

Multi-Scale Network Regression for Brain-Phenotype Associations

Cedric Huchuan Xia, Zongming Ma, Zaixu Cui, Danilo Bzdok, Bertrand Thirion, Danielle Bassett, Theodore D Satterthwaite, Russell T Shinohara, Daniela M Witten

▶ **To cite this version:**

Cedric Huchuan Xia, Zongming Ma, Zaixu Cui, Danilo Bzdok, Bertrand Thirion, et al.. Multi-Scale Network Regression for Brain-Phenotype Associations. Human Brain Mapping, In press, 10.1002/hbm.24982 . hal-02523974

HAL Id: hal-02523974
https://hal.science/hal-02523974

Submitted on 31 Mar 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Received: 15 October 2019 | Revised: 31 January 2020 | Accepted: 26 February 2020

DOI: 10.1002/hbm.24982

Check for updates

WILEY

TECHNICAL REPORT

Multi-scale network regression for brain-phenotype associations

Cedric Huchuan Xia[1] | Zongming Ma[2] | Zaixu Cui[1] | Danilo Bzdok[3,4,5,6] | Bertrand Thirion[5] | Danielle S. Bassett[1,7,8,9,10,11] | Theodore D. Satterthwaite[1] | Russell T. Shinohara[12,13] | Daniela M. Witten[14,15]

[1]Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania

[2]Department of Statistics, The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania

[3]Department of Psychiatry, Psychopathology and Psychosomatics, RWTH Aachen University, Aachen, Germany

[4]JARA-BRAIN, Jülich-Aachen Research Alliance, Jülich, Germany

[5]Université Paris-Saclay, CEA, Inria, Gif-sur-Yvette, France

[6]Department of Bioengineering, McGill University, Montreal, Canada

[7]Department of Bioengineering, School of Engineering and Applied Science, University of Pennsylvania, Philadelphia, Pennsylvania

[8]Department of Electrical and Systems Engineering, School of Engineering and Applied Science, University of Pennsylvania, Philadelphia, Pennsylvania

[9]Department of Neurology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania

[10]Department of Physics and Astronomy, School of Arts and Science, University of Pennsylvania, Philadelphia, Pennsylvania

[11]Santa Fe Institute, Santa Fe, New Mexico

[12]Penn Statistics and Visualization Center, Department of Biostatistics, Epidemiology, and Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania

[13]Center for Biomedical Imaging Computing and Analytics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania

[14]Department of Statistics, College of Arts and Science, University of Washington, Seattle, Washington

[15]Department of Biostatistics, School of Public Health, University of Washington, Seattle, Washington

Correspondence
Russell T. Shinohara, Penn Statistics and Visualization Center, Department of Biostatistics, Epidemiology, and Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA
Email: rshi@upenn.edu

Daniela M. Witten, Department of Statistics, College of Arts and Science, University of Washington, Seattle, WA
Email: dwitten@uw.edu

Funding information
Office of Naval Research; Army Research Laboratory; Paul Allen Foundation; Institute for Scientific Interchange Foundation; Alfred P. Sloan Foundation; MacArthur Foundation; Blavatnik Family Foundation; Amazon Web Service; Deutsche Forschungsgemeinschaft,

Abstract

Brain networks are increasingly characterized at different scales, including summary statistics, community connectivity, and individual edges. While research relating brain networks to behavioral measurements has yielded many insights into brain-phenotype relationships, common analytical approaches only consider network information at a single scale. Here, we designed, implemented, and deployed Multi-Scale Network Regression (MSNR), a penalized multivariate approach for modeling brain networks that explicitly respects both edge- and community-level information by assuming a low rank and sparse structure, both encouraging less complex and more interpretable modeling. Capitalizing on a large neuroimaging cohort (n = 1, 051), we demonstrate that MSNR recapitulates interpretable and statistically significant connectivity patterns associated with brain development, sex differences, and motion-related artifacts. Compared to single-scale methods, MSNR achieves a balance

Russell T. Shinohara and Daniela M. Witten authors contributed equally to this work.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2020 The Authors. *Human Brain Mapping* published by Wiley Periodicals, Inc.

Grant/Award Numbers: BZ2/4-1, BZ2/3-1, BZ2/2-1, IRTG2150; National Science Foundation, Grant/Award Numbers: DMS-1352060, DMS-1252624; NIH, Grant/Award Numbers: R01MH113550, R21MH106799, R01MH112847, R01NS060910, R01NS085211, DP5OD009145, R01 GM123993, MH089924, MH089983

between prediction performance and model complexity, with improved interpretability. Together, by jointly exploiting both edge- and community-level information, MSNR has the potential to yield novel insights into brain-behavior relationships.

KEYWORDS

functional connectivity, multivariate analysis, network neuroscience

1 | INTRODUCTION

Studying brain-phenotype relationships in high-dimensional connectomics is an active area of research in neuroscience (Bassett & Sporns, 2017; Bzdok et al., 2016). The advent of large neuroimaging datasets that provide measures of brain connectivity for unprecedented numbers of subjects has yielded novel insights into development and aging, cognition, and neuropsychiatric illnesses (Biswal et al., 2010; Bzdok & Yeo, 2017; Jernigan et al., 2016; Schumann et al., 2010; Van Essen et al., 2012). As the availability of datasets with rich neural, genetic, and behavioral measurements from large numbers of subjects continues to increase, there is a growing need for analytical tools that are tailored for the discovery of complex relationships between brain networks and phenotypes (Craddock, Tungaraza, & Milham, 2015; Varoquaux & Craddock, 2013; L. Wang, Durante, Jung, & Dunson, 2017; Zhang, Sun, & Li, 2018).

A typical brain network consists of hundreds of nodes that denote brain regions, and tens of thousands of edges that indicate connections between pairs of nodes (Rubinov & Sporns, 2009).The network can be viewed on the *micro-scale*, *meso-scale*, or *macro-scale*. The *micro*-scale of the network can be characterized by features of its edges. The *macro*-scale of the network can be characterized by summary statistics such as characteristic path length and global efficiency (Rubinov & Sporns, 2009). The *meso-scale* falls in between the micro-scale and macro-scale, and includes the communities that make up the network (Sporns & Betzel, 2016). A community refers to a collection of nodes that are highly connected to each other and have little connection to nodes outside of the community. Prior work has demonstrated that brain network architecture present on these different scales is associated with development and aging (Betzel et al., 2014; Gu et al., 2015; Power, Fair, Schlaggar, & Petersen, 2010), cognition (Bressler & Menon, 2010; Crossley et al., 2013; Park & Friston, 2013), and neuropsychiatric diseases (Bassett, Xia, & Satterthwaite, 2018; Braun et al., 2016; Fornito, Zalesky, & Breakspear, 2015; Grillon et al., 2013; Kernbach et al., 2018; Xia et al., 2018; Yu et al., 2019).

Despite increased appreciation that multi-scale organization of the brain may be responsible for some of its major functions (Bassett & Siebenhühner, 2013; Betzel & Bassett, 2017), thus far, common strategies for studying the relationship between brain connectivity and phenotypes consider network features at a single scale (Craddock et al., 2015; Varoquaux & Craddock, 2013). For example, a popular single-scale strategy focuses on group-level comparisons of individual connections in brain networks (Bressler & Menon, 2010; Fornito et al., 2015; Grillon et al., 2013). This approach typically involves performing a statistical test on each network edge. While this procedure is easy to implement, several drawbacks limit its effectiveness (Bzdok & Ioannidis, 2019). Two main limitations are the need to account for multiple comparisons, and a lack of interpretability (Craddock et al., 2015; Varoquaux & Craddock, 2013). To achieve high power while minimizing the risk of false discovery, alternative edge-based methods have been developed, such as the network-based statistic (Zalesky, Fornito, & Bullmore, 2010) and multivariate distance matrix regression (Zapala & Schork, 2012). While these strategies have yielded important insights, they nonetheless focus exclusively on the micro-scale without exploiting the multi-scale information present in brain networks, often producing results that are difficult to interpret.

Given the importance of community structure in brain networks and its interpretability in the context of neural and cognitive computations (Betzel, Medaglia, & Bassett, 2018; Sporns & Betzel, 2016), it might be tempting to conduct a mass-univariate analysis at the meso-scale, considering *within*- and *between*-community connectivity as the input features (Betzel et al., 2014; Braun et al., 2016; Gu et al., 2015; Yu et al., 2019). Such an approach dramatically reduces the dimensionality of the data, which in turn decreases the burden of multiple comparisons correction. A community-based approach also has the added benefit of not having to deconstruct the connectivity matrix into vectors, as in an edge-based approach, which inevitably disrupts the innate structure in the data. However, summarizing hundreds or thousands of edges as one single number to represent the connection within or between communities can be problematic. This is especially true for large communities such as the default mode network, whose edges are spatially distributed across the anterior and posterior portions of the brain (Raichle, 2015). Stated another way, extracting the mean connectivity at the community level risks mixing disparate signals.

In this paper, we introduce *Multi-Scale Network Regression (MSNR)*, which simultaneously incorporates information across multiple scales in order to reveal associations between high-dimensional connectomic data and phenotypes of interest. We first describe the MSNR model and introduce an algorithm to estimate the parameters. Next, we capitalize on one of the largest neurodevelopmental imaging cohorts, the Philadelphia Neurodevelopmental Cohort (PNC), to empirically assess the ability of MSNR in delineating brain connectivity patterns that are associated with a wide variety of phenotypes. Importantly, we conduct head-to-head comparisons between MSNR and common edge- and community-based analyses that are based on single-scale strategies, and show that MSNR achieves a balance

between prediction performance and interpretability by considering information at multiple network scales.

2 | METHODS

2.1 | A statistical model for multi-scale network regression

Given n subjects, let $A^1, \ldots, A^n \in R^{p \times p}$ denote the adjacency matrices corresponding to their brain connectivity networks, where p is the number of nodes. For instance, $A^i_{jj'}$ could represent the Pearson correlation of the mean activation timeseries of two brain regions, a common measure of functional connectivity, between the j-th and j'-th nodes for the i-th subject. Furthermore, we assume that the p nodes can be partitioned into K distinct communities $C_1, \ldots, C_K$ that are known *a priori*: $\cup_{k=1}^{K} C_k = \{1,\ldots,p\}, C_k \cap C_{k'} \neq \emptyset$ if $k \neq k'$. The notation $j \in C_k$ indicates that the j-th node is in the k-th community. Moreover, for each subject, q covariates have been measured, so that $X_i = \left(X_i^1 X_i^2 \ldots X_i^q\right)^T \in \mathbb{R}^q$ is a covariate vector for the i-th subject, $i = 1, \ldots, n$.

In what follows, we consider the model

$$A^i_{jj'} = \Theta_{jj'} + \sum_{f=1}^{q} X_i^f \cdot \left(\sum_{k=1}^{K} \sum_{k'=1}^{K} \Gamma^f_{kk'} 1_{(j \in C_k, j' \in C_{k'})} \right) + \epsilon^i_{jj'}, i = 1,\ldots,n, j, j' = 1,\ldots,p, \quad (1)$$

Where $\epsilon^i_{jj'}$ is a mean-zero noise term, and $\epsilon^i_{jj'} = \epsilon^i_{j'j}$. Θ is a symmetric $p \times p$ matrix that summarizes the mean connectivity, across all of the subjects, of each pair of nodes, in the absence of covariates. Finally, for $f = 1, \ldots, q$, Γ^f is a symmetric $K \times K$ matrix that quantifies the association between the f-th feature and the functional connectivity between each pair of communities. For instance, a one-unit increase in X^f is associated with a $\Gamma^f_{kk'}$ increase in the mean functional connectivity between nodes in the k-th and k'-th communities.

We now define a $p \times K$ matrix W for which $W_{jk} = 1_{(j \in C_k)}$, where $1_{(\cdot)}$ denotes an indicator variable. As such, Equation (1) can be rewritten as

$$A^i = \Theta + \sum_{f=1}^{q} X_i^f \cdot \left(W \Gamma^f W^T\right) + \epsilon^i, i = 1,\ldots,n. \quad (2)$$

In order to fit the Model (2), we make two assumptions about the structures of the unknown parameter matrices Θ and $\Gamma^1, \ldots, \Gamma^q$.

Assumption 1: Θ *has low rank* (Leonardi et al., 2013; K. Li, Guo, Nie, Li, & Liu, 2009; Smith et al., 2015). That is, $\Theta = VV^T$ where V is a $p \times d$ matrix, for a small positive constant d. This means that the p nodes effectively reside in a reduced subspace of d dimensions. The mean connectivity between any pair of nodes is simply given by their inner product in this low-dimensional subspace. Furthermore, to ensure identifiability, we assume that $Tr(W^T \Theta W) = 0$.

Assumption 2: $\Gamma^1, \ldots, \Gamma^q$ *are sparse* (Eavani et al., 2015; Ng, Varoquaux, Poline, & Thirion, 2012; Xia et al., 2018). That is, most of their elements are *exactly* equal to zero. If $\Gamma^f_{kk'} = 0$, then the value of the f-th feature is unassociated with the mean connectivity between nodes in the k-th and k'-th communities.

We note that Assumption 1: is closely related to the random dot product graph model and similar models (Durante & Dunson, 2018; Durante, Dunson, & Vogelstein, 2017; Fosdick & Hoff, 2015; Tang, Athreya, Sussman, Lyzinski, & Priebe, 2017; Young & Scheinerman, 2007), whereas Assumption 2: is a standard sparsity assumption for high-dimensional regression (Hastie, Tibshirani, & Friedman, 2008; Hastie, Tibshirani, Wainwright, Tibshirani, & Wainwright, 2015; Tibshirani, 1996). Under these two assumptions, a schematic of the Model (2) can be seen in Figure 1.

Model (2) is closely related to both the stochastic block model (Choi, Wolfe, & Airoldi, 2012) and the random dot product graph model (Young & Scheinerman, 2007). In particular, if $\Theta = 0$, $q = 1$, and $X_i^f = 1$ for $i = 1, \ldots, n$, then Equation (2) reduces to a stochastic block model with known communities $C_1, \ldots, C_K$. And if $\Gamma^1 = \ldots = \Gamma^q = 0$ and Assumption 1: holds, then Equation (2) reduces to a random dot product graph model. However, unlike those two models Equation (2) explicitly allows for the mean of the adjacency matrix to be a function of covariates, and effectively incorporates both edge- and community-level network information.

2.2 | Optimization problem

We now consider the task of fitting the Model (2), under Assumptions 1: and 2:. It is natural to consider the optimization problem

$$\underset{\Theta, \Gamma^1, \ldots, \Gamma^q}{minimize} \left\{ \sum_{i=1}^{n} \left\| A^i - \left(\Theta + \sum_{f=1}^{q} X_i^f \cdot \left(W \Gamma^f W^T\right)\right) \right\|_F^2 + \lambda_1 \text{rank}(\Theta) + \lambda_2 \sum_{f=1}^{q} \left\|\Gamma^f\right\|_0 \right\} \quad (3)$$

where the notation $||\cdot||_F^2$ indicates the squared *Frobenius* norm of a matrix, that is, $\|D\|_F^2 = \sum_{j=1}^{p} \sum_{j'}^{p} D_{jj'}^2$, and the notatiton $\|\cdot\|_0$ indicates the element-wise cardinality (or l_0 norm) of a matrix, that is, $\|D\|_0 = \sum_{j=1}^{p} \sum_{j'=1}^{p} 1_{(D_{jj'} \neq 0)}$. In Equation (3), λ_1 and λ_2 are non-negative tuning parameter values that control the rank of Θ and the sparsity of $\Gamma^1 = \ldots = \Gamma^q$, respectively.

Unfortunately, due to the presence of the rank and l_0 penalties, the optimization problem (3) is highly nonconvex, and no efficient algorithms are available to solve it. Therefore, in what follows, we will consider an alternative to Equation (3), which results from replacing

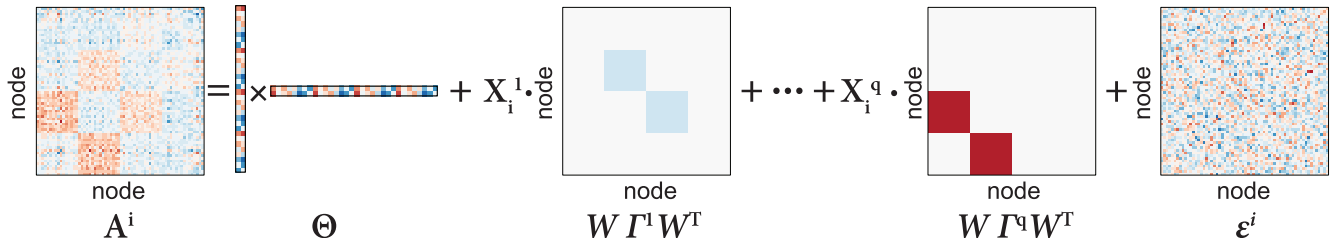


FIGURE 1 A schematic for Multi-Scale Network Regression (MSNR). We developed a penalized multivariate approach for modeling brain networks that explicitly respects both edge- and community-level information. We specified the MSNR model in Equation (2), which is visually represented here. Under the model, A^i is the connectivity matrix for the i-th subject, Θ is a low-rank matrix representing the mean connectivity across all subjects, Γ^1, ..., Γ^q are sparse matrices representing the community-level connectivity associated with the covariates $(X_i^1,...,X_i^q)$, and ϵ^i is the noise

the nonconvex rank and l_0 penalties in Equation (3) with their convex relaxations. This leads to the optimization problem

$$\underset{\Theta,\Gamma^1,...,\Gamma^q}{\text{minimize}}\left\{\sum_{i=1}^{n}\left\|A^i-\left(\Theta+\sum_{f=1}^{q}X_i^f\cdot\left(W\Gamma^f W^T\right)\right)\right\|_F^2 + \lambda_1\|\Theta\|_* + \lambda_2\sum_{f=1}^{q}\left\|\Gamma^f\right\|_1\right\} \quad (4)$$

In Equation (4), the notation $\|\cdot\|_*$ indicates the *nuclear norm* of a matrix, that is, the sum of its singular values (Bien & Witten, 2016; Fazel, 2002; Recht, Fazel, & Parrilo, 2010). The nuclear norm is a convex surrogate for the rank of a matrix. The notation $\|\cdot\|_1$ indicates the element-wise l_1 (or *lasso)* norm of a matrix, that is, $\|D\|_1 = \sum_{j=1}^{p}\sum_{j'=1}^{p}|D_{jj'}|$; this is a convex relaxation of the l_0 norm (Hastie et al., 2015; Hastie et al., 2008; Tibshirani, 1996). In Equation (4), the non-negative tuning parameters λ_1 and λ_2 encourage Θ and Γ^1, ..., Γ^q to be low-rank and sparse, respectively.

Importantly, the optimization problem (4) is convex, and so fast algorithms are available to solve it for the global optimum. In Supporting Information, we have derived a block coordinate descent algorithm for solving Equation (4). Simulation studies indicated that MSNR behaved in the manner that was dependent on the signal-to-noise ratio and the observation-to-feature ratio, particularly in its ability to model underlying connectivity patterns (see Figure S1).

2.3 | Philadelphia neurodevelopmental cohort

Resting-state functional magnetic resonance imaging (rs-fMRI) datasets were acquired as part of the Philadelphia Neurodevelopmental Cohort (PNC), a large community-based study of brain development (Satterthwaite et al., 2014). In total, 1,601 participants completed the cross-sectional neuroimaging protocol. Of these participants, 154 were excluded for meeting any of the following criteria: gross radiological abnormalities, history of medical problems that might affect brain function, history of inpatient psychiatric hospitalization, use of psychoactive medications at the time of data acquisition. Of the remaining 1,447 participants, 51 were excluded for low quality or incomplete FreeSurfer reconstruction of T1-weighted images. Of the remaining 1,396 participants, 381 were excluded for missing rs-fMRI, voxelwise coverage or excessive motion, which is defined as having an average framewise motion more than 0.20 mm and more than 20 frames exhibiting over 0.25 mm movement (using calculation from Jenkinson, Bannister, Brady, & Smith, 2002). These exclusions produced a final sample consisting of 1,015 youths (mean age 15.78, *SD* = 3.34; 461 males, and 554 females).

2.4 | Imaging acquisition

Structural and functional imaging data were acquired on a 3T Siemens Tim Trio scanner with a 32-channel head coil (Erlangen, Germany), as previously described (Satterthwaite et al., 2014, 2016). High-resolution structural images were acquired in order to facilitate alignment of individual subject images into a common space. Structural images were acquired using a magnetization-prepared, rapid-acquisition gradient-echo (MPRAGE) T1-weighted sequence (T_R = 1, 810 ms; T_E = 3.51 ms; *FoV* = 180 × 240 mm; resolution 0.9375 × 0.9375 × 1 mm). Approximately 6 min of task-free functional data were acquired for each subject using a blood oxygen level-dependent (BOLD-weighted) sequence (T_R = 3, 000 ms, T_E = 32 ms; *FoV* = 192 × 192 mm; resolution 3 mm isotropic; 124 volumes). Prior to scanning, in order to acclimatize subjects to the MRI environment and to help subjects learn to remain still during the actual scanning session, a mock scanning session was conducted using a decommissioned MRI scanner and head coil. Mock scanning was accompanied by acoustic recordings of the noise produced by gradient coils for each scanning pulse sequence. During these sessions, feedback regarding head movement was provided using the MoTrack motion tracking system (Psychology Software Tools, Inc., Sharpsburg, PA). Motion feedback was only given during the mock scanning session. In order to further minimize motion, prior to data acquisition subjects' heads were stabilized in the head coil using one foam pad over each ear and a third over the top of the head. During the resting-state scan, a fixation cross was displayed as images were acquired. Subjects were instructed to stay awake, keep their eyes open, fixate on the displayed crosshair, and remain still.

2.5 | Structural preprocessing

A study-specific template was generated from a sample of 120 PNC subjects balanced across sex, race, and age using the `buildTemplateParallel` procedure in ANTs (Avants, Tustison, Song, et al., 2011). Study-specific tissue priors were created using a multi-atlas segmentation procedure (H. Wang et al., 2013). Next, each subject's high-resolution structural image was processed using the ANTs Cortical Thickness Pipeline (Tustison et al., 2014). Following bias field correction (Tustison et al., 2010), each structural image was diffeomorphically registered to the study-specific PNC template using the top-performing `SyN` deformation (Klein et al., 2009). Study-specific tissue priors were used to guide brain extraction and segmentation of the subject's structural image (Avants, Tustison, Wu, et al., 2011).

2.6 | Functional preprocessing

Task-free functional images were processed using the XCP Engine (Ciric et al., 2017), which was configured to execute a top-performing pipeline for removal of motion-related variance (Ciric et al., 2018). Preprocessing steps included (a) correction for distortions induced by magnetic field inhomogeneities using FSL's `FUGUE` utility, (b) removal of the four initial volumes of each acquisition, (c) realignment of all volumes to a selected reference volume using `mcflirt` (Jenkinson et al., 2002), (d) removal of and interpolation over intensity outliers in each voxel's time series using AFNI's `3Ddespike` utility, (e) demeaning and removal of any linear or quadratic trends, and (f) co-registration of functional data to the high-resolution structural image using boundary-based registration (Greve & Fischl, 2009). Confounding signals in the data were modeled using a total of 36 parameters, including the six framewise estimates of motion, the mean signal extracted from eroded white matter and cerebrospinal fluid compartments, the mean extracted from the entire brain, the derivatives of each of these nine parameters, and quadratic terms of each of the nine parameters and their derivatives. Both the BOLD-weighted time series and the artefactual model time series were temporally filtered using a first-order Butterworth filter with a passband between 0.01 and 0.08 Hz (Hallquist, Hwang, & Luna, 2013).

2.7 | Network construction

We used a common parcellation of cortial and subcortical tissue into 264 regions (Power et al., 2011). The functional connectivity between any pair of brain regions was operationalized as the Fisher-transformed Pearson correlation coefficient between the mean activation timeseries extracted from those regions. Connectomes were computed across all regions within a common parcellation with 264 nodes and 13 communities (Power et al., 2011). We encoded the pattern of functional connectivity in a formal network model in which nodes represent regions and edges represent functional connections. We assigned each region to one of 13 *a priori* communities (Power et al., 2011) that were delineated using the `Infomap` algorithm (Rosvall & Bergstrom, 2008) and replicated in an independent sample. We excluded 28 nodes that were not sorted into any community, therefore resulting in the final $p = 236$ and $K = 13$ (Figure 2a). This parcellation was selected for our analysis as it has been previously used for studying individual differences in brain connectivity, including those related to brain development (Gu et al., 2015; Satterthwaite et al., 2012), sex differences (Satterthwaite et al., 2015), and in-scanner motion (Ciric et al., 2018).

2.8 | Cross-validation

We first randomly selected 20% of the total sample ($n = 1,015$) to serve as the left-out validation set ($n = 202$). We then performed five-fold cross validation on the remaining 80% of the sample ($n = 813$) to select the values of the tuning parameters λ_1 and λ_2 for MSNR (James, Witten, Hastie, & Tibshirani, 2013, Figure 2b). To ensure the results were not due to any single random data split, we repeated the entire procedure above five times. The age distributions of subjects between training and validation sets across these five data partitions were similar to each other (Figure S2). In each fold, the independent variables ($X_{n \times q}$) were centered to a mean of zero and scaled by each column's standard deviation. The prediction error used in cross-validation was the *Frobenius* norm of the difference between estimated and true connectivity matrices in the test set, $\left\| A^i - \hat{A}^i \right\|_F^2$ (Figure 2c). We ensured the prediction error was independent of the sample size by using the average prediction error over all subjects in the test set.

2.9 | Permutation procedure

To estimate the distribution of prediction error under the null hypothesis of no association between functional connectivity and phenotype, we permuted the rows of the covariate matrix $X_{n \times q}$. For each permutation, we tuned λ_1 and λ_2 using cross-validation, and calculated the prediction error in the left-out validation set. The p-value was defined to be the proportion of prediction errors among the 1,000 permuted datasets that are smaller than the prediction error on the observed data,

$$p_{\text{permutation}} = \frac{\sum_1^{1,000} 1_{e_i \leq \bar{e}}}{1,000} \quad (5)$$

where e_1, ..., $e_{1,000}$ denote the prediction errors on the 1,000 permuted data sets, and $\bar{e}$ denotes the prediction error on the original data. Here, $1_{(A)}$ is an indicator variable that equals 1 if the event A, and 0 otherwise.

2.10 | Comparison to single-scale approaches

We compared the performance of MSNR to two of the most commonly used single-scale network regression strategies, namely the individual edge model (Grillon et al., 2013; Lewis, Baldassarre, Committeri, Romani, & Corbetta, 2009) and community mean model

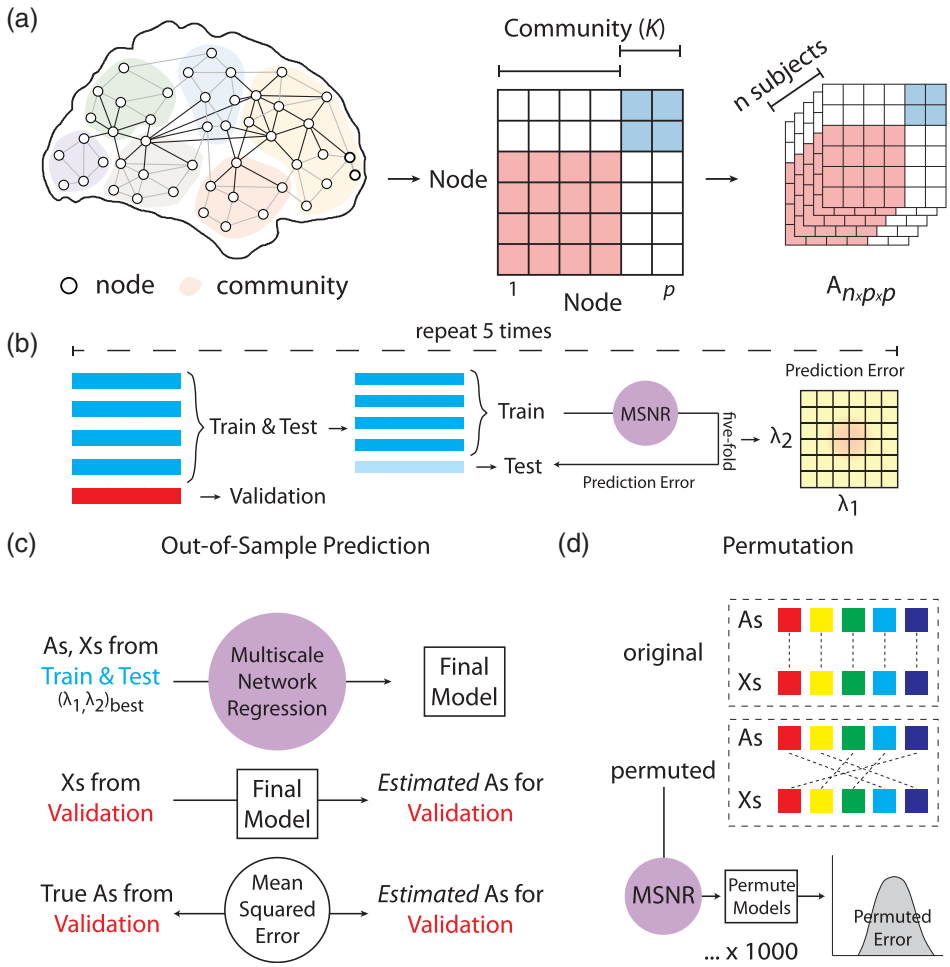


FIGURE 2 A schematic for MSNR model training and evaluation. (a) MSNR is designed to study the brain connectivity-phenotype relationship by taking into account both edge- and community-level information. The model takes in an $n \times p \times p$ matrix, where n is the number of subjects and p is the number of nodes in each symmetric adjacency matrix. The nodes belong to K communities, determined *a priori*. (b) 20% (n = 202) of the total sample (n = 1, 015) were randomly selected as the left-out validation data. We conducted five-fold cross-validation to select the values of the tuning parameters λ_1 and λ_2. These two parameters represent the nuclear norm penalty on the mean connectivity matrix (Θ) and the l_1 norm of the community-level connectivity-covariate relationship matrices (Γ^1, ..., Γ^q), respectively. This entire procedure was repeated five times. (c) The model was then trained using the tuning parameters determined in (b) on the rest 80% of the total data set (n = 813). Out-of-sample prediction error was then calculated as the *Frobenius* norm of the difference between the known and estimated connectivity matrices on the validation set. (d) We also evaluated the final model through a permutation procedure, where we disrupted the linkage between brain connectivity and covariate data to generate a null distribution of out-of-sample prediction error

(Betzel et al., 2014; King et al., 2018; Yan et al., 2019; Yu et al., 2019). These two approaches have been commonly used to study connectivity-phenotype relationships (Craddock et al., 2015; Varoquaux & Craddock, 2013) and differ primarily in terms of the scale of brain network examined (Figure S3). Details are as follows:

2.10.1 | Individual edge model

We vectorized the upper triangle of the adjacency matrix A^i for the i-th subject, i = 1, ..., n, in order to create a $n \times p(p-1)/2$ matrix. For each of the $p(p-1)/2$ columns of this matrix, we fit a linear regression to model that column using three covariates: age, sex, and in-scanner motion (Figure S3a). Specifically, we built a linear model for each edge in R, with the formula edge ~ age + sex + motion (Chen et al., 2016; Fjell et al., 2017; Wood, 2017; Xia et al., 2018, Figure S3b). We corrected the results for multiple comparisons using the false discovery rate (FDR, q < 0.05, Storey, 2002) and reshaped the $p(p-1)/2$ columns to a $p \times p$ matrix for visualizing significant coefficients. To calculate out-of-sample prediction error, we used linear models fit for all edges. The prediction error was calculated in the same way as in MSNR.

2.10.2 | Community mean model

Community-based linear models were built with mean *within-* and *between*-community connectivity as the dependent variables. The within-community connectivity is defined as

$$\frac{\sum_{j,j' \in C_k} A^i_{jj'}}{|C_k| \times (|C_k| - 1)} \quad (6)$$

where $A^i_{jj'}$ is the weighted edge strength between the node j and node j', both of which belong to the same community C_k, for the i-th subject. The cardinality of the community assignment vector, $|C_k|$, represents the number of nodes in the k-th community. The between-community connectivity is defined as

$$\frac{\sum_{j \in C_k, j' \in C_{k'}} A^i_{jj'}}{|C_k| \times |C_{k'}|} \quad (7)$$

where C_k and $C_{k'}$ represent two different communities, and $|C_k|$ and $|C_{k'}|$ are the number of nodes in each community, respectively.

By applying Equations (6) and (7) to each subject, we created a $n \times \left[\frac{K(K-1)}{2} + K\right]$ matrix. For each of the $\frac{K(K-1)}{2} + K$ columns of this matric, we fit a linear model to predict that column using three covariates: age, sex, and in-scanner motion. Similar to the edge-based model, we built a linear model for each edge in R, with the formula community ~ age + sex + motion (Chen et al., 2016; Fjell et al., 2017; Wood, 2017; Xia et al., 2018 Figure S3b). We corrected the results for multiple comparisons using the false discovery rate (FDR, $q < 0.05$, Storey, 2002) and reshaped the $\frac{K(K-1)}{2} + K$ columns to a $K \times K$ matrix to visualize significant coefficients. To calculate out-of-sample prediction error, we used linear models fit for all communities. The prediction error was calculated in the same way as in MSNR.

2.11 | Data and code availability

The data reported in this article have been deposited in database of Genotypes and Phenotypes (dbGaP): accession no. phs000607.v3.p2 [https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000607.v3.p2].

An implementation of the algorithm is available at bitbucket.org/rshinohara/networkregression.

3 | RESULTS

3.1 | MSNR shows high accuracy in a large developmental sample

We applied MSNR to data from the Philadelphia Neurodevelopmental Cohort (PNC) (Satterthwaite et al., 2014) in order to delineate meaningful brain-phenotype relationships. In total, we studied n = 1, 015 participants aged 8–22, who completed resting state functional neuroimaging as part of the PNC. We constructed functional connectivity matrices from a commonly-used parcellation scheme (=236 nodes) and community membership assignment (K = 13 communities; Power et al., 2011, Figure 2a). We first randomly selected 20% of the total sample as the left-out validation set (n = 202), with which we assessed the prediction performance of all subsequent models Figure 2b). The prediction performance was defined as the *Frobenius* norm of the difference between the observed and estimated adjacency matrices in the validation set (Figure 2c). For this proof-of-concept empirical study, we examined the association of functional connectivity with age, sex, and in-scanner motion. On the remaining 80% of the observations, we selected tuning parameters, λ_1 and λ_2, through five-fold cross-validation (Figure 2b). We iteratively refined the cross-validation grid (Figure 3a,b) in order to obtain the best possible tuning parameter values. Importantly, no boundary effect was observed in any of the iterations during successive grid searches, revealing a smooth convex landscape for the objective (Figure 3c).

We subsequently evaluated the model's out-of-sample prediction error on the validation set. The prediction error on the unseen data was comparable to the average error in the cross-validation procedure, indicating MSNR did not overfit to the training data. In addition, to determine the statistical significance of the model, we performed a permutation test to compare the model's prediction error to the distribution of prediction error under the null hypothesis of no association between brain networks and the predictors (Figure 3d). Specifically, we permuted the rows of the covariate data matrix 1,000 times. In each permutation, we disrupted the linkage between functional connectivity and phenotypes, while preserving the covariance structure of the covariates. We repeated the process of selecting tuning parameter values by cross-validation. Using these permuted data, we created a null distribution of prediction error. We then compared the true MSNR prediction error against this null distribution to estimate the p-value. We found that MSNR fit to the originally data outperformed any of the 1,000 null model when there was no linkage between covariates and connectivity ($p < .001$).

3.2 | MSNR recapitulates known individual differences in functional connectivity

Next, we investigated connectivity-phenotype relationships are summarized in the matrices Γ^1, Γ^2, and Γ^3 in the MSNR model. These coefficient matrices corresponded to the multivariate connectivity associated with age, sex, and motion, respectively (Figure 4). Of note, these matrices were relatively sparse, with 20.1, 19.5, and 16.6% entries exactly equaling to zero. Among the nonzero entries, many were previously reported in the literature. For example, stronger DMN connectivity was associated with age (Bluhm et al., 2008; Staffaroni et al., 2018) and with female sex (Bluhm et al., 2008). To summarize the multivariate patterns extracted by MSNR, We counted the number of positive and negative coefficients within each estimated matrix. These represent, respectively, positive and negative associations between community membership and age, sex, and in-scanner motion (Figure 4). Consistent with the previous literature (Gu et al., 2015; Satterthwaite et al., 2013), we found that as age increased, there were more within-community, rather than between-community connectivity, that strengthened with age (Figure 4a). Conversely, as age increased, there were more between-community, rather than within-community connectivity, that weakened with age.

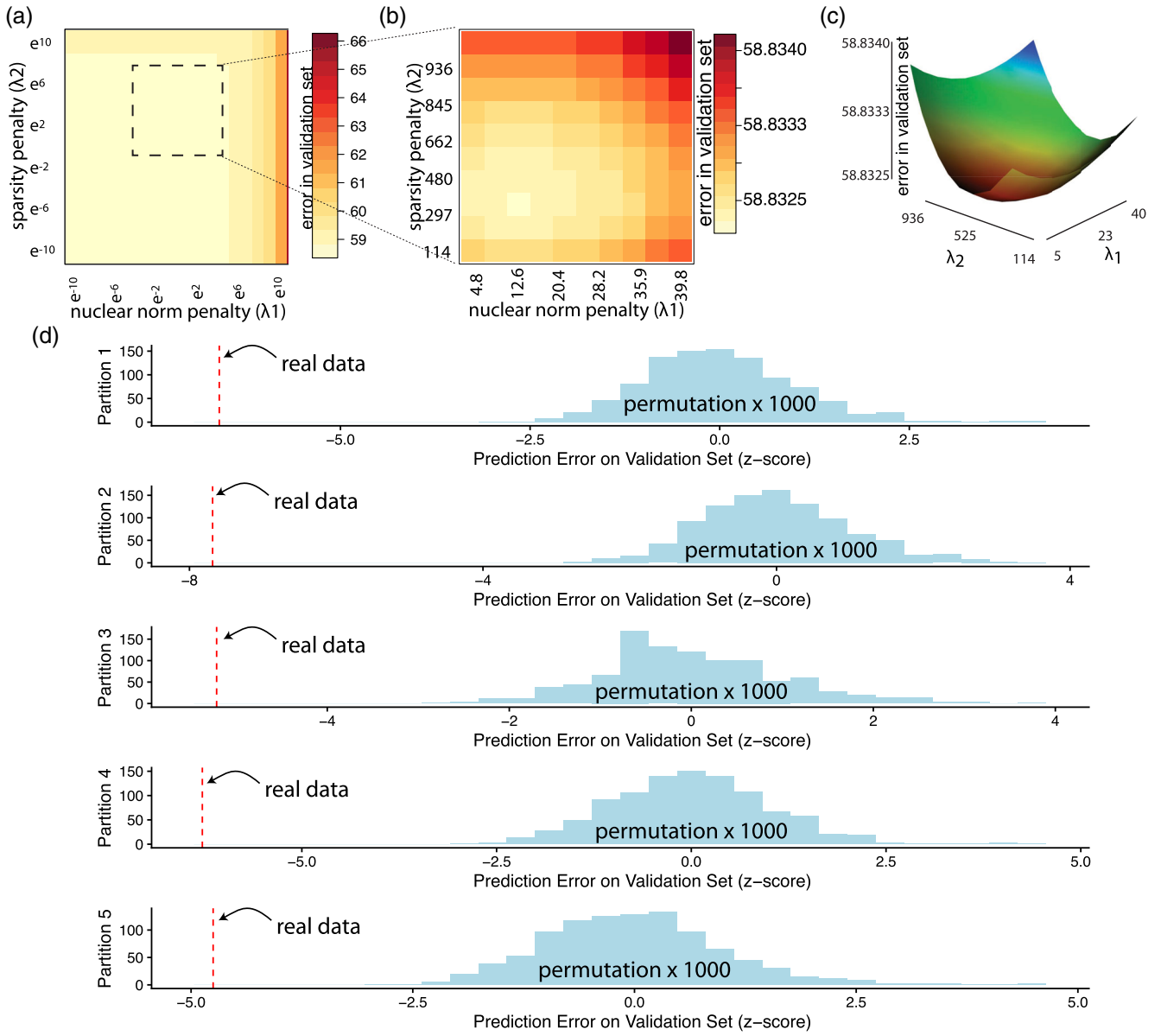


FIGURE 3 Tuning parameter selection and model evaluation of MSNR in a large neuroimaging dataset. (a) We used five-fold cross-validation in each data partition to estimate the test prediction error associated with various values of λ_1 and λ_2. The matrix here represents the average error across five different data partition. (b) After the initial search, we repeated the search on a finer scale, focusing on the range of λ_1 *and* λ_2 indicated by the dashed-line box. (c) As visualized, no boundary effect was observed in the grid search, revealing a smooth convex landscape for the objective, with warmer color indicating lower prediction error. (d) In each data partition, a permutation procedure showed that the MSNR fit to the original data significantly outperformed that to the permutated data with regards to prediction error on the validation set ($p < .001$). Consistent across five data partitions, the prediction error was consistently multiple standard deviations (*z*-score) below the mean of the null distributions

This pattern of results suggests that functional brain networks tend to segregate during normative brain development. Replicating findings from a previous report using mass-univariate analyses (Satterthwaite et al., 2015), here we observed that stronger within-community connectivity, rather than between-community, was more representative of functional brain networks in males; whereas stronger between-community connectivity, rather than within-community, was more representative of functional brain networks in females (Figure 4b). Finally, following on prior studies, we evaluated the degree to which the association between in-scanner motion and connectivity varies by inter-node distance, defined as the Euclidean distance between two spherical brain parcellations in the MNI space (Brett, Johnsrude, & Owen, 2002, Figure 4c). As expected, the MSNR coefficients for in-scanner motion in relation to functional connectivity were negatively correlated with the distances between pairs of communities. In other words, when two brain regions were close together, the presence of in-scanner motion was typically associated with an increase in their connectivity. This finding is consistent with prior reports that in-

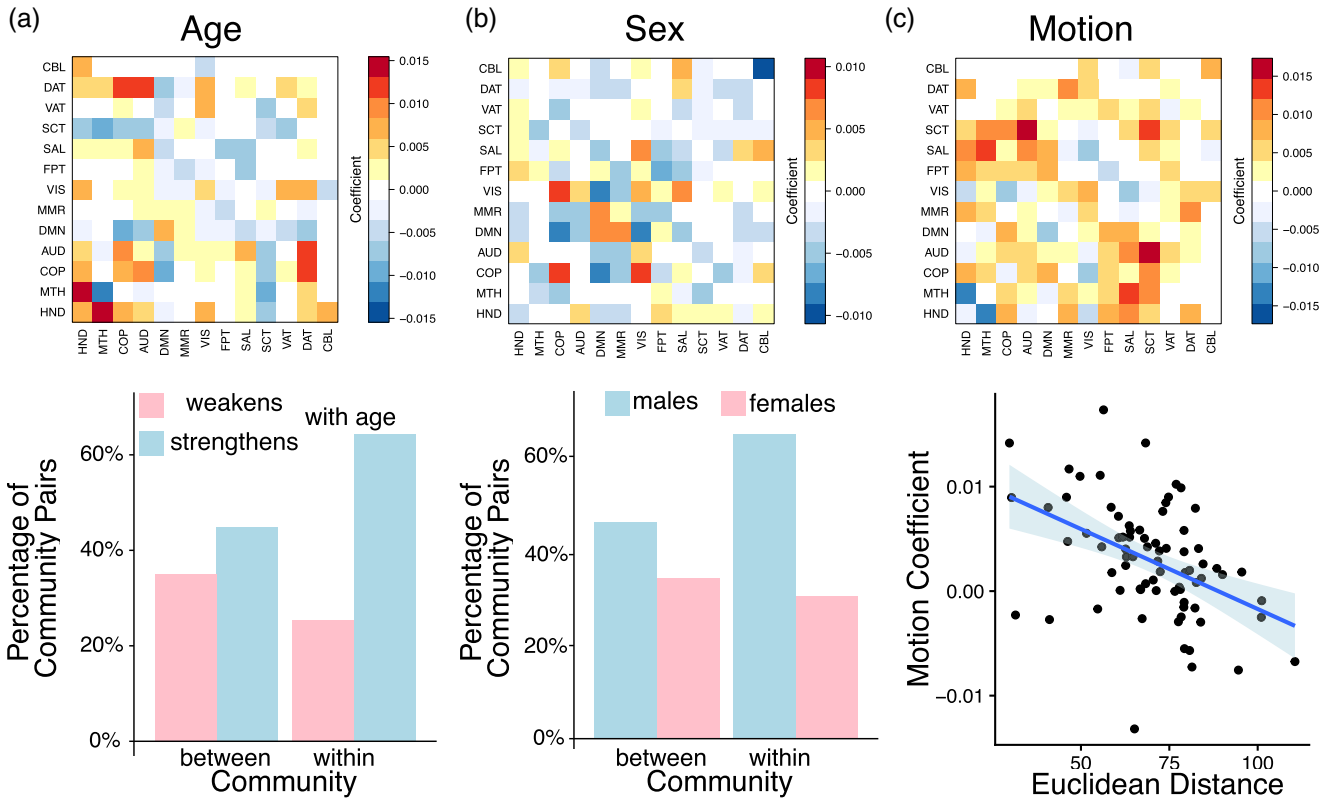


FIGURE 4 MSNR describes meaningful individual differences in brain connectivity. Top row represents the coefficient matrix Γ for each of the three phenotypes modeled in the MSNR. (a) We counted the number of positive and negative coefficients related to age. More within-community, rather than between-community, connectivity strengthened as the age increased. Conversely, more between-community, rather than within-community, connectivity weakened over age. (b) Stronger within-community than between-community connectivity was more representative of male functional brain networks, whereas stronger between-community than within-community connectivity was more representative of female functional brain networks. (c) Coefficient for in-scanner motion was negatively correlated with the average Euclidean distance between communities ($p < .001$)

scanner motion induces a distance-dependent bias in estimation of functional connectivity (Ciric et al., 2017; Satterthwaite et al., 2012). Of note, we verified that the Θ matrix was indeed low-rank (Figure S4).

3.3 | Comparison with typical mass-univariate single-scale strategies

Next, we compared MSNR to common single-scale mass-univariate approaches that make use of linear models at the edge-level or the community-level (Figure 5). We computed the out-of-sample performances of the two single-scale approaches using the left-out validation set. The prediction error of the community-based model on the validation set was poor, whereas that of the edge-based model was similar to MSNR (Figure 5a). Our estimation of prediction error of edge- and community-based models were likely to be overly optimistic, since we used all fitted models for the purpose of out-sample prediction.

Next, we examined the interpretability of coefficients obtained in each model after applying FDR correction to control for multiple comparisons in single-scale approaches (Storey, 2002). We found that while the edge-based model and MSNR achieved similar out-of-sample prediction, coefficients estimated in MSNR (Figure 5b) were more interpretable than the coefficients from edge-based models (Figure 5c). The number of coefficients in edge-based models for each covariate exceeded that of MSNR by three orders of magnitude. On the other hand, at the expense of higher prediction error, community-based models exhibited a level of interpretability that was similar to that exhibited by MSNR (Figure 5d).

4 | DISCUSSION

In the past decade, the neuroscience community has begun to complement the study of localized regions of the brain toward studying inter-regional relationships, or connectivity (Bassett & Sporns, 2017; Bzdok et al., 2016). The association of network architecture with development and aging throughout the lifespan (Betzel et al., 2014; Gu et al., 2015; Power et al., 2010), cognition (Bressler & Menon, 2010; Crossley et al., 2013; Park & Friston, 2013), and neuropsychiatric disorders (Bassett et al., 2018; Xia et al., 2018; Yu et al., 2019) is of

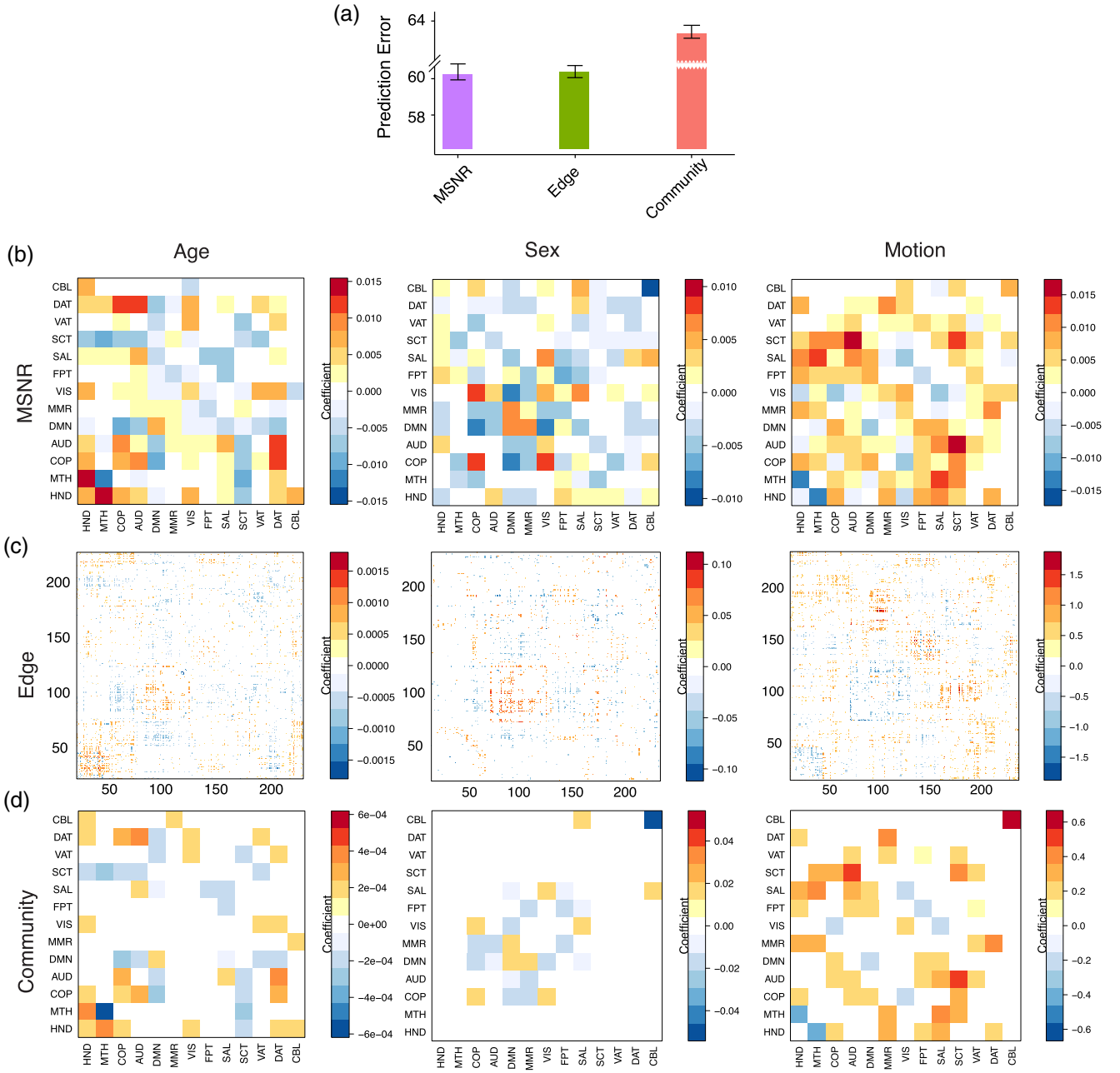


FIGURE 5 MSNR achieves a balance between out-of-sample prediction performance and model interpretability compared to common single-scale mass-univariate approaches. (a) We compared out-of-sample prediction performance of MSNR to common single-scale mass univariate analysis such as edge- and community-based methods. Among the three methods, the community-based approach had the highest prediction error. In contrast, MSNR had similar prediction error as the edge-based approach. Error bar represents the standard deviation across five different data partitions. (b) MSNR coefficients in Γ describe the multivariate connectivity-phenotype relationships. These correspond to age, sex, and in-scanner motion, respectively. Results from single-scale models were visualized in (c) for edge-based and in (d) for community-based approaches. Multiple comparisons were corrected using FDR

profound interest to the burgeoning network neuroscience literature. These brain-phenotype associations can be studied on the scale of individual edges (*micro*-scale), communities (*meso*-scale), or the network as a whole (*macro*-scale), with most existing approaches for analyzing networks, such as mass-univariate analyses, operate on a single scale.

In recent years, interest has centered on multi-scale modeling approaches (Jenatton et al., 2012; Y. Li et al., 2011, 2013), which aim to integrate information across homogeneous regions in the brain while still modeling data on finer scales. These methods have mainly focused on the problem of smoothing without prior knowledge of anatomical or functional parcellations of the brain, and have been

adapted for both classification (Romberg, Choi, Baraniuk, & Kingbury, 2000) and regression (Y. Li et al., 2011) as well as in longitudinal settings (Y. Li et al., 2013).

Building upon this recent work, we developed MSNR to study relationships between high-dimensional brain networks and variables of interest. Specifically, our proposal modeled the connectivity matrix for each subject by integrating both *micro*- and *meso*-scale network information. By applying a low-rank assumption to the mean connectivity network (Leonardi et al., 2013; K. Li et al., 2009; Smith et al., 2015) and a sparsity assumption to the community-level network (Meunier, Lambiotte, & Bullmore, 2010; Newman, 2006; Xia et al., 2018), we substantially decreased the number of parameters and encouraged the detection of interpretable brain-phenotype relationships.

Leveraging a large neuroimaging dataset of over 1000 youth, we demonstrated that MSNR recapitulated known individual differences in functional connectivity, including those related to development (Gu et al., 2015; Satterthwaite et al., 2013), sex differences (Satterthwaite et al., 2015), and in-scanner motion (Ciric et al., 2017; Satterthwaite et al., 2012). Notably, compared to common single-scale mass-univariate regression methods, MSNR achieved a balance between prediction performance and model complexity, with improved interpretability. Together, MSNR represents a new method for identifying individual differences in high-dimensional brain networks.

Several limitations of the MSNR approach should be noted. First, the term "scale" does not have a single definition. In fact, as pointed out by Betzel and Bassett (2017), scale can represent at least three different entities depending on the context: multi-scale topological structure, multi-scale temporal structure, and multi-scale spatial structure. In MSNR, we only considered multi-scale topological structure. Incorporating additional information from multiple scales beyond network topology will likely generate more nuanced and richer models for brain networks. Second, while we carefully conducted a permutation test to assess the statistical significance of the entire model, we did not provide an inferential procedure for determining the association between brain networks and each variable of interest. In particular, MSNR makes no claim of statistical significance for the coefficients in the matrices Γ^1, ..., Γ^q, which describe the community-level relationships with the covariates. Due to the inclusion of penalty terms in the MSNR framework, making such inferential statements is a challenging open problem.

In summary, by explicitly modeling variability at the edge and community levels, we developed a multi-scale network regression approach that achieved a balance between the trade-off of prediction and model complexity, potentially offering enhanced interpretability. Empirically, we demonstrated its advantages over alternative methods and illustrated its ability to uncover meaningful signals in a large neuroimaging dataset. Approaches such as MSNR have the potential to yield novel insights into brain-behavior relationships that incorporate realistic multi-scale network architecture.

ACKNOWLEDGMENTS

This project was supported in part by NIH R01 GM123993 (D.M.W.), DP5OD009145 (D.M.W.), R01NS085211 (R.T.S.), R01NS060910 (R.T.S.), R01MH112847 (R.T.S., T.D.S.), R21MH106799 (D.S.B., T.D.S.), and R01MH113550 (T.D.S., D.S.B.) from the National Institutes of Health, DMS-1252624 (D.M.W.) and DMS-1352060 (Z.M.) from the National Science Foundation, a Blavatnik Fellowship (C.H.X.), a Simons Investigator Award (D.M.W.), a Sloan Fellowship (Z.M.), and a pilot grant from the Center for Biomedical Computing and Analytics at the University of Pennsylvania (R.T.S.). D.B. was funded by BZ2/2-1, BZ2/3-1, and BZ2/4-1 from the Deutsche Forschungsgemeinschaft, IRTG2150 and Amazon Web Service grants. The PNC was supported by MH089983 and MH089924. D.S.B. also acknowledges support from the MacArthur Foundation, the Alfred P. Sloan Foundation, the ISI Foundation, the Paul Allen Foundation, the Army Research Laboratory, and the Office of Naval Research.

CONFLICT OF INTEREST

R.T.S. received consulting income from Genentech/Roche and income for editorial duties from the American Medical Association and Research Square. All other authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data reported in this paper have been deposited in database of Genotypes and Phenotypes (dbGaP): accession no. phs000607.v3.p2 [https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000607.v3.p2]. An implementation of the algorithm is available at bitbucket.org/rshinohara/networkregression.

ORCID

Cedric Huchuan Xia https://orcid.org/0000-0002-9703-333X
Bertrand Thirion https://orcid.org/0000-0001-5018-7895
Danielle S. Bassett https://orcid.org/0000-0002-6183-4493
Theodore D. Satterthwaite https://orcid.org/0000-0001-7072-9399

REFERENCES

Avants, B. B., Tustison, N. J., Song, G., Cook, P. A., Klein, A., & Gee, J. C. (2011). A reproducible evaluation of ANTs similarity metric performance in brain image registration. *NeuroImage*, *54*(3), 2033–2044. https://doi.org/10.1016/J.NEUROIMAGE.2010.09.025

Avants, B. B., Tustison, N. J., Wu, J., Cook, P. A., & Gee, J. C. (2011). An open source multivariate framework for n-tissue segmentation with evaluation on public data. *Neuroinformatics*, *9*(4), 381–400. https://doi.org/10.1007/s12021-011-9109-y

Bassett, D. S., & Siebenhühner, F. (2013). Multiscale network organization in the human brain. In *Multiscale analysis and nonlinear dynamics* (pp. 179–204). Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA. https://doi.org/10.1002/9783527671632.ch07

Bassett, D. S., & Sporns, O. (2017). Network neuroscience. *Nature Neuroscience*, *20*(3), 353–364. https://doi.org/10.1038/nn.4502

Bassett, D. S., Xia, C. H., & Satterthwaite, T. D. (2018). Understanding the emergence of neuropsychiatric disorders with network neuroscience. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *3*(9), 742–753. https://doi.org/10.1016/J.BPSC.2018.03.015

Betzel, R. F., & Bassett, D. S. (2017). Multi-scale brain networks. *NeuroImage*, *160*, 73–83. https://doi.org/10.1016/J.NEUROIMAGE.2016.11.006

Betzel, R. F., Byrge, L., He, Y., Goñi, J., Zuo, X.-N., & Sporns, O. (2014). Changes in structural and functional connectivity among resting-state

networks across the human lifespan. *NeuroImage*, *102*, 345–357. https://doi.org/10.1016/J.NEUROIMAGE.2014.07.067

Betzel, R. F., Medaglia, J. D., & Bassett, D. S. (2018). Diversity of mesoscale architecture in human and non-human connectomes. *Nature Communications*, *9*(1), 346. https://doi.org/10.1038/s41467-017-02681-z

Bien, J., & Witten, D. (2016). Penalized estimation in complex models. In P. Buhlmann, M. Kane, P. Drineas, & M. van der Laan (Eds.), *Handbook of big data* (pp. 285–299). Boca Raton, FL: CRC Press.

Biswal, B. B., Mennes, M., Zuo, X.-N., Gohel, S., Kelly, C., Smith, S. M., … Milham, M. P. (2010). Toward discovery science of human brain function. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(10), 4734–4739. https://doi.org/10.1073/pnas.0911855107

Bluhm, R. L., Osuch, E. A., Lanius, R. A., Boksman, K., Neufeld, R. W. J., Théberge, J., & Williamson, P. (2008). Default mode network connectivity: Effects of age, sex, and analytic approach. *Neuroreport*, *19*(8), 887–891. https://doi.org/10.1097/WNR.0b013e328300ebbf

Braun, U., Schäfer, A., Bassett, D. S., Rausch, F., Schweiger, J. I., Bilek, E., … Tost, H. (2016). Dynamic brain network reconfiguration as a potential schizophrenia genetic risk mechanism modulated by NMDA receptor function. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(44), 12568–12573. https://doi.org/10.1073/pnas.1608819113

Bressler, S. L., & Menon, V. (2010). Large-scale brain networks in cognition: Emerging methods and principles. *Trends in Cognitive Sciences*, *14*(6), 277–290. https://doi.org/10.1016/J.TICS.2010.04.004

Brett, M., Johnsrude, I. S., & Owen, A. M. (2002). The problem of functional localization in the human brain. *Nature Reviews Neuroscience*, *3*(3), 243–249. https://doi.org/10.1038/nrn756

Bzdok, D., & Ioannidis, J. P. A. (2019). Exploration, inference, and prediction in neuroscience and biomedicine. *Trends in Neurosciences*, *42*(4), 251–262. https://doi.org/10.1016/j.tins.2019.02.001

Bzdok, D., Varoquaux, G., Grisel, O., Eickenberg, M., Poupon, C., & Thirion, B. (2016). Formal models of the network co-occurrence underlying mental operations. *PLoS Computational Biology*, *12*(6), e1004994. https://doi.org/10.1371/journal.pcbi.1004994

Bzdok, D., & Yeo, B. T. T. (2017). Inference in the age of big data: Future perspectives on neuroscience. *NeuroImage*, *155*, 549–564. https://doi.org/10.1016/j.neuroimage.2017.04.061

Chen, H., Zhao, B., Cao, G., Proges, E. C., O'Shea, A., Woods, A. J., & Cohen, R. A. (2016). Statistical approaches for the study of cognitive and brain aging. *Frontiers in Aging Neuroscience*, *8*, 176. https://doi.org/10.3389/fnagi.2016.00176

Choi, D. S., Wolfe, P. J., & Airoldi, E. M. (2012). Stochastic blockmodels with a growing number of classes. *Biometrika*, *99*(2), 273–284. https://doi.org/10.1093/biomet/asr053

Ciric, R., Rosen, A. F. G., Erus, G., Cieslak, M., Adebimpe, A., Cook, P. A., … Satterthwaite, T. D. (2018). Mitigating head motion artifact in functional connectivity MRI. *Nature Protocols*, *13*(12), 2801–2826. https://doi.org/10.1038/s41596-018-0065-y

Ciric, R., Wolf, D. H., Power, J. D., Roalf, D. R., Baum, G. L., Ruparel, K., … 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. https://doi.org/10.1016/J.NEUROIMAGE.2017.03.020

Craddock, R. C., Tungaraza, R. L., & Milham, M. P. (2015). Connectomics and new approaches for analyzing human brain functional connectivity. *GigaScience*, *4*(1), 13. https://doi.org/10.1186/s13742-015-0045-x

Crossley, N. A., Mechelli, A., Vertes, P. E., Winton-Brown, T. T., Patel, A. X., Ginestet, C. E., … Bullmore, E. T. (2013). Cognitive relevance of the community structure of the human brain functional coactivation network. *Proceedings of the National Academy of Sciences*, *110*(28), 11583–11588. https://doi.org/10.1073/pnas.1220826110

Durante, D., & Dunson, D. B. (2018). Bayesian inference and testing of group differences in brain networks. *Bayesian Analysis*, *13*(1), 29–58. https://doi.org/10.1214/16-BA1030

Durante, D., Dunson, D. B., & Vogelstein, J. T. (2017). Nonparametric Bayes modeling of populations of networks. *Journal of the American Statistical Association*, *112*(520), 1516–1530. https://doi.org/10.1080/01621459.2016.1219260

Eavani, H., Satterthwaite, T. D., Filipovych, R., Gur, R. E., Gur, R. C., & Davatzikos, C. (2015). Identifying sparse connectivity patterns in the brain using resting-state fMRI. *NeuroImage*, *105*(2015), 286–299. https://doi.org/10.1016/j.neuroimage.2014.09.058

Fazel, M. (2002). *Matrix rank minimization with applications*. Palo Alto, CA: Stanford University.

Fjell, A. M., Sneve, M. H., Grydeland, H., Storsve, A. B., Amlien, I. K., Yendiki, A., & Walhovd, K. B. (2017). Relationship between structural and functional connectivity change across the adult lifespan: A longitudinal investigation. *Human Brain Mapping*, *38*(1), 561–573. https://doi.org/10.1002/hbm.23403

Fornito, A., Zalesky, A., & Breakspear, M. (2015). The connectomics of brain disorders. *Nature Reviews Neuroscience*, *16*(3), 159–172. https://doi.org/10.1038/nrn3901

Fosdick, B. K., & Hoff, P. D. (2015). Testing and modeling dependencies between a network and nodal attributes. *Journal of the American Statistical Association*, *110*(511), 1047–1056. https://doi.org/10.1080/01621459.2015.1008697

Greve, D. N., & Fischl, B. (2009). Accurate and robust brain image alignment using boundary-based registration. *NeuroImage*, *48*(1), 63–72. https://doi.org/10.1016/j.neuroimage.2009.06.060

Grillon, M.-L., Oppenheim, C., Varoquaux, G., Charbonneau, F., Devauchelle, A.-D., Krebs, M.-O., … Huron, C. (2013). Hyperfrontality and hypoconnectivity during refreshing in schizophrenia. *Psychiatry Research: Neuroimaging*, *211*(3), 226–233. https://doi.org/10.1016/J.PSCYCHRESNS.2012.09.001

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. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(44), 13681–13686. https://doi.org/10.1073/pnas.1502829112

Hallquist, M. N., Hwang, K., & Luna, B. (2013). The nuisance of nuisance regression: Spectral misspecification in a common approach to resting-state fMRI preprocessing reintroduces noise and obscures functional connectivity. *NeuroImage*, *82*, 208–225. https://doi.org/10.1016/J.NEUROIMAGE.2013.05.116

Hastie, T., Tibshirani, R., Wainwright, M., Tibshirani, R., & Wainwright, M. (2015). *Statistical learning with sparsity*. Boca Raton, FL: Chapman and Hall/CRC. https://doi.org/10.1201/b18401

Hastie, T., Tibshirani, R., & Friedman, J. (2008). *The elements of statistical learning* (2nd ed.). Stanford, CA: Springer.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning* (8th ed.). New York, NY: Springer. https://doi.org/10.1007/978-1-4614-7138-7

Jenatton, R., Gramfort, A., Michel, V., Obozinski, G., Eger, E., Bach, F., & Thirion, B. (2012). Multiscale mining of fMRI data with hierarchical structured sparsity. *SIAM Journal on Imaging Sciences*, *5*(3), 835–856. https://doi.org/10.1137/110832380

Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, *17*(2), 825–841. https://doi.org/10.1006/NIMG.2002.1132

Jernigan, T. L., Brown, T. T., Hagler, D. J., Akshoomoff, N., Bartsch, H., Newman, E., … Pediatric Imaging, Neurocognition and Genetics Study. (2016). The pediatric imaging, neurocognition, and genetics (PING) data repository. *NeuroImage*, *124*(Pt B), 1149–1154. https://doi.org/10.1016/j.neuroimage.2015.04.057

Kernbach, J. M., Satterthwaite, T. D., Bassett, D. S., Smallwood, J., Margulies, D., Krall, S., … Bzdok, D. (2018). Shared endo-phenotypes of default mode dysfunction in attention deficit/hyperactivity disorder and autism spectrum disorder. *Translational Psychiatry*, *8*(1), 133. https://doi.org/10.1038/s41398-018-0179-6

King, J. B., Prigge, M. B. D., King, C. K., Morgan, J., Dean, D. C., Freeman, A., … Anderson, J. S. (2018). Evaluation of differences in temporal synchrony between brain regions in individuals with autism and typical development. *JAMA Network Open*, *1*(7), e184777. https://doi.org/10.1001/jamanetworkopen.2018.4777

Klein, A., Andersson, J., Ardekani, B. A., Ashburner, J., Avants, B., Chiang, M.-C., … Parsey, R. V. (2009). Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *NeuroImage*, *46*(3), 786–802. https://doi.org/10.1016/J.NEUROIMAGE.2008.12.037

Leonardi, N., Richiardi, J., Gschwind, M., Simioni, S., Annoni, J.-M., Schluep, M., … Van De Ville, D. (2013). Principal components of functional connectivity: A new approach to study dynamic brain connectivity during rest. *NeuroImage*, *83*, 937–950. https://doi.org/10.1016/J.NEUROIMAGE.2013.07.019

Lewis, C. M., Baldassarre, A., Committeri, G., Romani, G. L., & Corbetta, M. (2009). Learning sculpts the spontaneous activity of the resting human brain. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(41), 17558–17563. https://doi.org/10.1073/pnas.0902455106

Li, K., Guo, L., Nie, J., Li, G., & Liu, T. (2009). Review of methods for functional brain connectivity detection using fMRI. *Computerized Medical Imaging and Graphics*, *33*(2), 131–139. https://doi.org/10.1016/J.COMPMEDIMAG.2008.10.011

Li, Y., Gilmore, J. H., Shen, D., Styner, M., Lin, W., & Zhu, H. (2013). Multiscale adaptive generalized estimating equations for longitudinal neuroimaging data. *NeuroImage*, *72*, 91–105. https://doi.org/10.1016/J.NEUROIMAGE.2013.01.034

Li, Y., Zhu, H., Shen, D., Lin, W., Gilmore, J. H., & Ibrahim, J. G. (2011). Multiscale adaptive regression models for neuroimaging data. *Journal of the Royal Statistical Society. Series B, Statistical Methodology*, *73*(4), 559–578. https://doi.org/10.1111/j.1467-9868.2010.00767.x

Meunier, D., Lambiotte, R., & Bullmore, E. T. (2010). Modular and hierarchically modular organization of brain networks. *Frontiers in Neuroscience*, *4*, 200. https://doi.org/10.3389/fnins.2010.00200

Newman, M. E. J. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, *103*(23), 8577–8582. https://doi.org/10.1073/pnas.0601602103

Ng, B., Varoquaux, G., Poline, J. B., & Thirion, B. (2012). A novel sparse graphical approach for multimodal brain connectivity inference. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), LNCS* (Vol. 7510 (pp. 707–714). New York, NY: Springer Verlag. https://doi.org/10.1007/978-3-642-33415-3_87

Park, H. J., & Friston, K. (2013). Structural and functional brain networks: From connections to cognition. *Science*, *342*(November), 579–588. https://doi.org/10.1126/science.1238411

Power, J. D., Cohen, A. L., Nelson, S. M., Wig, G. S., Barnes, K. A., Church, J. A., … Petersen, S. E. (2011). Functional network organization of the human brain. *Neuron*, *72*(4), 665–678. https://doi.org/10.1016/J.NEURON.2011.09.006

Power, J. D., Fair, D. A., Schlaggar, B. L., & Petersen, S. E. (2010). The development of human functional brain networks. *Neuron*, *67*(5), 735–748. https://doi.org/10.1016/J.NEURON.2010.08.017

Raichle, M. E. (2015). The Brain's default mode network. *Annual Review of Neuroscience*, *38*(1), 433–447. https://doi.org/10.1146/annurev-neuro-071013-014030

Recht, B., Fazel, M., & Parrilo, P. A. (2010). Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM Review*, *52*(3), 471–501. https://doi.org/10.1137/070697835

Romberg, J., Hyeokho Choi, Baraniuk, R., & Kingbury, N. (2000). *Multiscale classification using complex wavelets and hidden Markov tree models*. Proceedings 2000 International Conference on Image Processing (Cat. No.00CH37101). Vol. 2. pp. 371–374. IEEE https://doi.org/10.1109/ICIP.2000.899396

Rosvall, M., & Bergstrom, C. T. (2008). Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(4), 1118–1123.

Rubinov, M., & Sporns, O. (2009). Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage*, *52*(3), 1059–1069. https://doi.org/10.1016/j.neuroimage.2009.10.003

Satterthwaite, T. D., Connolly, J. J., Ruparel, K., Calkins, M. E., Jackson, C., Elliott, M. A., … Gur, R. E. (2016). The Philadelphia neurodevelopmental cohort: A publicly available resource for the study of normal and abnormal brain development in youth. *NeuroImage*, *124*, 1115–1119. https://doi.org/10.1016/J.NEUROIMAGE.2015.03.056

Satterthwaite, T. D., Elliott, M. A., Ruparel, K., Loughead, J., Prabhakaran, K., Calkins, M. E., … Gur, R. E. (2014). Neuroimaging of the Philadelphia neurodevelopmental cohort. *NeuroImage*, *86*(2014), 544–553. https://doi.org/10.1016/j.neuroimage.2013.07.064

Satterthwaite, T. D., Wolf, D. H., Loughead, J., Ruparel, K., Elliott, M. A., Hakonarson, H., … 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*(1), 623–632. https://doi.org/10.1016/j.neuroimage.2011.12.063

Satterthwaite, T. D., Wolf, D. H., Roalf, D. R., Ruparel, K., Erus, G., Vandekar, S., … Gur, R. C. (2015). Linked sex differences in cognition and functional connectivity in youth. *Cerebral Cortex*, *25*(9), 2383–2394. https://doi.org/10.1093/cercor/bhu036

Satterthwaite, T. D., Wolf, D. H., Ruparel, K., Erus, G., Elliott, M. A., Eickhoff, S. B., … Gur, R. C. (2013). Heterogeneous impact of motion on fundamental patterns of developmental changes in functional connectivity during youth. *NeuroImage*, *83*(2013), 45–57. https://doi.org/10.1016/j.neuroimage.2013.06.045

Schumann, G., Loth, E., Banaschewski, T., Barbot, A., Barker, G., Büchel, C., … IMAGEN consortium. (2010). The IMAGEN study: Reinforcement-related behaviour in normal brain function and psychopathology. *Molecular Psychiatry*, *15*(12), 1128–1139. https://doi.org/10.1038/mp.2010.4

Smith, S. M., Nichols, T. E., Vidaurre, D., Winkler, A. M., Behrens, J., Glasser, M. F., … Miller, K. L. (2015). A positive-negative mode of population covariation links brain connectivity, demographics and behavior. *Nature Neuroscience*, *18*(September), 1–7. https://doi.org/10.1038/nn.4125

Sporns, O., & Betzel, R. F. (2016). Modular brain networks. *Annual Review of Psychology*, *67*(1), 613–640. https://doi.org/10.1146/annurev-psych-122414-033634

Staffaroni, A. M., Brown, J. A., Casaletto, K. B., Elahi, F. M., Deng, J., Neuhaus, J., … Kramer, J. H. (2018). The longitudinal trajectory of default mode network connectivity in healthy older adults varies as a function of age and is associated with changes in episodic memory and processing speed. *Journal of Neuroscience*, *38*(11), 2809–2817. https://doi.org/10.1523/JNEUROSCI.3067-17.2018

Storey, J. D. (2002). A direct approach to false discovery rates. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *64*(3), 479–498. https://doi.org/10.1111/1467-9868.00346

Tang, M., Athreya, A., Sussman, D. L., Lyzinski, V., & Priebe, C. E. (2017). A nonparametric two-sample hypothesis testing problem for random graphs. *Bernoulli*, *23*(3), 1599–1630. https://doi.org/10.3150/15-BEJ789

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, *58*(1), 267–288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

Tustison, N. J., Avants, B. B., Cook, P. A., Zheng, Y., Egan, A., Yushkevich, P. A., & Gee, J. C. (2010). N4ITK: Improved N3 bias

correction. *IEEE Transactions on Medical Imaging*, *29*(6), 1310–1320. https://doi.org/10.1109/TMI.2010.2046908

Tustison, N. J., Cook, P. A., Klein, A., Song, G., Das, S. R., Duda, J. T., … Avants, B. B. (2014). Large-scale evaluation of ANTs and FreeSurfer cortical thickness measurements. *NeuroImage*, *99*, 166–179. https://doi.org/10.1016/j.neuroimage.2014.05.044

Van Essen, D. C., Ugurbil, K., Auerbach, E., Barch, D., Behrens, T. E. J., Bucholz, R., … Yacoub, E. (2012). The human connectome project: A data acquisition perspective. *NeuroImage*, *62*(4), 2222–2231. https://doi.org/10.1016/J.NEUROIMAGE.2012.02.018

Varoquaux, G., & Craddock, R. C. (2013). Learning and comparing functional connectomes across subjects. *NeuroImage*, *80*, 405–415. https://doi.org/10.1016/J.NEUROIMAGE.2013.04.007

Wang, H., Suh, J. W., Das, S. R., Pluta, J. B., Craige, C., & Yushkevich, P. A. (2013). Multi-atlas segmentation with joint label fusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *35*(3), 611–623. https://doi.org/10.1109/TPAMI.2012.143

Wang, L., Durante, D., Jung, R. E., & Dunson, D. B. (2017). Bayesian network-response regression. *Bioinformatics*, *33*(12), 1859–1866. https://doi.org/10.1093/bioinformatics/btx050

Wood, S. N. (2011). Fast stable REML and ML estimation of semi-parametric GLMs. *Journal of the Royal Statistical Society, Series B (Statistical Methodology)*, *73*(1), 3–36. https://doi.org/10.1111/j.1467-9868.2010.00749.x

Wood, S. N. (2017). *Generalized additive models: An introduction with R*. Boca Raton, FL: Chapman and Hall/CRC.

Wood, S. N., Pya, N., & Säfken, B. (2016). Smoothing parameter and model selection for general smooth models. *Journal of the American Statistical Association*, *111*(516), 1548–1563. https://doi.org/10.1080/01621459.2016.1180986

Xia, C. H., Ma, Z., Ciric, R., Gu, S., Betzel, R. F., Kaczkurkin, A. N., … Satterthwaite, T. D. (2018). Linked dimensions of psychopathology and connectivity in functional brain networks. *Nature Communications*, *9*(1), 3003. https://doi.org/10.1038/s41467-018-05317-y

Yan, C.-G., Chen, X., Li, L., Castellanos, F. X., Bai, T.-J., Bo, Q.-J., … Zang, Y.-F. (2019). Reduced default mode network functional connectivity in patients with recurrent major depressive disorder. *Proceedings of the National Academy of Sciences of the United States of America*, *116*, 9083. https://doi.org/10.1073/pnas.1900390116

Young, S. J., & Scheinerman, E. R. (2007). Random dot product graph models for social networks. In *Algorithms and models for the web-graph* (pp. 138–149). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-77004-6_11

Yu, M., Linn, K. A., Shinohara, R. T., Oathes, D. J., Cook, P. A., Duprat, R., … Sheline, Y. I. (2019). Childhood trauma history is linked to abnormal brain connectivity in major depression. https://doi.org/10.1073/pnas.1900801116.

Zalesky, A., Fornito, A., & Bullmore, E. T. (2010). Network-based statistic: Identifying differences in brain networks. *NeuroImage*, *53*(4), 1197–1207. https://doi.org/10.1016/J.NEUROIMAGE.2010.06.041

Zapala, M. A., & Schork, N. J. (2012). Statistical properties of multivariate distance matrix regression for high-dimensional data analysis. *Frontiers in Genetics*, *3*, 190. https://doi.org/10.3389/fgene.2012.00190

Zhang, J., Sun, W. W., & Li, L. (2018). Network response regression for modeling population of networks with covariates. https://www.researchgate.net/publication/328161028_Network_Response_Regression_for_Modeling_Population_of_Networks_with_Covariates

SUPPORTING INFORMATION

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

How to cite this article: Xia CH, Ma Z, Cui Z, et al. Multi-scale network regression for brain-phenotype associations. *Hum Brain Mapp*. 2020;1–14. https://doi.org/10.1002/hbm.24982