.096	0.084	0.115

TABLE 4 Summary statistics for intra-method modularity similarity between the two fMRI runs.

		Newman	Louvain	nSBM	SBM	Ricci	Infomap
Adjusted Rand index	Median	0.281	0.341	0.174	0.198	0.146	0.271
	IQR	0.148	0.133	0.066	0.071	0.165	0.216
Normalized mutual information	Median	0.321	0.462	0.605	0.537	0.480	0.413
	IQR	0.112	0.116	0.050	0.052	0.335	0.187

NMI) and degree assortativity and smallworldness were in the opposite direction. Specifically, there was a moderate positive correlation between degree assortativity and ARI, and similarly for NMI ($\rho = 0.49$–0.70) and similarly for small-worldness ($\rho = 0.54$–0.80) for most method pairs. Across methods, median connectivity was weakly correlated with ARI and NMI (positively or negatively, $\rho = -0.08$ to 0.15). These statistics are summarized in Figure 1. Similar correlations were found using the second fMRI run, and corresponding correlograms are shown in Supplemental Figure S1.

ARI and NMI values were also clustered based on quartiles of each topological measure, to assess the impact of network topology on method agreement at a more granular level. In other words, NMI and ARI quartile membership was determined by the statistics of the network property of interest, rather than the NMI or ARI statistics. Boxplots for NMI based on each property-specific statistic are shown in Figure 2 for the first run, and Figure 3 for the second second run. Corresponding box plots for ARI are shown in Supplemental Figures S2 and S3. NMI for SBM–nSBM was consistently highest (compared to other methods) across quartiles and topological properties, followed by Newman-Louvain. In the third and fourth quartiles of degree assortativity and small-worldness, NMI for the nSBM-Ricci was also high (comparable to that for SBM–nSBM), followed by NMI for SBM-Ricci, and the same was the case for the first and second quartile of topological robustness, stability, efficiency, and global clustering. Similar patterns were estimated in data from the second run. No other consistent patterns of NMI as a function of topological property statistics were identified, and NMI was overall lower (≤ 0.5) for most other pairs of comparisons. ARI for the Newman–Louvain comparison was highest across quartiles and properties, followed by the SBM–nSBM comparison. ARI for the nSBM-Ricci and SBM-Ricci comparisons was highest in the bottom quartile of topological robustness, stability, efficiency, and global clustering.

4.2 | Impact of method choice and inter-method agreement on associations between functional network modularity and other participant data

4.2.1 | Demographic, anthropometric, and physiological data

ARI and NMI were grouped as a function of demographic data, including sex [first run: $n = 2509$ males and $n = 2741$ females; second run: $n = 1782$ males and $n = 2038$ females], race [dichtomized as white (first run: $n = 3499$; second run: $n = 2587$) vs. nonwhite (first run: $n = 1680$; second run: $n = 1187$), given the unbalanced ABCD cohort in terms of race], and ethnicity [Hispanic (first run: $n = 1032$; second run: $n = 737$) vs. non-Hispanic (first run: $n = 4162$; second run: $n = 3052$)]. Summary statistics for ARI and NMI in each of these groups were separately estimated. In addition, in previous work, body mass index (BMI) status has also been associated with the topological organization of functional networks in a sample from the ABCD cohort (Brooks et al., 2021, 2023). Participants were thus grouped by BMI status as well (based on growth curves for age and sex (Centers for Disease Control)), first dichotomized as with obesity (first run: $n = 681$; second run: $n = 456$) versus without obesity (first run: $n = 4570$; second run: $n = 3364$), but also based on a more granular categorization: with underweight (first run: $n = 279$; second run: $n = 209$), normal BMI (first run: $n = 3545$; second run: $n = 2611$), overweight (but not with obesity; first run: $n = 746$; second run: $n = 544$), and with obesity. Participants were also classified based on sleep length [recommended amount for age (first run: $n = 2582$; second run: $n = 1897$) vs. less than recommended (first run: $n = 2669$; second run: $n = 1923$)], aspects of sleep quality, specifically frequency of snoring and gasping for air, which in prior work has been shown to impact the topologies of functional networks in this cohort (Brooks et al., 2022), and pubertal stage (pre early, mid puberty). Prior work in this cohort has identified differences in modularity as a function of pubertal stage (Brooks et al., 2021). Median (IQR) values for pairwise method similarity in each of these groups are summarized in Supplemental Tables S2 and S3.

To assess statistical differences in ARI and NMI between groups, simple comparisons were first performed using the nonparametric Wilcoxon rank sum test for unpaired samples, and then simple ordinary linear regression models. ARI was statistically associated with race for all but three method pairs (Ricci–Newman, Ricci–Louvain, and SBM–nSBM), $p < .01$, and NMI was statistically associated with race for all method pairs $p < .01$. ARI and NMI were higher for Ricci-SBM, and Ricci-nSBM comparisons in nonwhite participants, and higher for Ricci–Newman, Ricci–Louvain, and Newman–Louvain comparisons in white participants. Similar associations with race were estimated for ARI and NMI estimated from method comparisons based on the second fMRI run.

Based on the first run, NMI for all (but Newman–Louvain) method pairs was statistically higher in youth with obesity ($p<.02$). ARI was also correlated with obesity status but only for some pairs ($p<.01$). Based on the second run, consistent associations between obesity status and ARI for some methods were estimated, but none for NMI.

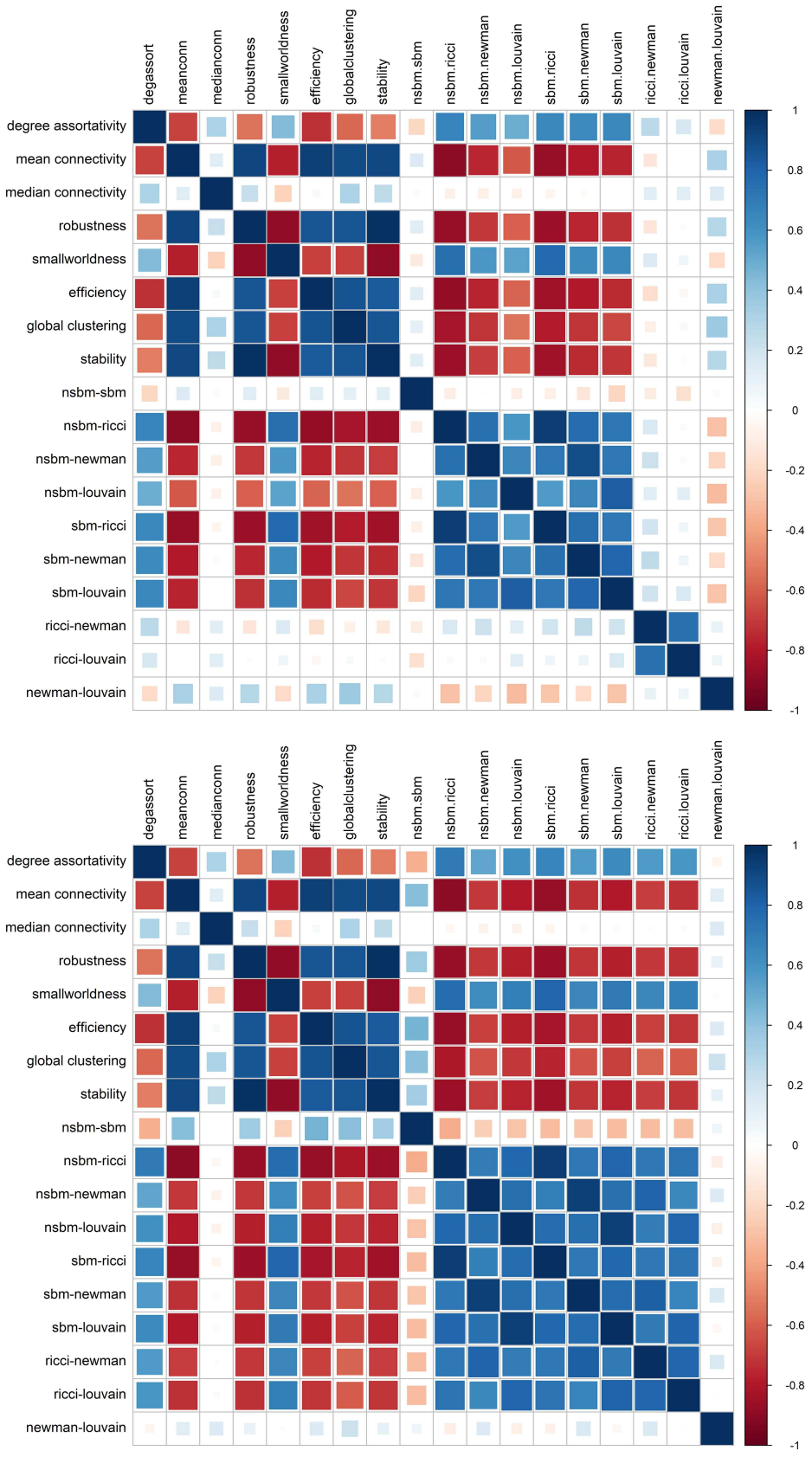

FIGURE 1 Correlograms showing correlations (estimated using Spearman's ρ) between topological network properties and method agreement based on ARI (top) and NMI (bottom), using best-run data. Red and blue colors represent positive and negative correlations, respectively, and the size of squares reflects correlation strength.

Statistical associations between sleep length and ARI and NMI were estimated for multiple method pairs ($p < .03$) and were fairly consistent across runs (though fewer associations were estimated for ARI in the second run). There were no differences between Hispanic and non-Hispanic participants for any comparison using ARI ($p > .32$). All adjusted p-values are provided in Tables 5 and 6. Corresponding p-values based on statistical models are summarized in Tables S4 and S5.

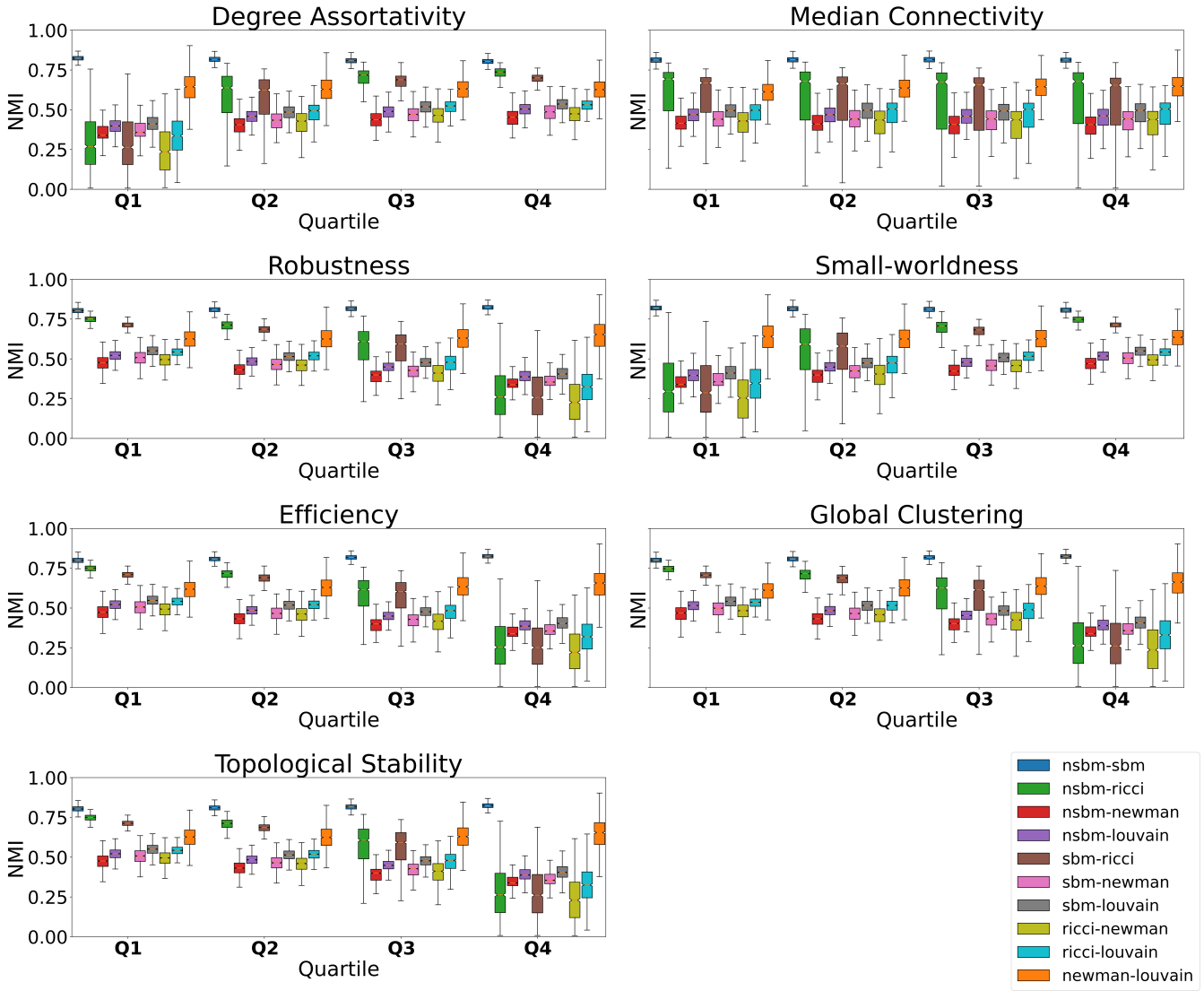

FIGURE 2 Boxplots for NMI (based on data from the best fMRI run), at each quartile of the estimated topological properties.

To assess the impact of method selection on biological associations of interest, ordinary linear regression models were developed to examine the relationship between method-specific modularity estimates (the dependent variables) and independent parameters related to sleep length and quality, BMI (a continuous variable), BMI category, and obesity status (dichotomous). Across methods, and both fMRI runs, modularity was statistically associated with BMI and obesity status ($p \leq .01$). Similarly consistent (across methods) statistical associations were estimated between modularity and sleep length, difficulty breathing, snoring, gasping for breath, and daytime sleepiness, in the fMRI run ($p < .03$) but not the second, possibly an issue of the smaller sample size for that run. These results are summarized in Tables 7 and 8.

4.2.2 | Cognitive outcomes

The ultimate goal of network analyses is to map the topology of functional brain circuits onto the cognitive processes they support. Therefore, associations between method-specific modularity estimates and cognitive outcomes were investigated using a subset of tasks from the ABCD neurocognitive battery (Luciana et al., 2018), including the Flanker, List Sorting Working, Dimensional Card Sort, Cash Choice, Matrix Reasoning, and Rey Auditory Verbal Learning Tasks. Together, these measure cognitive flexibility, information processing, attention, working memory, learning, inhibitory control, impulsivity, and problem-solving, all processes that continue to develop throughout adolescence and may be associated with community structure in the brain. Depending on the task outcome, linear or logistic regression models were developed, with modularity estimated by each method as the primary independent variable and task scores/performance measures as the dependent variables. Models were adjusted for age, sex, race, ethnicity, and family income. *p*-values were adjusted for false discovery, across methods. Modularity estimated with all methods (except nSBM) was statistically associated with performance in the List Sorting task ($p \leq .04$). Modularity estimated with SBM and nSBM but no other methods was

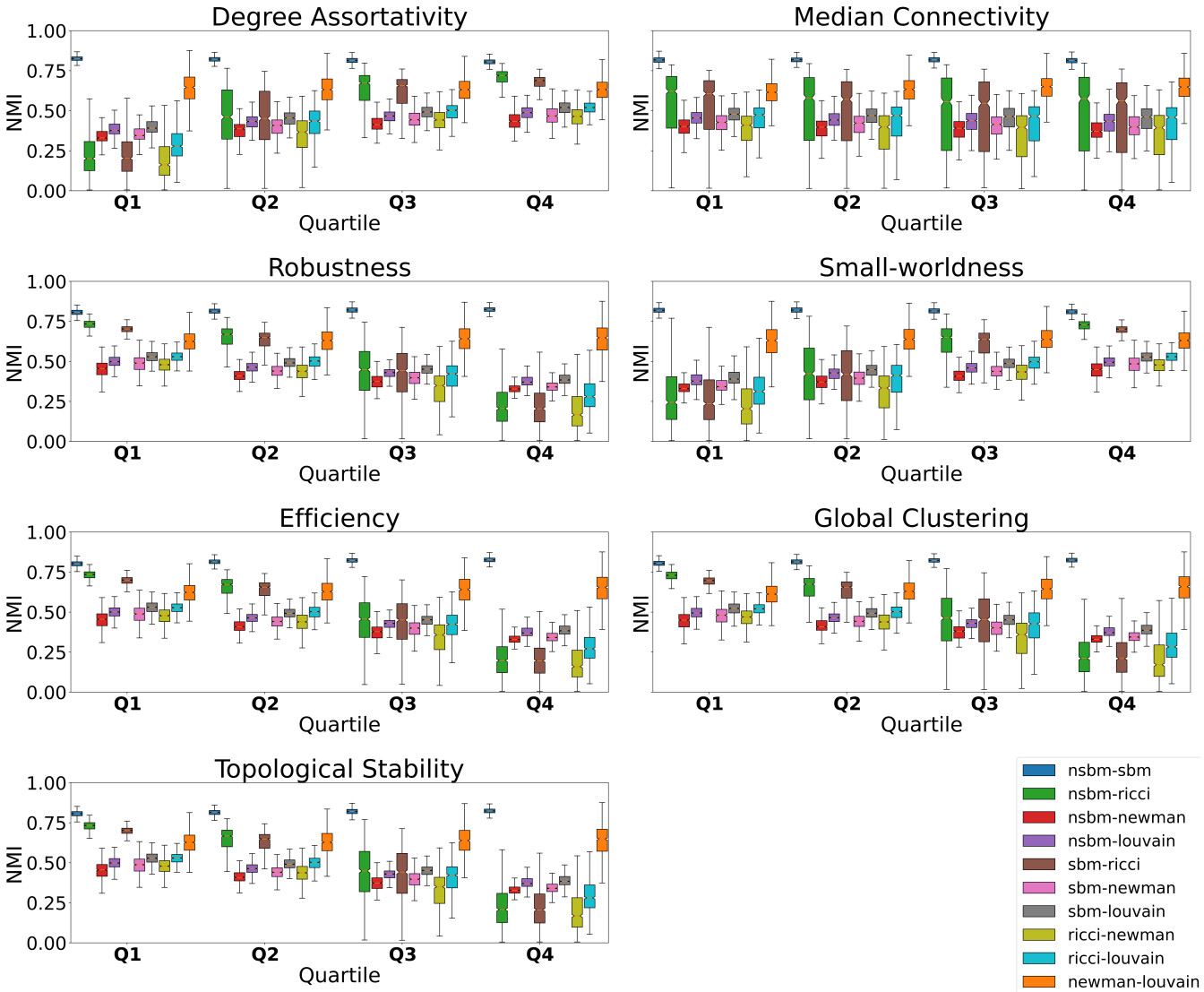

FIGURE 3 Boxplots for NMI (based on data from the second-best fMRI run), at each quartile of the estimated topological properties.

associated with performance in the dimensional card sort task ($p \leq .01$).

4.3 | Method comparison based on synthetic graphs

Synthetic graph modularity (the ground truth) was compared across methods, as model parameters were systematically varied. Of particular interest was the impact of varying μ on method performance. This is the model parameter that can be directly varied in a way that simulates developmental changes in the connections within and across communities, as the brain acquires its small-world topology, a hallmark of the fully-developed connectome in adulthood. In the real data, median (IQR) values of μ were 0.14 (0.12) and 0.17 (0.14) for the two runs. Figure 4 shows the variation of ARI as a function of μ, for each method compared to the ground truth (top plot). Pairwise ARI for comparisons of Ricci, SBM, and Louvain methods with all others (left, middle, and right panels, respectively) are also shown (bottom plots).

Overall, at $\mu < \sim 0.14$ (which approximately corresponds to connectomes that are more developed than those in the real dataset, and are highly connected within communities with relatively sparse connections between communities—a characteristic of the adult brain), SBM, nSBM, Newman, and Louvain performed similarly with median ARI with the ground truth $> \sim 0.95$, and Ricci flow had median ARI ~ 0.7. Modularity estimates based on SBM/nSBM had high agreement with the ground truth at higher μ ranges (up to ~ 0.3) as well. Beyond these μ values, their agreement progressively decreased to median ARI ~ 0.6, which remained fairly constant across high μ values. Performance of Newman and Louvain methods decreased more rapidly in the μ range ~ 0.3 to ~ 0.5, with ARI ≤ 0.5 at higher μ values, and almost no agreement with the ground truth (ARI < 0.1) at very high μs. Ricci followed a similar trend, but its almost complete disagreement with the ground truth occurred at lower μ values (~ 0.4).

TABLE 5 *P*-values (adjusted for the false discovery rate) for Wilcoxon rank sum test comparisons of ARI and NMI (first run) across distinct groups of participants, dichotomized based on sex (male vs. female), race (white vs. nonwhite), ethnicity (Hispanic vs. non-Hispanic), obesity status (with vs. without obesity) and sleep length (recommended amount for age vs. less than recommended).

	Ricci and SBM	Ricci and nSBM	Ricci and Newman	Ricci and Louvain	SBM and nSBM	SBM and Newman	SBM and Louvain	nSBM and Newman	nSBM and Louvain	Newman and Louvain
Adjusted Rand index (*ARI*)										
Sex	-	-	-	0.037	-	-	-	-	0.022	-
Race	<0.001	<0.001	-	-	-	<0.001	<0.001	<0.001	<0.001	<0.001
Ethnicity	-	-	-	-	-	-	-	-	-	-
Obesity Status	0.003	0.003	-	-	-	0.002	0.003	0.002	-	-
Sleep Length	<0.001	<0.001	0.005	0.033	-	<0.001	<0.001	<0.001	0.005	-
Normalized mutual information (*NMI*)										
Sex	-	-	-	-	-	-	-	-	-	-
Race	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
Ethnicity	-	-	-	-	-	-	-	-	-	-
Obesity status	0.015	0.015	0.015	0.015	0.018	0.017	0.015	0.020	0.016	-
Sleep length	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	-

TABLE 6 *P*-values (adjusted for false discovery) for Wilcoxon rank sum test comparisons of ARI and NMI (second run) across distinct groups of participants, dichotomized based on sex (male vs. female), race (white vs. nonwhite), ethnicity (Hispanic vs. non-Hispanic), obesity status (with vs. without obesity) and sleep length (recommended amount for age vs. less than recommended).

	Ricci and SBM	Ricci and nSBM	Ricci and Newman	Ricci and Louvain	SBM and nSBM	SBM and Newman	SBM and Louvain	nSBM and Newman	nSBM and Louvain	Newman and Louvain
Adjusted Rand index (*ARI*)										
Sex	-	-	-	-	-	-	-	-	-	-
Race	0.002	<0.001	-	-	0.034	0.002	<0.001	0.009	0.029	<0.001
Ethnicity	-	-	-	-	-	-	-	-	-	-
Obesity status	0.036	0.036	0.049	0.036	-	0.036	0.036	0.036	-	-
Sleep length	-	0.027	-	-	-	0.021	0.024	0.022	-	0.021
Normalized mutual information (*NMI*)										
Sex	-	-	-	-	-	-	-	-	-	-
Race	<0.001	<0.001	0.002	<0.001	<0.001	-	<0.001	-	<0.001	<0.001
Ethnicity	-	-	-	-	-	-	-	-	-	-
Obesity status	-	-	-	-	-	-	-	-	-	-
Sleep length	0.024	0.017	0.017	0.024	0.008	0.046	0.017	-	0.013	0.046

Statistics on estimated number of communities across methods (and ground truth) are summarized in Table 9. In these simulations, the number of communities was inherently constrained by the range of communities in the real data, based on the Newman method. Median number of communities estimated with SBM and nSBM was closest to the ground truth, followed by the Louvain method. In contrast, in the real (and thus unconstrained in number of communities) data, SBM and nSBM estimated large, and potentially biologically implausible, numbers of communities. In synthetic graphs, both Newman and Ricci methods underestimated the number of communities. In the real data, Newman estimated a relatively low number, but Ricci estimated almost 10 times as many communities. On average, Infomap also underestimated the number of communities in the synthetic graphs (median = 1 community, range = 1–33).

Based on inter-method ARI estimates, agreement between SBM and all but Ricci methods, and similarly for Louvain was high at $\mu \leq \sim 0.3$, but rapidly decreased in the μ range $0.3 - 0.6$, with very low agreement at higher μ values. High μ values reflect networks topologies in which communities are difficult to detect, corresponding to connectomes at earlier developmental stages, in which the brain has a high number of redundant connections and not well-defined modular organization. The performance of the Ricci flow method was lower than other methods across μ values, and poor at high μ regimes. Increasing μ values are reflected in topologies where the difference in density of edges within and across communities is decreasing, ultimately resulting in networks that have weak or no community structure. For such topologies, the Ricci flow method keeps all vertices in a single community, recognizing that there is no meaningful community

TABLE 7 Statistics (regression coefficients and corresponding *p*-values—adjusted for false discovery) for linear regression models testing the association between method-specific modularity estimates from the first fMRI run and sleep quantity and quality parameters (from the Sleep Disturbance Scale for Children (SDSC)), BMI and obesity status.

		Newman	Louvain	nSBM	SBM	Ricci
Sleep length	Beta	0.040	0.037	0.043	0.046	0.049
	p-value	0.009	0.013	0.007	0.006	0.006
Difficulty breathing	Beta	0.035	0.036	0.037	0.037	0.030
	p-value	0.015	0.015	0.015	0.015	0.030
Gasping for breath	Beta	0.041	0.040	0.035	0.040	0.040
	p-value	0.006	0.006	0.013	0.006	0.006
Snoring	Beta	0.047	0.045	0.046	0.047	0.044
	p-value	0.002	0.002	0.002	0.002	0.002
Daytime sleepiness	Beta	0.039	0.038	0.037	0.038	0.034
	p-value	0.010	0.010	0.011	0.010	0.015
BMI	Beta	0.066	0.064	0.065	0.064	0.071
	p-value	<0.001	<0.001	<0.001	<0.001	<0.001
Obesity status (binary)	Beta (non-standardized)	0.019	0.019	0.013	0.019	0.023
	p-value	<0.001	<0.001	<0.001	<0.001	<0.001
BMI category	Beta	0.058	0.054	0.058	0.056	0.062
	p-value	<0.001	<0.001	<0.001	<0.001	<0.001

TABLE 8 Statistics (regression coefficients and corresponding *p*-values—adjusted for false discovery) for linear regression models testing the association between method-specific modularity estimates from the second fMRI run and sleep quantity and quality parameters (from the Sleep Disturbance Scale for Children (SDSC)), BMI and obesity status.

		Newman	Louvain	nSBM	SBM	Ricci
BMI	Beta	0.052	0.053	0.060	0.057	0.043
	p-value	0.002	0.002	0.002	0.002	0.012
Obesity status (binary)	Beta	0.021	0.021	0.011	0.018	0.022
	p-value	<0.001	<0.001	<0.001	<0.001	0.002
BMI category	Beta	0.058	0.057	0.057	0.058	0.052
	p-value	<0.001	<0.001	<0.001	<0.001	0.002

partitioning. Median number of communities estimated with this method in synthetic graphs was indeed 1 (IQR = 4). This statistic was further impacted by the distribution of μ in synthetic graphs. Finally, when a method estimates only one community in a network that what more than one community, by definition, the ARI value is 0, although, by the power law distribution of the community size in the generative model, there may be a large overlap of the (single) community with the largest community in the graph.

To examine the impact of other model parameters on method accuracy and inter-method similarity, synthetic graphs were dichotomized based on μ, $\mu \leq 0.4$ versus $\mu > 0.4$. The cutoff was based on the ARI curves in Figure 4. Each method was then compared to the ground truth as a function of the number of communities and median node degree. The variation of ARI as a function of these parameters is shown in Figure 5. For $\mu \leq 0.4$, agreement with the ground truth increased for all methods for graphs with 2 to ~10 communities. For those with a higher number of communities (up to 20 communities), ARI was ≥ 0.9 for Louvain, SBM, and nSBM, ~0.5 to ~0.8 for Newman, and ~0.5 to ~0.6 for Ricci flow. The opposite trend was identified as a function of median node degree. Louvain, SBM, and nSBM had higher agreement (ARI > 0.8) with the ground truth for median node degree ≤ ~60 − 70, and lower agreement for higher node degree. Similar trends were estimated for Newman and Ricci flow, but both methods performed statistically worse even in graphs with lower degree < 50. SBM and nSBM performed better overall even in highly connected graphs (median node degree 100), followed by Louvain, Newman, and Ricci flow. For $\mu > 0.4$, all methods except SBM and nSBM had very low agreement with the ground truth (ARI < 0.2), across number of communities and median node degree. SBM and nSBM-based modularity estimates were in moderate agreement with the ground truth (ARI ~0.7 to ~0.8) for ≤6 communities, but ARI rapidly decreased for higher numbers of communities. For graphs with more than ~10 communities, SBM and nSBM had very low agreement with the ground truth, statistically similar to the other methods. The opposite trend was identified for ARI as a function of median node degree. All methods had low ARI for graphs with degree ~ < 50. SBM and nSBM had significantly higher ARI in graphs with higher median node degree (up to ~0.7 even for degree > 100).

Finally, ARI was examined as a function of median inter-community connection strength, dichotomized as low versus high,

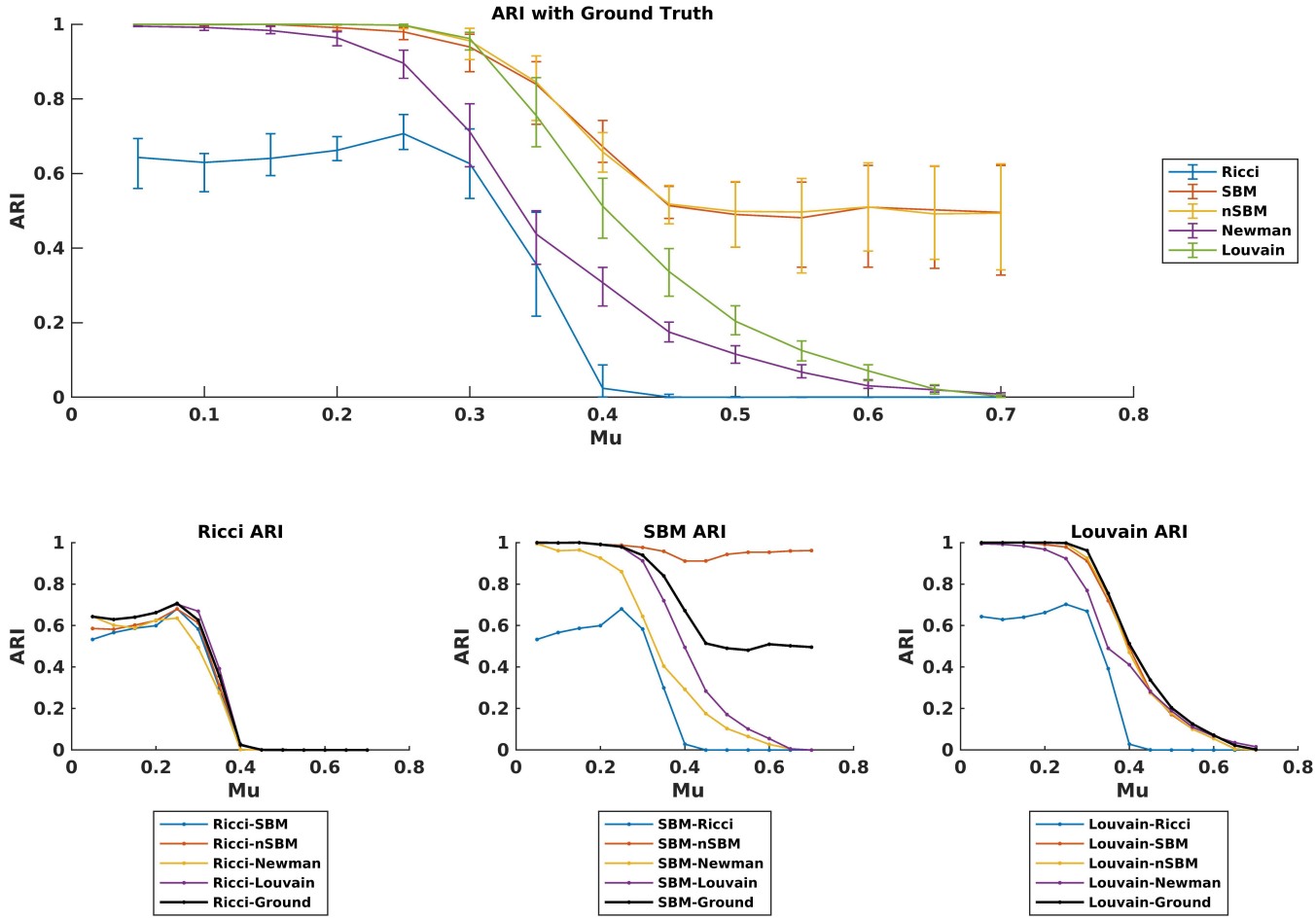

FIGURE 4 Median ARI as a function of μ. Top Plot: ARI comparing each method to ground-truth communities. Bottom plots: ARI comparing Ricci, SBM, Newman, and Louvain (respectively) to each other and to the ground truth.

based on cutoffs estimated from the real data (bottom 10% and top 10% of connectivity values in the real adjacency matrices). The parameter μ was held constant at 0.35. In the low connectivity group, median (IQR) ARI was 0.84 (0.40) for SBM, 0.84 (0.38) for nSBM, 0.76 (0.50) for Louvain and 0.44 (0.37) for Newman. In the high connectivity group, median ARI was approximately the same for SBM, slightly higher for nSBM (0.87 (0.39)), but lower for Louvain (0.59 (0.59)) and Newman (0.32 (0.31)), suggesting an inverse relationship between similarity with ground truth and connectivity for these two methods. SBM of nSBM performed equally well in both connectivity regimes. Ricci could not be compared to these methods, since the implementation uses only graph topology, not the connection strength (i.e., was applied to binary graphs).

TABLE 9 Summary statistics (minimum, maximum, median, and inter-quartile range (IQR)) for ground truth number of communities as well as those detected with each method from synthetic graphs.

Number of communities			
Method	**Minimum**	**Maximum**	**Median (IQR)**
Ground truth	2	20	7 (7)
Newman	1	22	2 (1)
Louvain	1	20	5 (4)
SBM	1	23	6 (4)
nSBM	1	21	6 (4)
Ricci flow	1	38	1 (4)

5 | DISCUSSION

Graph theoretic analyses are increasingly used to investigate the topological characteristics of structural and functional brain networks. These analyses have repeatedly revealed a modular connectome, in which community structure represents a fundamental characteristic of the brain's organization that is critical for efficient information processing and integration, robustness to perturbations (e.g., stressors), rapid response to cognitive demands, and learning (Bullmore & Sporns, 2012; Sporns & Betzel, 2016). Modularity of brain networks has been correlated with cognitive function across domains and may be abnormally altered by neurological, neuropsychiatric, and neurodevelopmental diseases and disorders (Fornito et al., 2015;

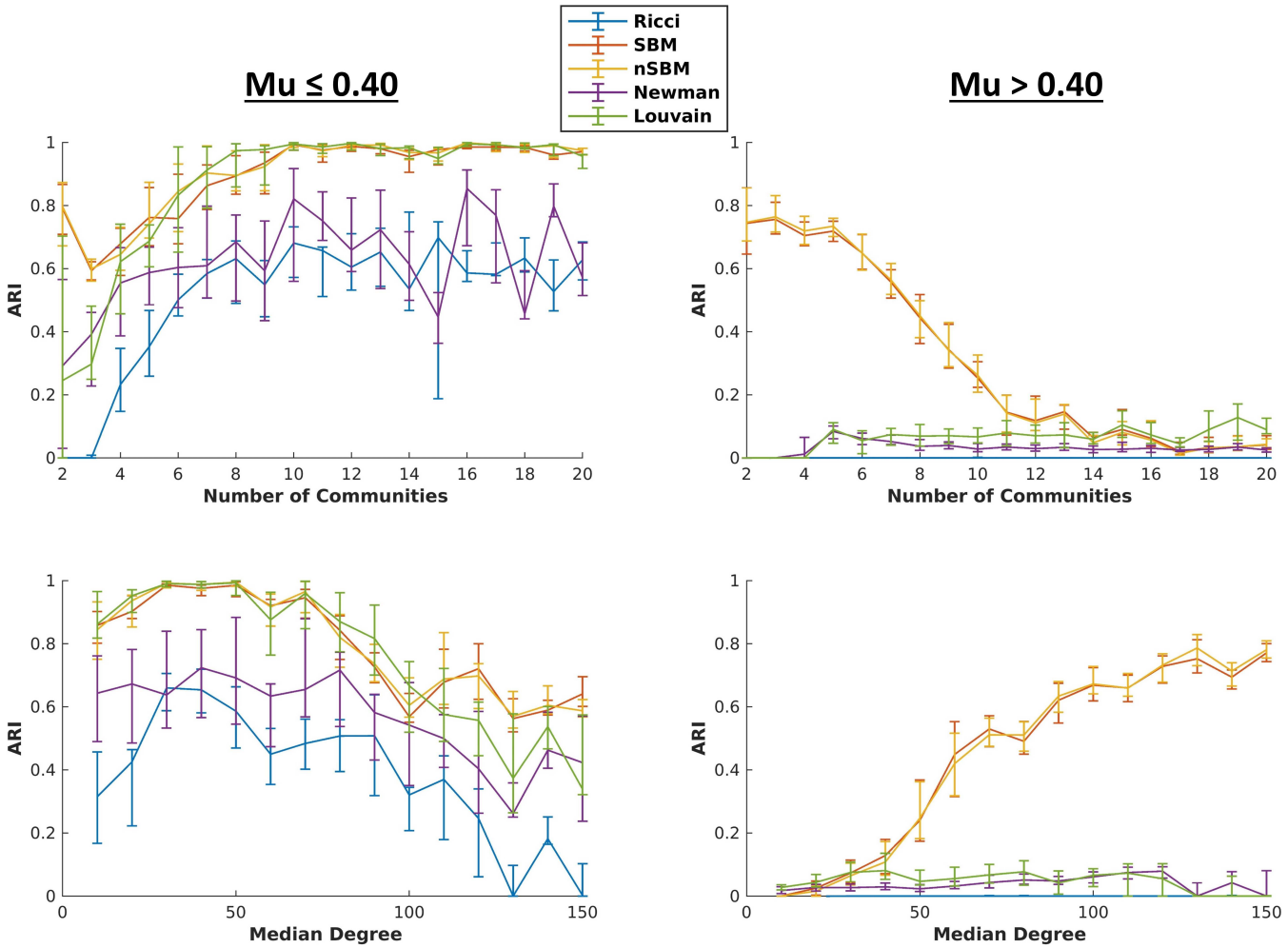

FIGURE 5 Median ARI (compared to ground truth) as a function of true number of communities and degree for synthetic graphs with $\mu \leq 0.40$ (left) and all graphs with $\mu > 0.40$ (right).

Griffa et al., 2013; Koubiyr et al., 2019; Ma et al., 2017). Despite the role of the connectome's community structure in brain function and cognition, there is neither a methodological gold standard for its estimation nor a systematic approach for selecting an optimal method. This has contributed to the reproducibility crisis in the Neuroscience field in general, and that of connectomics more specifically, and has limited the interpretation of findings based on method-dependent brain community structure and modularity. It is also unclear how different methods perform in terms of absolute accuracy and reliability, as well as relative to each other, as a function of graph topology. It is further unclear how method-dependent modularity estimates are mapped onto differences in associations between this topological property and cognitive, behavioral, and other individual data.

This study has systematically addressed these gaps in knowledge using a historically large dataset of developing brain connectomes estimated from over 5000 early adolescents in the ABCD study, and a large dataset of data-inspired synthetic graphs with controlled parameters. It has focused on three classes of methods, including state-of-the-art and widely used approaches in Neuroscience (the Newman and Louvain methods, which maximize modularity) and more novel to the field probabilistic (Bayesian inference within the SBM framework) and geometric (Ricci flow) methods. In secondary comparisons, it has also applied Infomap, an increasingly popular method, to the data. Using synthetic graphs that together represented the topologies of connectomes from early life (underdeveloped) to adulthood (optimally developed), the study has investigated the topological parameters that may impact method performance, and their differential effects on inference. Using the real data, which captured the inherent heterogeneity of the developing connectome, the study has compared modularity estimates from different methods and the consistency of their associations with physiological, demographic, and cognitive data. To assess reliability of the findings, all methods were applied to rs data from two fMRI runs.

The primary study finding is that different types of community detection methods can yield highly dissimilar modularity estimates, which raises the issue of method selection, as well accuracy and reproducibility of results with different approaches. Although there was modest similarity between modularity estimates obtained with the Newman and Louvain methods, both of which use modularity maximization as the basis for community detection, modest similarity

between Louvain and Infomap, and, as expected, higher similarity between SBM and its nSBM variant, there was low agreement between all other pairwise method comparisons. These results were consistent across fMRI runs. In part, this dissimilarity may be attributed to different approaches and metrics used by these methods, but also to the underlying characteristics of the graphs. Prior work on comparison of community detection methods in synthetic graphs and real networks (mostly unrelated to the brain) has shown that performance depends on multiple parameters, including node degree distribution, network connectivity, and size and distribution of communities (Funke & Becker, 2019; Orman & Labatut, 2009; Taya et al., 2016; Yang et al., 2016). The synthetic brain graph results are aligned with prior findings and confirm this dependence.

To understand the role of topology on community estimation, agreement between methods was examined as a function of each brain's network topological properties. Results were consistent across fMRI runs. Moderate positive correlations between method agreement (ARI and NMI) and degree assortativity and small-worldness were estimated for most methods, which suggests that the higher the modular organization (reflected in both the affinity of nodes connecting to similar nodes, e.g., within the same community, and high local (within community) connectivity but sparse inter-community connectivity), the higher the similarity of community structure identified with different approaches. In turn, this indicates that in developed brain networks, which are likely to have high assortativity and a small-world organization, these methods may yield similar results. For such topologies, community detection analyses are likely to have relatively low method bias. However, in underdeveloped (or neurodegenerating) brain networks where, these properties are either not yet optimized or are no longer optimal, sensitivity and specificity in detecting communities may vary substantially between methods. In weakly and/or redundantly connected networks, communities may be harder to detect, and method bias is likely to be high. A range of weak to moderate negative correlations was estimated between method agreement and global efficiency, clustering, and natural connectivity (a measure of topological robustness). The latter increases with the number of network connections, and may thus favor edge redundancy, which, in turn, makes it difficult to reliably identify communities. The negative correlation between global clustering coefficient and method agreement was surprising. This coefficient is based on node triplets—the ratio of closed triplets to all triplets, which reflects clustering within the network. The negative correlation suggests that for networks that cluster to a higher degree, community detection methods are more likely to differ in the partitions they generate. It is possible that small clusters may negatively impact method agreement, that is, a higher number of small communities may lower the probability of community overlap and matching between methods.

In a dataset of over 5000 developing connectomes, topological network heterogeneity is expected to be high. Thus, method agreement was also examined in stratified subcohorts, based on demographic, anthropometric, and physiological parameters previously reported as significant contributors to this heterogeneity, and modulators of the organization of the early adolescence connectome (Brooks et al., 2021, 2022). Several method pairs had higher agreement in subcohorts of participants with obesity and than those who obtained less than the recommended sleep amount. In prior work, we have reported lower global clustering and topological robustness in early adolescents who do not obtain sufficient sleep or have unhealthy excess BMI. Given the negative correlations between these properties and method agreement, these results are not surprising. Similarly, sex- and race-related differences in method agreement may reflect differences in sleep and BMI in these groups. For example, we have previously identified racial disparities in sleep quantity and quality in nonwhite participants from this cohort, as well as statistically higher BMI in some racial groups (Brooks et al., 2022). These findings suggest that increased cohort homogeneity may be correlated with higher inter-method agreement.

Given substantial dissimilarity between method types, the impact of method choice on associations between method-specific modularity and cognitive measures, sleep quantity and quality, BMI, obesity status, and performance in multiple cognitive tasks that require the activation of distributed brain networks (and domain-specific communities) was next assessed. Consistent modularity associations with sleep, BMI-related measures, and performance in the list sorting task were estimated across methods (and some of them for both fMRI runs), which indicates method invariance, at least for these associations. Inconsistent associations were estimated for the dimensional card sort task (only with SBM and nSBM-based modularity), which suggests that method selection could impact the identification of relationships between modularity and cognitive performance. It is also possible that some methods may better resolve smaller clusters (i.e., finer structure and thus a large number of communities), and these higher-resolution communities correlate with some cognitive outcomes. However, the Ricci method also estimated a relatively large number of communities, but the Ricci-based modularity association with this particular task was nonsignificant. Without a ground truth, the accuracy of any of these higher-resolution communities are impossible to assess.

To investigate method accuracy and impact of the underlying network topology on it, a large dataset of synthetic graphs was analyzed. Topological properties estimated from the data were used to systematically vary the parameters of the LFR model used to generate the graphs. The number of communities was constrained to vary within the range estimated from the real data, using the Newman method. Method accuracy and agreement were examined as a function of within-community connectivity relative to the remaining network, node degree, and number of communities, all properties that in the real data analyses emerged as important contributors to method agreement. Although each parameter was varied separately, the resulting graphs included complex combinations of topological characteristics. Due to issues of convergence of some graphs, the median parameter μ in the synthetic graphs was higher than that in the real data but did span ranges that would correspond to underdeveloped (early life) to fully developed (young adulthood) brain networks.

The synthetic graph analysis led to several important findings. First, there are topological regimes, related to the local community

connections relative to the rest of the network, in which most methods performed similarly for networks with high numbers of communities and low median node degree, with the exception of Ricci flow. These regimes correspond to fairly well-developed connectomes but also overlapped with the μ range in the real data (< 0.2). In other topological regimes, corresponding to underdeveloped or degenerating connectomes, and thus networks with less clearly identifiable community structure, SBM and its nSBM variant performed significantly better than all other methods. SBM uses Bayesian inference for community detection. The advantage of using statistical inference within the SBM framework is that the algorithm aims to identify both the node community assignment and interaction between communities based on the latent network structure. Being the only probabilistic method investigated in this study, it may also be more robust to the uncertainty associated with community structure in noisy (as a result of incomplete maturation or degeneration) and/or redundantly or weakly connected networks. The synthetic graph results also further highlight differences between methods based on their underlying assumptions. For example, Ricci flow assumes that edges within communities are denser than edges across communities. When this assumption does not hold (or is reversed), Ricci flow cannot detect multiple communities. In contrast, SBM does not require such an assumption and finds a separation when the intra-community and inter-community probabilities differ.

In real graphs, the estimation of community structure was not constrained for any of the methods, so that they can be compared fairly given that Newman and Louvain are entirely unsupervised approaches. Thus, no upper limit on the estimated number of communities was imposed on SBM/nSMB and Ricci flow. This led to a median number of communities by these methods that was on average ~8 to ~18 times higher than the number estimated by Newman and Louvain. Prior work has shown that the number of communities in brain networks in youth is typically low (Tooley et al., 2022), and close to the number estimated by Newman and Louvain. In synthetic graphs, the number of communities was inherently constrained by the Newman-based estimated range in the real data. In these graphs, SBM and nSBM estimated numbers of communities that were closest to the ground truth, whereas all other methods underestimated this number, with Ricci producing the lowest number of communities. Therefore, if constrained, by imposing a biologically reasonable upper limit on the number of communities, SBM may be the most accurate community detection approach in brain networks.

Despite its strengths, including the investigation of several types of methods and application to large and topologically heterogeneous real and synthetic network datasets, this study also had some limitations. First, an exhaustive comparison of available community detection methods was not performed and was beyond the scope of this study. Prior work has compared a large number of methods (Dimitriadis et al., 2021). Instead, the goal of the study was to compare classes of methods. Representative approaches from each class were selected based on their popularity in the Neurosciences (Newman and Louvain), increasing applications in other fields, and type (probabilistic—SBM, and geometric—Ricci flow). There are other emerging methods, for example, based on deep learning (Jin et al., 2021; Kipf & Welling, 2017; Yang et al., 2018), that hold promise to improve community detection in complex networks such as brain connectomes. These methods continue to be optimized but may also have shortcomings due to the volume of data needed to train the associated models. Furthermore, deep learning techniques often have limited interpretability and transparency. In secondary comparisons, we also examined the Infomap algorithm. Results from both the real data and synthetic graphs further highlighted the high disparity of community detection estimates and method dependence.

Another limitation of this study is that given that the generative model used to generate the synthetic graphs is inherently similar to SBM (though not the same). Thus, the superior performance of SBM could be biased by the generative model choice. However, there are differences between the LFR and SMB models, and the former was chosen specifically because its parameters could be mapped onto topological properties that could be estimated from the real data. Furthermore, despite a systematic, data-driven approach for varying the graph parameters, resulting models had complex topologies, with topological characteristics that were inherently correlated with each others. This is a limitation of all generative graph models. Also, the performance of Ricci flow rapidly decreases when ground truth community structure drops. A context-specific optimization of this approach at the low modularity range could be separately designed, but this optimization was outside the scope of the present study. Finally, it is possible that the metrics of similarity (ARI and NMI) used in this study may, in some cases, be influenced by the skewness of community size and number of communities distributions. Although these are well-established and widely used metrics, additional future investigations using other similarity measures could be valuable (Gates et al., 2019).

Interrogation and characterization of the brain's circuit topology across scales has been the focus of a wide range of human and animal studies. Circuit-level experimental investigations and, more recently, manipulations, have called for appropriate analytic tools to characterize the organizational complexity and fundamental properties (e.g., community structure) of neural circuits. However, these tools need to be carefully selected, to maximize reliability of estimated topological properties. To the best of our knowledge, this is the first study to systematically assess performance of different classes of community detection methods using a large, topologically heterogeneous dataset of incompletely maturated human connectomes from early adolescents, and a large dataset of data-inspired synthetic graphs. Beyond method comparisons, this study also has assessed the impact of method selection on associations between method-specific modularity and individual data of interest, including cognitive measures. Thus, despite some limitations, this study makes a significant contribution to the field and raises awareness of potential advantages and shortcomings of different types of community detection methods. It also provides quantitative evidence that some community detection methods may yield relatively similar results in networks with specific characteristics, including clearly identifiable communities and parsimonious connections between communities, but disparate results in

other topological regimes, particularly those corresponding to highly underdeveloped brain networks.

This study has also provided new insights into similarities of community estimates as a function of graph topology, particularly degree assortativity, small-worldness, and sample homogeneity (based on individual demographic, physiological, and/or anthropometric characteristics), as well as insights into method performance as a function of connectivity, number of connections within and across communities and overall number of communities in a graph. However, it has also provided evidence that associations between method-specific modularity estimates and other data (demographic, physiological, cognitive) could be method-dependent. Associations between some individual characteristics, such as sleep quantity and quality, BMI and select cognitive measures and modularity were found to be method-invariant, but one association with a cognitive outcome was method-dependent. On one hand, this may be due to higher sensitivity of some methods to identify smaller communities that are associated with a parameter of interest. However, method dependence may also suggest lack of reproducibility and thus limited biological relevance. These findings highlight the importance of reliability analysis of all modularity-related findings, using multiple methods to ensure their consistency. Finally, it also highlights the robustness and accuracy of Bayesian inference within the framework of SBM (and nSBM), and thus the potential utility of this method for community detection in brain networks, across topological regimes.

6 | CONCLUSION

The topological organization and characteristics of brain networks plays a critical role in the accuracy, reliability, and similarity of community detection methods, with relatively low bias, high similarity, and comparable accuracy in networks with well-defined community structure and sparse inter-community connections, but low similarity and differential accuracy in redundantly and/or weakly connected networks with difficult to detect communities. A probabilistic approach, such as Bayesian inference in the framework of SBM, may provide robust estimates of community structure independent of topological characteristics and may be more appropriate than widely used methods in Neuroscience, such as Newman and Louvain. Finally, to ensure biologically meaningful inferences, findings based on analyses of brain network community structure should be confirmed with multiple reliable detection methods.

ACKNOWLEDGEMENTS

The authors would like to thank Calli Smith for her assistance with this manuscript.

FUNDING INFORMATION

This study was supported by the National Science Foundation awards #1940094, 1649865, 2116707 and 2207733 (SJB, CS), #1938914 (VOJ, PD), #2126582 (HW, CD, JG).

CONFLICT OF INTEREST STATEMENT

The authors have no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available through the National Institute of Mental Health Data Archive (NDA): https://nda.nih.gov/. The assession number (DOI) associated with study in NDA repository is: https://doi.org/10.15154/1523041. Codes used in the data analyses are available at https://github.com/cstamoulis1/Next-Generation-Neural-Data-Analysis-NGNDA/, https://github.com/victoria-oceanna/SBM-brain-clustering/, and https://github.com/SBUhaotian/BrainRicciFlow/.

ORCID

Catherine Stamoulis https://orcid.org/0000-0001-9814-1802

REFERENCES

Alexander-Bloch, A., Lambiotte, R., Roberts, B., Giedd, J., Gogtay, N., & Bullmore, E. (2012). The discovery of population differences in network community structure: New methods and applications to brain functional networks in schizophrenia. *NeuroImage*, *59*, 3889–3900.

Alexander-Bloch, A. F., Gogtay, N., Meunier, D., Birn, R., Clasen, L., Lalonde, F., Lenroot, R., Giedd, J., & Bullmore, E. T. (2010). Disrupted modularity and local connectivity of brain functional networks in childhood-onset schizophrenia. *Frontiers in Systems Neuroscience*, *4*, 147.

Amini, A. A., Paez, M. S., & Lin, L. (2019). Hierarchical stochastic block model for community detection in multiplex networks. http://arxiv.org/abs/1904.05330

Arenas, A., Fernandez, A., & Gomez, S. (2008). Analysis of the structure of complex networks at different resolution levels. *New Journal of Physics*, *10*, 053039.

Arnemann, K. L., Chen, A. J. W., Novakovic-Agopian, T., Gratton, C., Nomura, E. M., & D'Esposito, M. (2015). Functional brain network modularity predicts response to cognitive training after brain injury. *Neurology*, *84*, 1568–1574.

Ball, G., Aljabar, P., Zebari, S., Tusor, N., Arichi, T., Merchant, N., Robinson, E. C., Ogundipe, E., Rueckert, D., Edwards, A. D., & Counsell, S. J. (2014). Rich-club organization of the newborn human brain. *Proceedings of the National Academy of Sciences*, *111*, 7456–7461.

Baniqued, P. L., Gallen, C. L., Voss, M. W., Burzynska, A. Z., Wong, C. N., Cooke, G. E., Duffy, K., Fanning, J., Ehlers, D. K., Salerno, E. A., Aguiñaga, S., McAuley, E., Kramer, A. F., & D'Esposito, M. (2018). Brain network modularity predicts exercise-related executive function gains in older adults. *Frontiers in Aging Neuroscience*, *9*, 426.

Bassett, D. S., & Bullmore, E. (2006). Small-world brain networks. *The Neuroscientist*, *12*, 512–523.

Bassett, D. S., Porter, M. A., Wymbs, N. F., Grafton, S. T., Carlson, J. M., & Mucha, P. J. (2013). Robust detection of dynamic community structure in networks. Chaos: An interdisciplinary. *Journal of Nonlinear Science*, *23*, 013142.

Bastos, A., & Schoffelen, J. (2016). A tutorial review of functional connectivity analysis methods and their interpretational pitfalls. *Frontiers in Systems Neuroscience*, *9*, 1–23.

Baum, G. L., Ciric, R., Roalf, D. R., Betzel, R. F., Moore, T. M., Shinohara, R. T., Kahn, A. E., Vandekar, S. N., Rupert, P. E., Quarmley, M., Cook, P. A., Elliott, M. A., Ruparel, K., Gur, R. E., Gur, R. C., Bassett, D. S., & Satterthwaite, T. D. (2017). Modular segregation of structural brain networks supports the development of executive function in youth. *Current Biology*, *27*, 1561–1572.

Belmonte, M. K., Allen, G., Beckel-Mitchener, A., Boulanger, L. M., Carper, R. A., & Webb, S. J. (2004). Autism and abnormal development of brain connectivity. *Journal of Neuroscience*, *24*, 9228–9231.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B*, *57*, 289–300.

Bertolero, M. A., Yeo, B., Bassett, D. S., & D'Esposito, M. (2018). A mechanistic model of connector hubs, modularity and cognition. *Nature Human Behaviour*, *2*, 765–777.

Bertolero, M. A., Yeo, B. T., & D'Esposito, M. (2015). The modular and integrative functional architecture of the human brain. *Proceedings of the National Academy of Sciences*, *112*, E6798–E6807.

Betzel, R. F., Bertolero, M. A., & Bassett, D. S. (2018). Non-assortative community structure in resting and task-evoked functional brain networks. bioRxiv, 355016 https://doi.org/10.1101/355016

Betzel, R. F., Byrge, L., He, Y., Goñi, J., Zuo, X. N., & Sporns, O. (2014). Changes in structural and functional connectivity among resting-state networks across the human lifespan. *NeuroImage*, *102*, 345–357.

Bielczyk, N., Fabian Walocha, F., Ebel, P., Haak, K., Llera, A., Buitelaar, J., Glennon, J., & Beckmann, C. (2018). Thresholding functional connectomes by means of mixture modeling. *NeuroImage*, *171*, 402–414.

Blondel, V. D., Guillaume, J. L., Lambiotte, R., & Lefebvre, E. (2008). Fast Unfolding of Communities in Large Networks. *Journal of Statistical Mechanics: Theory and Experiment*, *2008*, P10008.

Bordier, C., Nicolini, C., & Bifone, A. (2017). Graph analysis and modularity of brain functional connectivity networks: Searching for the optimal threshold. *Frontiers in Neuroscience*, *11*, 1–9.

Brandes, U. (2008). On variants of shortest-path betweenness centrality and their generic computation. *Social Networks*, *30*, 136–145. https://doi.org/10.1016/j.socnet.2007.11.001

Brier, M. R., Thomas, J. B., Fagan, A. M., Hassenstab, J., Holtzman, D. M., Benzinger, T. L., Morris, J. C., & Ances, B. M. (2014). Functional connectivity and graph theory in preclinical alzheimer's disease. *Neurobiology of Aging*, *35*, 757–768.

Brooks, S., Katz, E., & Stamoulis, C. (2022). Widespread positive direct and indirect effects of regular physical activity on the developing functional connectome in early adolescence. *Cereb Cortex Commun*, *3*, 1–19.

Brooks, S., Parks, S., & Stamoulis, C. (2021). Widespread positive direct and indirect effects of regular physical activity on the developing functional connectome in early adolescence. *Cerebral Cortex*, *31*, 4840–4852.

Brooks, S., Smith, C., & Stamoulis, C. (2023). Excess BMI in early adolescence adversely impacts maturating functional circuits supporting high-level cognition and their structural correlates. *International Journal of Obesity*, *47*, 590–605.

Bullmore, E., & Sporns, O. (2009). Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, *10*, 186–198.

Bullmore, E., & Sporns, O. (2012). The economy of brain network organization. *Nature Reviews Neuroscience*, *13*, 336–349.

Cao, M., Wang, J. H., Dai, Z. J., Cao, X. Y., Jiang, L. L., Fan, F. M., Song, X. W., Xia, M. R., Shu, N., Dong, Q., Milham, M. P., Castellanos, F. X., Zuo, X. N., & He, Y. (2014). Topological organization of the human brain functional connectome across the lifespan. *Developmental Cognitive Neuroscience*, *7*, 76–93.

Casey, B. J., Cannonier, T., Conley, M. I., Cohen, A. O., Barch, D. M., Heitzeg, M. M., Soules, M. E., Teslovich, T., Dellarco, D. V., Garavan, H., Orr, C. A., Wager, T. D., Banich, M. T., Speer, N. K., Sutherland, M. T., Riedel, M. C., Dick, A. S., Bjork, J. M., Thomas, K. M., … ABCD Imaging Acquisition Workgroup. (2018). The adolescent brain cognitive development (ABCD) study: Imaging acquisition across 21 sites. *Developmental Cognitive Neuroscience*, *32*, 43–54.

Chaddock-Heyman, L., Weng, T. B., Kienzler, C., Weisshappel, R., Drollette, E. S., Raine, L. B., Westfall, D. R., Kao, S. C., Baniqued, P., Castelli, D. M., Hillman, C. H., & Kramer, A. F. (2020). Brain network modularity predicts improvements in cognitive and scholastic performance in children involved in a physical activity intervention. *Frontiers in Human Neuroscience*, *14*, 346.

Chase, H. W., & Phillips, M. L. (2016). Elucidating neural network functional connectivity abnormalities in bipolar disorder: Toward a harmonized methodological approach. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *1*, 288–298.

Chen, M., & Deem, M. W. (2015). Development of modularity in the neural activity of children's brains. *Physical Biology*, *12*, 016009.

Chong, J. S. X., Ng, K. K., Tandi, J., Wang, C., Poh, J. H., Lo, J. C., Chee, M. W., & Zhou, J. H. (2019). Longitudinal changes in the cerebral cortex functional organization of healthy elderly. *Journal of Neuroscience*, *39*, 5534–5550.

Cohen, J. R., & D'Esposito, M. (2016). The segregation and integration of distinct brain networks and their relationship to cognition. *Journal of Neuroscience*, *36*, 12083–12094.

Crossley, N. A., Mechelli, A., Vértes, P. E., Winton-Brown, T. T., Patel, A. X., Ginestet, C. E., McGuire, P., & Bullmore, E. T. (2013). Cognitive relevance of the community structure of the human brain functional coactivation network. *Proceedings of the National Academy of Sciences*, *110*, 11583–11588.

Diedrichsen, J., Balsters, J. H., Flavell, J., Cussans, E., & Ramnani, N. (2009). A probabilistic mr atlas of the human cerebellum. *NeuroImage*, *46*, 39–46.

Dimitriadis, S., Messaritaki, E., & Jones, D. K. (2021). The impact of graph construction scheme and community detection algorithm on the repeatability of community and hub identification in structural brain networks. *Human Brain Mapping*, *42*, 4261–4280.

Fair, D., Cohen, A., Power, J., Dosenbach, N., Church, J., Miezin, F., Schlaggar, B., & Petersen, S. E. (2009). Functional brain networks develop from a "local to distributed" organization. *PLoS Computational Biology*, *5*, e1000381.

Faskowitz, J., & Sporns, O. (2020). Mapping the community structure of the rat cerebral cortex with weighted stochastic block modeling. *Brain Structure and Function*, *225*, 71–84. https://doi.org/10.1007/s00429-019-01984-9

Faust, K., & Wasserman, S. (1992). Blockmodels: Interpretation and evaluation. *Social Networks*, *14*, 5–61.

Fornito, A., Zalesky, A., & Breakspear, M. (2015). The connectomics of brain disorders. *Nature Reviews. Neuroscience*, *16*, 159–172.

Fransson, P., Schiffler, B. C., & Thompson, W. H. (2018). Brain network segregation and integration during an epoch-related working memory fMRI experiment. *NeuroImage*, *178*, 147–161.

Friston, K. J., Ashburner, J., Kiebel, S. J., Nichols, T., & Penny, W. (2007). *Statistical parametric mapping*. Academic Press.

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J. P., Frith, C. D., & Frackowiak, R. S. (1994). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, *2*, 189–210.

Funke, T., & Becker, T. (2019). Stochastic block models: A comparison of variants and inference methods. *PLoS One*, *14*, e0215296.

Gallen, C. L., & D'Esposito, M. (2019). Brain modularity: A biomarker of intervention-related plasticity. *Trends in Cognitive Sciences*, *23*, 293–304.

Garcia, J., Ashourvan, A., Muldoon, S., Vettel, J., & Bassett, D. (2018). Applications of community detection techniques to brain graphs: Algorithmic considerations and implications for neural function. *Proceedings of the IEEE*, *106*, 846–867.

Gates, A., Wood, I., Hetrick, W., & Ahn, Y. (2019). Element-centric clustering comparison unifies overlaps and hierarchy. *Scientific Reports*, *9*, 8574.

Geerligs, L., Renken, R. J., Saliasi, E., Maurits, N. M., & Lorist, M. M. (2015). A brain-wide study of age-related changes in functional connectivity. *Cerebral Cortex*, *25*, 1987–1999.

Göttlich, M., Münte, T. F., Heldmann, M., Kasten, M., Hagenah, J., & Krämer, U. M. (2013). Altered resting state brain networks in parkinson's disease. *PLoS One*, *8*, e77336.

Greicius, M., Krasnow, B., Reiss, A., & Menon, V. (2003). Functional connectivity in the resting brain: A network analysis of the default mode hypothesis. *Proceedings of the National Academy of Sciences of the United States of America*, *100*, 253–258.

Griffa, A., Baumann, P., Thiran, J., & Hagmann, P. (2013). Structural connectomics in brain diseases. *NeuroImage*, *80*, 515–526.

Guimera, R., & Sales-Pardo, M. (2009). Missing and spurious interactions and the reconstruction of complex networks. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 22073–22078. https://doi.org/10.1073/pnas.0908366106

Hagler, D. J., Hatton, S., Cornejo, M., Makowski, C., Fair, D., Dick, A., Sutherland, M., Casey, B., Barch, D., Harms, M., Watts, R., Bjork, J. M., Garavan, H. P., Hilmer, L., Pung, C. J., Sicat, C. S., Kuperman, J., Bartsch, H., Xue, F., … Dale, A. M. (2019). Image processing and analysis methods for the adolescent brain cognitive development study. *NeuroImage*, *202*, 116091.

Hamilton, R. S. (1982). Three-manifolds with positive ricci curvature. *Journal of Differential Geometry*, *17*, 255–306.

Harush, U., & Baruch, B. (2017). Dynamic patterns of information flow in complex networks. *Nature Communications*, *8*, 2181.

Holland, P. W., Laskey, K. B., & Leinhardt, S. (1983). Stochastic blockmodels: First steps. *Social Networks*, *5*, 109–137.

Huang, H., Shu, N., Mishra, V., Jeon, T., Chalak, L., Wang, Z. J., Rollins, N., Gong, G., Cheng, H., Peng, Y., Dong, Q., & He, Y. (2015). Development of human brain structural networks through infancy and childhood. *Cerebral Cortex*, *25*, 1389–1404.

Iordan, A. D., Cooke, K. A., Moored, K. D., Katz, B., Buschkuehl, M., Jaeggi, S. M., Jonides, J., Peltier, S. J., Polk, T. A., & Reuter-Lorenz, P. A. (2018). Aging and network properties: Stability over time and links with learning during working memory training. *Frontiers in Aging Neuroscience*, *9*, 419.

Ji, J. L., Spronk, M., Kulkarni, K., Repovš, G., Anticevic, A., & Cole, M. W. (2019). Mapping the human brain's cortical-subcortical functional network organization. *NeuroImage*, *185*, 35–57.

Jin, D., Huo, C., Liang, C., & Yang, L. (2021). Heterogeneous graph neural network via attribute completion. *Proceedings of WWW*, *1*, 391–400.

Karrer, B., & Newman, M. E. (2011). Stochastic blockmodels and community structure in networks. *Physical Review E*, *83*, 016107.

Khambati, A. N., Davis, K. A., Oommen, B. S., Chen, S. H., Lucas, T. H., Litt, B., & Bassett, D. S. (2015). Dynamic network drivers of seizure generation, propagation and termination in human epilepsy. *PLoS Computational Biology*, *11*, 1–19.

Kipf, T., & Welling, M. (2017). Semi-supervised classification with graph convolutional networks. *Proceedings of ICLR*, *1*, 1–14.

Konrad, K., & Eickhoff, S. (2010). Is the adhd brain wired differently? A review on structural and functional connectivity in attention deficit hyperactivity disorder. *Human Brain Mapping*, *31*, 904–916.

Koubiyr, I., Besson, P., Deloire, M., Charre-Morin, J., Saubusse, A., Tourdias, T., Brochet, B., & Ruet, A. (2019). Dynamic modular-level alterations of structural-functional coupling in clinically isolated syndrome. *Brain*, *142*, 3428–3439.

Laender, A. H. F., Leão, J. C., & Melo, P. O. S. V. D. (2020). Overcoming bias in community detection evaluation. *Journal of Information and Data Management*, *11*, 265–280. https://www.researchgate.net/publication/346098433

Lancichinetti, A., Fortunato, S., & Radicchi, F. (2008). Benchmark graphs for testing community detection algorithms. *Physical Review E*, *78*, 046110.

Lange, R., Benjamin, A., Haefner, R., & Pitkow, X. (2022). Interpolating between sampling and variational inference with infinite stochastic mixtures. *Proceedings of Machine Learning Research*, *180*, 1063–1073.

Lee, C., & Wilkinson, D. J. (2019). A review of stochastic block models and extensions for graph clustering. *Applied Network Science*, *4*, 1–50. https://doi.org/10.1007/s41109-019-0232-2

Leskovec, J., Lang, K. J., & Mahoney, M. W. (2010). Empirical comparison of algorithms for network community detection. In *Proceedings of the 19th international conference on world wide web* (pp. 631–640). Association for Computing Machinery. http://arxiv.org/abs/1004.3539

Li, K., Guo, L., Nie, J., Li, G., & Liu, T. (2009). Review of methods for functional brain connectivity detection using fMRI. *Computerized Medical Imaging and Graphics*, *22*, 131–139.

Lopez-Madrona, V., Matias, F., Mirasso, C., Canals, S., & Pereda, E. (2019). Inferring correlations associated to causal interactions in brain signals using autoregressive models. *Scientific Reports*, *9*, 17041.

Lorenz, D. M., Jeng, A., & Deem, M. W. (2011). The emergence of modularity in biological systems. *Physics of Life Reviews*, *8*, 129–160.

Luciana, M., Bjork, J., Nagel, B., Barch, D., Gonzalez, R., Nixon, S., & Banich, M. T. (2018). Adolescent neurocognitive development and impacts of substance use: Overview of the adolescent brain cognitive development (ABCD) baseline neurocognition battery. *Developmental Cognitive Neuroscience*, *32*, 67–79.

Ma, Q., Huang, B., Wang, J., Seger, C., Yang, W., Li, C., Wang, J., Feng, J., Weng, L., Jiang, W., & Huang, R. (2017). Altered modular organization of intrinsic brain functional networks in patients with Parkinson's disease. *Brain Imaging and Behavior*, *11*, 430–443.

Meunier, D., Achard, S., Morcom, A., & Bullmore, E. (2009). Age-related changes in modular organization of human brain functional networks. *NeuroImage*, *44*, 715–723.

Newman, M. (2016). Equivalence between modularity optimization and maximum likelihood methods for community detection. *Physical Review E*, *94*, 052315.

Newman, M. E. (2004). Fast algorithm for detecting community structure in networks. *Physical Review E*, *69*, 066133.

Ng, A. S. L., Wang, J., Ng, K. K., Chong, J. S. X., Qian, X., Lim, J. K. W., Tan, Y. J., Yong, A. C. W., Chander, R. J., Hameed, S., Ting, S. K. S., Kandiah, N., & Zhou, J. H. (2021). Distinct network topology in alzheimer's disease and behavioral variant frontotemporal dementia. *Alzheimer's Research & Therapy*, *13*, 1–16.

Ni, C. C., Lin, Y. Y., Gao, J., Gu, X. D., & Saucan, E. (2015). Ricci curvature of the internet topology. In *2015 IEEE conference on computer communications (INFOCOM)* (pp. 2758–2766). IEEE.

Ni, C. C., Lin, Y. Y., Luo, F., & Gao, J. (2019). Community detection on networks with ricci flow. *Scientific Reports*, *9*, 1–12.

Ollivier, Y. (2007). Ricci curvature of metric spaces. *Comptes Rendus Mathematique*, *345*, 643–646.

Onoda, K., & Yamaguchi, S. (2013). Small-worldness and modularity of the resting-state functional brain network decrease with aging. *Neuroscience Letters*, *556*, 104–108.

Orman, G., & Labatut, V. (2009). A comparison of community detection algorithms on artificial networks. *Scientific Reports*, *5805*, 242–256.

Pedersen, M., Omidvarnia, A. H., Walz, J. M., & Jackson, G. D. (2015). Increased segregation of brain networks in focal epilepsy: An fMRI graph theory finding. *NeuroImage: Clinical*, *8*, 536–542.

Peixoto, T. P. (2014a). The graph-tool python library. Figshare. http://figshare.com/articles/graph_tool/1164194 https://doi.org/10.6084/m9.figshare.1164194

Peixoto, T. P. (2014b). Hierarchical block structures and high-resolution model selection in large networks. *Physical Review X*, *4*, 011047.

Peixoto, T. P. (2016). Nonparametric bayesian inference of the microcanonical stochastic block model. *Physical Review E*, *95*, 012317.

Peng, D., Shi, F., Shen, T., Peng, Z., Zhang, C., Liu, X., Qiu, M., Liu, J., Jiang, K., Fang, Y., & Shen, D. (2014). Altered brain network modules induce helplessness in major depressive disorder. *Journal of Affective Disorders*, *168*, 21–29.

Power, J. D., Mitra, A., Laumann, T. O., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2014). Methods to detect, characterize, and remove motion artifact in resting state fMRI. *NeuroImage*, *84*, 320–341.

Pradhan, N., Dasgupta, S., & Sinha, S. (2011). Modular organization enhances the robustness of attractor network dynamics. *Europhysics Letters*, *94*, 38004.

Qian, X., Castellanos, F. X., Uddin, L. Q., Loo, B. R. Y., Liu, S., Koh, H. L., Poh, X. W. W., Fung, D., Guan, C., Lee, T. S., Lim, C. G., & Zhou, J. (2019). Large-scale brain functional network topology disruptions underlie symptom heterogeneity in children with attention-deficit/hyperactivity disorder. *NeuroImage: Clinical*, *21*, 101600.

Richardt, J., & Leone, M. (2008). (un)detectable cluster structure in sparse networks. *Physical Review Letters*, *101*, 1–4.

Rossini, P., di Iorio, R., Bentivoglio, M., Bertini, G., Ferreri, F., Gerloff, C., Ilmoniemi, R. J., Miraglia, F., Nitsche, M. A., Pestilli, F., Rosanova, M., Shirota, Y., Tesoriero, C., Ugawa, Y., Vecchio, F., Ziemann, U., & Hallett, M. (2019). Methods for analysis of brain connectivity: An IFCN-sponsored review. *Clinical Neurophysiology*, *130*, 1833–1858.

Rossval, M., & Bergstrom, C. (2008). Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 1118–1123.

Rudie, J. D., Brown, J., Beck-Pancer, D., Hernandez, L., Dennis, E., Thompson, P., Bookheimer, S., & Dapretto, M. (2013). Altered functional and structural brain network organization in autism. *NeuroImage: Clinical*, *2*, 79–94.

Schaefer, A., Kong, R., Gordon, E. M., Laumann, T. O., Zuo, X. N., Holmes, A. J., Eickhoff, S. B., & Yeo, B. T. (2018). Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity mri. *Cerebral Cortex*, *28*, 3095–3114.

Seitzman, B., Gratton, C., Marek, S., Raut, R., Dosenbach, N., Schlaggar, B., Petersen, S., & Greene, D. (2020). A set of functionally-defined brain regions with improved representation of the subcortex and cerebellum. *NeuroImage*, *206*, 116290.

Sia, J., Jonckheere, E., & Bogdan, P. (2019). Ollivier-ricci curvature-based method to community detection in complex networks. *Scientific Reports*, *9*, 9800.

Sia, J., Zhang, W., Jonckheere, E., Cook, D., & Bogdan, P. (2022). Inferring functional communities from partially observed biological networks exploiting geometric topology and side information. *Scientific Reports*, *12*, 9800.

Siegel, J. S., Power, J. D., Dubis, J. W., Vogel, A. C., Church, J. A., Schlaggar, B. L., & Petersen, S. E. (2014). Statistical improvements in functional magnetic resonance imaging analyses produced by censoring high-motion data points. *Human Brain Mapping*, *35*, 1981–1996.

Simon, H. (1962). The architecture of complexity. *Proceedings of the National Academy of Sciences*, *106*, 467–482.

Song, J., Birn, R. M., Boly, M., Meier, T. B., Nair, V. A., Meyerand, M. E., & Prabhakaran, V. (2014). Age-related reorganizational changes in modularity and functional connectivity of human brain networks. *Brain Connectivity*, *4*, 662–676.

Sporns, O., & Betzel, R. F. (2016). Modular brain networks. *Annual Review of Psychology*, *67*, 613–640.

Stephan, K., & Friston, K. (2010). Analyzing effective connectivity with functional magnetic resonance imaging. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*, 446–459.

Stevens, A. A., Tappon, S. C., Garg, A., & Fair, D. A. (2012). Functional brain network modularity captures inter-and intra-individual variation in working memory capacity. *PLoS One*, *7*, e30468.

Takeda, K., Matsuda, H., Miyamoto, Y., & Yamamoto, H. (2017). Structural brain network analysis of children with localization-related epilepsy. *Brain and Development*, *39*, 678–686.

Taya, F., de Souza, J., Thakor, N., & Bezerianos, A. (2016). Comparison method for community detection on brain networks from neuroimaging data. *Applied Network Science*, *1*, 1–20.

Tian, Y., Margulies, D. S., Breakspear, M., & Zalesky, A. (2020). Topographic organization of the human subcortex unveiled with functional connectivity gradients. *Nature Neuroscience*, *23*, 1421–1432.

Tooley, U., Basset, D., & Mackey, A. P. (2022). Functional brain network community structure in childhood: Unfinished territories and fuzzy boundaries. *NeuroImage*, *247*, 118843.

Torres, M., Colominas, M., Schlotthauer, G., & Flandrin, P. (2011). A complete ensemble empirical mode decomposition with adaptive noise. *ICASSP*, *1*, 4144–4147.

Traag, V. A., Waltman, L., & van Eck, N. J. (2019). From louvain to leiden: Guaranteeing well-connected communities. *Scientific Reports*, *9*, 1–12.

Vaessen, M., Braakman, H., Heerink, J., Jansen, J., Debeij-van Hall, M., Hofman, P., Aldenkamp, A., & Backes, W. (2013). Abnormal modular organization of functional networks in cognitively impaired children with frontal lobe epilepsy. *Cerebral Cortex*, *23*, 1997–2006.

van den Heuvel, M. P., & Sporns, O. (2013). Network hubs in the human brain. *Trends in Cognitive Sciences*, *17*, 683–696.

Venon, M. (2013). Developmental pathways to functional brain networks: Emerging principles. *Trends in Cognitive Sciences*, *17*, 627–638.

Wass, S. (2011). Distortions and disconnections: Disrupted brain connectivity in autism. *Brain and Cognition*, *75*, 18–28.

Weber, M., Stelzer, J., Saucan, E., Naitsat, A., Lohmann, G., & Jost, J. (2017). Curvature-based methods for brain network analysis. arXiv preprint arXiv:1707.00180.

Wen, X., Zhang, H., Li, G., Liu, M., Yin, W., Lin, W., Zhang, J., & Shen, D. (2019). First-year development of modules and hubs in infant brain functional networks. *NeuroImage*, *185*, 222–235.

Wu, Z., & Huang, N. (2009). Ensemble empirical mode decomposition: A noise-assisted data analysis method. *Advances in Adaptive Data Analysis*, *1*, 1–41.

Yan, X., Shalizi, C., Jensen, J. E., Krzakala, F., Moore, C., Zdeborová, L., Zhang, P., & Zhu, Y. (2014). Model selection for degree-corrected block models. *Journal of Statistical Mechanics: Theory and Experiment*, *2014*, P05007.

Yang, Y., Chu, L., Zhang, Y., Wang, Z., Pei, J., & Chen, E. (2018). Mining density contrast subgraphs. *Proceedings of ICDE*, *1*, 221–232.

Yang, Z., Algesheimer, R., & Tessone, C. (2016). A comparative analysis of community detection algorithms on artificial networks. *Scientific Reports*, *6*, 30750.

Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., Roffman, J. L., Smoller, J. W., Zöllei, L., Polimeni, J. R., Fischl, B., Liu, H., & Buckner, R. L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, *106*, 1125–1165.

Yuen, N. H., Osachoff, N., & Chen, J. J. (2019). Intrinsic frequencies of the resting-state fMRI signal: The frequency dependence of functional connectivity and the effect of mode mixing. *Frontiers in Neuroscience*, *900*, 1–17.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Brooks, S. J., Jones, V. O., Wang, H., Deng, C., Golding, S. G. H., Lim, J., Gao, J., Daoutidis, P., & Stamoulis, C. (2024). Community detection in the human connectome: Method types, differences and their impact on inference. *Human Brain Mapping*, *45*(5), e26669. https://doi.org/10.1002/hbm.26669