# Evaluating the sensitivity of functional connectivity measures to motion artifact in resting-state fMRI data

Arun S. Mahadevan [a], Ursula A. Tooley [f], Maxwell A. Bertolero [a], Allyson P. Mackey [g],
Danielle S. Bassett [a,b,c,d,e,h,∗]

[a] *Department of Bioengineering, School of Engineering & Applied Science, University of Pennsylvania, Philadelphia, PA 19104, USA*
[b] *Department of Electrical & Systems Engineering, School of Engineering & Applied Science, University of Pennsylvania, Philadelphia, PA 19104, USA*
[c] *Department of Physics & Astronomy, College of Arts & Sciences, University of Pennsylvania, Philadelphia, PA 19104, USA*
[d] *Department of Neurology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA*
[e] *Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA*
[f] *Neuroscience Graduate Group, Perelman School of Medicine, University of Pennsylvania, Pennsylvania, PA 19104, USA*
[g] *Department of Psychology, College of Arts & Sciences, University of Pennsylvania, Philadelphia, PA 19104, USA*
[h] *Santa Fe Institute, 1399 Hyde Park Rd, Santa Fe, NM 87501, USA*

## ARTICLE INFO

## ABSTRACT

Functional connectivity (FC) networks are typically inferred from resting-state fMRI data using the Pearson correlation between BOLD time series from pairs of brain regions. However, alternative methods of estimating functional connectivity have not been systematically tested for their sensitivity or robustness to head motion artifact. Here, we evaluate the sensitivity of eight different functional connectivity measures to motion artifact using resting-state data from the Human Connectome Project. We report that FC estimated using full correlation has a relatively high residual distance-dependent relationship with motion compared to partial correlation, coherence, and information theory-based measures, even after implementing rigorous methods for motion artifact mitigation. This disadvantage of full correlation, however, may be offset by higher test-retest reliability, fingerprinting accuracy, and system identifiability. FC estimated by partial correlation offers the best of both worlds, with low sensitivity to motion artifact and intermediate system identifiability, with the caveat of low test-retest reliability and fingerprinting accuracy. We highlight spatial differences in the sub-networks affected by motion with different FC metrics. Further, we report that intra-network edges in the default mode and retrosplenial temporal sub-networks are highly correlated with motion in all FC methods. Our findings indicate that the method of estimating functional connectivity is an important consideration in resting-state fMRI studies and must be chosen carefully based on the parameters of the study.

## 1. Introduction

Ever since the initial observation of correlations in spontaneous functional magnetic resonance imaging (fMRI) blood-oxygen level dependent (BOLD) signals acquired from subjects at rest, the field of resting-state functional connectivity has grown exponentially (Buckner et al., 2013). Functional connectivity has been used as a tool to explore large-scale features of human brain organization (Achard et al., 2006; Achard and Bullmore, 2007; Morgan et al., 2017; van den Heuvel and Hulshoff Pol, 2010; Vértes et al., 2012), how this organization changes over the course of development (Gu et al., 2015; Morgan et al., 2018; Thomason, 2020; Tooley et al., 2020; van den Heuvel et al., 2018; Vértes and Bullmore, 2015; Wheelock et al., 2019), and the association of such organization with individual behavior (Finn et al., 2015;

Gratton et al., 2018). However, head motion artifact is a pervasive problem in functional connectivity analysis, decreasing certainty of findings and impacting subsequent interpretations. In-scanner head movements result in structured noise that leads to the spurious identification of putative functional connections, a problem further compounded by the fact that some individuals move systematically more than others (Power et al., 2015). Consequently, a number of research groups have developed statistical preprocessing methods to mitigate the impact of motion artifact; the development of such methods is an ongoing and active field of research in its own right (Burgess et al., 2016; Caballero-Gaudes and Reynolds, 2017; Ciric et al., 2017; Parkes et al., 2018).

While much effort has been directed toward developing effective denoising pipelines to mitigate motion artifact, the subsequent estimation of functional connectivity in fMRI data has remained fairly con-

stant. Functional connectivity (FC) between brain regions is typically estimated through a Pearson correlation between the BOLD time series of two regions of interest (ROIs). However, functional connectivity is formally defined as any statistical relation between time series (Friston, 2011) and there exist many other statistical methods to compute similarity between time series. A few examples include coherence methods (Beauchene et al., 2018; Bullmore et al., 2004; Sun et al., 2004), which compute similarity in frequency space, and methods based on information theory (Quiroga et al., 2001; Salvador et al., 2005; Shannon, 1948), which quantify the amount of shared information between signals.

The performance of different FC estimation methods has been evaluated using generative models for a variety of neurophysiological data, including simulated BOLD signals (Smith et al., 2011; Wang et al., 2014). These studies have typically focused on the ability of FC estimation methods to recover the underlying network structure from simulated BOLD data. Major findings include that full or partial correlation successfully recovers the underlying network structure in simulated data (Smith et al., 2011). It is perhaps due to these results, and the associated ease of implementation, that correlation-based methods are so popular in the field. Yet, very few studies have used real fMRI data to compare the differential sensitivity or robustness of different FC estimation methods to motion artifact. The field awaits an appraisal of different FC estimation approaches with regard to their ability to overcome the specific type of noise introduced by motion artifact in fMRI data (Ciric et al., 2018; Power et al., 2015).

In the present study, we used resting-state fMRI data from the Human Connectome Project (HCP) to evaluate eight different FC estimation methods: Pearson correlation, Spearman correlation, partial correlation, Tikhonov partial correlation, coherence, wavelet coherence, mutual information in the time domain, and mutual information in the frequency domain. The sensitivity of each of these methods to subject motion and their success in identifying network structure was evaluated using five benchmarks: (a) correlations of subject motion with edge weights after denoising (QC-FC correlations), (b) the distance-dependence of QC-FC correlations, (c) the degree to which canonical brain systems could be identified through modularity maximization (system identifiability), (d) the extent to which FC estimates from one scan session could be used to identify subjects in other scan sessions (fingerprinting accuracy), and (e) the extent to which the FC estimates could be reproduced in successive scans (test-retest reliability). Collectively, these efforts serve to inform our usage of FC estimation methods, and their relative strengths and weaknesses.

## 2. Methods

In order to evaluate the differential sensitivity of different FC estimation methods to motion, we first applied common denoising pipelines to a large resting-state dataset, estimated functional connectivity matrices using eight different methods, and finally compared the performance of each of these estimates using a set of common quality control (QC) measures. Details of data preprocessing, FC estimation, and QC measures are described below.

### 2.1. Data and preprocessing

In this study, we leveraged data from the S1200 release of the Human Connectome Project (HCP) (Van Essen et al., 2013), a multi-site consortium that collected extensive MRI, behavioral, and demographic data from a large cohort of over 1000 subjects. As part of the HCP protocol, subjects underwent four separate resting-state scans, which included both left-right (REST1_LR, REST2_LR) and right-left (REST1_RL, REST2_RL) phase encoding directions. All functional connectivity data analyzed in this report came from these scans.

We used the ICA-FIX resting-state data provided by the Human Connectome Project, which used 24-parameter regression fol-

lowed by ICA+FIX denoising to remove nuisance and motion signals (Griffanti et al., 2014; Salimi-Khorshidi et al., 2014). In addition, we removed the mean global signal and bandpass filtered the time series using a fifth-order Butterworth filter in both forward and reverse directions to retain frequencies between 0.009 and 0.08 Hz. However, all results in the main text are reported for 4 preprocessing variants defined by the inclusion and exclusion of GSR and bandpass filtering. Further, we did not analyze subjects for whom greater than 50% of frames had a framewise displacement above 0.2 millimeters or a derivative root mean square above 75, leaving 778 subjects from REST1_LR, 800 from REST1_RL, 776 from REST2_LR, and 763 from REST2_RL. This threshold was chosen because it is typical for analyses of functional connectivity, and we wanted our conclusions about motion and functional connectivity to apply to common analysis pipelines (Bertolero et al., 2019; Schaefer et al., 2018; Vértes et al., 2016; Whitaker et al., 2016).

For each scan, we used the mean relative RMS (root-mean squared) displacement during realignment through MCFLIRT, provided by the Human Connectome Project, as our primary measure of motion. Recent studies have shown that multiband datasets with high temporal resolution like the HCP contain additional respiratory artifacts that manifest in the realignment parameters (RPs) typically used to calculate summary statistics of head motion (Agrawal et al., 2020; Fair et al., 2020; Power et al., 2019). To mitigate this possibility, we calculated mean framewise displacement (FD) after bandpass filtering RPs between 0.3 and 0.4 Hz, and used this measure as an additional summary statistic of motion (Fair et al., 2020). Summary statistics of the cohorts analyzed, and their head motion are shown in Table 1.

From the preprocessed data in CIFTI space, we estimated mean BOLD time series using two cortical parcellation schemes: the 333-node Gordon parcellation (Gordon et al., 2016) and the 100-node Schaefer parcellation (Schaefer et al., 2018) to test motion-related effects in networks of different sizes. We used sub-networks defined in the Gordon parcellation (Gordon et al., 2016) to evaluate motion-related effects within canonically defined sub-networks. For all scans, the MSMAll registration was used, and the mean time series of vertices on the cortical surface (fsL32K) in each parcel was calculated.

### 2.2. Description of functional connectivity measures

In the present study, we evaluated eight different methods for estimating functional connectivity from BOLD time series data. Here we provide a brief overview of the methods evaluated.

#### 2.2.1. Pearson's correlation coefficient

The Pearson correlation coefficient is a simple and commonly applied method to evaluate linear correlation between two time series. It is defined as the covariance between the two signals over time divided by the product of their standard deviations. Formally, the zero-order (without lag) Pearson correlation, $\rho_{ij}$, between the signals of regions $i$ and $j$ is given by

$$\rho_{ij} = \frac{cov_{ij}(t)}{\sqrt{var_i(t).var_j(t)}} \ .$$

Note that $\rho_{ij}$ varies in the interval [-1, 1] with positive values indicating positive correlation and negative values indicating negative correlation.

#### 2.2.2. Spearman's rank correlation coefficient

This method evaluates the rank correlation between two time series, providing an estimate of the extent to which one time series is a monotonic function of the other. The zero-order Spearman rank correlation, $r_{ij}$, between the signals of regions $i$ and $j$ is given by the Pearson correlation between their respective rankings,

$$r_{ij} = \rho_{R_i,R_j} = \frac{cov(R_i, R_j)}{\sqrt{var(R_i).var(R_j)}},$$

**Table 1**

Summary statistics of motion in a cohort taken from the HCP S1200 release.

| Scan | REST1_LR | REST1_RL | REST2_LR | REST2_RL |
|---|---|---|---|---|
| Number of subjects analyzed (n) | 778 | 800 | 776 | 763 |
| Mean relative RMS motion (average) | 0.079 | 0.077 | 0.078 | 0.079 |
| Mean relative RMS motion (standard deviation) | 0.022 | 0.021 | 0.021 | 0.023 |
| Mean FD from bandpass-filtered realignment parameters (average) | 0.081 | 0.078 | 0.080 | 0.082 |
| Mean FD from bandpass-filtered realignment parameters (standard deviation) | 0.022 | 0.020 | 0.020 | 0.023 |

where $R_i$ and $R_j$ are the rankings of the individual time series from low to high values. Note that $r_{ij}$ varies in the interval [-1, 1] with $r_{ij} = +1$ indicating a perfect monotonic relationship between the time series of region $i$ and the time series of region $j$.

### 2.2.3. Partial correlation coefficient

Partial correlation is a measure of the linear correlation between two time series after regressing out the time series of all other nodes in the network. Partial correlation has been proposed as an effective method to distinguish between direct and indirect links between nodes (Smith et al., 2011). Partial correlation is calculated by inverting the covariance matrix $\overset{\vee}{\Sigma}$ to obtain the precision matrix $\hat{\Omega}$, followed by flipping the signs of non-diagonal elements and normalization (Pervaiz et al., 2020). Note that partial correlation also varies between [-1, 1] with positive values indicating positive correlation and negative values indicating negative correlation, after accounting for all other time series in the network.

### 2.2.4. Tikhonov partial correlation coefficient

In this method, also referred to as $L_2$ ridge regression, subject precision matrices are obtained by inverting the covariance matrix after adding a regularization term $\Gamma = \alpha I$, where $I$ is the identity matrix. The subject precision matrix can be written as follows:

$$\bar{\Omega} = \left( \overset{\vee}{\Sigma} + \Gamma \right)^{-1} .$$

The scalar $\alpha$ determines the strength of the regularization applied. We optimized $\alpha$ separately for each preprocessing variant (Supplementary Table 1) by minimizing the root mean square distance between regularized subject precision matrices ($\bar{\Omega}$) and the group average of unregularized precision matrices ($\hat{\Omega}$) following prior work (Pervaiz et al., 2020).

### 2.2.5. Mutual information (time domain)

The mutual information is a statistical measure of the shared information between two time series. The information content of a given time series $X(t)$ can be defined through its Shannon entropy (Quiroga et al., 2001; Shannon, 1948), which is given by

$$I(X) = - \sum_{i=1}^{M} p_i . ln p_i,$$

where $X(t)$ is partitioned into $M$ bins, with $p_i$ representing the probability of the $i$-th bin. Now, the joint entropy between $X(t)$ and a second time series $Y(t)$ is defined as

$$I(X, Y) = - \sum_{i,j} p_{ij}^{XY} ln p_{ij}^{XY},$$

where $p_{ij}^{XY}$ is the joint probability of $X = X_i$ and $Y = Y_i$. The mutual information between $X$ and $Y$ is then,

$$MI(X, Y) = I(X) + I(Y) - I(X, Y) .$$

We estimated joint and marginal distributions empirically, using the Freedman-Diaconis rule (Freedman and Diaconis, 1981) to determine the optimal number of bins $M$. No corrections were applied for finite sample sizes (Quiroga et al., 2001), due to the relatively large sample size in the HCP dataset. In order to obtain values in the range [0, 1], we

computed the normalized mutual information (Strehl and Ghosh, 2002) as

$$NMI = \frac{MI(X, Y)}{\sqrt{I(X) . I(Y)}} .$$

Thus, the normalized mutual information between two independent signals is 0 and has a maximum of 1 for identical signals. While correlation coefficients measure linear relationships, the mutual information is a statistical measure of both linear and non-linear relationships between time series. However, mutual information operationalized in this manner does not exploit the autocorrelation property of time series data.

### 2.2.6. Coherence

Coherence is a measure of the cross-correlation between two signals in the frequency domain. At a given frequency $\lambda$, the coherence between the signal of region $i$ and the signal of region $j$ is given by

$$C_{ij}(\lambda) = \frac{\left| f_{ij}(\lambda) \right|^2}{f_{ii}(\lambda) . f_{jj}(\lambda)},$$

where $f_{ij}(\lambda)$ is the cross-spectral density between signals $i$ and $j$, and $f_{ii}(\lambda)$ and $f_{jj}(\lambda)$ are the auto-spectral densities of signal $i$ and $j$, respectively. Note that $C_{ij}(\lambda)$ varies in the interval [0, 1].

We evaluated coherence using the MATLAB toolbox for functional connectivity (Zhou et al., 2009), in which spectral densities are calculated using Welch's averaged, modified periodogram method for all frequencies. The average coherence in the frequency range [0.009Hz, 0.08Hz] was used to define connectivity.

### 2.2.7. Wavelet coherence

Wavelet coherence is a measure of the correlation between two signals in the time-frequency space. It is calculated in a similar manner to coherence, but spectral densities are calculated by convolving time series with wavelet functions such as the Morlet wavelet function (Zhang et al., 2016), which expand the signal in time-frequency space. We evaluated wavelet coherence using the Grinsted toolbox with the default Morlet wavelet function (Grinsted et al., 2004). Connectivity was determined by averaging coherence values in the frequency range [0.009Hz, 0.08Hz].

### 2.2.8. Mutual information (frequency domain)

Mutual information can also be evaluated based on coherence in the frequency domain (Salvador et al., 2005), defined for a given frequency range [$\lambda 1$, $\lambda 2$] as

$$\delta_{ij} = \frac{1}{2\pi} \int_{\lambda 1}^{\lambda 2} log \left( 1 - C_{ij}(\lambda) \right) d\lambda .$$

With a simple transformation, a normalized mutual information in the range [0, 1] can be obtained as

$$\emptyset_{ij} = \left[ 1 - exp(-2\delta_{ij}) \right]^{\frac{1}{2}} .$$

We used the implementation provided in the MATLAB toolbox for functional connectivity (Zhou et al., 2009). Connectivity was determined by averaging mutual information values in the frequency range [0.009Hz, 0.08Hz].

**Table 2**

Overview of functional connectivity estimation measures, their range and domain of operation.

| Domain | Metric | Range |
|--------|--------|-------|
| Time | Pearson correlation | [-1,1] |
| | Spearman correlation | [-1,1] |
| | Partial correlation | [-1,1] |
| | Tikhonov partial correlation | [-1,1] |
| | Mutual Information (time) | [0,1] |
| Frequency | Coherence | [0,1] |
| | Mutual Information (frequency) | [0,1] |
| Time-frequency | Wavelet Coherence | [0,1] |

## 2.3. Overview of functional connectivity estimation

Using the 8 measures described above, we estimated functional connectivity between each pair of preprocessed BOLD times series, resulting in $n \times n$ matrices for each subject, where $n$ is the number of parcels in the parcellation scheme. The estimated networks provide a description of interactions (edges) among brain regions (nodes) that can then be probed for various features of interest using network science (Bassett and Sporns, 2017).

The 8 metrics evaluated can be broadly classified into categories based on their mode of operation. Pearson correlation, Spearman correlation, partial correlation, Tikhonov partial correlation, and mutual information (time) operate in the time domain, whereas coherence and mutual information (frequency) operate in the frequency domain, and wavelet coherence operates in time-frequency space. For all the frequency-based methods, we evaluated the average connectivity in the frequency range [0.009Hz, 0.08Hz], which is the same frequency range at which all resting-state scans were bandpass filtered. Full and partial correlation metrics fall in the interval [-1, 1], and all the other metrics fall in the interval [0, 1] (Table 2).

Correlation matrices are typically subjected to Fisher's r-to-z transformation to normalize the range of values. Fig. S1 shows that results do not change significantly when the transform is applied to the correlation matrices. Therefore, we report results in the main text without performing Fisher transforms on any of the matrices, in order to facilitate a more direct comparison between methods.

## 2.4. Overview of outcome measures

We evaluated the sensitivity of each FC metric to subject motion using 5 benchmarks: residual QC-FC correlations, distance-dependence of QC-FC correlations, test-retest reliability, fingerprinting accuracy, and system identifiability.

### 2.4.1. Residual QC-FC correlations

Quality control-functional connectivity (QC-FC) correlations are a widely used benchmark to evaluate the efficacy of denoising pipelines applied in resting-state fMRI connectivity analysis (Ciric et al., 2017; Parkes et al., 2018). Here we used this benchmark to evaluate residual motion artifact for each of the functional connectivity estimates after application of a common denoising pipeline. First, we computed functional connectivity using the 8 metrics described in the previous section, for the 333-node Gordon and 100-node Schaefer parcellation schemes. We then computed the partial correlation between functional connectivity estimates for each edge and the relative mean RMS motion of each subject, controlling for subject age and sex, thus obtaining a distribution of edge-specific correlations with subject motion. From this distribution, we computed the percentage of edges for which the QC-FC correlations were statistically significant (p<0.05, no correction for multiple comparisons).

### 2.4.2. Distance-dependence of QC-FC correlations

Motion artifact has been known to have a distance-dependent effect on FC estimates. Short-range connections are affected by a combination of local and global motion artifact, while long-range connections tend to be affected only by global artifact (Power et al., 2015; Satterthwaite et al., 2012). To quantify this effect, we measured the correlation between the absolute values of QC-FC correlations (see Section 2.4.1) and the Euclidean distance separating the centroids of the node pair associated with each edge. This correlation served as a benchmark for the distance-dependence of the residual motion artifact.

### 2.4.3. Test-retest reliability

To evaluate the reliability of functional connectivity estimates, we computed the intra-class correlation (ICC) across different resting-state scans performed on the same subject in the HCP dataset. The intra-class correlation coefficient $\rho$ is defined (Shrout and Fleiss, 1979) as

$$\rho = \frac{MS_b - MS_w}{MS_b + (n-1)MS_w},$$

where $MS_b$ is the between-subject mean squared strength of each edge, $MS_w$ is the within-subject mean squared strength of each edge, and $n$ is the number of scans per subject, which in this case is 4.

### 2.4.4. Fingerprinting accuracy

We estimated the ability of resting-state functional connectivity profiles to accurately identify subjects in the HCP cohort using the approach described by Finn and colleagues (Finn et al., 2015). First, we evaluated all possible combinations of target and database pairs for the 4 available resting-state scans. Next, for each individual scan in the target set, we computed the similarity of estimated edge weights to all scans in the database set based on Pearson correlation. The predicted identity was that with the maximal Pearson correlation. Overall fingerprinting accuracy for each target-database pair was then calculated as the fraction of subjects whose identity was successfully predicted by their target resting-state scans.

### 2.4.5. System identifiability

To evaluate the possibility that more motion-resilient FC metrics might enable better detection of signals of interest, we consider the outcome measure of system identifiability (Ciric et al., 2017; Girvan and Newman, 2002; Newman, 2006). We use the term *system* to refer to a set of brain regions that are strongly functionally connected; and we use the phrase *system identifiability* to refer to the ease with which such systems can be detected from functional connectivity matrices. We employed the modularity quality index, $Q$, as a measure of system identifiability. The modularity quality index is a quantification of the extent to which a network can be subdivided into groups or modules characterized by strong intramodular connectivity and weak intermodular connectivity. Such modularity is indicative of the assortative community structure commonly observed in functional brain networks (Fortunato, 2010; Porter et al., 2009).

We estimated the modularity quality index for each subject's network by maximizing the modularity quality function originally defined by Newman (2006) and subsequently extended to weighted and signed networks by Rubinov and Sporns (2011), among others (Gómez et al., 2009; Traag and Bruggeman, 2009). For FC metrics resulting in weights falling within the interval [0,1], or results estimated from the absolute value of edge weights, we employed the weighted generalization of the modularity quality index. We first let the weight of a positive edge between nodes $i$ and $j$ be given by $a_{ij}^+$, and the strength of a node $i$, $s_i^\pm = \int_j a^{\pm ij}$, be given by the sum of the positive edge weights of $i$. We denote the chance expected within-module edge weights as $e_{ij}^+$ for positive weights where $e_{ij}^\pm = \frac{s_i^\pm s_j^\pm}{v^\pm}$. We let the total weight, $v^\pm = \int_{ij} a_{ij}^\pm$, be

the sum of all positive edge weights in the network. Then the weighted generalization of the modularity quality index is given by

$$Q^+ = \frac{1}{v^+} \sum_{ij} \left( a_{ij}^+ - \gamma e_{ij}^+ \right) \delta_{M_i M_j},$$

where $M_i$ is the community to which node $i$ is assigned, and $M_j$ is the community to which node $j$ is assigned. The Kronecker delta function, $\delta_{M_i M_j}$, takes on a value of 1 when $M_i = M_j$ and a value of 0 when $M_i \neq M_j$. The tunable structural resolution parameter, $\gamma$, scales the relative importance of the expected within-module weights (the null model) and in practice, the size of the communities; smaller or larger values of $\gamma$ result in correspondingly larger or smaller communities. We used a Louvain-like locally greedy algorithm (Blondel et al., 2008) as a heuristic to maximize this modularity quality index subject to a partition $M$ of nodes into communities.

For FC metrics resulting in weights falling in the interval [-1,1], specifically full and partial correlations, we employed the asymmetrically weighted generalization of $Q$ suitable for networks containing negative weights (Rubinov and Sporns, 2011). Specifically, we follow Rubinov and Sporns (2011) by first letting the weight of a positive edge between nodes $i$ and $j$ be given by $a_{ij}^+$, the weight of a negative edge between nodes $i$ and $j$ be given by $a_{ij}^-$, and the strength of a node $i$, $s_i^\pm = \int_j a^{\pm ij}$, be given by the sum of the positive or negative edge weights of $i$. We denote the chance expected within-module edge weights as $e_{ij}^+$ for positive weights and $e_{ij}^-$ for negative weights, where $e_{ij}^\pm = \frac{s_i^\pm s_j^\pm}{v^\pm}$. We let the total weight, $v^\pm = \int_{ij} a^\pm_{ij}$, be the sum of all positive or negative edge weights in the network. Then an asymmetric signed generalization of the modularity quality index can be written as

$$Q^* = \frac{1}{v^+} \sum_{ij} \left( a_{ij}^+ - \gamma e_{ij}^+ \right) \delta_{M_i M_j} - \frac{1}{v^+ + v^-} \sum_{ij} \left( a_{ij}^- - \gamma e_{ij}^- \right) \delta_{M_i M_j},$$

where $M_i$, $M_j$, $\delta_{M_i M_j}$, and $\gamma$ are defined as above.

We examined average $Q$ for each FC metric as a measure of system identifiability, as well as the partial correlation between $Q$ and mean relative RMS for each subject while controlling for average network weight, age, and sex. Additionally, we addressed two potential confounds that have not been previously addressed in work examining $Q$ as a measure of system identifiability: the number of communities $k$ detected during modularity maximization, and the mean and distribution of edge weights in a given network (see Supplementary Methods for details).

## 3. Results

### 3.1. Characteristics of FC matrices computed using different methods

We first characterized the functional connectivity edge weights estimated using different methods. Fig. 1 shows pairwise scatterplots between edge weights computed using all 8 methods. These plots show the non-linear relationships between edges estimated using full or partial correlation and non-correlation based methods. Of particular interest is the mapping from negative edges in full correlation to others. For instance, the weights of negative edges in Pearson matrices have an inverse relationship with the weights of edges in wavelet coherence matrices – the more negative a Pearson edge weight, the higher its wavelet coherence edge weight. In Fig. 2, FC matrices estimated using different methods are displayed as heatmaps, with canonical systems in the Gordon parcellation highlighted in the x and y color bars. Modular structure can clearly be seen in all matrices, with clean delineation of canonical systems in all matrices. Further, in full correlation matrices, and to a lesser extent in partial correlation matrices, well known negative associations are apparent, for instance between the default mode and dorsal attention systems (Buckner and DiNicola, 2019).

### 3.2. Full correlation shows high residual QC-FC correlations

Next, we evaluated the sensitivity of edge weights (computed using different methods) to subject motion. We used the residual QC-FC corre-
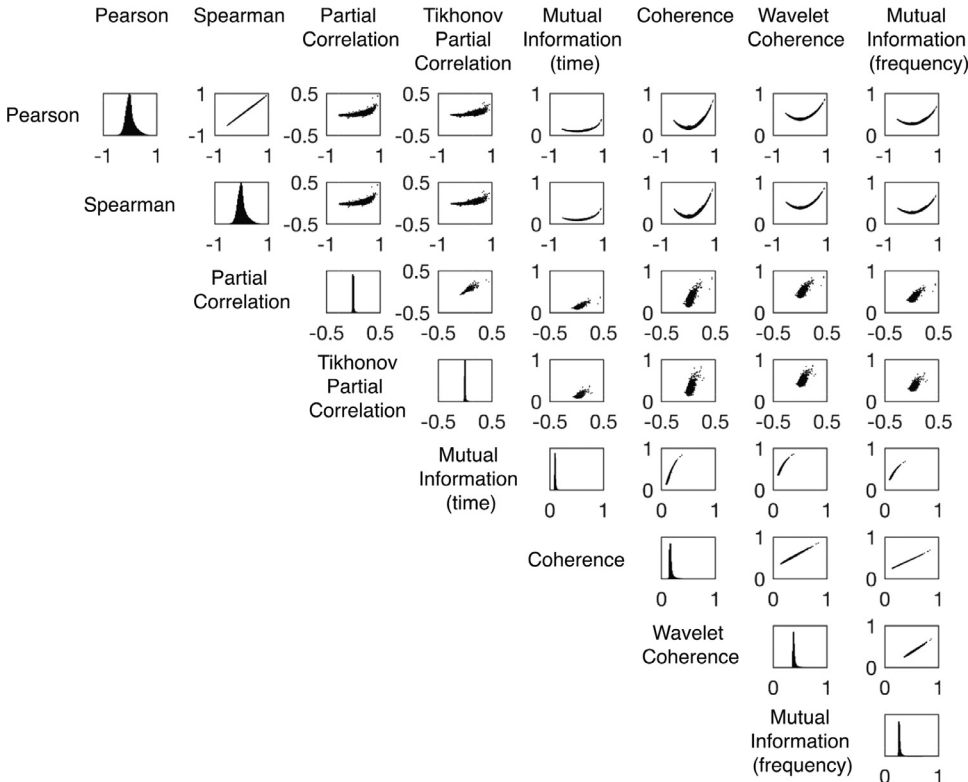


**Fig. 1.** Edge weight correlations for different FC metrics. Pairwise scatter plots between edge weights calculated using different FC methods. Diagonal entries show histograms of edge weights. Results are shown for the REST1_LR scan, with FC estimated using the 333-node Gordon parcellation.
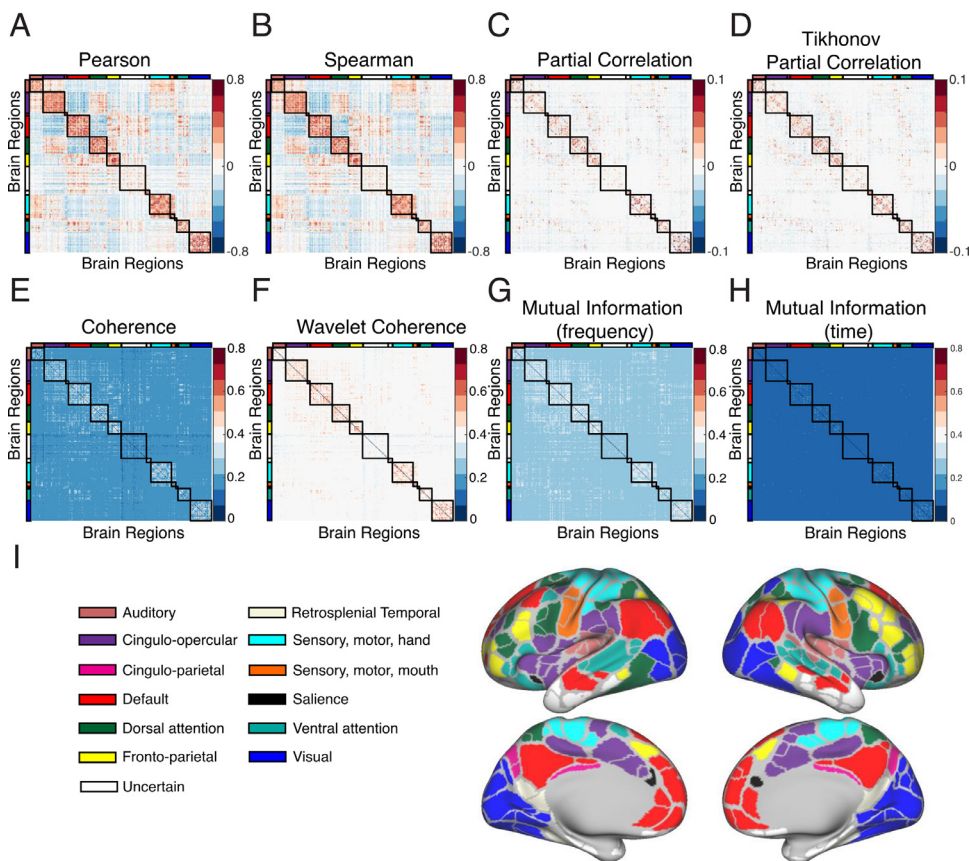
**Fig. 2.** FC matrices using different estimation methods. Edge weight heatmaps are shown for the 333-node Gordon parcellation, with canonical systems labeled as colored bars. Each entry in the heatmap is the average edge weight across 778 subjects in the REST1_LR scans, estimated using (A) Pearson correlation, (B) Spearman correlation, (C) partial correlation, (D) Tikhonov partial correlation, (E) coherence, (F) wavelet coherence, (G) mutual information (frequency), and (H) mutual information (time). (I) Canonical systems for the Gordon parcellation displayed on the HCP S1200 group average cortical surface.

lation benchmark, which measures the edgewise relationship between the relative mean RMS motion of each subject and their estimated edge weights, after the application of denoising pipelines.

Our findings are summarized in Fig. 3. Specifically, panels A–D of Fig. 3 show the fraction of edges for which the QC-FC correlations are statistically significant (p<0.05, no correction for multiple comparisons). Significantly, partial correlation and Tikhonov partial correlation are the best-performing FC methods, full correlation methods including Pearson and Spearman correlation are the worst-performing FC methods, and coherence- and mutual information-based metrics fall in between. Results largely remain similar when GSR and bandpass filtering are not applied, with the caveat that all FC methods other than partial correlation perform equally badly without the application of GSR. If only the absolute values of edge weights are taken or if negative edges are set to zero, the performance of Pearson correlation improves, but is still worse than other measures; similar trends are seen when correlations are not corrected for age and sex (Fig. S1). Results remain largely similar when motion is estimated using bandpass-filtered mean framewise displacement (Fig. S2). Further, the phase-encoding direction does not appear to have much impact on the number of edges significantly affected by motion (Fig. S3). Fig. 3E shows the distribution of QC-FC correlations for each FC estimation method. The distribution for full correlation methods is wider than for other methods, confirming that more edges are significantly associated with motion.

### 3.3. Motion differentially affects putative cognitive systems

Next, we analyzed the amount of motion artifact in edges connecting regions within specific putative cognitive systems. Fig. 4 shows heatmaps of QC-FC correlations for all edges in the Gordon parcellation, arranged by the associated *a priori* defined systems (Gordon et al., 2016). Heatmaps of QC-FC correlations averaged for each system pair are shown in Fig. S4. We also computed pairwise inter- and intra-

community QC-FC correlations and rank-ordered them by their median values. The six highest ranked inter-community QC-FC correlations are shown in Fig. 5. QC-FC correlations only within intra-system connections along with intra-network edge lengths and network sizes of all sub-networks in the Gordon parcellation are shown in Fig. S5. We did not include the 'uncertain' network (representing brain regions not assigned to any network due to susceptibility artifact) in this analysis.

Our analysis reveals a number of interesting details about the differential vulnerability of brain systems to motion artifact. Edges within the default mode (D-D) and the retrosplenial temporal systems (RT-RT) appear to be especially vulnerable to motion artifact in most FC estimation methods except partial correlation methods (Figs. 5 and S5). The increased vulnerability of the default and retrosplenial temporal networks to motion artifact does not appear to be due to the particular distribution of edge lengths within those networks. Intra-network edges in other sub-networks with similar intra-network edge lengths like the auditory and frontoparietal networks are notably less correlated with motion (Fig. S5). For full correlation methods, edges between cingulo-opercular and visual (CO-V), and between default and visual (D-V) systems have high QC-FC correlations but are not affected as much in the other FC measures.

### 3.4. Distance-dependence of motion artifact

Next, we evaluated the distance-dependence of motion artifact by measuring edgewise correlations between the Euclidean distance between nodes and the edge's absolute QC-FC correlation value. We find that edges estimated using full correlation have higher positive distance-dependence than other methods for most preprocessing variants, implying that long-distance edges are more affected by motion than short-distance edges (Fig. 6). It is to be noted that the distance-dependence of QC-FC correlations is overall quite low (Ciric et al., 2017), which likely reflects the stringent preprocessing protocol employed.
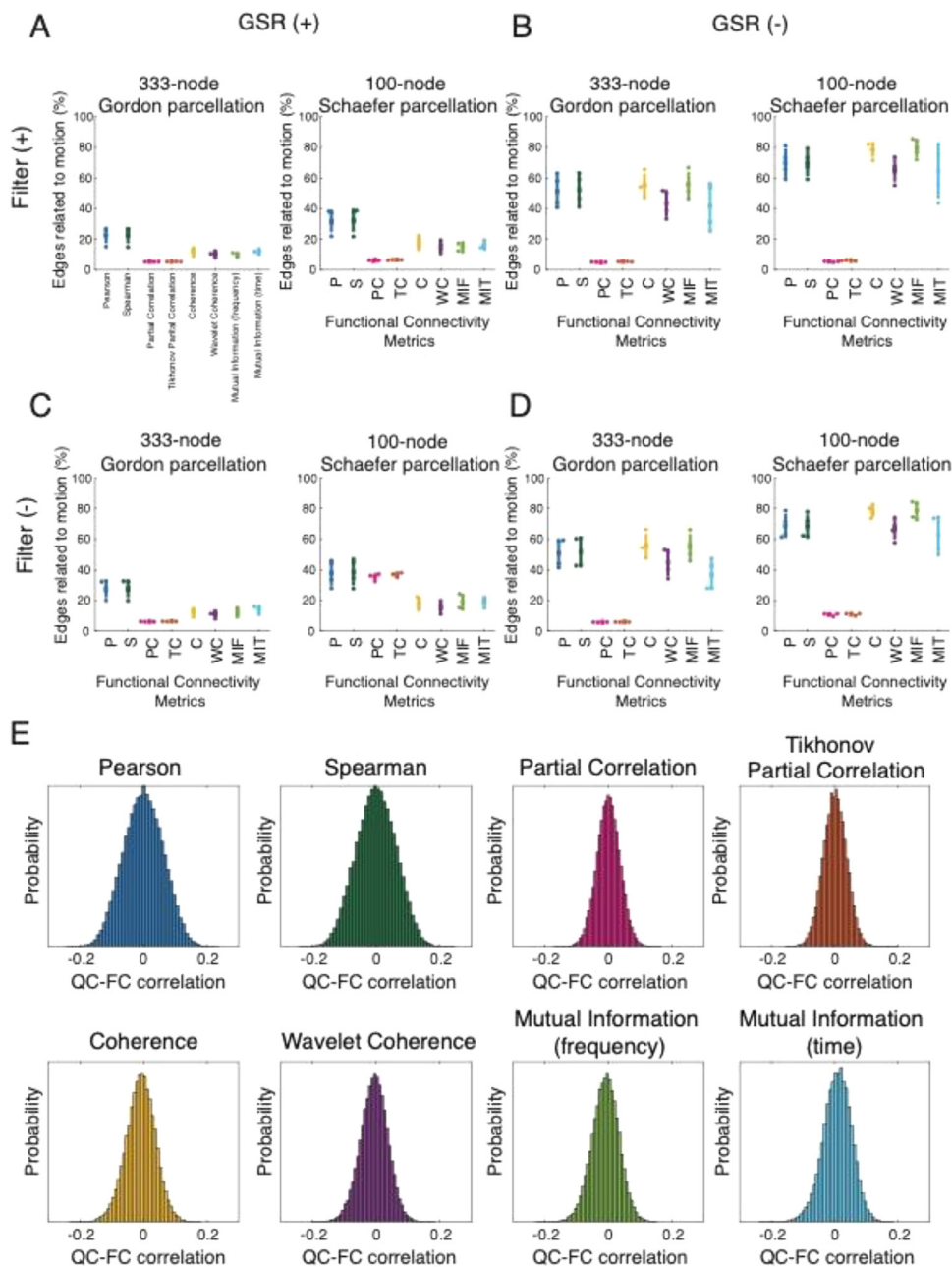
**Fig. 3.** Full correlation shows higher residual QC-FC correlations. (A–D) Fraction of edges significantly associated with motion for all 4 resting-state scans, estimated using the 333-node Gordon parcellation and the 100-node Schaefer parcellation. Different panels show results from different preprocessing variants defined by the inclusion or exclusion of global signal regression (GSR) and bandpass filtering. Metric names are shortened as follows: P=Pearson, S=Spearman, PC=Partial Correlation, TC=Tikhonov Partial Correlation, C=Coherence, WC=Wavelet Coherence, MIF=Mutual Information (frequency), and MIT=Mutual Information (time). Each plot contains 4 data points, one for each resting-state scan. Notches represent the mean and error bars show the standard deviation. (E) Histograms of QC-FC correlations for all estimation methods. Note the wider distributions for edge weights computed using full correlation (Pearson and Spearman correlation).

### 3.5. Reliability of functional connectivity estimates

To estimate the reproducibility in functional connectivity estimates with different methods, we measured the intra-class correlation across 4 resting-state scans in the HCP dataset. Fig. 7 shows that the intra-class correlations are highest for edges estimated using full correlation and lowest for edges estimated using partial correlation, for most preprocessing variants. The relatively high reliability of full correlation edges could be due to accurate estimates of trait-like biology or could be due to sensitivity to a highly reliable third-party variable, such as motion. To determine whether the latter could be the case, we evaluated the reliability of subject motion. We found that the intra-class correlation for relative RMS motion was also high (0.72), indicating that motion itself is reproducible across scans. In order to separate the reproducibility of motion from reproducibility of FC edges, we re-computed the intra-class correlations for edges that were in the bottom 20% of absolute QC-FC correlation values in all 4 scans. This analysis showed that results re-

mained largely similar after mitigating the influence of reliable motion, indicating that edges estimated using full correlation are more reproducible over this time scale than edges estimated using other methods (Fig. S6).

Next, we evaluated the ability of resting-state functional connectivity estimates to reliably predict subject identity using the measure of fingerprinting accuracy. Fig. 8 shows that fingerprinting accuracy is comparably high among full correlation, coherence, and mutual information metrics, but is quite low among partial correlation metrics when bandpass filtering is applied. Interestingly, partial correlation metrics show high fingerprinting accuracy when bandpass filtering is not applied.

### 3.6. System Identifiability

Functional brain networks are organized into systems, or modules, that are indicated by strong intramodular connectivity and weak in-
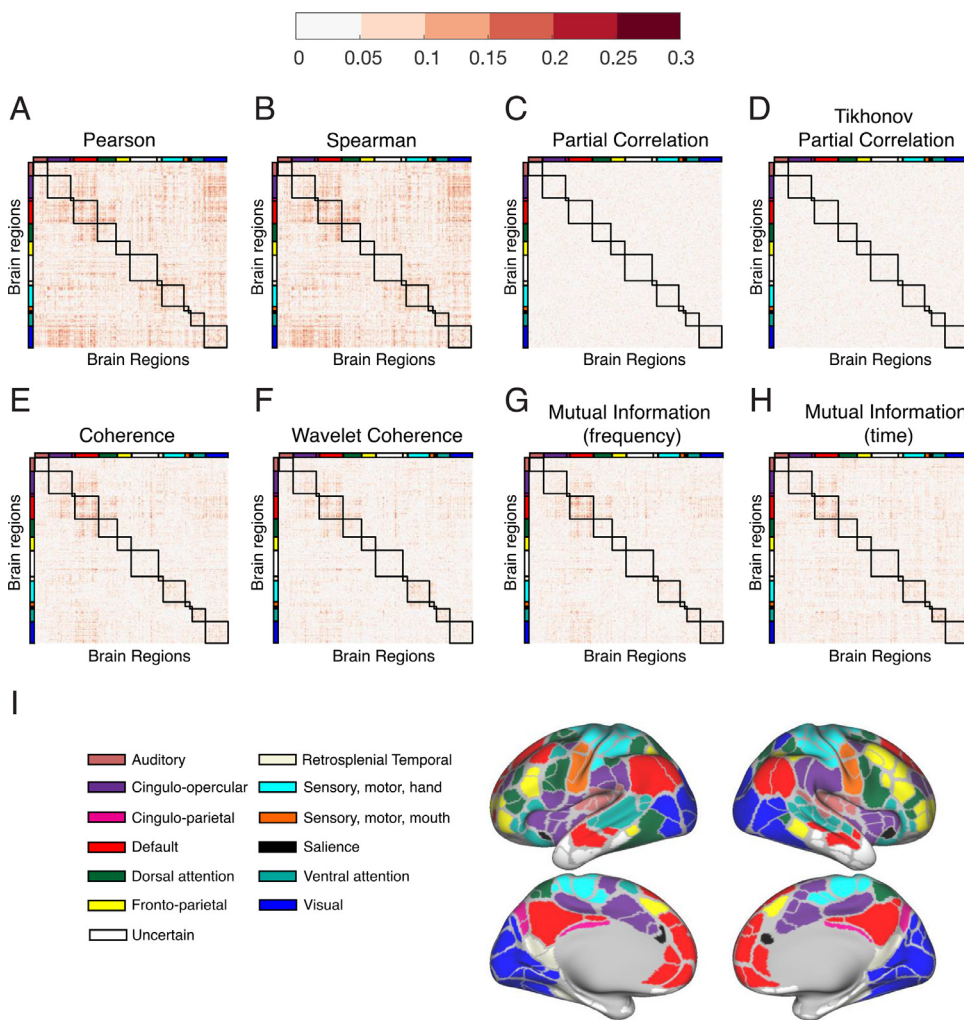
**Fig. 4.** Distribution of QC-FC correlations across functional brain networks. QC-FC correlation heatmaps are shown for the REST1_LR scan, for the 333-node Gordon parcellation, and with canonical *a priori* defined systems labeled as colored bars. Each entry in the heatmap is the absolute value of the QC-FC correlation (across subjects), with edge weights estimated using (A) Pearson correlation, (B) Spearman correlation, (C) partial correlation, (D) Tikhonov partial correlation, (E) coherence, (F) wavelet coherence, (G) mutual information (frequency), and (H) mutual information (time). (I) Canonical systems for the Gordon parcellation displayed on the HCP S1200 group average cortical surface.
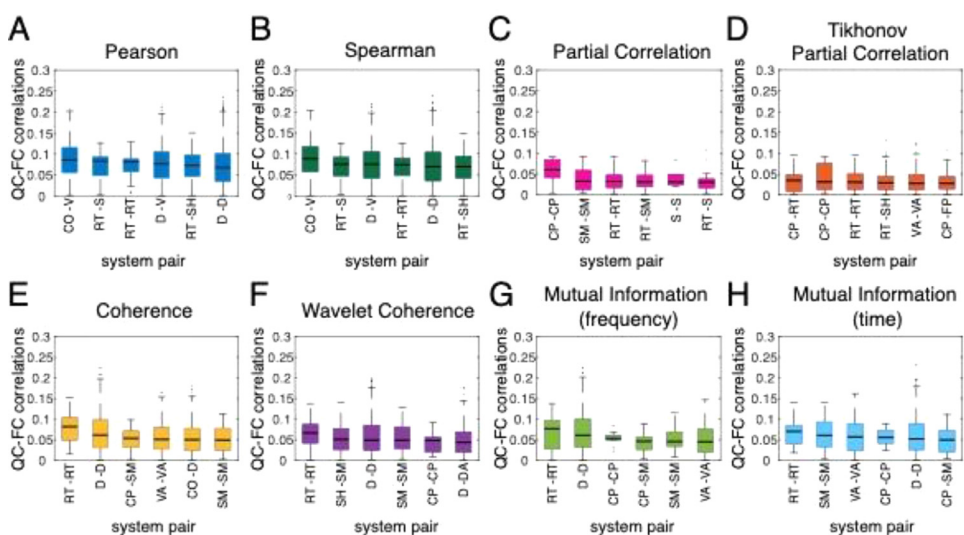


**Fig. 5.** System affiliations of edges highly impacted by motion. Boxplots of inter-system QC-FC correlations are shown for the REST1_LR scan and the 333-node Gordon parcellation. All inter-system edges were rank-ordered by their median absolute QC-FC correlation, and the top 6 ranking distributions are shown for each FC metric. CO=cingulo-opercular; CP=cingulo-parietal; D=default; DA=dorsal attention; FP=fronto-parietal; RT=retrosplenial temporal; SH=sensory, motor, hand; SM=sensory, motor, mouth; S=salience: VA=ventral attention; V=visual. Note that the 'None' system representing nodes not assigned to any sub-network due to susceptibility artifact is not included in this analysis.

termodular connectivity (Fortunato, 2010; Porter et al., 2009). A large body of literature has characterized this quintessential principle of brain network organization, commonly detecting between 6 and 8 systems (Smith et al., 2009; Thomas Yeo et al., 2011; Uddin et al., 2019). Thus, we examined the extent to which different metrics of FC resulted in different levels of system identifiability, as operationalized by the mod-

ularity quality index Q, which captures how separable a network is into coherent modules or systems (Bertolero et al., 2015; Newman, 2006; Sporns and Betzel, 2016). Fig. 9 shows that on average, our measure of system identifiability, the modularity quality index Q, is highest in systems estimated using full correlation, lowest in all non-correlation based methods and intermediate for partial correlation.
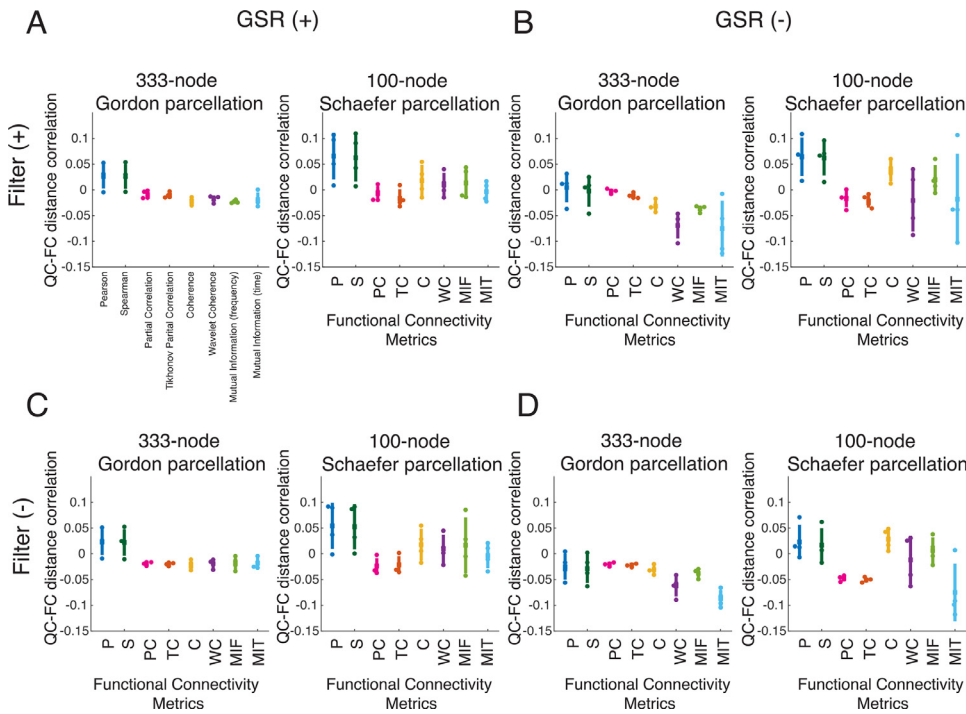
**Fig. 6.** Distance-dependence of QC-FC correlations. Residual distance-dependence of motion artifact for all 4 resting-state scans, estimated using the 333-node Gordon parcellation and the 100-node Schaefer parcellation. Panels (A)–(D) show results from different preprocessing variants defined by the inclusion or exclusion of global signal regression (GSR) and bandpass filtering. Metric names are shortened as follows: P=Pearson, S=Spearman, PC=Partial Correlation, TC=Tikhonov Partial Correlation, C=Coherence, WC=Wavelet Coherence, MIF=Mutual Information (frequency), and MIT=Mutual Information (time). Each plot contains 4 data points, one for each resting-state scan. Notches represent the mean and error bars show the standard deviation.
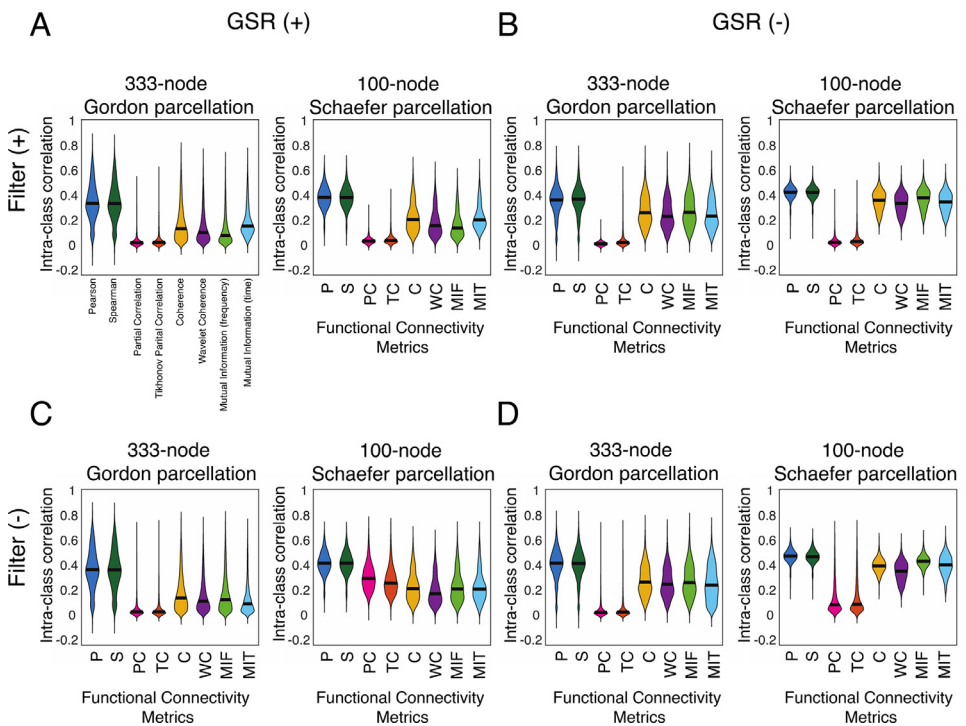


**Fig. 7.** Test-retest reliability of edge weights for different FC metrics. Intra-class correlation for all edge weights across 4 resting-state scans, estimated using the 333-node Gordon parcellation and the 100-node Schaefer parcellation. Violin plots show the distribution of intra-class correlation values across all subjects. Panels A–D show results from different preprocessing variants defined by the inclusion or exclusion of global signal regression (GSR) and bandpass filtering. Metric names are shortened as follows: P=Pearson, S=Spearman, PC=Partial Correlation, TC=Tikhonov Partial Correlation, C=Coherence, WC=Wavelet Coherence, MIF=Mutual Information (frequency), and MIT=Mutual Information (time).

The relatively high system identifiability in metrics based on full correlation could be due to an accurate sensitivity to the biology of putative cognitive systems or could be due to motion impacting regions of the brain in a spatially heterogeneous manner that partially drives the data-driven identification of systems. To address these possibilities, we studied the relation between $Q$ and motion. We find that the relationship between $Q$ and motion is also strongest in full correlation compared to all other FC metrics, even when controlling for average weight, age, and sex (Fig. 10). Without the application of bandpass filtering, partial correlation methods also show a strong relationship between $Q$ and motion, but the relationship between $Q$

and motion was similar among all FC methods without application of GSR.

We also estimated $Q$ from networks containing only the absolute values of edge weights, which reduces but does not eliminate differences in system identifiability between correlation-based metrics and other metrics (Fig. S7). To ensure that differences in system identifiability were not driven by differences in the functional systems detected when maximizing the modularity quality index, we also calculated $Q$ using the canonical system partition associated with each of our 2 parcellations and obtained similar results (Fig. S7). Further, relationships between $Q$ and motion remained similar when not corrected for age and sex, and
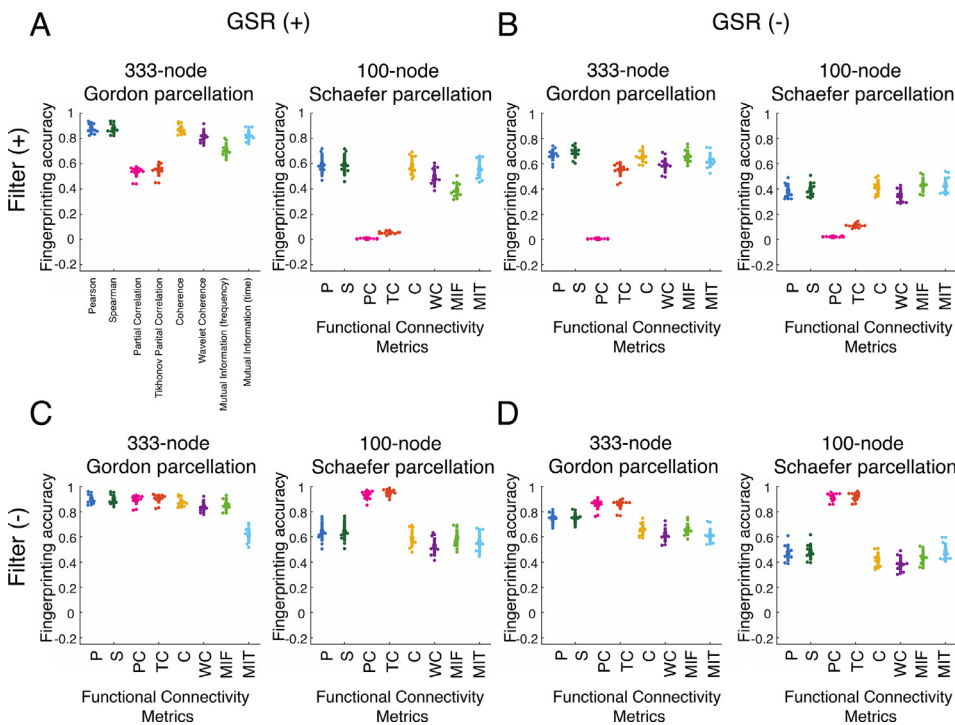
**Fig. 8.** Resting-state functional connectivity fingerprinting accuracy for different FC metrics. Accuracy of fingerprinting using 12 combinations of database and target pairs from 4 resting-state scans, estimated using the 333-node Gordon parcellation and the 100-node Schaefer parcellation. Panels A–D show results from different preprocessing variants defined by the inclusion or exclusion of global signal regression (GSR) and bandpass filtering. Metric names are shortened as follows: P=Pearson, S=Spearman, PC=Partial Correlation, TC=Tikhonov Partial Correlation, C=Coherence, WC=Wavelet Coherence, MIF=Mutual Information (frequency), and MIT=Mutual Information (time). Each plot contains 12 data points, one for each combination of database and target pairs. Notches represent the mean and error bars show the standard deviation.
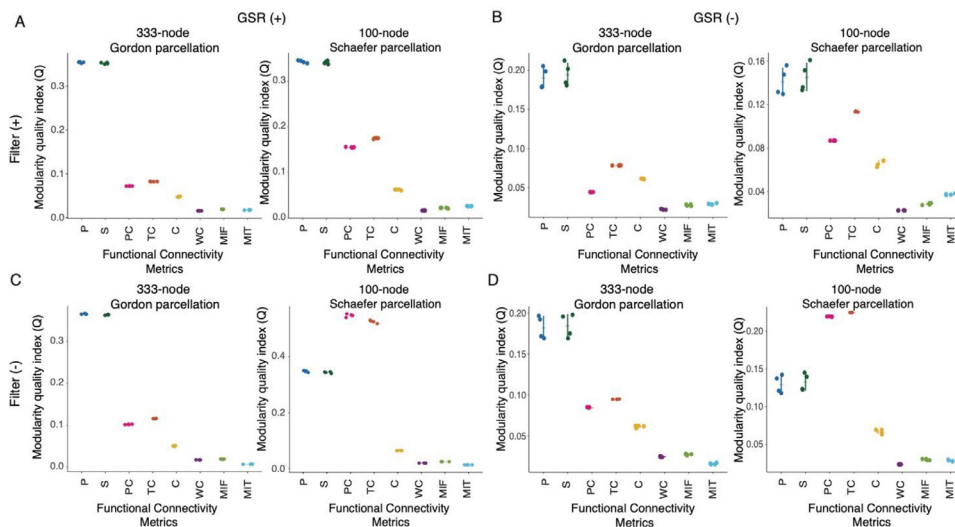


**Fig. 9.** System identifiability is highest for full correlation. The modularity quality index Q is highest in networks estimated using full correlation, lowest for non-correlation based metrics and intermediate for partial correlation. Modularity maximization was tuned across subjects to estimate between 6 and 8 communities for each metric. Panels A-D show results from different preprocessing variants defined by the inclusion or exclusion of global signal regression (GSR) and bandpass filtering. P=Pearson, S=Spearman, PC=Partial Correlation, TC=Tikhonov Partial Correlation, MIT=Mutual Information (time), C=Coherence, WC=Wavelet Coherence, MIF=Mutual Information (frequency). Each plot contains 4 data points, one for each resting-state scan. Notches represent the mean and error bars show the standard deviation.

when motion was calculated as mean FD using bandpass filtered realignment estimates (Fig. S7).

To further ensure that results were not driven by variability in edge weight distributions across FC metrics, we examined two boundary cases from prior results as the FC metrics of interest: Pearson correlation and wavelet coherence. We reordered the edge weight values in our *values* matrix (see Supplementary Methods) to reflect the rank order of weights in the *ranks* matrix. This process preserves the topology of the network contained in the *ranks* matrix while ensuring that both the average edge weight and distribution of edge weights are the same as those in the *values* network. We estimated the modularity quality index *Q* of these reordered matrices and found that *Q* was consistently higher when the rank ordering matrix was derived from Pearson correlations (Fig. 11). Taken together, these results suggest that correlation-based FC metrics consistently result in higher levels of system identifiability, and that they may better reflect the modular architecture of functional brain networks than other methods.

## 4. Discussion

In this report, we systematically investigated the sensitivity to motion of 8 different FC estimation measures drawn from the correlation, coherence, and mutual information families, based on their performance on commonly used benchmarks. The context, implications, and limitations of our results are discussed below.

### 4.1. Clear distinction between correlation-based FC measures and other measures

Our main finding is that FC edges estimated using full correlation result in a high fraction of edges significantly correlated with motion and a relatively high distance-dependence of motion artifact compared to all other methods. By contrast, using partial correlation methods almost completely eliminated the relationship between estimated edge weights and motion as well as their distance-dependence. These results were
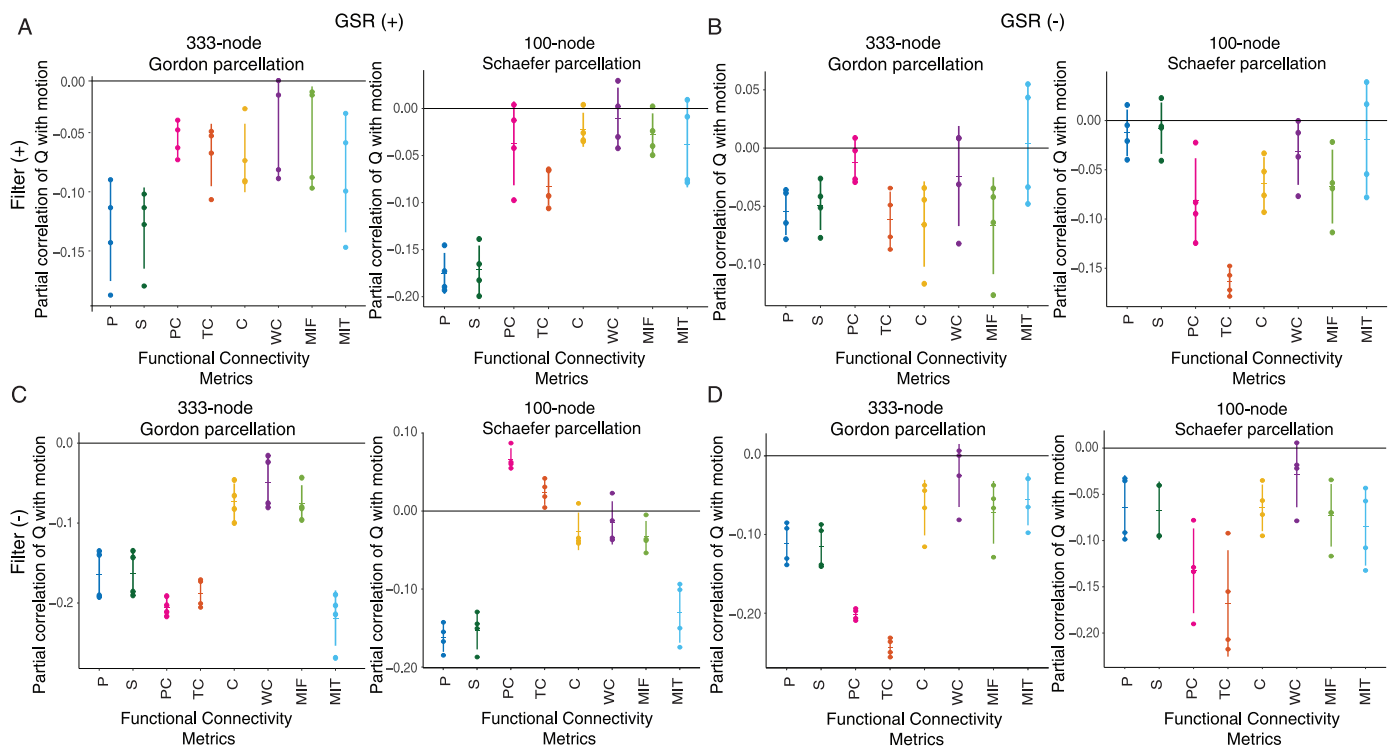
**Fig. 10.** Modularity quality index is highly related to motion in full correlation metrics. Partial correlation between Q and mean relative RMS, controlling for average network weight, age, and sex. Edges estimated using full correlation show large negative correlations between subject motion and Q. Panels A–D show results from different preprocessing variants defined by the inclusion or exclusion of global signal regression (GSR) and bandpass filtering. P=Pearson, S=Spearman, PC=Partial Correlation, TC=Tikhonov Partial Correlation, MIT=Mutual Information (time), C=Coherence, WC=Wavelet Coherence, MIF=Mutual Information (frequency). Each plot contains 4 data points, one for each resting-state scan. Notches represent the mean and error bars show the standard deviation.

largely maintained when global signal regression and bandpass filtering were not applied.

Head motion artifact predominantly manifests as spurious signal fluctuations in BOLD signal across multiple voxels in the brain (Ciric et al., 2018; Power et al., 2015). Since traditional functional connectivity based on full correlation measures temporal covariance, it follows that such measures of connectivity are directly impacted by artifactual covariance introduced by head motion. Partial correlation estimates, on the other hand, measure the temporal covariance between time series after regressing out all other time series in the network, thereby eliminating shared covariance induced by head motion artifact. Among the other FC methods we evaluated, coherence-based FC measures quantify statistical dependencies in the frequency domain, including phase locking and correlation in power spectra. We also evaluated mutual information in the frequency domain (Salvador et al., 2005), which is an information theoretic measure of relationships in the frequency domain. Frequency-based FC measures are less likely to be influenced by short-lived temporal fluctuations in the BOLD signal. Taken together, our findings indicate that the statistical properties of full correlation render them relatively more sensitive to temporal outliers introduced by head motion, an effect that is reduced by using frequency-based connectivity estimation, and effectively eliminated by regressing out common sources of artifact through partial correlation.

In our study, we averaged frequency-based connectivity estimates (coherence, wavelet coherence, and mutual information in frequency) within a low frequency band (0.009–0.08Hz). Although information on the power spectral properties of motion artifact is limited, some prior studies have shown that motion affects the spectral power and connectivity estimates mainly at high frequencies (Kim et al., 2014; Salvador et al., 2008; Satterthwaite et al., 2013a). It is therefore possible that averaging connectivity estimates within a low frequency band reduced the impact of high-frequency motion artifact in these measures.

Further, if motion artifact manifested in any one given frequency, the process of averaging in multiple frequency bands may have diluted the overall impact of motion on the FC estimates.

While much previous work has shown the utility of bandpass filtering fMRI time series to mitigate motion effects (Caballero-Gaudes and Reynolds, 2017; Power et al., 2014; Satterthwaite et al., 2013b), the shape and design of the implemented filter are important considerations, especially while evaluating FC measures in the frequency domain. For instance, Butterworth filters have been known to introduce edge effects, artifacts that can be mitigated by discarding volumes before and after the cutoff bands (Caballero-Gaudes and Reynolds, 2017; Power et al., 2014). Bandpass filtering using higher-order filters may also introduce spurious oscillations close to the cutoff frequencies (de Cheveigné and Nelken, 2019). Care must be taken to ensure that such artifacts do not adversely impact the calculation of frequency-based estimates of connectivity.

We found that the performance of Pearson correlation on the QC-FC benchmark improved when taking the absolute values of edge weights, and when negative edges were set to zero (Fig. S1). Analysis of fully connected complex networks with positive and negative weights can be rigorously performed (Rubinov and Sporns, 2011). However, the interpretation of negative correlations is controversial, especially in the context of global signal regression (Anderson et al., 2011; Chai et al., 2012; Murphy and Fox, 2017). As a result, many studies omit negative edges from analyses of functional and dynamic connectivity (Chan et al., 2014; Grady et al., 2016). Our results indicate that omitting negative edges or taking their absolute values might also reduce the susceptibility to motion artifact, although it is also possible that this step decreases sensitivity to individual differences.

The systems whose edges were most affected by motion differed among FC estimation methods. For instance, connections between large-scale systems such as the default mode, cingulo-opercular and visual
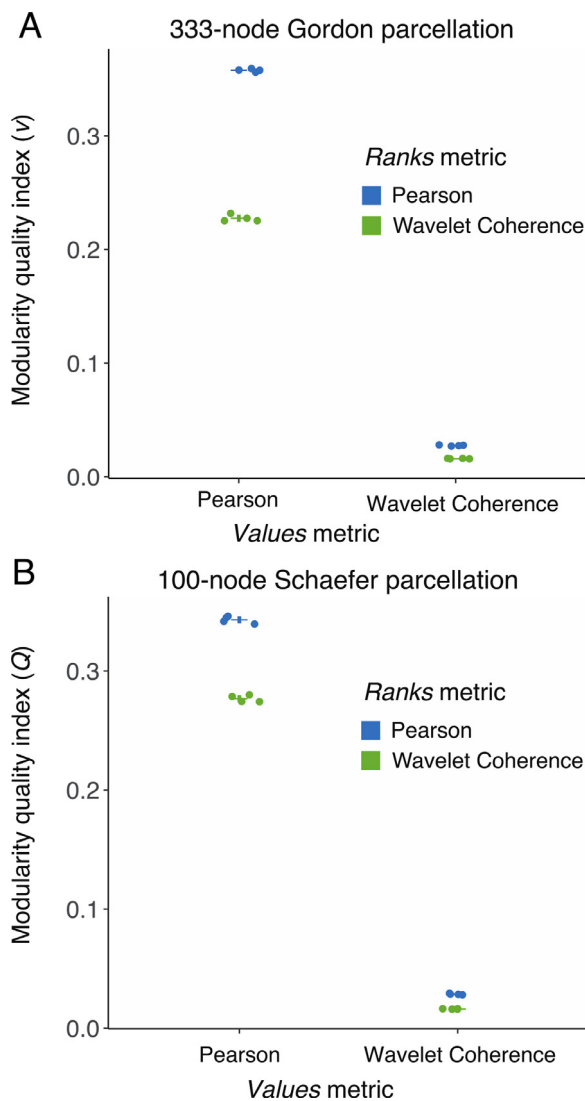
**Fig. 11.** A correlation-based metric results in higher system identifiability even when using edge weight ranks rather than edge weight values. (A, B) Holding the edge weight value distribution constant, higher Q is found when using the ranks of edge weights of a time-based metric (Pearson) than a time-frequency-based metric (Wavelet Coherence). Datapoints where values network and ranks network are the same (e.g. Pearson values-Pearson ranks) recapitulate results in Fig. 9A and are shown for ease of comparison. Notches represent the mean and error bars show the standard deviation.

systems were correlated with motion in full correlation but not in other FC methods. In contrast, within-system edges in the default mode system and retrosplenial temporal cortex were strongly related to motion in all FC estimation methods other than partial correlation methods. This last result deserves attention, given the large number of scientific hypotheses surrounding the brain's default mode system (Buckner and DiNicola, 2019; Fox and Raichle, 2007).

The strong relationship between motion and default mode connectivity is unlikely solely due to geometry because other networks whose edges are less correlated with motion, including the frontoparietal system, are similarly distributed with anterior and posterior nodes on both medial and lateral surfaces, with similar intra-network edge lengths. It is possible that individuals with specific patterns of default mode connectivity find it more difficult to remember to stay still as their minds wander (Christoff et al., 2009; Golchert et al., 2017; Kajimura et al., 2016). Retrosplenial cortex, another system that we found to have a strong relationship with motion, is often functionally integrated with the default

mode system to support memory processes, but also plays a role in spatial navigation and locomotion (Fischer et al., 2020; Mao et al., 2020; Vann et al., 2009). Specifically, retrosplenial cortex integrates vestibular input, which encodes head position, with visual cortex, to calibrate self-motion with visual motion signals (Chaplin and Margrie, 2020; Vélez-Fort et al., 2018). Our findings highlight the need to carefully consider the confounding effects of motion, as well as the causes of motion, while studying these systems.

Significantly, full correlation methods scored highest and partial correlation methods scored lowest on test-retest reliability and fingerprinting accuracy. While the high reliability of edges estimated using full correlation was partly caused by the well-known reproducibility of motion itself (Noble et al., 2019; van Dijk et al., 2012), we observed similar results when restricting our analysis to the 20% of edges with the lowest absolute QC-FC correlations for all 4 scans. This additional finding suggests that full correlation, while highly sensitive to motion, leads to relatively reproducible functional connectivity estimates. On the other hand, edges estimated using partial correlation, while not associated with motion, vary highly between scans.

Perhaps surprisingly, full correlation resulted in significantly higher system identifiability than other measures. This observation suggests that findings of higher system identifiability in full correlation methods are not solely driven by motion, and that these metrics, while highly motion-sensitive, might excel in detecting coherent community structure. Alternatively, these findings may hint that the well-established finding of a modular architecture in human functional brain networks may be relatively metric-dependent (Bertolero et al., 2015; Sporns and Betzel, 2016).

The success with which correlation-based measures detect modular architecture may be due to the presence of negative edge weights or anticorrelations in these measures, which contribute to reducing inter-module connections in calculations of modularity quality. A negative edge calculated using Pearson correlation, for instance between the default mode and dorsal attention systems, reduces the overall inter-module connectivity, crystallizing the boundaries between modules. However, the same edge calculated using coherence is highly positive, increasing inter-module connectivity and obfuscating boundaries between modules. Further, taking absolute values of correlation-based edges acts as a similar transformation, converting negative edges to positive edges, reducing the modularity quality index. The presence of negative edges or anticorrelation between internal and external attention systems has been argued to reflect a functional toggle between systems (Chai et al., 2012; Clare Kelly et al., 2008; Gao and Lin, 2012; Murphy et al., 2019; Owens et al., 2018). It is therefore unclear which method best captures true interactions between systems. Indeed, both could reflect complementary aspects of network dynamics if dorsal attention activation lags default mode activation at a consistent delay.

### 4.2. Implications for researchers

We have shown that moving away from standard FC metrics based on full correlation can improve the robustness of FC estimates to head motion. However, full correlation excels at detecting community structure and is highly reliable. Our findings indicate that the FC estimation method should be chosen carefully based on the nature of the study. For instance, studies on group comparisons, where motion artifact can introduce systematic bias in connectivity estimates (Power et al., 2015), could benefit from using frequency-based FC estimation methods like coherence. The appropriate choice for studies on modular brain architecture, in contrast, might remain full correlation. Notably, partial correlation offers a best-of-both-worlds option – low contamination by motion artifact and relatively high system identifiability, with the caveat of low reliability. Our results also highlight a spatial heterogeneity in the impact of motion. FC edges in the default mode and retrosplenial cortex were especially sensitive to the effects of motion. Studies that explore the fine-scale organization and function of these networks could bene-

fit from exploring different choices of FC estimation. Finally, full and partial correlation methods are computationally less expensive than the rest of the measures reported here and might therefore be preferable for large datasets where computational time is at a premium.

### 4.3. Limitations

It is prudent to mention several limitations of our study. First, the lack of a noise-free ground truth is a challenge while estimating the impact of motion artifact in real fMRI data. It is difficult to separate out true signal from noise in fMRI data, a challenge further complicated by findings that head motion is a stable trait, and likely related to an individual's physiology and neural dynamics (Beyer et al., 2020; Engelhardt et al., 2017). Even our estimation of system identifiability lacks a true ground truth, as much prior work on systems in functional brain networks has employed correlation-based metrics, though newer studies using direct intracranial recordings also recover coherent modular structure in functional brain networks (Betzel et al., 2019). Second, due to the lack of ground truth, it is necessary to rely on indirect benchmarks such as QC-FC correlations and QC-FC distance-dependence. Central to the computation of these benchmarks is the estimation of an average measure of head motion for the whole scan from realignment estimates (Power et al., 2015). This average measure can miss important spatiotemporal details of motion. Future studies could use voxel-wise displacement maps to extract more detailed information about motion and its impact on FC (Satterthwaite et al., 2013a). Third, we used data from the Human Connectome Project that was preprocessed using ICA-FIX. This denoising approach has been shown to be particularly effective with HCP data (Salimi-Khorshidi et al., 2014). In the future, it might be beneficial to investigate the effect of varying FC estimation methods with more noisy datasets with different denoising pipelines. Further, because we imposed a fairly stringent motion exclusion threshold, it is unclear whether our results generalize to samples with higher motion including pediatric, geriatric, or psychiatric samples. Finally, we did not evaluate FC estimation methods from many statistical families, including Bayes nets, Granger causality, and generalized synchronization. Future studies could investigate additional families of FC estimation methods omitted from this study.

### 5. Citation diversity statement

Recent work in several fields of science has identified a bias in citation practices such that papers from women and other minority scholars are under-cited relative to the number of such papers in the field (Caplar et al., 2017; Dion et al., 2018; Dworkin et al., 2020; Maliniak et al., 2013; Mitchell et al., 2013). Here we sought to proactively consider choosing references that reflect the diversity of the field in thought, form of contribution, gender, race, ethnicity, and other factors. First, we obtained the predicted gender of the first and last author of each reference by using databases that store the probability of a first name being carried by a woman (Dworkin et al., 2020; Zhou et al., 2020). By this measure (and excluding self-citations to the first and last authors of our current paper), our references contain 12.34% woman(first)/woman(last), 10.28% man/woman, 22.62% woman/man, and 54.76% man/man. This method is limited in that (a) names, pronouns, and social media profiles used to construct the databases may not, in every case, be indicative of gender identity and (b) it cannot account for intersex, non-binary, or transgender people. Second, we obtained predicted racial/ethnic category of the first and last author of each reference by databases that store the probability of a first and last name being carried by an author of color (Ambekar et al., 2009; Sood and Laohaprapanon, 2018). By this measure (and excluding self-citations), our references contain 7.42% author of color (first)/author of color(last), 11.39% white author/author of color, 27.14% author of color/white author, and 54.05% white author/white author. This method is limited in that (a) names and Florida Voter Data to make

the predictions may not be indicative of racial/ethnic identity, and (b) it cannot account for Indigenous and mixed-race authors, or those who may face differential biases due to the ambiguous racialization or ethnicization of their names. We look forward to future work that could help us to better understand how to support equitable practices in science.

### Declaration of Competing Interest

The authors declare no competing interests.

### Credit authorship contribution statement

**Arun S. Mahadevan:** Methodology, Software, Formal analysis, Visualization, Writing – original draft. **Ursula A. Tooley:** Methodology, Software, Formal analysis, Visualization, Writing – original draft. **Maxwell A. Bertolero:** Data curation, Writing – original draft. **Allyson P. Mackey:** Supervision, Writing – review & editing. **Danielle S. Bassett:** Conceptualization, Supervision, Formal analysis, Writing – review & editing.

### Data/code availability statement

All data used in this manuscript comes from the open-source Human Connectome Project. Code associated with this manuscript is available at https://github.com/BassettLab/motion-FC-metrics.git.

### Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2021.118408.

### References

Achard, S., Bullmore, E., 2007. Efficiency and cost of economical brain functional networks. PLoS Comput. Biol. 3, 0174–0183. doi:10.1371/journal.pcbi.0030017.

Achard, S., Salvador, R., Whitcher, B., Suckling, J., Bullmore, E., 2006. A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs. J. Neurosci. 26, 63–72. doi:10.1523/JNEUROSCI.3874-05.2006.

Agrawal, U., Brown, E.N., Lewis, L.D., 2020. Model-based physiological noise removal in fast fMRI. Neuroimage 205, 116231. doi:10.1016/j.neuroimage.2019.116231.

Ambekar, A., Ward, C., Mohammed, J., Male, S., Skiena, S., 2009. Name-ethnicity classification from open sources. In: Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 49–57. doi:10.1145/1557019.1557032.

Anderson, J.S., Druzgal, T.J., Lopez-Larson, M., Jeong, E.K., Desai, K., Yurgelun-Todd, D., 2011. Network anticorrelations, global regression, and phase-shifted soft tissue correction. Hum. Brain Mapp. 32, 919–934. doi:10.1002/hbm.21079.

Bassett, D.S., Sporns, O., 2017. Network neuroscience. Nat. Neurosci. 20, 353–364. doi:10.1038/nn.4502.

Beauchene, C., Roy, S., Moran, R., Leonessa, A., Abaid, N., 2018. Comparing brain connectivity metrics: a didactic tutorial with a toy model and experimental data. J. Neural Eng. 15, 56031. doi:10.1088/1741-2552/aad96e.

Bertolero, M.A., Blevins, A.S., Baum, G.L., Gur, R.C., Gur, R.E., Roalf, D.R., Satterthwaite, T.D., Bassett, D.S., 2019. The network architecture of the human brain is modularly encoded in the genome. arXiv Prepr. arXiv:1905.07606.

Bertolero, M.A., Thomas Yeo, B.T., D'Esposito, M., 2015. The modular and integrative functional architecture of the human brain. Proc. Natl. Acad. Sci. U. S. A. 112, E6798–E6807. doi:10.1073/pnas.1510619112.

Betzel, R.F., Medaglia, J.D., Kahn, A.E., Soffer, J., Schonhaut, D.R., Bassett, D.S., 2019. Structural, geometric and genetic factors predict interregional brain connectivity patterns probed by electrocorticography. Nat. Biomed. Eng. 3, 902–916. doi:10.1038/s41551-019-0404-5.

Beyer, F., Prehn, K., Wüsten, K.A., Villringer, A., Ordemann, J., Flöel, A., Witte, A.V., 2020. Weight loss reduces head motion: Revisiting a major confound in neuroimaging. Hum. Brain Mapp. 1–5. doi:10.1002/hbm.24959.

Blondel, V.D., Guillaume, J.L., Lambiotte, R., Lefebvre, E., 2008. Fast unfolding of communities in large networks. J. Stat. Mech. Theory Exp. 2008, 0–12. doi:10.1088/1742-5468/2008/10/P10008.

Buckner, R.L., DiNicola, L.M., 2019. The brain's default network: updated anatomy, physiology and evolving insights. Nat. Rev. Neurosci. doi:10.1038/s41583-019-0212-7.

Buckner, R.L., Krienen, F.M., Yeo, B.T.T., 2013. Opportunities and limitations of intrinsic functional connectivity MRI. Nat. Neurosci. 16, 832–837. doi:10.1038/nn.3423.

Bullmore, E., Fadili, J., Maxim, V., Şendur, L., Whitcher, B., Suckling, J., Brammer, M., Breakspear, M., 2004. Wavelets and functional magnetic resonance imaging of the human brain. Neuroimage 23, 234–249. doi:10.1016/j.neuroimage.2004.07.012.

Burgess, G.C., Kandala, S., Nolan, D., Laumann, T.O., Power, J.D., Adeyemo, B., Harms, M.P., Petersen, S.E., Barch, D.M., 2016. Evaluation of denoising strategies to address motion-correlated artifacts in resting-state functional magnetic resonance imaging data from the human connectome project. Brain Connect. 6, 669–680. doi:10.1089/brain.2016.0435.

Caballero-Gaudes, C., Reynolds, R.C., 2017. Methods for cleaning the BOLD fMRI signal. Neuroimage 154, 128–149. doi:10.1016/j.neuroimage.2016.12.018.

Caplar, N., Tacchella, S., Birrer, S., 2017. Quantitative evaluation of gender bias in astronomical publications from citation counts. Nat. Astron. 1. doi:10.1038/s41550-017-0141.

Chai, X.J., Castañán, A.N., Öngür, D., Whitfield-Gabrieli, S., 2012. Anticorrelations in resting state networks without global signal regression. Neuroimage 59, 1420–1428. doi:10.1016/j.neuroimage.2011.08.048.

Chan, M.Y., Park, D.C., Savalia, N.K., Petersen, S.E., Wig, G.S., 2014. Decreased segregation of brain systems across the healthy adult lifespan. Proc. Natl. Acad. Sci. U. S. A. 111, E4997–E5006. doi:10.1073/pnas.1415122111.

Chaplin, T.A., Margrie, T.W., 2020. Cortical circuits for integration of self-motion and visual-motion signals. Curr. Opin. Neurobiol. 60, 122–128. doi:10.1016/j.conb.2019.11.013.

Christoff, K., Gordon, A.M., Smallwood, J., Smith, R., Schooler, J.W., 2009. Experience sampling during fMRI reveals default network and executive system contributions to mind wandering. Proc. Natl. Acad. Sci. U. S. A. 106, 8719–8724. doi:10.1073/pnas.0900234106.

Ciric, R., Rosen, A.F.G., Erus, G., Cieslak, M., Adebimpe, A., Cook, P.A., Bassett, D.S., Davatzikos, C., Wolf, D.H., Satterthwaite, T.D., 2018. Mitigating head motion artifact in functional connectivity MRI. Nat. Protoc. 13, 2801–2826. doi:10.1038/s41596-018-0065-y.

Ciric, R., Wolf, D.H., Power, J.D., Roalf, D.R., Baum, G.L., Ruparel, K., Shinohara, R.T., Elliott, M.A., Eickhoff, S.B., Davatzikos, C., Gur, R.C., Gur, R.E., Bassett, D.S., Satterthwaite, T.D., 2017. Benchmarking of participant-level confound regression strategies for the control of motion artifact in studies of functional connectivity. Neuroimage 154, 174–187. doi:10.1016/j.neuroimage.2017.03.020.

Clare Kelly, A.M., Uddin, L.Q., Biswal, B.B., Castellanos, F.X., Milham, M.P., 2008. Competition between functional brain networks mediates behavioral variability. Neuroimage 39, 527–537. doi:10.1016/j.neuroimage.2007.08.008.

de Cheveigné, A., Nelken, I., 2019. Filters: When, Why, and How (Not) to Use Them. Neuron 102, 280–293. doi:10.1016/j.neuron.2019.02.039.

Dion, M.L., Sumner, J.L., Mitchell, S.M.L., 2018. Gendered citation patterns across political science and social science methodology fields. Political Anal. 26, 312–327. doi:10.1017/pan.2018.12.

Dworkin, J.D., Linn, K.A., Teich, E.G., Zurn, P., Shinohara, R.T., Bassett, D.S., 2020. The extent and drivers of gender imbalance in neuroscience reference lists. Nat. Neurosci. 23, 918–926. doi:10.1038/s41593-020-0658-y.

Engelhardt, L.E., Roe, M.A., Juranek, J., DeMaster, D., Harden, K.P., Tucker-Drob, E.M., Church, J.A., 2017. Children's head motion during fMRI tasks is heritable and stable over time. Dev. Cogn. Neurosci. 25, 58–68. doi:10.1016/j.dcn.2017.01.011.

Fair, D.A., Miranda-Dominguez, O., Snyder, A.Z., Perrone, A., Earl, E.A., Van, A.N., Koller, J.M., Feczko, E., Tisdall, M.D., van der Kouwe, A., Klein, R.L., Mirro, A.E., Hampton, J.M., Adeyemo, B., Laumann, T.O., Gratton, C., Greene, D.J., Schlaggar, B.L., Hagler, D.J., Watts, R., Garavan, H., Barch, D.M., Nigg, J.T., Petersen, S.E., Dale, A.M., Feldstein-Ewing, S.W., Nagel, B.J., Dosenbach, N.U.F., 2020. Correction of respiratory artifacts in MRI head motion estimates. Neuroimage 208, 116400. doi:10.1016/j.neuroimage.2019.116400.

Finn, E.S., Shen, X., Scheinost, D., Rosenberg, M.D., Huang, J., Chun, M.M., Papademetris, X., Constable, R.T., 2015. Functional connectome fingerprinting: Identifying individuals using patterns of brain connectivity. Nat. Neurosci. 18, 1664–1671. doi:10.1038/nn.4135.

Fischer, L.F., Mojica Soto-Albors, R., Buck, F., Harnett, M.T., 2020. Representation of visual landmarks in retrosplenial cortex. Elife 9, 1–25. doi:10.7554/eLife.51458.

Fortunato, S., 2010. Community detection in graphs. Phys. Rep. 486, 75–174. doi:10.1016/j.physrep.2009.11.002.

Fox, M.D., Raichle, M.E., 2007. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. Nat. Rev. Neurosci. 8, 700–711. doi:10.1038/nrn2201.

Freedman, D., Diaconis, P., 1981. On the histogram as a density estimator:L2 theory. Z. für Wahrscheinlichkeitstheorie und Verwandte Geb. 57, 453–476. doi:10.1007/BF01025868.

Friston, K.J., 2011. Functional and effective connectivity: a review. Brain Connect. 1, 13–36. doi:10.1089/brain.2011.0008.

Gao, W., Lin, W., 2012. Frontal parietal control network regulates the anti-correlated default and dorsal attention networks. Hum. Brain Mapp. 33, 192–202. doi:10.1002/hbm.21204.

Girvan, M., Newman, M.E.J., 2002. Community structure in social and biological networks. Proc. Natl. Acad. Sci. U. S. A. 99, 7821–7826. doi:10.1073/pnas.122653799.

Golchert, J., Smallwood, J., Jefferies, E., Seli, P., Huntenburg, J.M., Liem, F., Lauckner, M.E., Oligschläger, S., Bernhardt, B.C., Villringer, A., Margulies, D.S., 2017. Individual variation in intentionality in the mind-wandering state is reflected in the integration of the default-mode, fronto-parietal, and limbic networks. Neuroimage 146, 226–235. doi:10.1016/j.neuroimage.2016.11.025.

Gómez, S., Jensen, P., Arenas, A., 2009. Analysis of community structure in networks of correlated data. Phys. Rev. E Stat. Nonlinear Soft Matter Phys. 80, 1–5. doi:10.1103/PhysRevE.80.016114.

Gordon, E.M., Laumann, T.O., Adeyemo, B., Huckins, J.F., Kelley, W.M., Petersen, S.E., 2016. Generation and evaluation of a cortical area parcellation from resting-state correlations. Cereb. Cortex 26, 288–303. doi:10.1093/cercor/bhu239.

Grady, C., Sarraf, S., Saverino, C., Campbell, K., 2016. Age differences in the functional interactions among the default, frontoparietal control, and dorsal attention networks. Neurobiol. Aging 41, 159–172. doi:10.1016/J.NEUROBIOLAGING.2016.02.020.

Gratton, C., Laumann, T.O., Nielsen, A.N., Greene, D.J., Gordon, E.M., Gilmore, A.W., Nelson, S.M., Coalson, R.S., Snyder, A.Z., Schlaggar, B.L., Dosenbach, N.U.F., Petersen, S.E., 2018. Functional brain networks are dominated by stable group and individual factors, not cognitive or daily variation. Neuron 98, 439–452. doi:10.1016/j.neuron.2018.03.035, e5.

Griffanti, L., Salimi-Khorshidi, G., Beckmann, C.F., Auerbach, E.J., Douaud, G., Sexton, C.E., Zsoldos, E., Ebmeier, K.P., Filippini, N., Mackay, C.E., Moeller, S., Xu, J., Yacoub, E., Baselli, G., Ugurbil, K., Miller, K.L., Smith, S.M., 2014. ICA-based artefact removal and accelerated fMRI acquisition for improved resting state network imaging. Neuroimage 95, 232–247. doi:10.1016/j.neuroimage.2014.03.034.

Grinsted, A., Moore, J.C., Jevrejeva, S., 2004. Application of the cross wavelet transform and wavelet coherence to geophysical time series. Nonlinear Process. Geophys. 11, 561–566. doi:10.5194/npg-11-561-2004.

Gu, S., Satterthwaite, T.D., Medaglia, J.D., Yang, M., Gur, R.E., Gur, R.C., Bassett, D.S., 2015. Emergence of system roles in normative neurodevelopment. Proc. Natl. Acad. Sci. U. S. A. 112, 13681–13686. 10.1073/pnas.1502829112

Kajimura, S., Kochiyama, T., Nakai, R., Abe, N., Nomura, M., 2016. Causal relationship between effective connectivity within the default mode network and mind-wandering regulation and facilitation. Neuroimage 133, 21–30. doi:10.1016/j.neuroimage.2016.03.009.

Kim, J., Van Dijk, K.R.A., Libby, A., Napadow, V., 2014. Frequency-dependent relationship between resting-state functional magnetic resonance imaging signal power and head motion is localized within distributed association networks. Brain Connect. 4, 30–39. doi:10.1089/brain.2013.0153.

Maliniak, D., Powers, R., Walter, B.F., 2013. The gender citation gap in international relations. Int. Organ. doi:10.1017/S0020818313000209.

Mao, D., Molina, L.A., Bonin, V., McNaughton, B.L., 2020. Vision and locomotion combine to drive path integration sequences in mouse retrosplenial cortex. Curr. Biol. 1–9. doi:10.1016/j.cub.2020.02.070.

Mitchell, S.M., Lange, S., Brus, H., 2013. Gendered citation patterns in international relations journals. Int. Stud. Perspect. 14, 485–492. doi:10.1111/insp.12026.

Morgan, S.E., Achard, S., Termenon, M., Bullmore, E.T., Vértes, P.E., 2017. Low-dimensional morphospace of topological motifs in human fMRI brain networks. Netw. Neurosci. 2, 285–302. doi:10.1162/netn_a_00038.

Morgan, S.E., White, S.R., Bullmore, E.T., Vértes, P.E., 2018. A network neuroscience approach to typical and atypical brain development. Biol. Psychiatry Cogn. Neurosci. Neuroimaging doi:10.1016/j.bpsc.2018.03.003.

Murphy, A.C., Bertolero, M.A., Papadopoulos, L., Lydon-Staley, D.M., Bassett, D.S., 2020. Multimodal network dynamics underpinning working memory. Nat. Commun. 11, 113. doi:10.1038/s41467-020-15541-0.

Murphy, K., Fox, M.D., 2017. Towards a consensus regarding global signal regression for resting state functional connectivity MRI. Neuroimage 154, 169–173. doi:10.1016/j.neuroimage.2016.11.052.

Newman, M., 2006. Modularity and community structure in networks, 103, pp. 8577–8582. doi:10.1073/pnas.0601602103.

Noble, S., Scheinost, D., Constable, R.T., 2019. A decade of test-retest reliability of functional connectivity: a systematic review and meta-analysis. Neuroimage 203, 116157. doi:10.1016/j.neuroimage.2019.116157.

Owens, M.M., Duda, B., Sweet, L.H., MacKillop, J., 2018. Distinct functional and structural neural underpinnings of working memory. Neuroimage 174, 463–471. doi:10.1016/j.neuroimage.2018.03.022.

Parkes, L., Fulcher, B., Yücel, M., Fornito, A., 2018. An evaluation of the efficacy, reliability, and sensitivity of motion correction strategies for resting-state functional MRI. Neuroimage 171, 415–436. doi:10.1016/j.neuroimage.2017.12.073.

Pervaiz, U., Vidaurre, D., Woolrich, M.W., Smith, S.M., 2020. Optimising network modelling methods for fMRI. Neuroimage 211, 116604. doi:10.1016/j.neuroimage.2020.116604.

Porter, M.A., Onnela, J.P., Mucha, P.J., 2009. Communities in Networks. Not. AMS 56, 1082–1097.

Power, J.D., Lynch, C.J., Silver, B.M., Dubin, M.J., Martin, A., Jones, R.M., 2019. Distinctions among real and apparent respiratory motions in human fMRI data. Neuroimage 201, 116041. doi:10.1016/j.neuroimage.2019.116041.

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. doi:10.1016/j.neuroimage.2013.08.048.

Power, J.D., Schlaggar, B.L., Petersen, S.E., 2015. Recent progress and outstanding issues in motion correction in resting state fMRI. Neuroimage 105, 536–551. doi:10.1016/j.neuroimage.2014.10.044.

Quiroga, R.Q., Kraskov, A., Kreuz, T., Grassberger, P., 2001. On the performance of different synchronization measures in real data: a case study on EEG signals. Phys. Rev. E 65, 1–14. doi:10.1103/PhysRevE.65.041903.

Rubinov, M., Sporns, O., 2011. Weight-conserving characterization of complex functional brain networks. Neuroimage 56, 2068–2079. doi:10.1016/j.neuroimage.2011.03.069.

Salimi-Khorshidi, G., Douaud, G., Beckmann, C.F., Glasser, M.F., Griffanti, L., Smith, S.M., 2014. Automatic denoising of functional MRI data: Combining independent component analysis and hierarchical fusion of classifiers. Neuroimage 90, 449–468. doi:10.1016/j.neuroimage.2013.11.046.

Salvador, R., Martínez, A., Pomarol-Clotet, E., Gomar, J., Vila, F., Sarró, S., Capdevila, A., Bullmore, E., 2008. A simple view of the brain through a frequency-specific functional connectivity measure. Neuroimage 39, 279–289. doi:10.1016/j.neuroimage.2007.08.018.

Salvador, R., Suckling, J., Schwarzbauer, C., Bullmore, E., 2005. Undirected graphs of frequency-dependent functional connectivity in whole brain networks. Philos. Trans. R. Soc. B Biol. Sci. 360, 937–946. doi:10.1098/rstb.2005.1645.

Satterthwaite, T.D., Elliott, M.A., Gerraty, R.T., Ruparel, K., Loughead, J., Calkins, M.E., Eickhoff, S.B., Hakonarson, H., Gur, R.C., Gur, R.E., Wolf, D.H., 2013a. An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. Neuroimage 64, 240–256. doi:10.1016/j.neuroimage.2012.08.052.

Satterthwaite, T.D., Elliott, M.A., Gerraty, R.T., Ruparel, K., Loughead, J., Calkins, M.E., Eickhoff, S.B., Hakonarson, H., Gur, R.C., Gur, R.E., Wolf, D.H., 2013b. An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. Neuroimage 64, 240–256. doi:10.1016/j.neuroimage.2012.08.052.

Satterthwaite, T.D., Wolf, D.H., Loughead, J., Ruparel, K., Elliott, M.A., Hakonarson, H., Gur, R.C., Gur, R.E., 2012. Impact of in-scanner head motion on multiple measures of functional connectivity: relevance for studies of neurodevelopment in youth. Neuroimage 60, 623–632. doi:10.1016/j.neuroimage.2011.12.063.

Schaefer, A., Kong, R., Gordon, E.M., Laumann, T.O., Zuo, X.-N., Holmes, A.J., Eickhoff, S.B., Yeo, B.T.T., 2018. Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. Cereb. Cortex 28, 3095–3114. doi:10.1093/cercor/bhx179.

Shannon, C.E., 1948. A mathematical theory of communication. Bell Syst. Tech. J. 27, 379–423. doi:10.1002/j.1538-7305.1948.tb01338.x.

Shrout, P.E., Fleiss, J.L., 1979. Intraclass correlations: Uses in assessing rater reliability. Psychol. Bull. 86, 420–428. doi:10.1037/0033-2909.86.2.420.

Smith, S.M., Fox, P.T., Miller, K.L., Glahn, D.C., Fox, P.M., Mackay, C.E., Filippini, N., Watkins, K.E., Toro, R., Laird, A.R., Beckmann, C.F., 2009. Correspondence of the brain's functional architecture during activation and rest. Proc. Natl. Acad. Sci. U. S. A. 106, 13040–13045. doi:10.1073/pnas.0905267106.

Smith, S.M., Miller, K.L., Salimi-Khorshidi, G., Webster, M., Beckmann, C.F., Nichols, T.E., Ramsey, J.D., Woolrich, M.W., 2011. Network modelling methods for FMRI. Neuroimage 54, 875–891. doi:10.1016/j.neuroimage.2010.08.063.

Sood, G., Laohaprapanon, S., 2018. Predicting race and ethnicity from the sequence of characters in a name. arXiv Prepr. arXiv:1805.02109.

Sporns, O., Betzel, R.F., 2016. Modular brain networks. Annu. Rev. Psychol. 67, 613–640. doi:10.1146/annurev-psych-122414-033634.

Strehl, A., Ghosh, J., 2002. Cluster ensembles -a knowledge reuse framework for combining multiple partitions. J. Mach. Learn. Res. 3, 583–617.

Sun, F.T., Miller, L.M., D'Esposito, M., 2004. Measuring interregional functional connectivity using coherence and partial coherence analyses of fMRI data. Neuroimage 21, 647–658. doi:10.1016/j.neuroimage.2003.09.056.

Thomas 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., Fisch, B., Liu, H., Buckner, R.L., 2011. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. J. Neurophysiol. 106, 1125–1165. doi:10.1152/jn.00338.2011.

Thomason, M.E., 2020. Development of brain networks in utero: relevance for common neural disorders. Biol. Psychiatry doi:10.1016/j.biopsych.2020.02.007.

Tooley, U.A., Mackey, A.P., Ciric, R., Ruparel, K., Moore, T.M., Gur, R.C., Gur, R.E., Satterthwaite, T.D., Bassett, D.S., 2020. Associations between neighborhood SES and functional brain network development. Cereb. Cortex 30, 1–19. doi:10.1093/cercor/bhz066.

Traag, V.A., Bruggeman, J., 2009. Community detection in networks with positive and negative links. Phys. Rev. E Stat. Nonlinear Soft Matter Phys. 80, 1–6. doi:10.1103/PhysRevE.80.036115.

Uddin, L.Q., Yeo, B.T.T., Spreng, R.N., 2019. Towards a universal taxonomy of macro-scale functional human brain networks. Brain Topogr. 32, 926–942. doi:10.1007/s10548-019-00744-6.

van den Heuvel, M.I., Turk, E., Manning, J.H., Hect, J., Hernandez-Andrade, E., Hassan, S.S., Romero, R., van den Heuvel, M.P., Thomason, M.E., 2018. Hubs in the human fetal brain network. Dev. Cogn. Neurosci. 30, 108–115. doi:10.1016/j.dcn.2018.02.001.

van den Heuvel, M.P., Hulshoff Pol, H.E., 2010. Exploring the brain network: a review on resting-state fMRI functional connectivity. Eur. Neuropsychopharmacol. 20, 519–534. doi:10.1016/j.euroneuro.2010.03.008.

van Dijk, K.R.A., Sabuncu, M.R., Buckner, R.L., 2012. The influence of head motion on intrinsic functional connectivity MRI. Neuroimage 59, 431–438. doi:10.1016/j.neuroimage.2011.07.044.

Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E.J., Yacoub, E., Ugurbil, K., 2013. The WU-Minn human connectome project: an overview. Neuroimage 80, 62–79. doi:10.1016/j.neuroimage.2013.05.041.

Vann, S.D., Aggleton, J.P., Maguire, E.A., 2009. What does the retrosplenial cortex do? Nat. Rev. Neurosci. 10, 792–802. doi:10.1038/nrn2733.

Vélez-Fort, M., Bracey, E.F., Keshavarzi, S., Rousseau, C.V., Cossell, L., Lenzi, S.C., Strom, M., Margrie, T.W., 2018. A circuit for integration of head- and visual-motion signals in layer 6 of mouse primary visual cortex. Neuron 98, 179–191. doi:10.1016/j.neuron.2018.02.023, e6.

Vértes, P.E., Alexander-Bloch, A.F., Gogtay, N., Giedd, J.N., Rapoport, J.L., Bullmore, E.T., 2012. Simple models of human brain functional networks. Proc. Natl. Acad. Sci. U. S. A. 109, 5868–5873. doi:10.1073/pnas.1111738109.

Vértes, P.E., Bullmore, E.T., 2015. Annual research review: growth connectomics-the organization and reorganization of brain networks during normal and abnormal development. J. Child Psychol. Psychiatry Allied Discip. 56, 299–320. doi:10.1111/jcpp.12365.

Vértes, P.E., Rittman, T., Whitaker, K.J., Romero-Garcia, R., Váša, F., Kitzbichler, M.G., Wagstyl, K., Fonagy, P., Dolan, R.J., Jones, P.B., Goodyer, I.M., Bullmore, E.T., 2016. Gene transcription profiles associated with inter-modular hubs and connection distance in human functional magnetic resonance imaging networks. Philos. Trans. R. Soc. B Biol. Sci. 371. doi:10.1098/rstb.2015.0362.

Wang, H.E., Bénar, C.G., Quilichini, P.P., Friston, K.J., Jirsa, V.K., Bernard, C., 2014. A systematic framework for functional connectivity measures. Front. Neurosci. 8, 1–22. doi:10.3389/fnins.2014.00405.

Wheelock, M.D., Hect, J.L., Hernandez-Andrade, E., Hassan, S.S., Romero, R., Eggebrecht, A.T., Thomason, M.E., 2019. Sex differences in functional connectivity during fetal brain development. Dev. Cogn. Neurosci. 36, 100632. doi:10.1016/j.dcn.2019.100632.

Whitaker, K.J., Vértes, P.E., Romero-Garciaa, R., Váša, F., Moutoussis, M., Prabhu, G., Weiskopf, N., Callaghan, M.F., Wagstyl, K., Rittman, T., Tait, R., Ooi, C., Suckling, J., Inkster, B., Fonagy, P., Dolan, R.J., Jones, P.B., Goodyer, I.M., Bullmore, E.T., 2016. Adolescence is associated with genomically patterned consolidation of the hubs of the human brain connectome. Proc. Natl. Acad. Sci. U. S. A. 113, 9105–9110. doi:10.1073/pnas.1601745113.

Zhang, Z., Telesford, Q.K., Giusti, C., Lim, K.O., Bassett, D.S., 2016. Choosing wavelet methods, filters, and lengths for functional brain network construction. PLoS One 11, 1–24. doi:10.1371/journal.pone.0157243.

Zhou, D., Stiso, J., Cornblath, E., Teich, E., Blevins, A.S., Oudyk, K., Michael, C., Virtualmaster; Camp, C., 2020. dalejn/cleanBib: v1.1.1. Zenodo. doi:10.5281/ZENODO.4104748.

Zhou, D., Thompson, W.K., Siegle, G., 2009. MATLAB toolbox for functional connectivity. Neuroimage 47, 1590–1607. doi:10.1016/j.neuroimage.2009.05.089.