A Comparative Study of Brain Structural and Functional Connectivity: Graph Topology, Individual Fingerprinting, and Predictive Modeling

Yuhong Li

A thesis
submitted in partial fulfillment of the
requirements for the degree of

Master of Science

University of Washington
2025

Committee:
Eardi Lila
Ali Shojaie

Program Authorized to Offer Degree:
Biostatistics

University of Washington

Abstract

A Comparative Study of Brain Structural and Functional Connectivity: Graph Topology, Individual Fingerprinting, and Predictive Modeling

Yuhong Li

Co-Chairs of the Supervisory Committee:
Eardi Lila
Ali Shojaie
Biostatistics

Brain connectivity analyses using neuroimaging data provide insights into the structural and functional organization of the human brain. Several approaches have thus been proposed for modeling structural and functional connectivity, each with its own strengths and limitations. This thesis compares functional connectivity (FC) estimates derived based on correlation and partial correlation, evaluating their graph topology and performance in individual identification and prediction tasks, while also contrasting them with structural connectivity (SC) networks. We begin by estimating FC networks based on marginal and partial correlation and further explore low-order partial correlation graphs as an intermediate approach. Motivated by studies suggesting that brain structural hubs are closely related to functional activity, we also propose an alternative FC construction method by regressing out the temporal activity of SC hubs. Our analysis shows that SC emphasizes within-hemisphere connections and exhibits small-world properties, while FC consistently reveals strong interhemispheric connections, regardless of the methodology. However, other network properties vary depending on the estimation method used. Finally, we assess these networks' abilities to capture individual-specific features through subject identification and behavioral prediction tasks. We observe that partial correlation-based FC network performs well in subject identification when the sample size is large, but its performance deteriorates sharply with smaller samples. Additionally, regardless of the estimation method, FC networks consistently outperform SC in predicting behavioral variables, and combining both typically improves predictive accuracy, except for partial correlation-based FC. Sample size and number of scan sessions used in FC estimation also have a non-trivial impact on predictive performance. Our study highlights the methodological implications of FC estimation strategies for brain network analysis.

Acknowledgements

I want to express my deepest gratitude to my thesis advisors, Drs. Ali Shojaie and Eardi Lila, for their dedicated time and guidance throughout this thesis project.

Looking back on this two-year journey, I am particularly grateful to Eardi for introducing me to statistical methods in brain imaging analysis and for his generous support with the intricate details I encountered along the way. I appreciate his rigorous approach to research, which has left a meaningful impact on me.

I extend my sincere thanks to Dr. Kevin Lin for his direct mentorship and thoughtful support for both my academic and personal development.

I also appreciate Drs. Yiliu Wang and Mariano Gabitto at the Allen Institute for welcoming me into the fascinating scientific world and for the insightful discussions that enriched my understanding of statistical application in addressing real-world problems.

I would like to thank my peers, especially Jenna, Zhilong, and Tati, for their companionship and the shared experiences that made this journey more enjoyable and meaningful.

Finally, genuine thanks to my parents for their wholehearted support of my academic pursuits.

Contents

1 Introduction

Neuroimaging techniques have been widely used to study the brain's anatomical and functional organizations. Diffusion magnetic resonance imaging (dMRI) enables the estimation of neural white matter pathways between brain regions by tracing streamlines, which represent bundles of anatomical connections. In contrast, functional MRI (fMRI) captures neural activity by recording blood oxygen level-dependent (BOLD) signals over time as an indirect measure of neuronal activation. In particular, resting-state fMRI measures brain activity under resting conditions without external stimuli, capturing spontaneous fluctuations in blood oxygenation levels. Studies have shown that these signals exhibit organized patterns rather than random noise, reflecting the brain's intrinsic functional organizations [6]. Therefore, fMRI time-series are used to define a notion of functional connectivity, typically based on a measure of association between brain activity in different regions over time.

Network analysis provides a natural framework for modeling and understanding brain connectivity, where nodes are defined as spatial regions of interest (ROIs), while edges capture the interactions between them. Structural connectivity (SC) refers to the physical white matter pathways linking pairs of ROIs, with their strength most typically quantified by the count of streamlines between regions, as derived from dMRI. In contrast, functional connectivity (FC) typically refers to a measure of association between the time-series activities in ROI pairs. The more similar the temporal dynamics of two regions, the more likely they are functionally connected [18]. Analyzing these constructed connectivity networks through various summary statistics can reveal their topological properties and provide insights into the brain's organization.

Marginal correlations are commonly used to construct FC networks from fMRI time-series data. However, this approach cannot distinguish between direct and indirect associations. For example, if two nodes are each connected to a third node but not directly to one another, the marginal correlation between the two will still be high due to shared activity patterns. In contrast, partial correlations, which are computed as the correlation between two nodes after regressing out the activities in all other nodes, isolate direct relationships and provide a more accurate depiction of the true network structure [18, 15]. Despite this advantage, partial correlations become challenging to estimate in high-dimensional settings where the number of nodes exceeds the number of observations, which is a common scenario in fMRI studies. To address this, methods like graphical lasso impose sparsity assumptions on the network by applying an L1 penalty to the precision matrix, which is the inverse of the covariance matrix and encodes conditional dependencies between variables [16, 24]. Alterna-

tively, low-order partial correlation methods have been proposed for analyzing complex biological systems [23, 13, 17]. These methods estimate dependencies by conditioning on only a small subset of variables, rather than all variables simultaneously, making them computationally feasible for systems with limited sample sizes and high dimensionality. However, their use has not been as extensively explored in the context of FC estimation. The first contribution of this work is to adopt the idea of low-order conditional dependence to construct FC networks and compare them to networks derived from marginal and full partial correlations.

Furthermore, research has shown that the brain's structural network is organized around highly interconnected hub cores, which are regions serving critical roles for global information integration and efficient network communication [19, 10]. These SC hubs are thought to support FC [10, 3]. The second contribution of this work involves defining FC networks by specifically regressing out the activity of SC hubs, as opposed to regressing out the time series of all nodes, as is done in partial correlation methods. Specifically, we compute FC networks using SC hub-adjusted partial correlations and compare them to other notions of FC networks. To assess the influence of different FC estimation methods, we compare the resulting graphs with the graph derived from SC.

FC has increasingly been recognized as a unique brain fingerprint that captures the individual-specific features of connectivity patterns. Previous work has shown that these unique FC profiles can reliably distinguish one subject from another and predict cognitive and behavioral outcomes at the subject level [5, 21]. The third contribution of this work is to examine how methodological choices in FC estimation affect their ability to capture individual-specific connectivity signatures, both for behavioral variable prediction and for subject-level identification. Additionally, we explore how sample size impacts the robustness of FC estimation across different methods.

In the rest of this thesis, we first describe the data, then outline the methods used for estimating functional connectivity, including correlation-based and graphical lasso approaches, as well as strategies for constructing FC networks using magnitude-thresholded and significance-thresholded thresholding. Next, we present analyses of FC network edge selection variability, graph topology, subject identification, and behavioral prediction, comparing these FC estimates with each other and with SC. We conclude with a discussion of our results.

2 Methods

2.1 Data description

The structural and functional connectivity data used in this study are derived from the Human Connectome Project (HCP) [20], which provides high-quality neuroimaging data from a large cohort of healthy adults. Our analysis includes 957 subjects who have both structural connectivity information and complete recordings from all four resting-state fMRI sessions, which were obtained on two days and two on each day. Each fMRI session consists of 1200 time points (one every 0.72s) recording dynamic brain activity across time. In this work, brain regions are defined using the Desikan-Killiany 68-region parcellation scheme (34 regions per hemisphere), with consistent anatomical correspondence across subjects [4]. Given that this work focuses on the graph structure implied by SC and FC, we next describe the process of constructing a connectivity graph with non-trivial topological features from continuous measures of structural and functional connectivity.

2.2 Structural connectivity (SC) construction

To obtain representations of SC with a comparable number of edges to the FC, we fix the number of edges in each network to twice the number of regions, i.e., $2 \times 68 = 136$, for all methods. For structural connectivity (SC), binarizing based solely on zero or non-zero entries would lead to overly dense networks, as many entries are non-zero despite having small values. Instead, we construct SC networks using global thresholding with the fixed number of edges, as described below.

Given the pre-specified number of edges (136 edges), we will construct the SC networks through global thresholding, where we retain 136 edges, which correspond to the strongest connection between region pairs, and set to zero all other pairs. We first threshold each individual matrix, sum over them, and keep 136 edges that are selected most frequently across all the subjects.

2.3 Functional connectivity (FC) construction

To investigate the effect of sample size and autocorrelation in FC estimation, we will apply these methods on two versions of the fMRI data: (1) using all of 1200 time points, (2) using observations at every 6 time points, starting from the first one. Each subject in the subsample has 200 observations in our dataset.

We construct FC networks using two types of methods: correlation-based and graphical lasso-based approaches. While correlation-based networks can be represented in both unthresholded and thresholded versions, the graphical lasso-based approaches we use produce only thresholded binary net-

works. As with the SC thresholded networks, we retain 136 edges when constructing the binarized FC networks. We provide a detailed description of each method within these two categories.

2.3.1 Correlation-based methods

For each correlation-based method, we first compute the correlation matrices of dimension 68×68 for each subject and each fMRI session. At this stage, each subject has four fMRI session-specific FC correlation matrices.

To obtain a population-level FC representation, we first binarize each individual correlation matrix in two ways: (a) by thresholding on the magnitude of correlations to keep $2 \times 68 = 136$ edges, and (b) by applying a Fisher transformation and retaining edges significant at level $\frac{0.05}{2278}$ (Bonferroni corrected), where 2278 is the total number of possible edges in the network. Specifically, let r_{ij} denote the correlation between brain regions i and j and n_{tp} denote the number of time points in the resting-state fMRI time series. The Fisher transformation converts r_{ij} into a normally distributed variable

$$z_{ij} = \frac{1}{2}\log\left(\frac{1+r_{ij}}{1-r_{ij}}\right) \sim \mathcal{N}\left(\rho_{ij}, \frac{1}{n_{tp}-d-3}\right),$$

where ρ_{ij} is the true correlation, d is the number of variables being conditioned on, and n_{tp} is the number of observations, which takes either 1200 or 200 in our study.

To test $H_0 : \rho_{ij} = 0$ that there is no correlation of brain activities between region i and j, we compare the test statistic $z_{ij}/\frac{1}{\sqrt{n_{tp}-d-3}} \sim \mathcal{N}(0,1)$ and obtain the p-value. We will refer to (a) as the *magnitude-thresholded* network and (b) as the *significance-thresholded* network.

Using the binarized networks, we calculate the frequency of each edge being selected across the four session-specific FC networks for each subject. This results in subject-specific edge selection rate matrices, where each entry ranges from 0 to 4. The final population-level FC matrix is obtained by averaging these matrices entry-wise and retaining the 136 edges that are selected most frequently. Note that, due to ties in the selection rates, some networks may have fewer than 136 edges.

We next outline each of the correlation methods below.

Marginal correlation *(mcor)*

Let $\mathbf{X}$ denote the fMRI data matrix for one subject in a single scan session, with dimension $n_{tp} \times p$, where $p = 68$ is the number of brain regions. Let X_i denote the i-th column in $\mathbf{X}$, which represents the activities in region i. Suppose $\mathbf{X}$ is centered by subtracting the column-wise means, so that each column has zero mean. The sample covariance matrix is then defined as $\Sigma = \frac{1}{n_{tp}-1}\mathbf{X}^\top\mathbf{X}$, and the marginal correlation between region i and j is calculated as $r_{ij} = \frac{\Sigma_{ij}}{\sqrt{\Sigma_{ii}\Sigma_{jj}}}$, where Σ_{ij} represents the sample covariance between regions i and j. To test whether r_{ij} is significantly different from 0, we

compute the p-value using the test statistic $\frac{1}{2}\log\left(\frac{1+r_{ij}}{1-r_{ij}}\right)/\frac{1}{\sqrt{n_{tp}-3}}$. An edge is drawn if and only if $p_{ij} < \frac{0.05}{2278}$.

Partial correlation *(pcor)*

The partial correlation between regions i and j is calculated by $r_{ij} = -\frac{\Omega_{ij}}{\sqrt{\Omega_{ii}\Omega_{jj}}}$ where $\Omega = \Sigma^{-1}$ is the precision matrix and Ω_{ij} is its (i, j) entry. The p-value p_{ij} is obtained using the test statistic $\frac{1}{2}\log\left(\frac{1+r_{ij}}{1-r_{ij}}\right)/\frac{1}{\sqrt{n_{tp}-66-3}}$, since each r_{ij} conditions on the other 66 variables.

In our analysis, partial correlations are estimated using R package *ppcor* [11].

First-order partial correlation *(pcor_ord01)*

Following the idea of low-order partial correlation-based 0-1 graph in [23], the first-order partial correlation for a pair of regions (i, j) is computed as:

1. Compute the marginal correlation and its p-value using the Fisher transformation. If it is not significant, retain it as the correlation.

2. If the marginal correlation is significant, compute the first-order partial correlations by conditioning on each of the remaining 66 nodes individually (excluding i and j) and their corresponding p-values. This procedure gives 66 first-order partial correlations and p-values.

3. Compare these 66 p-values with the marginal correlation p-value and assign the correlation between i and j to the value corresponding to the largest p-value.

SC hub-adjusted partial correlation *(pcor_sc_hub01, pcor_sc_hub02)*

Hub nodes are regions that are structurally highly connected to other regions, serving as central pathways for information flow across the brain. Conditioning on these hub nodes allows us to account for their potential influence on correlations between other regions and helps eliminate indirect associations that arise from shared connections with these hubs.

Building on this idea, SC hub-adjusted partial correlation provides another low-order partial correlation estimate by conditioning on the hub node(s) from a subject's SC data matrix, where a hub node is defined as the brain region with the highest hub score, which will be formally introduced later. We will construct two such estimates: *pcor_sc_hub01* computes the first-order partial correlation by conditioning on one hub node, and *pcor_sc_hub02* computes the second-order partial correlation by conditioning on two hub nodes. The calculation follows these steps:

1. For a region pair (i, j), compute the marginal correlation and its p-value using the Fisher transformation. If the marginal correlation is not significant, retain it as the correlation estimate.

2. If the marginal correlation is significant: For *pcor_sc_hub01*, we partial out the highest-scoring hub node in the subject's unthresholded SC matrix from i, j's fMRI observations. For *pcor_sc_hub02*, we

partial out the two nodes with the highest hub scores from i, j's fMRI observations. If either i or j is a hub node, exclude it and select the next highest-scoring hub node.

3. The final correlation between (i, j) is chosen as either the marginal or low-order partial correlation with the larger p-value.

2.3.2 Graphical lasso-based methods

In addition to the methods described, we also used graphical lasso to estimate the connectivity graph by imposing sparsity constraints on the precision matrix. We use two variants of the graphical lasso method to estimate the FC networks using the R package *glasso* [8], namely the exact graphical lasso (*gl_aprxF*) and the approximated graphical lasso (*gl_aprxT*). Unlike the correlation-based networks, our usage of the graphical lasso methods does not produce significance-thresholded networks.

The method requires selecting a hyperparameter λ to control the sparsity penalty. For consistency with the correlation-based methods, we fix the number of edges at 136, and for each subject and each fMRI session, we choose λ such that the resulting graph contains exactly 136 non-zero entries in the upper triangular part of the precision matrix.

The estimates obtained from the graphical lasso are binarized, where an entry is set to 1 if it is non-zero, and 0 otherwise.

Exact graphical lasso *(gl_aprxF)*

The graphical lasso aims to solve the following optimization problem,

$$\hat{\Theta} = \underset{\Theta \succeq 0}{\arg\min} \left(tr(\Sigma\Theta) - \log det(\Theta) + \lambda \sum_{j \neq k} |\Theta_{jk}| \right),$$

where Σ is the sample covariance and λ is the penalizing parameter [24]. *gl_aprxF* computes the solution and binarizes it based on whether an entry is zero or nonzero to estimate the FC network.

Approximated graphical lasso *(gl_aprxT)*

This approach estimates a sparse graph by identifying the neighborhood of each variable through nodewise regression with a lasso penalty, where each variable is regressed on all the other variables. A regressor with a non-zero coefficient is considered a neighbor of the response variable [16].

This process can be expressed as follows. Suppose there are p variables $X_1, \ldots, X_p$ and let $\mathbf{X} \in \mathbb{R}^{n \times p}$ denote the matrix containing n independent observations of each X. In our analysis, X is the time series data matrix for one subject in a single session, n is the number of time points (1200 or 200), and

p is the number of brain regions. For each $X_k, k \in \{1, \ldots, p\}$,

$$\hat{\theta}^{k,\lambda} = \underset{\theta \in \mathbb{R}^p : \theta_k = 0}{\arg\min} \frac{1}{n} ||X_k - \mathbf{X}\theta||_2^2 + \lambda ||\theta||_1.$$

The neighborhood of X_k is defined as

$$\widehat{\text{ne}}_k^{\lambda} = \{l \in \{1, \ldots, p\} \setminus \{k\} : \hat{\theta}_l^{k,\lambda} \neq 0\}, \text{where } \theta_l^{k,\lambda} \text{ is the } l\text{-th element in } \theta.$$

While there is no guarantee that this procedure yields a symmetric neighborhood structure, which occurs when only one of $\hat{\theta}_l^{k,\lambda}, \hat{\theta}_k^{l,\lambda}$ equals 0 for some k, l. Therefore, we use an OR-rule to ensure the symmetry of our FC networks: if at least one of (i, j) or (j, i) is 1, we set both entries to 1.

2.4 Variability in FC networks

As previously mentioned, we construct the population-level FC network by selecting the most frequently chosen edges across fMRI sessions and subjects. Based on this approach, we can examine the variability in FC edge selection across different methods. Specifically, for each subject, we compute four FC networks corresponding to the four fMRI sessions and then sum these networks. This results in an aggregated FC network for each subject, where each entry (i, j) ranges from 0 to 4, indicating the number of sessions in which an edge between brain regions i and j is present. For each subject-level network, we count the frequency of each nonzero value and visualize the distribution of these frequencies across subjects using boxplots.

2.5 Graph structure analysis

To examine the topological organization of the thresholded SC and FC networks, we apply four graph-theoretic criteria. These will be formally defined later, but can be intuitively described as follows: (1) graph edit distance, which quantifies architectural dissimilarities between networks; (2) graph component analysis, which reveals global integration and fragmentation; (3) cliquishness and clustering, which measure the tendency for nodes to form triangles and the prevalence of fully interconnected subgroups; and (4) centrality measures, which identify nodes critical to global integration. These metrics offer valuable insight into the topological structure of SC and FC networks, highlighting both their shared properties and distinctive organizational principles.

To ensure clarity, we begin by defining relevant concepts and notations.

2.5.1 Definitions and notations

Let $G = (V, E)$ be an undirected graph, where V is the set of vertices and $E \subseteq V \times V$ is the set of edges. An *edge* $e_{ij} = (v_i, v_j) \in E$ is a connection between vertices $v_i, v_j \in V$. Note that since G is an

undirected graph, $e_{ij} = e_{ji}\ \forall e_{ij} \in E$.

A *path*, or a *simple path*, between two vertices $v_i, v_j \in V$ is a sequence of edges that join a set of distinct vertices, which can be written as $w = (e_{i,k_1}, e_{k_1,k_2}, ..., e_{k_{m-1},k_m}, e_{k_m,j})$ for some vertices $v_{k_1}, ..., v_{k_m}$. A *shortest path* between two vertices is the path that contains a minimal set of distinct vertices. The *length* of a path is the number of edges in it, also known as the distance between the two end vertices connected by the path.

A graph is *connected* if there exists at least one path between every pair of vertices within it. A *component* is a connected subgraph that is not part of any larger connected subgraph, which includes any isolated vertex which does not connect to any other vertex.

An undirected graph can be fully characterized by an *adjacency matrix* A, a symmetric binary matrix where every entry $A_{ij} = 1$ if and only if $e_{ij} \in E$ for $v_i, v_j \in V$, and 0 otherwise.

A *clique* in a graph is a subset of nodes in which every pair is directly connected by an edge. If a clique consists of k nodes, it is called a *k-clique*. In particular, a 3-clique is also known as a *closed triangle*.

2.5.2 Graph edit distance

We compare these networks using graph edit distance, a metric that quantifies the dissimilarity between graphs by measuring the minimum number of operations required to transform one graph into another. In the context of thresholded connectivity networks, which are binary and symmetric, the graph edit distance corresponds to the absolute difference between the upper triangular elements of two networks. A larger distance indicates more distinct graph structures, and a smaller distance suggests greater similarity.

2.5.3 Graph component analysis

A *component* in an undirected graph is a subset of vertices and edges where any two vertices are connected by at least one path, and no additional vertex in the graph is connected to the component. We summarize and compare the patterns of components for each of the SC and FC network, in terms of the size of the largest component, the number of non-trivial components, and the number of isolated vertices. Note that a non-trivial component is a component containing more than one node, and an isolated vertex is a node disconnected from all other nodes.

In addition, we generate graphs from Erdős-Rényi (ER) models $G(68, 136)$, which are random graphs with 68 nodes and 136 edges as in the SC and FC networks, with the edges uniformly distributed across all possible node pairs. We then use the statistics computed from ER graphs as a

baseline to evaluate SC and FC patterns, helping to highlight any non-random patterns or distinctive topological features present in the SC and FC networks.

2.5.4 Cliquishness and clustering

Cliquishness

Cliquishness quantifies the propensity of nodes to organize into such fully connected subgraphs, reflecting both the prevalence of these structures and the strength of local cohesion. For each of the SC and FC networks, we will analyze the size of the largest clique, the number of maximal cliques, and the number of 3-cliques. These results will again be compared to those obtained from Erdős-Rényi models $G(68, 136)$.

Global clustering coefficient

Clustering in graphs and networks describes the tendency of nodes to form local neighborhoods. In particular, the global clustering coefficient [22] quantifies the overall tendency of nodes in a network to cluster together. It measures how often the neighbors of a node are also connected to one another, reflecting the overall local cohesiveness of the network. For a graph $G = (V, E)$, the global clustering coefficient is defined as the ratio of the number of closed triangles in the network to the number of triplets. Note that a closed triangle is a set of three nodes that are mutually connected, and a triplet consists of three nodes connected by either two or three edges. A high global clustering coefficient indicates that the network has a strong tendency for nodes to form clusters, meaning that most neighbors of a node are interconnected, suggesting the presence of groups or communities within the network. In contrast, a low value suggests that nodes are less likely to form clusters, implying less local cohesiveness within the network.

2.5.5 Centrality measures

Graph centrality measures capture the topological importance of nodes within a network. We examine four centrality measures: node degree, betweenness centrality, harmonic centrality, and hub score. Each measure is described below and computed for every node in the network. We report the distributions of these node-level centrality metrics.

Node degree

Node degree of a node is the number of edges incident to a given node, or equivalently, the number of nodes directly connected to the given node. A larger degree indicates that the node is more connected to other nodes, while a smaller degree suggests that it is more isolated within the network.

Bewteenness

Betweenness centrality [7] quantifies how often a node lies on the shortest paths between other nodes in a network. For a graph $G = (V, E)$ and $v \in V$, the betweenness is defined as $\sum_{s,t \in V, s \neq v \neq t} \frac{\sigma(s,t|v)}{\sigma(s,t)}$, where $\sigma(s,t)$ is the total number of shortest paths between $s, t \in V$, and $\sigma(s,t \mid v)$ is the number of shortest paths between s, t which pass through v.

A higher value indicates that the node acts as a critical bridge, facilitating connections and controlling information flow, whereas a lower value suggests that the node is more peripheral and less involved in linking other nodes.

Harmonic centrality

Harmonic centrality [14] quantifies the proximity of a node to all other nodes in a network, by summing the reciprocals of individual shortest path lengths from the given node to all other nodes. Formally, for an undirected graph $G = (V, E)$ and a given node $v \in V$, it is defined as $\sum_{u \in V, u \neq v} \frac{1}{d(v,u)}$ where $d(u, v)$ is the length of the shortest path between u, v. If u, v are disconnected, $d(u, v) = \infty$, and so $\frac{1}{d(u,v)} = 0$. A higher harmonic centrality value for v indicates that, on average, it is closer to other nodes and has stronger connections; on the contrary, a lower harmonic centrality value suggests that the node is farther from the rest of the network.

Hub score

Kleinberg's hub score [12] is defined as the principal eigenvector of $AA^\top$ associated with the highest eigenvalue of $AA^\top$, where A is the adjacency matrix of the network. Hub score quantifies the extent to which a node links to other influential nodes. A higher value indicates that the node is more connected to other central nodes, and a lower score suggests that the node is less central in linking other nodes.

2.6 FC fingerprint analysis

Within-subject similarity

In addition to cross-subject similarity, we measure the similarity of FC within each subject across fMRI sessions, in order to assess the robustness of each method in capturing subject-specific connectivity, similar to the experiment in [5]. For each subject and every pair of fMRI sessions, we vectorize the two FC estimates (using the upper triangles) and compute Pearson's correlation as the within-subject similarity metric for the specific pair of scan sessions. Note that this value is specific to each subject and session pair. Higher similarity values reflect strong consistency in a subject's FC estimates across sessions, whereas lower values indicate greater variability in connectivity patterns

between sessions.

Cross-subject similarity

To assess the similarity of FC estimates across subjects, we compute a cross-subject similarity measure. We begin by vectorizing the upper-triangular parts of the FC correlation matrices obtained from four fMRI sessions for all subjects. Then, we compute the Pearson correlation for these vectorized FC estimates between different subjects, from either the same session or different sessions. This procedure yields a distribution of similarity values comparing FC estimates of different subjects. We perform this analysis for all FC methods, using both 1200 and 200 time points.

Note that higher similarity values indicate different subjects share common connectivity patterns, whereas lower values suggest that the FC method retains information unique to each subject.

Identification

To conduct the identification task, we begin by selecting a reference FC estimate for each subject from a given fMRI session s_{ref}. For each reference, we identify the most similar FC estimate computed from a target session s_{tar}, where similarity is defined as the Pearson correlation between the vectorized FC estimates. An identification is deemed successful if the most similar FC from s_{tar} belongs to the same subject as the reference FC. This process is repeated for every subject in s_{ref}, and the overall success rate is calculated. Note that switching the role of reference and target session leads to different identification results, so there will be 12 trials of the identification task. We apply this analysis for every FC method, using both 1200 and 200 time points.

In addition, we perform a permutation test in which the subject labels in s_{tar} are shuffled to create a false correspondence between s_{ref} and s_{tar}. By repeating the identification procedure on this permuted data, we generate a null distribution that serves as a baseline for evaluating the significance of the observed success rate. We use each pair of sessions to perform 200 permuted identification runs, which leads to 1200 runs in total.

2.7 Predictive models

Outcomes

We use 19 behavioral outcomes measuring cognition, personality, and emotion. These outcome variables include: episodic memory using Picture Sequence Memory Test (*PicSeq_Unadj*), executive function using Dimensional Change Card Sort Test (*CardSort_Unadj*), attention and inhibitory control using the Flanker task (*Flanker_Unadj*), fluid intelligence using Raven's Progressive Matrices (*PMAT24_A_CR*), language and reading decoding using Oral Reading Recognition test (*Read-*

Eng_Unadj), language and vocabulary comprehension using Picture Vocabulary test (*PicVocab_Unadj*), processing speed using Pattern Completion Processing Speed Test (*ProcSpeed_Unadj*), self-regulation using Delay Discounting (*DDisc_AUC_40K*), spatial orientation using Variable Short Penn Line Orientation Test (*VSPLOT_TC*), sustained attention using Short Penn Continuous Performance Test (*SCPT_SPEC*), working memory using List Sorting test (*ListSort_Unadj*), sub-maximal cardiovascular endurance using 2-min walk endurance test (*Endurance_Unadj*), the grip strength test (*Strength_Unadj*), agreeableness, openness, and conscientiousness in Five Factor Model (*NEOFAC_A*, *NEOFAC_C*, *NEOFAC_O*), Pittsburgh sleep quality total scores (*PSQI_Score*), anger measured by Anger-Affect Survey (*AngAggr_Unadj*), life satisfaction measured by NIH Toolbox Meaning and Purpose Survey (*LifeSatisf_Unadj*).

Predictors

We use the FC correlations entrywise averaged across the four sessions as the FC predictor. For the SC predictor, we perform a standardization by dividing each subject's SC matrix by its maximum value. Furthermore, we combine both FC and SC in the ridge regression model to compare the performance of this integrated approach against models that use only SC or FC. We also examine the effect of data sample size by performing the predictions using both 1200-time-point FC and 200-time-point FC, and then comparing their performance.

Model and assessment

We fit ridge regression models for each outcome variable, using FC, SC predictors both separately and jointly. The model performance is evaluated by out-of-sample R^2 scores, obtained through a nested cross-validation procedure, as described below.

The nested cross-validation framework

The nested cross-validation framework consists of outer and inner loops. The outer loop first splits the entire dataset into a training and testing set, and the inner loops perform model training and validation using cross-validation on the training set in the outer loop. The best model fitted in the inner loop is applied to the outer loop testing set and computes the test set R^2. This procedure yields 10 test R^2 for each outcome and each step is described below.

1. Divide all the subjects ($N = 957$) into 10 folds of roughly equal size in the outer loop (approximately 96 per fold). Conduct 10 iterations, with each iteration using one fold as the test set in the outer loop, and the remaining folds are used for the inner loop.

2. In each of the 10 outer loop iterations, combine the remaining folds and further divide them into 10 folds for the inner loop. In each inner loop iteration, use 9 of the folds to train a ridge regression

model and compute the validation R^2 on the remaining fold.

3. The best model selected by the inner loop is then applied to the outer loop test set. A test set R^2 is computed.

4. After 10 outer loop iterations, we obtain a set of 10 test R^2's. We compare the mean test set R^2 for different FC methods.

This procedure is implemented using R package *glmnet* [9]. We report the mean out-of-sample R^2 for each outcome.

3 Results

3.1 Connectivity patterns

Unthresholded estimates

Figure 1 displays population-averaged unthresholded SC and FC matrices derived from 1200 time points. The SC reflects anatomical connections averaged across subjects, while FC represents averaged statistical dependencies in activity. All matrices share a consistent color scale (-1 to 1).

SC emphasizes a larger number of within-hemispheric anatomical connections, with a weaker cross-hemispheric organization. In contrast, all FC estimates reveal a band-like pattern, reflecting strong cross-hemispheric interactions between regions in correspondence, which are anatomically corresponding areas in the left and right hemispheres. While marginal correlation (*mcor*) and SC hub-adjusted partial correlation (*pcor_sc_hub*) exhibit stronger magnitudes, likely capturing indirect associations, partial correlation (*pcor*) and first-order partial correlation (*pcor_ord01*) show weaker magnitudes and a sparser connectivity pattern, suggesting they reflect more direct interactions after accounting for nodes with similar activity.

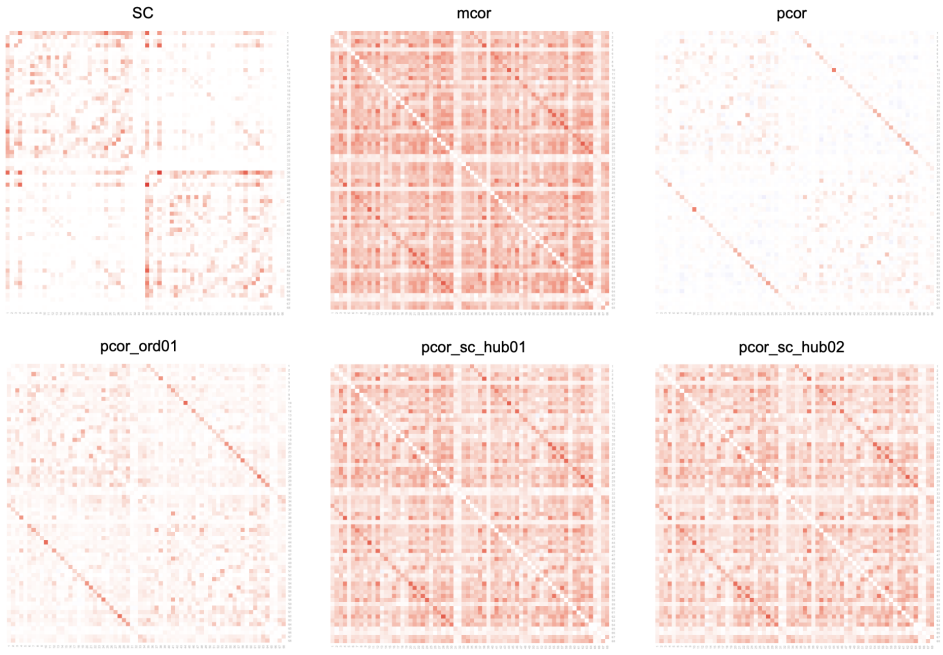

Figure 1: Population-level unthresholded SC and FC networks. These networks are obtained by averaging each subject's unthresholded SC and FC matrices. The FC networks contain mean correlations ranging from -1 to 1, computed from 1200 time points.

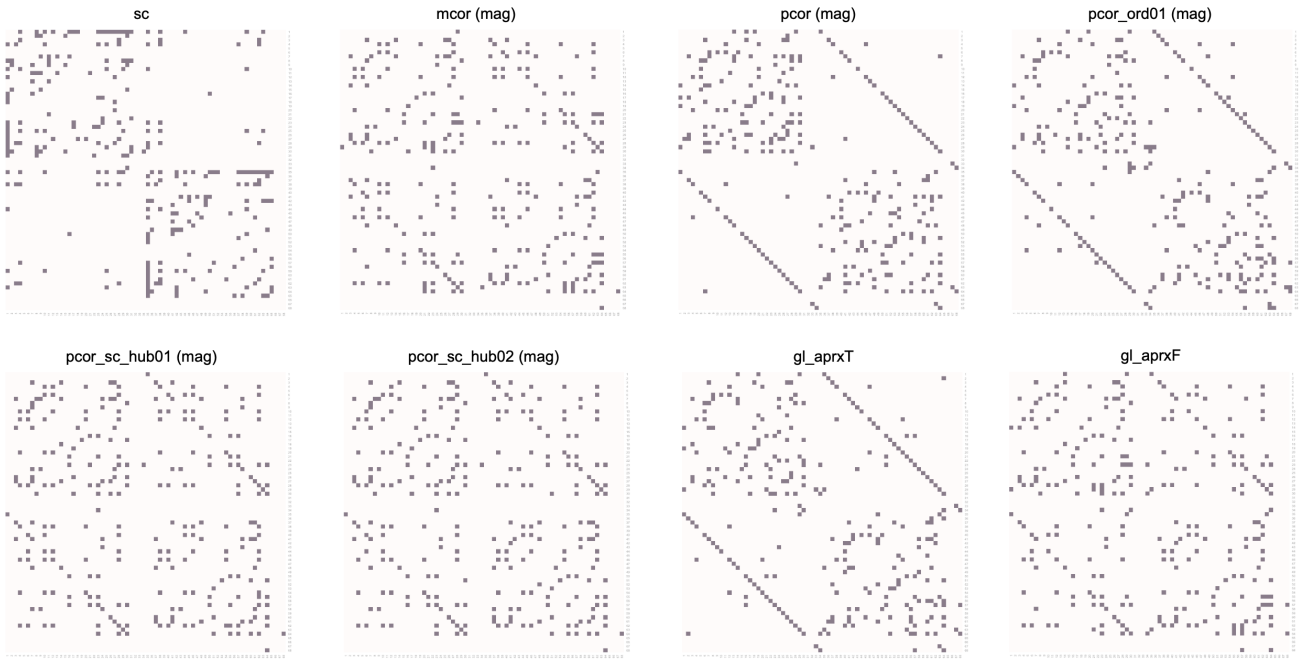

Figure 2: Population-level thresholded SC and FC magnitude networks. These networks are obtained by selecting the 136 edges most frequently retained across individual magnitude-thresholded networks.

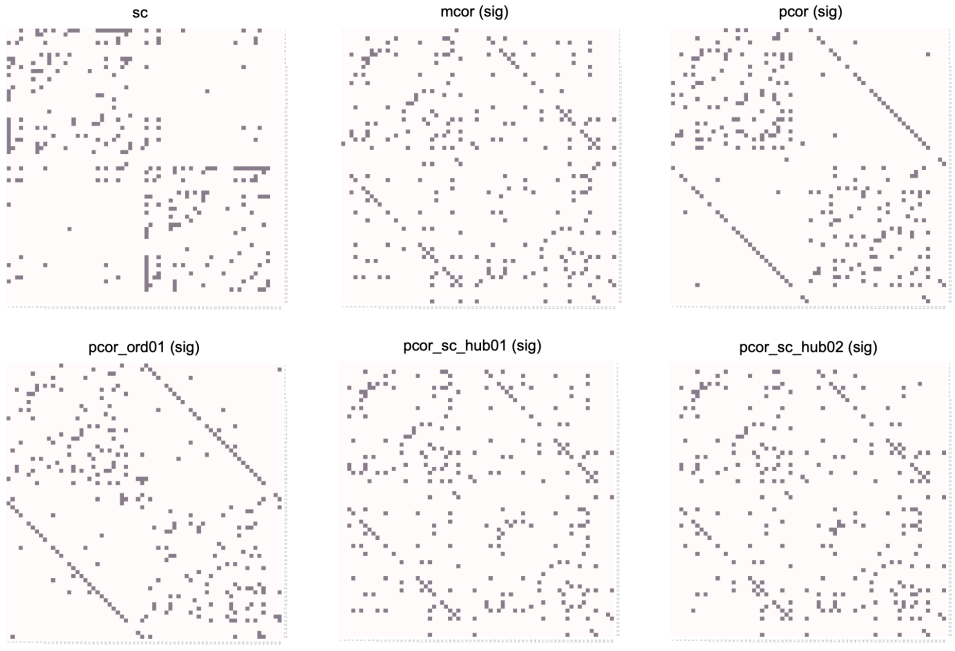

Figure 3: Population-level thresholded SC and FC significance networks. These networks are obtained by selecting the 136 edges most frequently retained across individual significance-thresholded networks. The SC network exhibits stronger within-hemispheric connections but sparser cross-hemispheric ones, whereas FC networks generally display a more diffuse connectivity pattern with a clear diagonal band-like structure.

Thresholded networks

Figure 2 displays population-level thresholded SC and FC magnitude networks with top 136 edges. Similar to the unthresholded images, the SC network demonstrates within-hemisphere anatomical connectivity, with sparse cross-hemispheric links. Thresholded FC networks retain diagonal band-like structures, reflecting strong connections between left- and right-hemispheric corresponding regions. Marginal correlation and SC hub-adjusted partial correlation FC graphs exhibit diffuse connectivity, likely capturing indirect associations, whereas partial correlation and first-order methods concentrate edges within hemispheres, isolating direct interactions. Approximated graphical lasso (*gl_aprxT*) mirrors partial correlation's pattern, while the exact method (*gl_aprxF*) aligns with marginal correlation's distributed topology. These trends persist in significance-thresholded FC networks, as shown in Figure 3.

Functional communities

Figure 4 displays SC and FC networks with nodes grouped by functional communities. While the SC network exhibits sparse within-community connectivity, all FC methods reveal denser within- than between-community connections, highlighting their capacity to capture functionally synchronized groups. This distinction between SC and FC networks is expected, as the communities are defined based on functional activity within each region. Among the FC methods, marginal correlation, SC hub-adjusted partial correlation, and exact graphical lasso show particularly strong within-community connectivity, likely due to their retention of indirect associations or hub-adjusted synchronization, where partial correlation and approximated graphical lasso produce relatively sparser within-group connectivity. We will examine more quantitative aspects of the networks in later sections.

The graph patterns derived from 200 time points closely resemble their 1200-time-point counterparts (see edit distance summary in the appendix), suggesting that sample size and autocorrelation do not substantially affect the resulting network structure.

3.2 Variability in FC networks

Figure 5 and 6 illustrate the edge selection stability of thresholded FC networks across four fMRI sessions. For each method, the x-axis represents the frequency at which an edge is selected across sessions (1–4 sessions, excluding edges never selected). The y-axis shows the proportion of edges selected at each frequency relative to the total possible edges per network (2278). Each boxplot summarizes the distribution of these proportions across 957 subjects. Specifically, a single data point in a

boxplot corresponds to the proportion of edges selected at a given frequency for one subject, and the boxplot aggregates these proportions across all 957 subjects. We discuss magnitude-thresholded and significance-thresholded FC networks separately below.

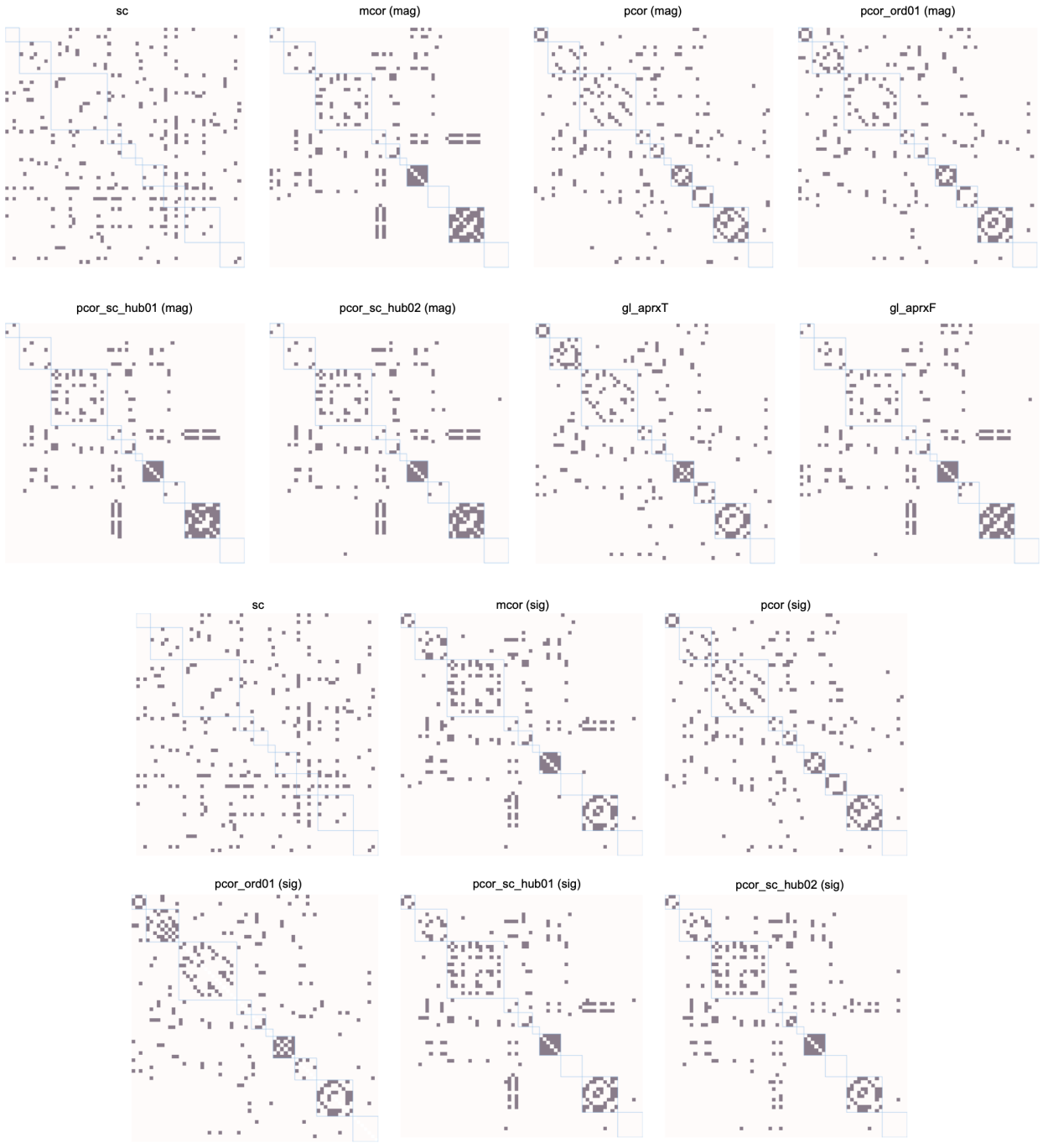

Figure 4: Population-level thresholded SC and FC networks, with nodes grouped by functional ROI communities. These networks are the same as in Figures 2,3, except that the nodes are grouped by functional communities.

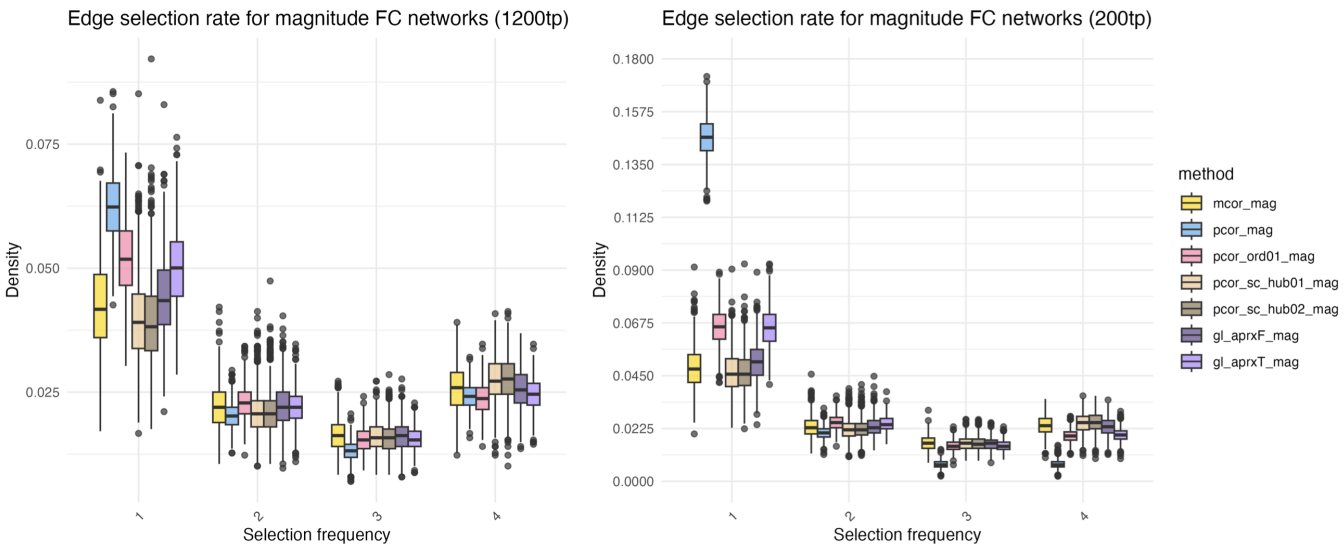

Figure 5: Edge selection rate for FC magnitude networks.

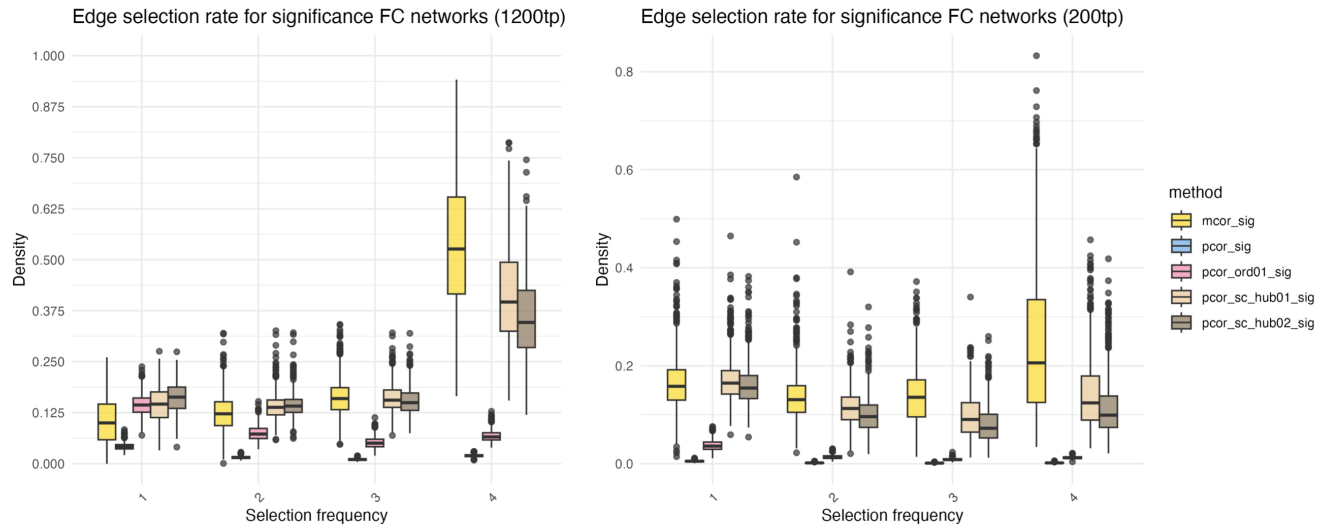

Figure 6: Edge selection rate for FC significance networks. The x-axis represents the frequency with which an edge is selected for a single subject, ranging from 1 to 4, and the y-axis shows the proportion of edges selected at each frequency. Partial correlation-based FC networks generally exhibit greater variability and increased sensitivity to sample size.

Magnitude-thresholded FC networks

From Figure 5, compared to other methods, partial correlation (*pcor_mag*) exhibits lower stability, with a higher proportion of edges selected only once (frequency=1) and fewer edges selected consistently (frequency=4). This pattern is exacerbated when using 200 rather than 1200 time points to construct the networks. First-order partial correlation (*pcor_ord01_mag*) and approximated graphical lasso (*gl_aprxT_mag*) show intermediate level of stability, likely due to a reduced number of nodes

in conditioning or the regularization in estimation. In contrast, marginal correlation (*mcor_mag*), exact graphical lasso (*gl_aprxF_mag*), and SC hub-adjusted methods (*pcor_sc_hub01/02_mag*) demonstrate higher consistency, with SC hub methods showing slightly elevated proportions at frequency = 4.

Significance-thresholded FC networks

In Figure 6, which illustrates edge selection stability for significance-thresholded FC networks, edges are retained based on statistical significance rather than magnitude, leading to different network densities across sessions and subjects. Graphical lasso methods are excluded, as the implementation we use generates magnitude-thresholded networks without significance testing.

Marginal correlation (*mcor*) and SC hub-adjusted methods (*pcor_sc_hub01/02*) consistently show higher edge selection proportions for subjects across all four frequencies compared to partial correlation (*pcor*) and first-order partial correlation (*pcor_ord01*). This occurs because marginal correlation captures both direct and indirect associations and inflates edge counts. Additionally, SC hub-adjusted methods condition on structurally hub nodes, which may not be directly involved in functional interactions and are insufficient to eliminate indirect associations.

Notably, in SC hub-adjusted methods, the first-order network (*pcor_sc_hub01*) selects more edges across all four sessions compared to the second-order network (*pcor_sc_hub02*). While both methods exhibit similar edge density patterns in magnitude-thresholded networks, their differences are amplified in significance-thresholded networks. This discrepancy arises possibly because the correlation magnitude in the two methods is comparable, but their statistical significance differs due to different variance of the Fisher-transformed correlations. Specifically, by conditioning on one extra node than *pcor_sc_hub01*, the variance of correlations in *pcor_sc_hub02* increases, leading to larger p-values and fewer edges surviving significance thresholds.

In addition, first-order partial correlation (*pcor_ord01*) exhibits similar behavior to full partial correlation (*pcor*), retaining fewer edges than *pcor_sc_hub01*, despite both conditioning on a single node. This difference likely stems from their distinct conditioning strategies: *pcor_ord01* selects the node that minimizes the residual correlation between two target regions, while *pcor_sc_hub01* conditions on structurally defined hub nodes, which may be less effective at capturing direct functional connectivity.

Comparing results from 1200 time points to 200 time points, all methods show a reduction in edge selection rates, with the effect being more pronounced in the significance-thresholded networks. This reduction reflects the increased variability in estimates due to fewer observations, regardless of whether conditioning on additional variables or not.

3.3 Connectivity network structures

3.3.1 Graph edit distance

When comparing the edit distance across population-level SC and FC networks derived from 1200 time points, we observe that SC exhibits substantial discrepancies with all FC networks (all greater than 240, with the largest possible value of 272), reflecting its anatomical basis in within-hemisphere connections, in contrast to FC's stronger cross-hemispheric interactions. Among the FC networks, partial correlation (*pcor*) and marginal correlation (*mcor*) show the largest edit distances (around 120 edges), reflecting their different target estimates, where *pcor* removes indirect paths through conditioning on other variables, whereas *mcor* retains both direct and indirect correlations. Notably,

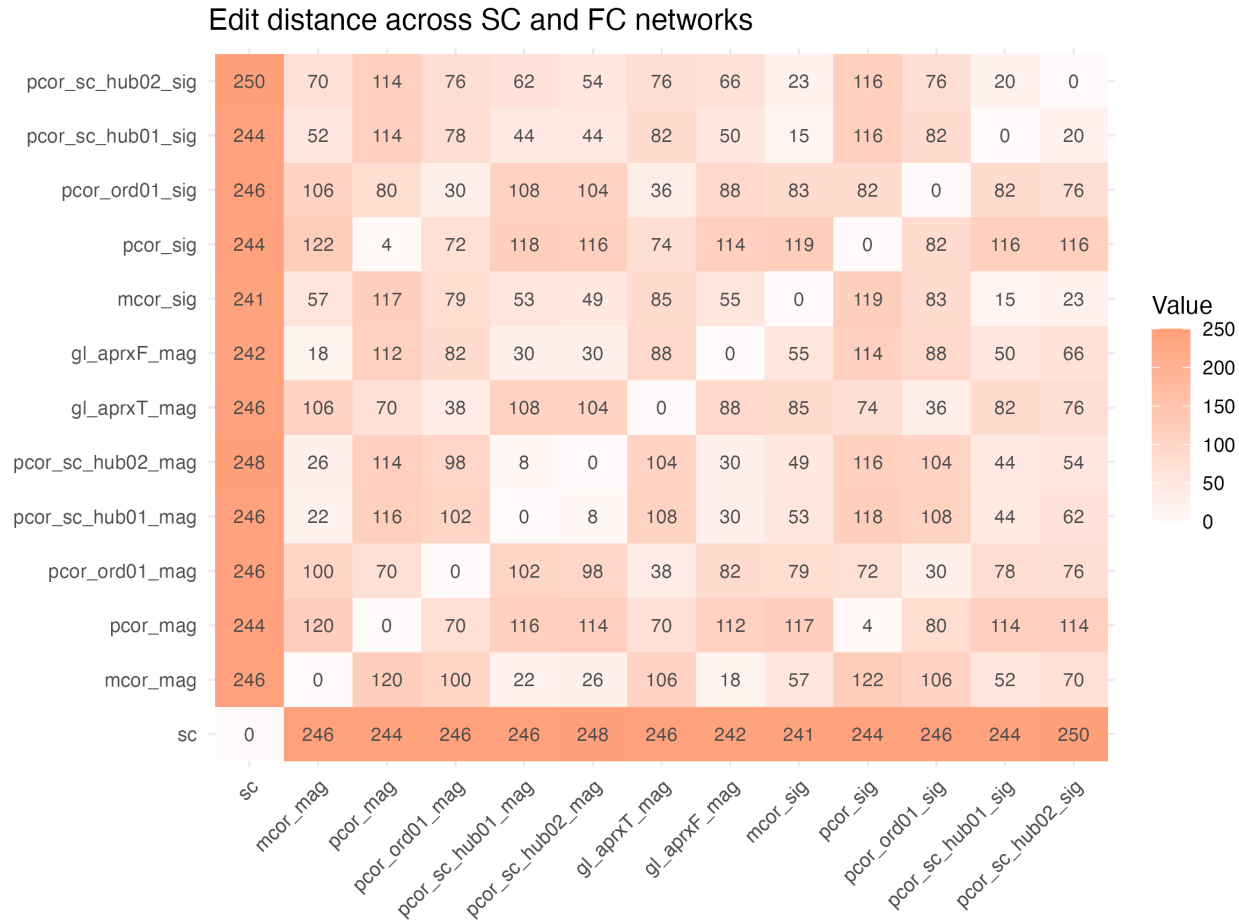

Figure 7: Graph edit distance across SC and FC networks. The graph edit distance for two binary and symmetric connectivity networks is the absolute difference between their corresponding upper triangular elements. Edit distances between SC and FC networks are larger than those between FC networks. Moreover, *mcor* exhibits a high edit distance from *pcor*, whereas its edit distance to *gl_aprxF_mag* is relatively low.

mcor and SC hub-adjusted partial correlation (*pcor_sc_hub*) networks resemble each other more closely (ranging from 15 to 26), especially when applying the same thresholding metric. This similarity likely results from both approaches preserving certain indirect associations, either due to the lack of adjustment in *mcor* or partial adjustment limited to structural hubs in *pcor_sc_hub*. Meanwhile, exact graphical lasso (*gl_aprxF*) networks show smaller edit distances to *mcor* (18 with *mcor_mag* and 55 with *mcor_sig*), aligning with their tendency to maintain distributed connectivity patterns. Finally, despite visual similarities in strong within-hemispheric connections and cross-hemispheric connections between corresponding regions, approximated graphical lasso (*gl_aprxT*) and *pcor* networks still exhibit moderate edit distances (70 with *pcor_mag* and 74 with *pcor_sig*), underscoring subtle differences in their within-hemisphere connection structure and their edge-selection processes. Networks derived from both 1200 and 200 time points using the same connectivity estimation method show a high degree of structural similarity, as demonstrated in the appendix, suggesting their robustness to variations in sample size and autocorrelation.

3.3.2 Graph component analysis

This section shows the results of graph component analysis on the thresholded SC and FC networks (using 1200 time points), including the largest component size, number of non-trivial components, and number of isolated nodes, shown in Figure 8.

We find that all SC and FC methods exhibit a dominant connected component with sizes exceeding 40 nodes, indicating a centralized large component alongside smaller clusters. Notably, partial correlation networks (*pcor_mag*, *pcor_sig*) and approximated graphical lasso (*gl_aprxT*) show particularly large components (around 60 nodes). SC networks, while slightly smaller in maximum component size, still highlight a centralized architecture. In contrast, *pcor_ord01_mag* uniquely displays up to four non-trivial components, while most methods (*sc*, *mcor_mag*, *gl_aprxF*, *pcor_ord01_sig*) feature only one non-trivial component, paired with numerous isolated nodes. These patterns persist when using 200 time points, as shown in the appendix.

To provide a baseline, we generate 1000 ER random graphs of the same node and edge counts as in SC and FC networks, and compute the same summary statistics for them. The random networks exhibit larger maximum component sizes (peaking at 67–68 nodes, with 381/1000 and 338/1000 frequencies respectively, surpassing all SC and FC networks), fewer isolated nodes (0 or 1 in 740/1000 cases), and comparable numbers of non-trivial components (1-2), except for *pcor_ord01_mag* and SC hub-adjusted methods, which shows more fragmented clusters.

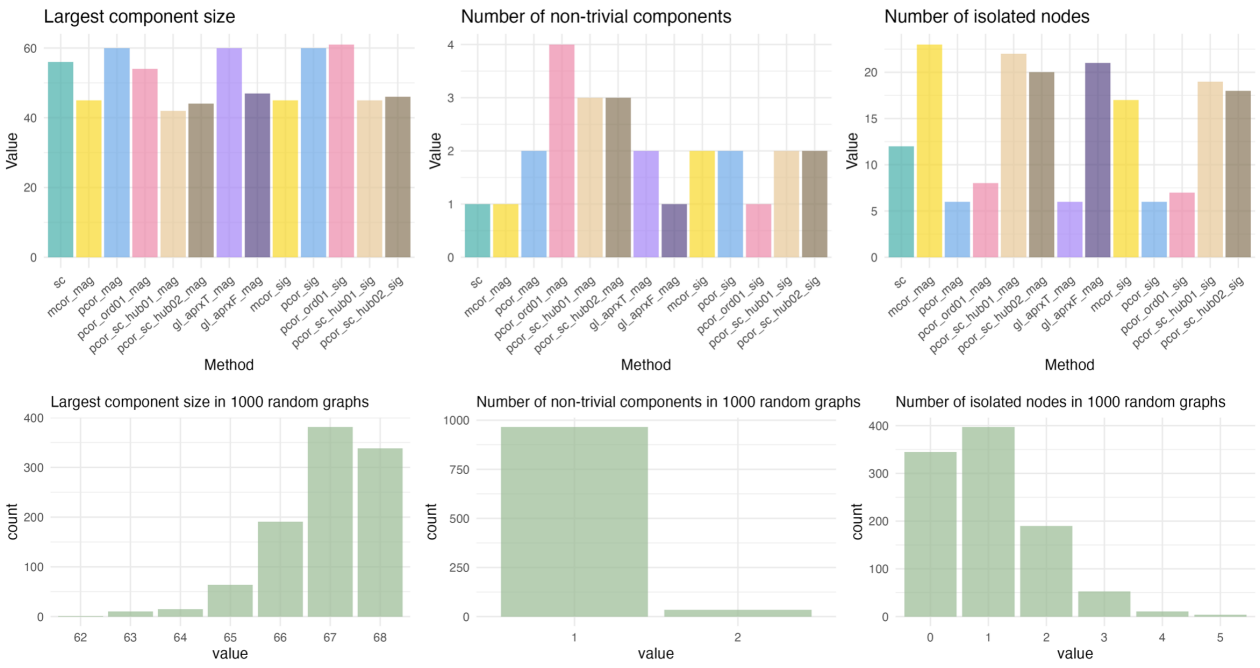

Figure 8: Summary of graph component analysis. We compare the size of the largest connected component, the number of non-trivial components, and the number of isolated nodes across SC and FC networks, as well as in the ER random graphs. While different SC and FC networks exhibit variable patterns, they overall differ from the random graphs by having a smaller largest component, more non-trivial components, and a greater number of isolated nodes.

These distinct graph component patterns reveal fundamental differences in connectivity between random and biological networks. ER random graphs, of the specific size and density, typically form a large connected component due to uniform edge connection probability, whereas SC and FC networks under the threshold of a limited number of edges highlight strong, biologically meaningful connections, at the expense of isolating weaker nodes.

3.3.3 Cliquishness and clustering

Figure 9 summarizes the cliquishness and clustering properties of SC and FC networks (using 1200 time points), including the size of the largest clique, the number of 3-cliques, and the global clustering coefficient. For reference, the same statistics are computed for 1000 ER random graphs.

All SC and FC (using 1200 time points) networks have the largest clique size between 4 and 7. Specifically, *pcor*, *pcor_ord01_sig*, and *gl_aprxT* yield the smallest cliques and fewest 3-cliques, likely reflecting the effectiveness in adjusting for common neighbors and indirect correlations between brain regions. In contrast, *mcor*, *gl_aprxF*, and *pcor_sc_hub* retain more 3-cliques and larger cliques, preserving many strongly connected clusters with indirect associations. SC occupies an intermediate

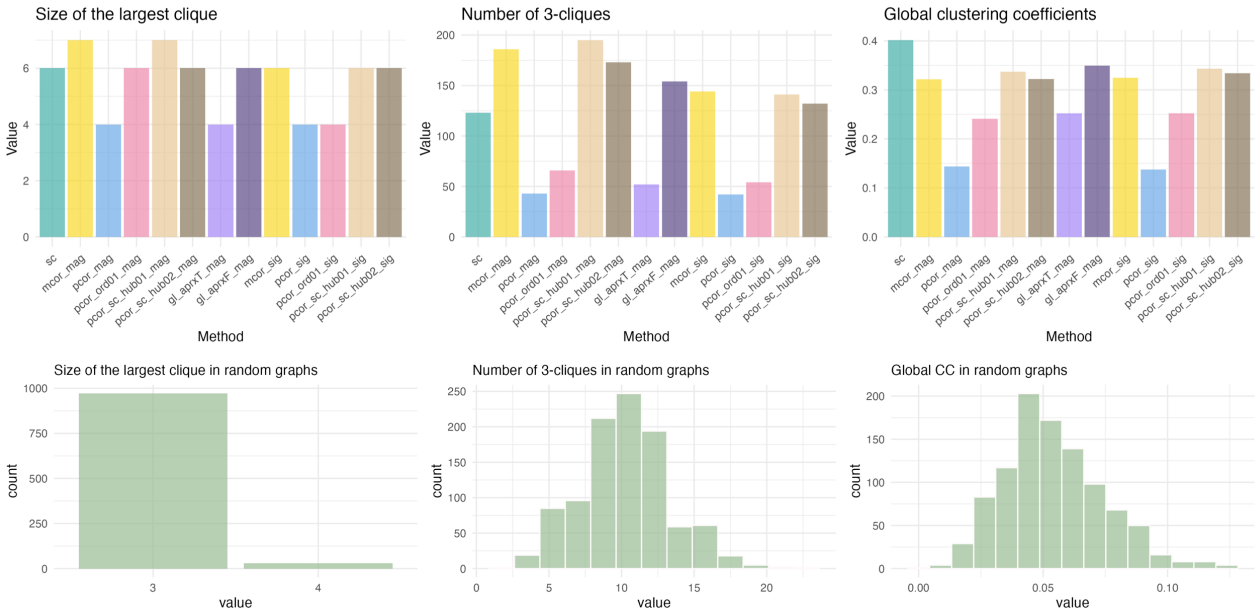

Figure 9: Summary of cliquishness and clustering. We compare the size of the largest clique, the number of 3-cliques, and global clustering coefficient across SC and FC networks, as well as in the ER random graphs. While different SC and FC networks exhibit variable patterns, they overall differ from the random graphs by having a larger size of the largest clique, more 3-cliques, and a larger global clustering coefficient.

position in clique size but exhibits the highest global clustering coefficient (0.402), highlighting the densely interconnected nature of anatomical structures; *pcor* has the lowest measure of clustering.

Random graphs, on the other hand, show smaller cliques (largest clique size $\leq$3 in 95% of 1000 cases), fewer 3-cliques ($\leq$22 vs $\geq$50 in SC/FC networks), and reduced clustering (max CC 0.128), when compared to all SC and FC networks. These findings highlight that, despite differences in cliquishness and clustering across SC and FC networks, they all show a stronger tendency to form densely connected local neighborhoods.

3.3.4 Centrality measures

Figure 10 shows the distributions of node-level centrality measures, including node degree, betweenness centrality, harmonic centrality, and hub score. Note that isolated nodes have zero values in their degrees, betweenness, and harmonic centrality measures.

Compared to FC networks, SC exhibits a wider range and smaller subsets of nodes with exceptionally high values in degree, betweenness centrality, and harmonic centrality. This pattern suggests the presence of strongly connected anatomical hubs, which serve as bridges between disparate regions and reduce the average path lengths across the network.

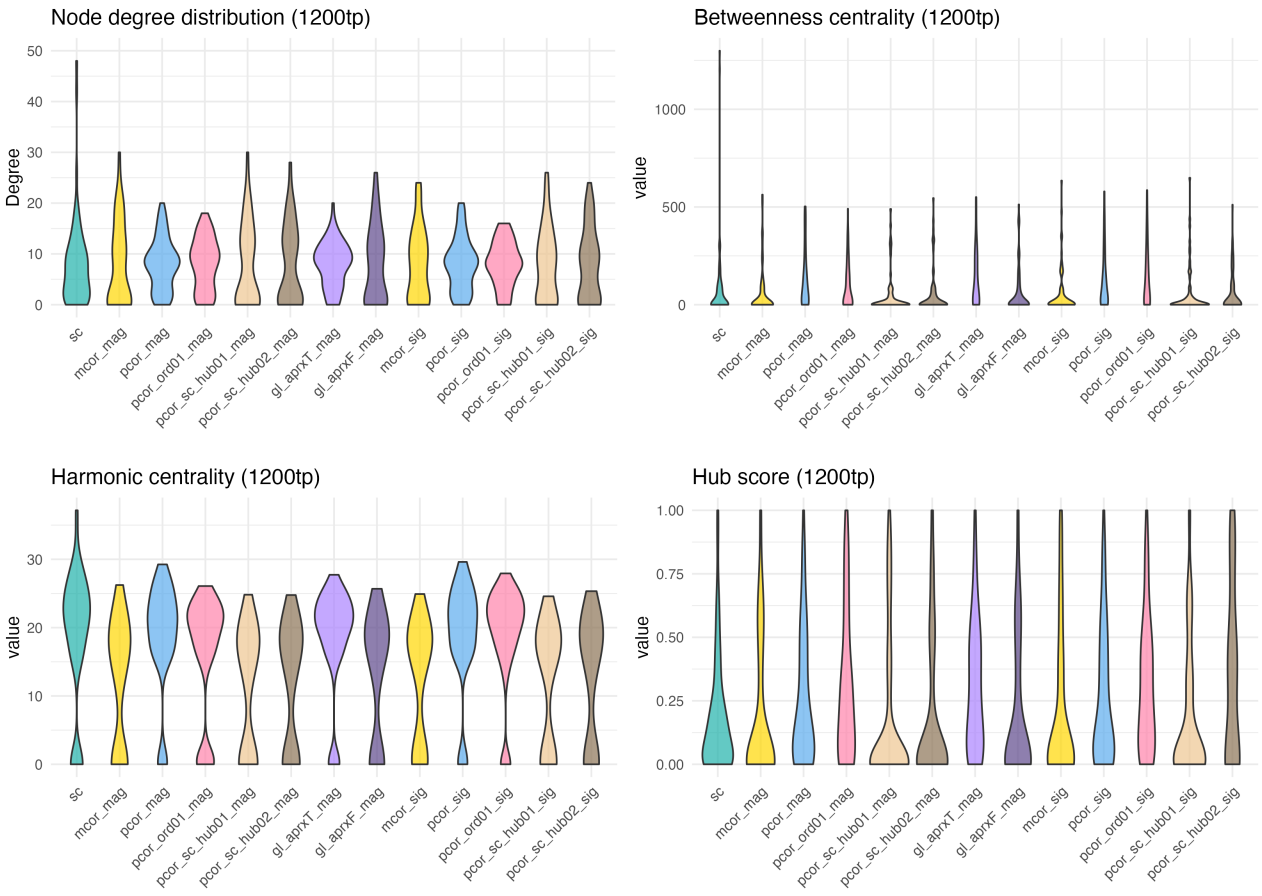

Figure 10: Summary of node-level centrality measures. The violin plots show the distribution of node degrees, betweenness centrality measures, harmonic centrality measures, and hub scores in SC and FC networks, where the FC networks are constructed using 1200 time points. These centrality measures quantify the topological importance of nodes and their connection strength with others within a network, with a larger value indicating a stronger connection to other nodes in the network.

Among the FC networks, *pcor*, *pcor_ord01*, and *gl_aprxT* contain some nodes with moderate degrees, while *mcor*, *gl_aprxF*, and *pcor_sc_hub* have more nodes with small degrees. The latter group also shows more nodes with low betweenness, indicating that they are less likely to act as intermediaries along paths. Additionally, these networks exhibit lower harmonic centrality overall, meaning that nodes in these networks require longer paths to reach others on average. Note that isolated nodes consistently show zero values in both betweenness and harmonic centrality. In contrast, networks such as *pcor*, *pcor_ord01*, and *gl_aprxT* have more nodes with higher harmonic centrality, reflecting closer interactions among nodes.

A similar distinction emerges in the hub score distribution: SC, *mcor*, *pcor_sc_hub*, and *gl_aprxF* dis-

play more values near zero, indicating the presence of a few highly influential hubs. Conversely, *pcor*, *pcor_ord01*, and *gl_aprxT* show a more balanced distribution of hub scores across nodes, suggesting a more even spread of connection strength throughout the network.

In addition, we compute Spearman's correlations for the node-level metrics across networks, as shown in the appendix. Based on these correlations, we again observe two clear families: one comprising *mcor*, *gl_aprxF*, and *pcor_sc_hub* networks, and another formed by *pcor*, *pcor_ord01*, and *gl_aprxT*. This grouping aligns with earlier findings from component, cliquishness, and clustering analyses, suggesting two groups of FC patterns. Notably, SC consistently shows lower correlations with all FC networks, though still positive, indicating a distinct node-level structure between anatomical and functional connectivity.

3.4 FC fingerprint analysis

FC within- and cross-subject similarity

Figure 11 displays the within- and cross-subject similarity across FC estimates, where similarity is defined as the Pearson's correlation between individual vectorized FC estimates. For reference, the cross-subject similarity of SC is also shown, which is identical in the 1200tp and 200tp plots.

For within-subject comparisons, all unthresholded FC estimates exhibit reduced similarity when transitioning from 1200 to 200 time points, with *pcor* showing the most pronounced decline (mean correlation 0.64 to 0.25). This sensitivity reflects *pcor*'s dependency on conditioning variables, which amplify noise with fewer time series.

On the other hand, *mcor* and *pcor_sc_hub* demonstrate relatively high within-subject similarity at 1200tp (mean correlation over 0.6). However, *pcor_sc_hub02*, which conditions on the largest number of variables, experiences the steepest drop when using 200 time points, suggesting a potentially higher sensitivity to sample size and autocorrelation.

For cross-subject comparisons, SC demonstrates high similarity across individuals (mean correlation of 0.94) and reflects anatomical consistency. In contrast, FC methods have more variable behaviors. *mcor* and *pcor_sc_hub* exhibit higher cross-subject similarity overall but wider variability, with longer tails toward lower correlations, whereas *pcor* and *pcor_ord01* show lower similarity, particularly at 200tp (*pcor* mean: 0.54 to 0.21). This instability in *pcor* suggests its vulnerability to limited sample size due to a large number of conditioning variables. Methods with fewer or no conditioning variables remain more stable across time points.

Notably, all methods exhibit higher within-subject than cross-subject similarity, regardless of the

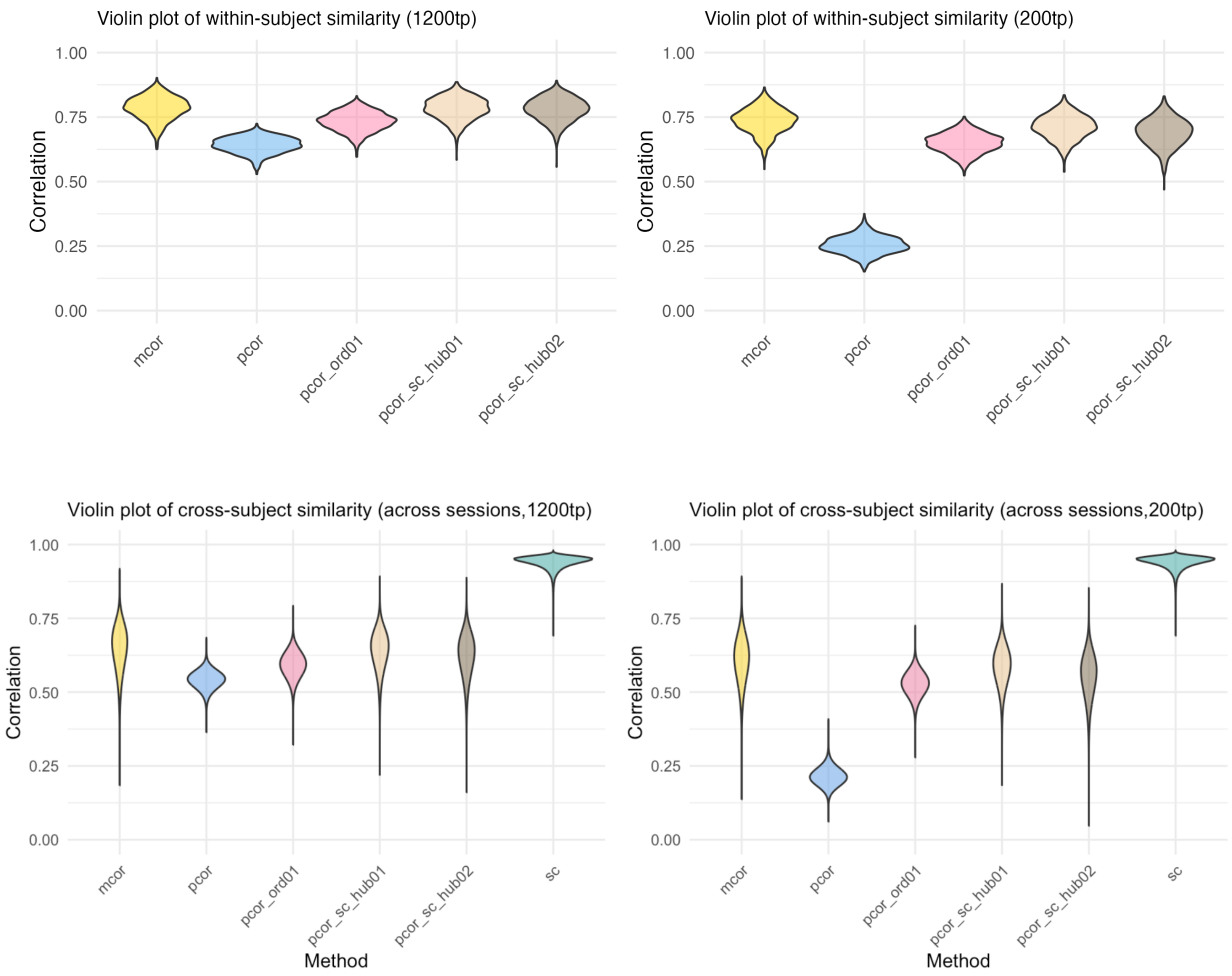

Figure 11: FC within- and cross-subject similarity. Within-subject similarity is quantified by Pearson correlation between the vectorized unthresholded upper-triangular elements of the FC matrices across all four scan sessions pairwise for each subject; cross-subject similarity is computed by taking the pairwise correlation between vectorized unthresholded FC estimates from different subjects and scan sessions.

time points used, indicating the inherent individual variations in functional connectivity patterns compared to structural connectivity's shared anatomy.

Identification accuracy

Figure 12 shows the identification accuracy across FC estimates, where each boxplot contains 12 values associated with the identification accuracy rate for each pair of two different sessions.

In the 1200tp scenario, *pcor* achieves the highest identification accuracy (mean accuracy 88.20%), followed by *pcor_ord01* (76.18%), while *mcor* and *pcor_sc_hub*, display lower accuracy (*pcor_sc_hub02* at 55.54%, *pcor_sc_hub01* at 50.04%, and *mcor* at 38.32% in the descending order). This ordering aligns with the number of conditioning nodes, likely suggesting that partialling out additional variables

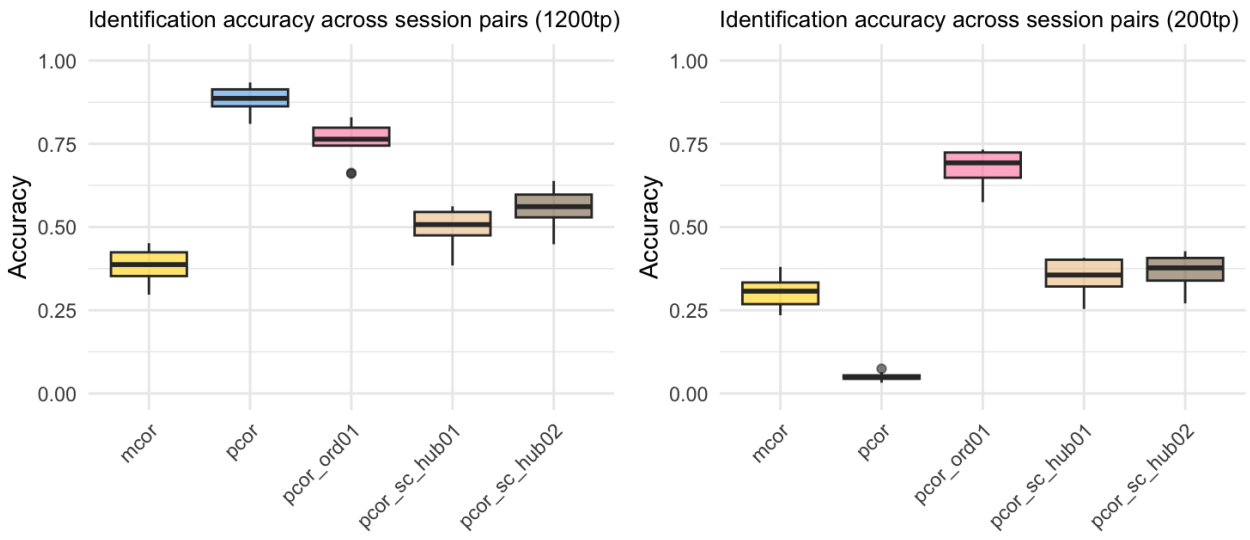

Figure 12: FC identification accuracy rate across session pairs (1200tp vs 200tp). Each FC method corresponds to a boxplot, which represents 12 identification accuracy rates, one for each pair of fMRI scan sessions. For each pair, one session's FC estimate is used as a reference, and the estimate from another session with the highest correlation is selected. The identification is considered accurate if the matched FC corresponds to the same subject in the other session, and the accuracy rate is computed as the proportion of correctly identified subjects. To compare identification performance across different sample sizes, we show results based on FC networks constructed from both 1200tp and 200tp.

enhances individual-specific features, although the effect is most pronounced in *pcor_ord01*, which relies on a different conditioning strategy.

Across all methods, identification accuracy decreases when reducing the number of time points from 1200 to 200. Notably, *pcor*, which performs best in 1200tp, drops dramatically to a mean accuracy of 4.97% in 200tp, making it the worst-performing method. In contrast, *pcor_ord01* remains relatively robust (67.49% in 200tp), indicating greater stability in maintaining individual uniqueness against reduced observations. Among *mcor* and *pcor_sc_hub* methods, the relative ordering persists in 200tp (*pcor_sc_hub02*: 37.04%, *pcor_sc_hub01*: 35.28%, *mcor*: 30.53%), but *mcor* exhibits the smallest decline while *pcor_sc_hub02* shows the largest.

Additionally, identification between session pairs with the longest interval (session 1 as the reference and session 4 as the target, or the reverse) typically shows the lowest accuracy rate across all methods, with the only exception in 200tp *pcor*, reflecting substantial variations in brain functional activity over time.

Furthermore, permutation tests confirm the significance of the identification accuracy. Across 1200

permuted identification runs, the maximum success rate is 0.522% (5 out of 957 subjects) for both 1200tp and 200tp, which is even below the weakest performance (accuracy 3.24% in 200tp *pcor* with 31 success out of 957 subjects, using session 1 as the reference and session 2 as the target). This result indicates that functional connectivity estimates capture identifiable individual-specific patterns far beyond random chance, suggesting their potential as reliable brain fingerprints.

3.5 Prediction performance of FC estimates

Overall, as Figure 13, 14 shows, FC-based predictions consistently outperform those based on SC, suggesting that FC captures behavioral features more effectively than SC. Among the various FC methods, *pcor* sometimes shows better performance for specific outcomes, such as *NEOFAC_C* and *Flanker_Unadj*, and generally performs slightly better than other approaches. However, the relative ranking of FC methods in terms of prediction R^2 varies depending on the outcome.

Figure 15 shows that reducing the number of time points from 1200 to 200 diminishes predictive performance across all methods, with *pcor* experiencing the most pronounced drop, which reiterates the importance of sufficient sample size and autocorrelation for reliable partial correlation estimation. In contrast, other methods exhibit more robustness to reduced time points, albeit with lower overall performance.

From Figure 16, when SC and FC are modeled jointly, most FC methods typically show an improvement in predicted R^2 in most outcomes, indicating that anatomical and functional interactions contain complementary information in explaining behavioral variance, even though SC alone cannot predict phenotypes as well. However, when SC is combined with *pcor* in the predictive models, *pcor*'s R^2 performance declines for most outcomes. This possibly suggests that, for some outcomes, including SC may introduce irrelevant information into the joint ridge regression model, which hampers its ability to effectively integrate and reconcile the distinct signals from both *pcor* and SC.

Additionally, predictions based on FC estimates averaged across four fMRI sessions outperform those derived from individual sessions, as illustrated by the results from the first session in Figure 17 an example. This observation underscores the noisy nature of fMRI data and the benefits of averaging in stabilizing FC estimates.

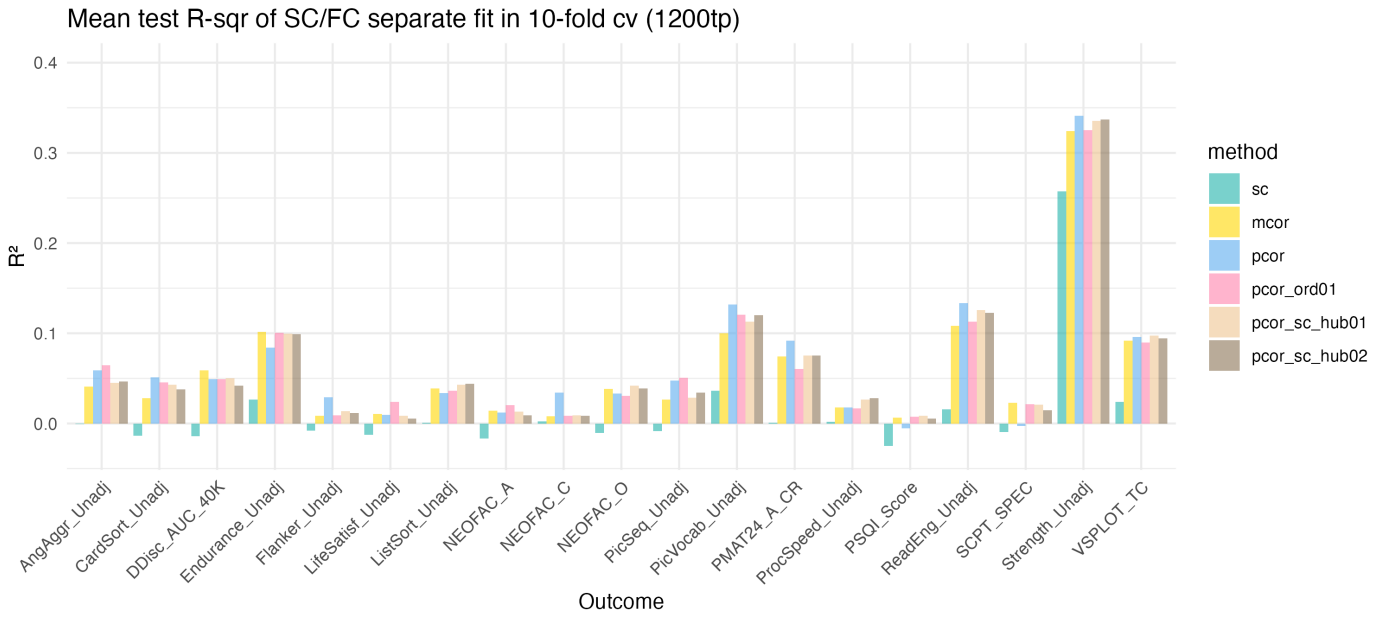

Figure 13: Mean test R^2, with SC and FC estimates fitted separately. Ridge regression models using FC predictors demonstrate better predictive performance than those using SC predictors, indicating that FC more effectively captures behaviorally relevant features.

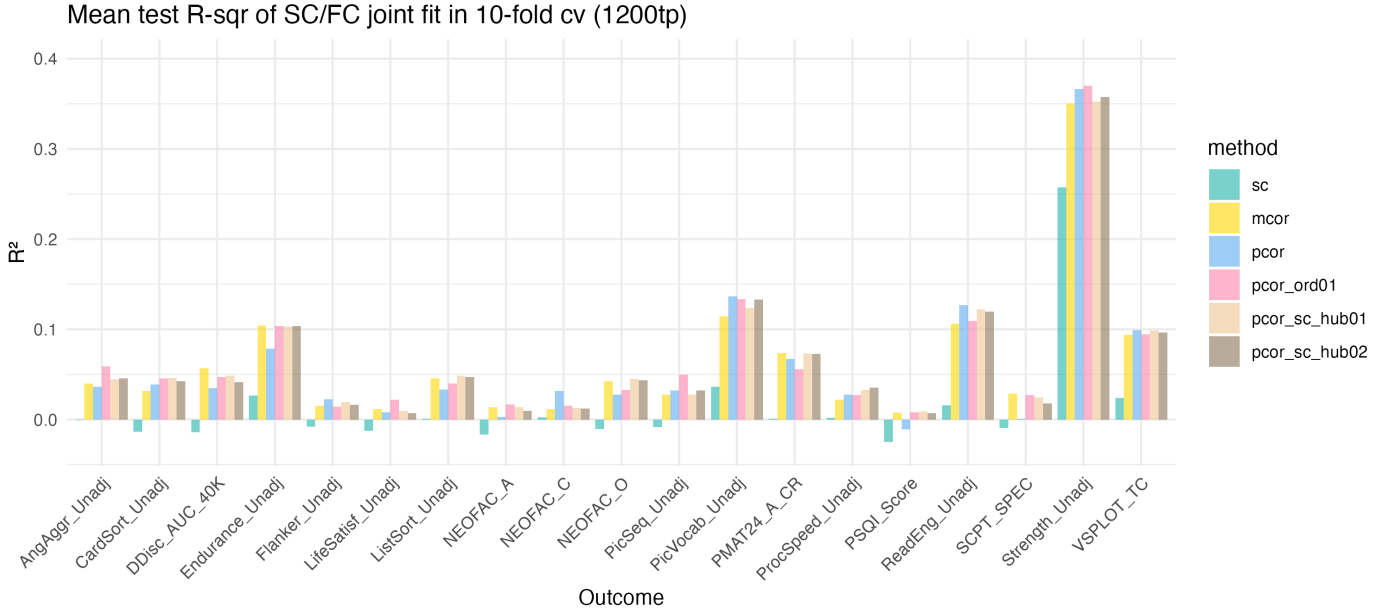

Figure 14: Mean test R^2, with SC and FC estimates fitted jointly. Compared to models using SC predictors alone, jointly incorporating both SC and FC predictors leads to improved predictive performance.

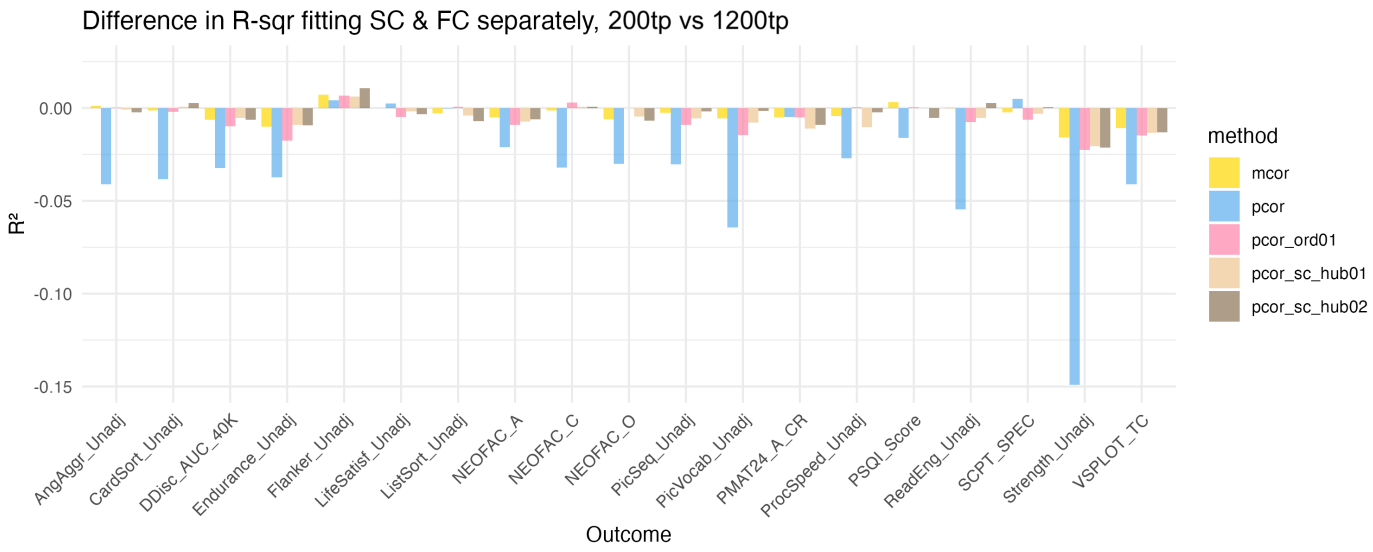

Figure 15: Difference in mean R^2 between using FC estimates derived from 200tp v.s. 1200tp. Compared to models using FC predictors derived from 1200 time points, the predictive performance of models based on 200 time points generally decreases across all methods. The largest drop is observed with pcor, suggesting greater sensitivity to sample size and autocorrelation.

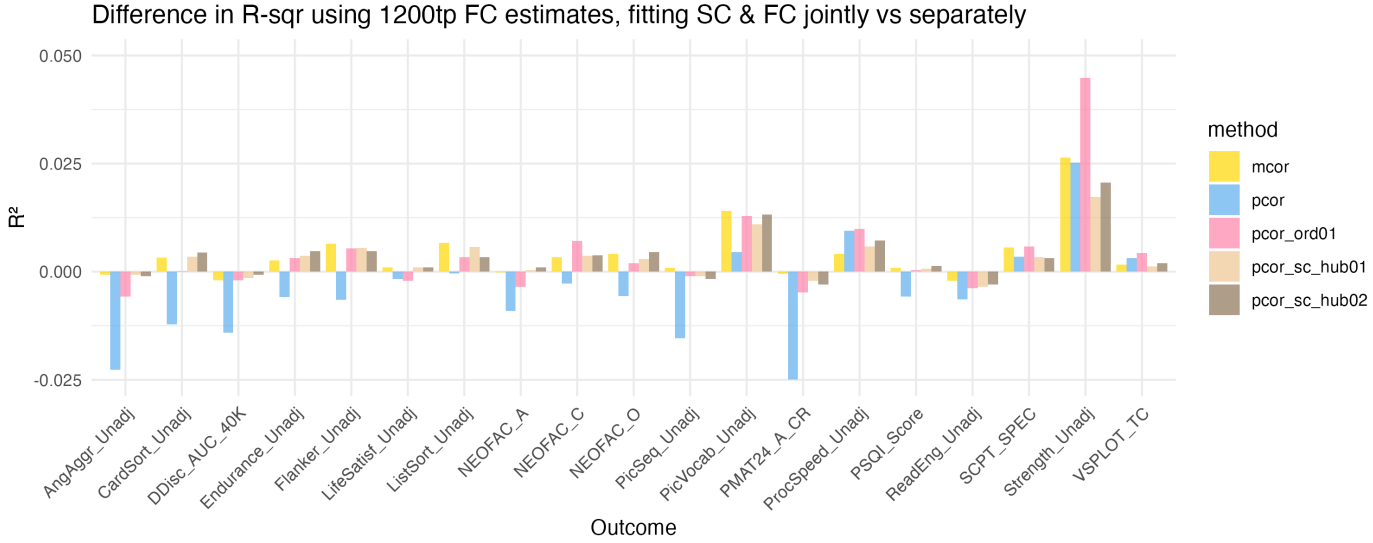

Figure 16: Difference in mean R^2 between fitting SC and FC jointly v.s. separately. Compared to models that include SC and FC predictors separately, models that incorporate them jointly generally achieve better predictive performance; except for pcor, which shows the largest performance drops across most outcomes.

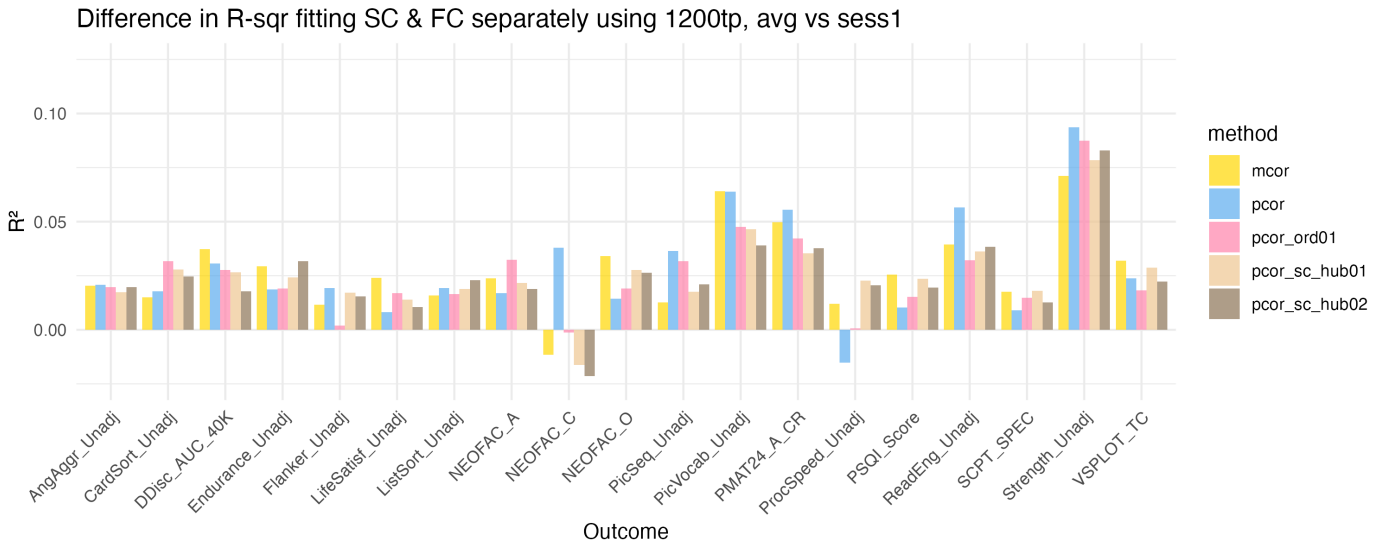

Figure 17: Difference in mean R^2 between using FC averaged across sessions v.s. from the 1st session. Compared to models using FC predictors derived from a single session, those using FC predictors based on correlations averaged across four sessions show improved predictive performance.

4 Discussion

In this work, we explore various methods to construct FC networks and compare them with SC, focusing on graph structure and topology of thresholded estimates, as well as fingerprinting and prediction analyses of unthresholded estimates. We summarize the patterns and findings comparing SC and various FC estimates, as well as limitations in this work.

Inherently different organizations in SC and FC estimates

We observe that thresholded SC and FC networks exhibit distinct organizational patterns. SC networks are characterized by sparse inter-hemispheric connections and dense intra-hemispheric wiring, whereas all FC networks display diagonal band-like structures, indicating strong interhemispheric coupling between anatomically corresponding regions. The large graph edit distances between SC and FC networks further confirm their structural differences.

Graph topology analysis further characterizes these differences. Thresholded SC networks exhibit small-world properties, characterized by high clustering coefficients and short average path lengths [1, 2], with its global clustering coefficient being larger than that of all FC methods and centrality measures. Specifically, SC's degree distribution is skewed towards large values, indicating the presence of highly connected nodes that act as potential information hubs. In contrast, FC networks show distinct degree patterns, as evidenced by low Spearman correlations (0.029-0.208) between SC and FC node degrees (shown in the appendix). Compared to FC networks, SC networks also demonstrate high harmonic centrality overall, reflecting shorter average path lengths and greater ease of information traversal. Similarly, SC's betweenness centrality distribution contains more large values, indicating that specific nodes serve as crucial bridges in the network. These findings collectively demonstrate that SC embodies a small-world network, a characteristic less apparent in FC networks. This structural difference possibly implies that SC serves as the basis for efficient communication, enabling more complex and integrative functional activities.

Furthermore, unthresholded SC and FC exhibit distinct patterns in cross-subject similarity and predictive performance. Vectorized FC estimates show varying degrees of cross-subject similarity, while SC demonstrates a much higher similarity across subjects. In ridge regression prediction models, SC performs notably worse, yielding lower mean test R^2 values across all outcomes and frequently producing negative out-of-sample R^2 values. However, combining SC with FC typically improves predictive performance, outperforming both individual SC models and most FC methods (except partial correlation). These findings suggest that SC and most FC measures provide complementary

information in explaining behavioral and cognitive processes.

Despite the distinctions between SC and FC networks, both exhibit patterns that contrast with random networks in terms of graph components, cliquishness, and clustering coefficients, supporting the presence of underlying biological significance.

Two groups of FC estimate patterns

Based on similarities in graph structure and topology, FC estimates can be categorized into two groups: (1) the marginal correlation group, which includes marginal correlation, SC hub-adjusted partial correlation, and approximated graphical lasso (thresholded network), and (2) the partial correlation group, consisting of partial correlation, first-order partial correlation, and exact graphical lasso (thresholded network). The marginal correlation group tends to capture both direct and indirect associations between brain regions, while the partial correlation group typically isolates direct relationships more effectively by removing indirect influences.

In unthresholded FC estimates, the marginal correlation group exhibits larger magnitudes compared to the partial correlation group. For thresholded networks, although both groups maintain prominent band-like structures between cross-hemispheric brain regions in correspondence, the marginal correlation group displays more distributed connections across the entire brain network and denser connections within functional communities. In contrast, the partial correlation group shows denser within-hemisphere connections and weaker co-activations specific to functional communities. Additionally, the marginal correlation group typically features a smaller largest component, more isolated regions, and a greater number of 3-cliques, suggesting that it captures clusters of brain regions strongly interconnected with each other while leaving some regions disconnected due to weaker correlations.

Taken together, these observations indicate that the marginal correlation group is more likely to preserve indirect associations, whereas the partial correlation group tends to filter them out, retaining only direct connections.

FC variability and the effect of conditioning in partial correlation

Among the FC correlation calculation methods, the resultant estimates exhibit varying levels of sensitivity to the number of observations and reflect the effects of different conditioning strategies.

Partial correlation, which conditions on the largest number of nodes, shows the greatest sensitivity to reduced observations. When time points decrease from 1200 to 200, partial correlation suffers significant decline in edge selection rate, within-subject and cross-subject similarity, individual identification accuracy, and its prediction R^2. However, with larger sample sizes, partial correlation typi-

cally performs well in these tasks, underscoring its heavy reliance on sufficient data to ensure stable and accurate estimation.

In contrast, marginal correlation, which does not condition on any variables, remains the most stable as the sample size decreases in terms of edge selection rate, within- and cross-subject similarity, identification accuracy, and prediction performance. While it is affected by a reduction in time points, the impact is minimal compared to other methods. This robustness likely comes at the cost of individual specificity, evidenced by a higher mean cross-subject similarity than other methods.

First-order partial correlation, which selects conditioning nodes by minimizing the correlation magnitude, exhibits intermediate sensitivity between marginal and partial correlation across all aforementioned situations. Notably, this method maintains relatively high subject identification accuracy despite changes in sample size, highlighting its potential for individual representation in brain fingerprinting.

SC hub-adjusted partial correlation, which conditions on nodes with high hub scores in each subject's structural connectivity, mirrors a sensitivity pattern similar to marginal correlation. Though it matches or exceeds first-order partial correlation in the number of conditioned nodes, its reliance on anatomically central hubs introduces less variability. This fact reveals that variability due to sample size is not solely determined by the number of conditioning nodes. Instead, the conditioning strategy plays a crucial role, with selection strategies aimed at minimizing correlations leading to greater variability.

Limitations

Here, we highlight a few limitations of our study. First, the binarization of SC and FC networks using a heuristic threshold, set at twice the number of nodes, introduces potential variability in network topology. While this threshold ensures an equal number of network edges across methods, it remains unclear how different thresholds might influence observed patterns. For instance, higher densities could amplify within-hemisphere connections in marginal correlation networks and inter-hemisphere connections in partial correlation networks, potentially blurring the current distinctions and interpretations.

Furthermore, the current predictive modeling relies on vectorized FC correlation matrices without dimensionality reduction or feature selection. While penalized regression mitigates some challenges of high-dimensional prediction, the vectorization operation discards the inherent network structure, leading to a loss of spatial and topological information. Introducing methods that preserve the network structure may provide additional insights into their predictive capabilities.

In addition, although first-order partial correlation demonstrates robustness to reduced time points and high individual identification accuracy, its estimation procedure is computationally expensive due to its exhaustive search, especially when the number of nodes is large. The selection of conditioning nodes can vary across subjects and even within subjects across sessions, which complicates the interpretation of the resultant partial correlations. Developing strategies that constrain the search space could reduce runtime and improve interpretability without sacrificing performance.

References

[1] Danielle S. Bassett and Edward T. Bullmore. "Small-World Brain Networks". In: *The Neuroscientist* 12 (2006), pp. 512–523.

[2] Danielle S. Bassett and Edward T. Bullmore. "Small-World Brain Networks Revisited". In: *The Neuroscientist* 23.5 (2017). PMID: 27655008, pp. 499–516. DOI: 10.1177/1073858416667720. eprint: https://doi.org/10.1177/1073858416667720.

[3] Jessica S. Damoiseaux and Michael D. Greicius. "Greater than the sum of its parts: a review of studies combining structural connectivity and resting-state functional connectivity". In: *Brain Structure and Function* 213.6 (June 2009), pp. 525–533. DOI: https://doi.org/10.1007/s00429-009-0208-6.

[4] Rahul S. Desikan et al. "An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest". In: *NeuroImage* 31.3 (2006), pp. 968–980. ISSN: 1053-8119. DOI: https://doi.org/10.1016/j.neuroimage.2006.01.021. URL: https://www.sciencedirect.com/science/article/pii/S1053811906000437.

[5] Emily S Finn et al. "Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity". In: *Nature Neuroscience* 18.11 (2015), pp. 1664–1671. DOI: 10.1038/nn.4135.

[6] Michael D. Fox and Marcus E. Raichle. "Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging". In: *Nature Reviews Neuroscience* 8.9 (Sept. 2007), pp. 700–711. DOI: https://doi.org/10.1038/nrn2201.

[7] Linton C. Freeman. "Centrality in social networks conceptual clarification". In: *Social Networks* 1.3 (1978), pp. 215–239. ISSN: 0378-8733. DOI: https://doi.org/10.1016/0378-8733(78)90021-7.

[8] Jerome Friedman, Trevor Hastie, and Robert Tibshirani. "Sparse inverse covariance estimation with the graphical lasso". In: *Biostatistics* 9.3 (Dec. 2007), pp. 432–441. ISSN: 1465-4644. DOI: 10.1093/biostatistics/kxm045.

[9] Jerome Friedman, Robert Tibshirani, and Trevor Hastie. "Regularization Paths for Generalized Linear Models via Coordinate Descent". In: *Journal of Statistical Software* 33.1 (2010), pp. 1–22. DOI: 10.18637/jss.v033.i01.

[10] Patric Hagmann et al. “Mapping the Structural Core of Human Cerebral Cortex”. In: *PLoS Biology* 6.7 (July 2008). Ed. by Karl J Friston, e159. DOI: https://doi.org/10.1371/journal.pbio.0060159.

[11] Seongho Kim. “ppcor: An R Package for a Fast Calculation to Semi-partial Correlation Coefficients”. In: *Communications for statistical applications and methods* 22.6 (Nov. 2015), pp. 665–674. DOI: https://doi.org/10.5351/CSAM.2015.22.6.665.

[12] Jon M. Kleinberg. “Authoritative sources in a hyperlinked environment”. In: *J. ACM* 46.5 (Sept. 1999), pp. 604–632. ISSN: 0004-5411. DOI: 10.1145/324133.324140.

[13] Paul M Magwene and Junhyong Kim. “Estimating genomic coexpression networks using first-order conditional independence”. In: *Genome Biology* 5.12 (2004), R100. DOI: https://doi.org/10.1186/gb-2004-5-12-r100.

[14] Massimo Marchiori and Vito Latora. “Harmony in the small-world”. In: *Physica A: Statistical Mechanics and its Applications* 285.3 (2000), pp. 539–546. ISSN: 0378-4371. DOI: https://doi.org/10.1016/S0378-4371(00)00311-3.

[15] Guillaume Marrelec et al. “Partial correlation for functional brain interactivity investigation in functional MRI”. In: *NeuroImage* 32.1 (Aug. 2006), pp. 228–237. DOI: https://doi.org/10.1016/j.neuroimage.2005.12.057.

[16] Nicolai Meinshausen and Peter Bühlmann. “High-dimensional graphs and variable selection with the Lasso”. In: *The Annals of Statistics* 34.3 (2006), pp. 1436–1462. DOI: 10.1214/009053606000000281

[17] Lei Nie et al. “Inferring functional connectivity in fMRI using minimum partial correlation”. In: *International Journal of Automation and Computing* 14.4 (June 2017), pp. 371–385. DOI: https://doi.org/10.1007/s11633-017-1084-9.

[18] Stephen M. Smith et al. “Network modelling methods for FMRI”. In: *NeuroImage* 54.2 (Jan. 2011), pp. 875–891. DOI: https://doi.org/10.1016/j.neuroimage.2010.08.063.

[19] Olaf Sporns. “Structure and function of complex brain networks”. In: *Static and Dynamic Imaging: Clinical and Therapeutic Implications* 15.3 (Sept. 2013), pp. 247–262. DOI: https://doi.org/10.31887/dcns.2013.15.3/osporns.

[20] David C. Van Essen et al. “The WU-Minn Human Connectome Project: An overview”. In: *NeuroImage* 80 (Oct. 2013), pp. 62–79. DOI: https://doi.org/10.1016/j.neuroimage.2013.05.041.

[21] Lea Waller et al. “Evaluating the replicability, specificity, and generalizability of connectome fingerprints”. In: *NeuroImage* 158 (Sept. 2017), pp. 371–377. DOI: https://doi.org/10.1016/j.neuroimage.2017.07.016.

[22] Stanley Wasserman and Katherine Faust. *Social Network Analysis: Methods and Applications*. Structural Analysis in the Social Sciences. Cambridge University Press, 1994.

[23] Anja Wille and Peter Bühlmann. “Low-Order Conditional Independence Graphs for Inferring Genetic Networks”. In: *Statistical Applications in Genetics and Molecular Biology* 5.1 (2006). DOI: doi:10.2202/1544-6115.1170.

[24] Ming Yuan and Yi Lin. “Model selection and estimation in the Gaussian graphical model”. In: *Biometrika* 94.1 (Mar. 2007), pp. 19–35. ISSN: 0006-3444. DOI: 10.1093/biomet/asm018.

Appendix

Graph edit distance between 1200tp and 200tp FC

Figure 18 summarizes the edit distances between FC networks constructed using 1200 and 200 time points. Networks obtained from the same method exhibit strong resemblance, indicating consistency across different sampling durations.

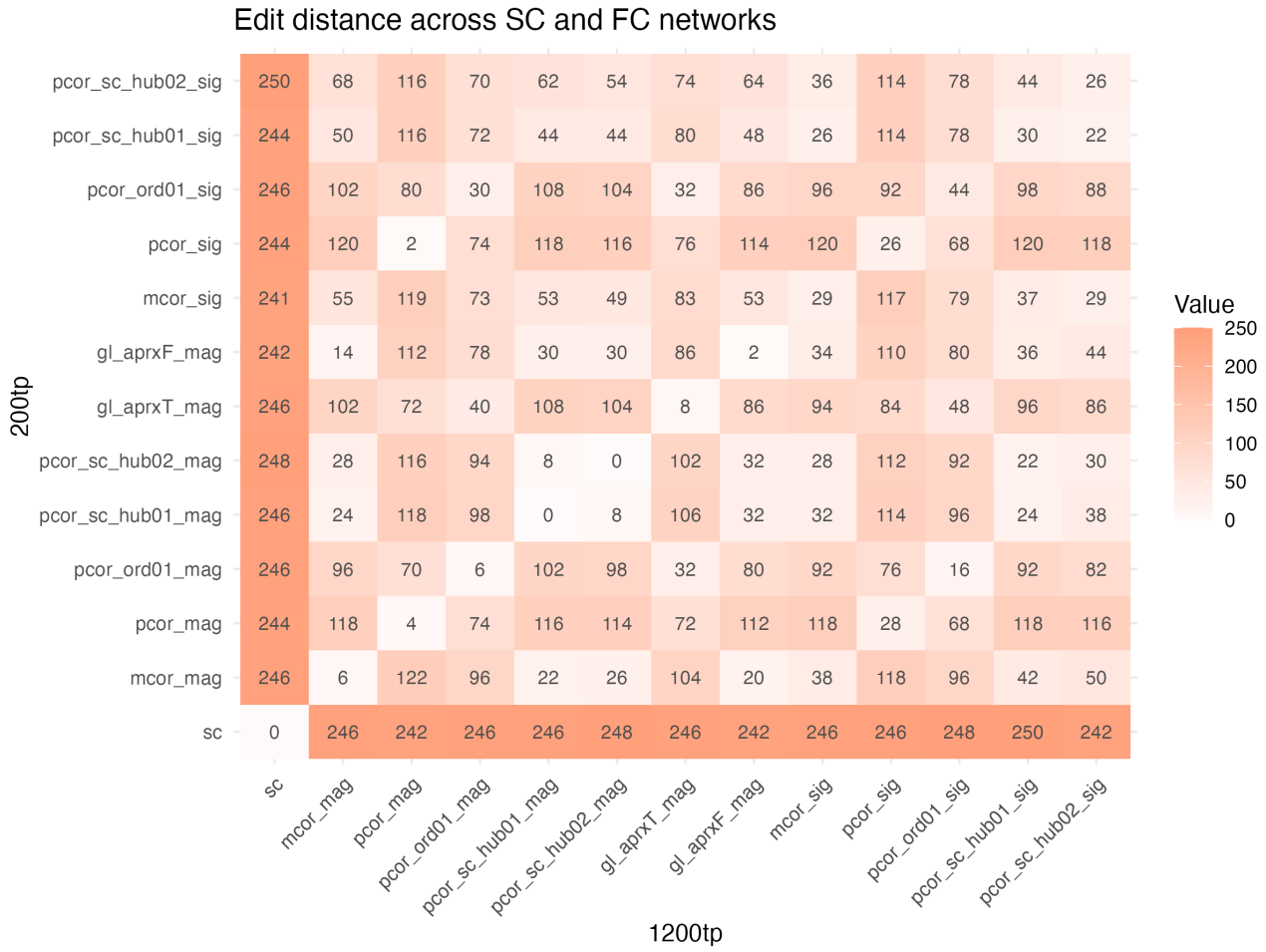

Figure 18: Edit distance of thresholded SC and FC networks (1200tp vs 200tp). The graph edit distances comparing FC networks derived from 1200 and 200 time points using the same estimation method show a high level of structural similarity.

Graph component, cliquishness and clustering

Figure 19 and 20 present the graph component cliquishness summary for thresholded SC and FC networks using 200 time points, respectively. The overall patterns remain consistent with those observed using 1200 time points (Figure 8 and 9), as described previously.

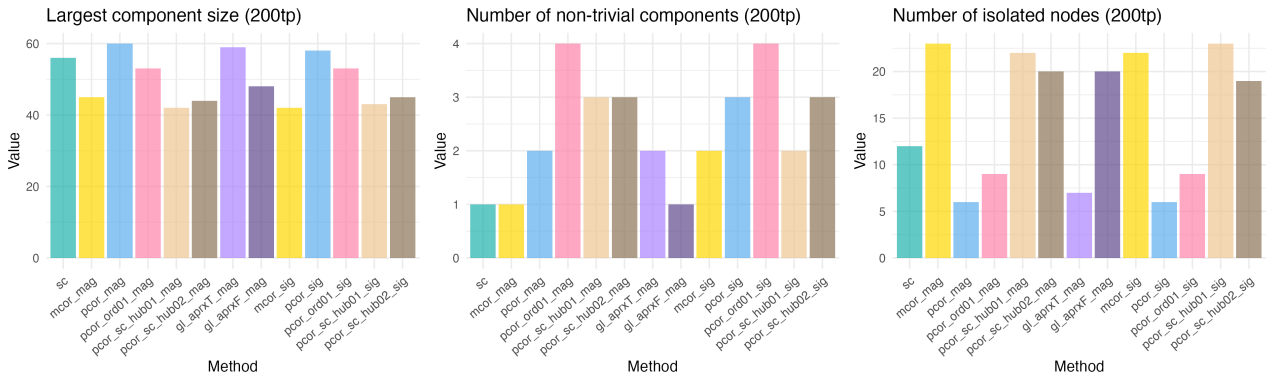

Figure 19: Summary of thresholded SC and FC graph components using 200 time points. We compare the size of the largest connected component, the number of non-trivial components, and the number of isolated nodes across SC and FC networks, with FC networks derived from 200 time points.

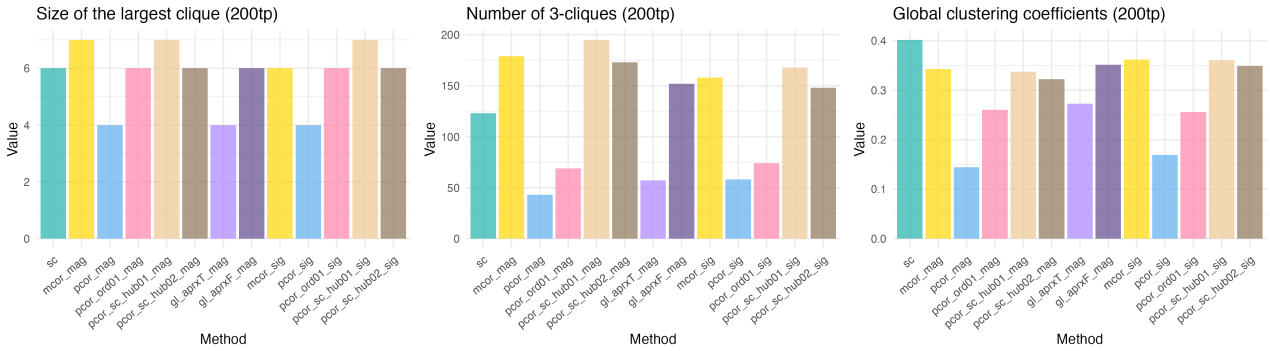

Figure 20: Summary of thresholded SC and FC graph cliquishness using 200 time points. We compare the size of the largest clique, the number of 3-cliques, and the global clustering coefficient across SC and FC networks, with FC networks derived from 200 time points.

Correlation of node-level centrality

Figure 21 show the Spearman's correlation for node-level centrality metrics, as described in the main text. High correlations are observed within two groups of networks: (1) *mcor*, *gl_aprxF*, *pcor_sc_hub*, and (2) *pcor*, *pcor_ord01*, *gl_aprxT*. Additionally, low correlations between SC and FC networks suggest a discordance between structural and functional connectivity at the node level.

Figure 21: Spearman's correlation of node-level measures for FC networks derived from 1200 time points. For a pair of connectivity networks, we calculate Spearman's correlation between upper triangular elements. We find high correlations within two groups of FC networks and low correlations between SC and FC networks in general.