



Graph auto-encoding brain networks with applications to analyzing large-scale brain imaging datasets

Meimei Liu^a, Zhengwu Zhang^b, David B. Dunson^{c,*}

^a Virginia Tech, Blacksburg, VA 24060, USA

^b The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

^c Duke University, Durham, NC 27705, USA

ARTICLE INFO

Keywords:

Brain networks
Non-linear factor analysis
Graph CNN
Replicated networks
Variational auto-encoder

ABSTRACT

There has been a huge interest in studying human brain connectomes inferred from different imaging modalities and exploring their relationships with human traits, such as cognition. Brain connectomes are usually represented as networks, with nodes corresponding to different regions of interest (ROIs) and edges to connection strengths between ROIs. Due to the high-dimensionality and non-Euclidean nature of networks, it is challenging to depict their population distribution and relate them to human traits. Current approaches focus on summarizing the network using either pre-specified topological features or principal components analysis (PCA). In this paper, building on recent advances in deep learning, we develop a nonlinear latent factor model to characterize the population distribution of brain graphs and infer their relationships to human traits. We refer to our method as Graph AuTo-Encoding (GATE). We applied GATE to two large-scale brain imaging datasets, the Adolescent Brain Cognitive Development (ABCD) study and the Human Connectome Project (HCP) for adults, to study the structural brain connectome and its relationship with cognition. Numerical results demonstrate huge advantages of GATE over competitors in terms of prediction accuracy, statistical inference, and computing efficiency. We found that the structural connectome has a stronger association with a wide range of human cognitive traits than was apparent using previous approaches.

1. Introduction

Understanding the brain connectome and how it relates to human traits and various clinical variables has drawn huge attention (Craddock et al., 2013; Fornito et al., 2013; Jones et al., 2013; Park and Friston, 2013). This has motivated large neuroimaging studies with thousands of subjects, such as the UK Biobank (UKB) (Miller et al., 2016), the Adolescent Brain Cognitive Development (ABCD) study (Casey et al., 2018), and the Human Connectome Project (HCP) (Van Essen et al., 2013). Through these studies, there have been dramatic improvements in the ability to reconstruct brain connectomes thanks to advanced hardware (Glasser et al., 2016), novel image acquisition protocols (Glasser et al., 2016; Tuch, 2004), and new reconstruction algorithms (Smith et al., 2012; Zhang et al., 2018b). In this paper, we are particularly interested in diffusion magnetic resonance imaging (dMRI), which is a commonly used technique that measures the movement of water molecules along major fiber bundles in white matter (WM) fiber tracts, enabling reconstruction of individual-level microstructural brain networks delineating anatomical connections between brain regions. This paper aims at devel-

oping advanced analysis methods for the brain structural connectomes recovered from diffusion MRI data.

Let A_i represent the structural connectivity recovered from subject i , with element $A_{i[uv]}$ measuring white matter connections between brain regions u and v . Using n individual brain networks, we are interested in (a) appropriately summarizing each individual brain network in a parsimonious manner, isolating unique features of the network without discarding valuable information, (b) inferring relationships between brain networks and human traits, and (c) characterizing variation across individuals in their network structure.

There are existing methods relevant to these goals. Based on a latent space characterization, Durante et al. (2017) proposed a random effects model to represent the population distribution of brain networks. Their approach clusters individuals based on brain structure and allows inferences on group differences (Durante and Dunson, 2018). Disadvantages include the highly computationally intensive implementation and coarse characterization of individual differences based on clustering. There is also literature on PCA-style approaches. One possibility is to simply stack the adjacency matrices A_i for individuals $i = 1, \dots, n$

* Corresponding author.

E-mail addresses: meimeiliu@vt.edu (M. Liu), zhengwu_zhang@unc.edu (Z. Zhang), dunson@duke.edu (D.B. Dunson).

<https://doi.org/10.1016/j.neuroimage.2021.118750>.

Received 18 May 2021; Received in revised form 21 October 2021; Accepted 20 November 2021

Available online 22 November 2021.

1053-8119/© 2021 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

into a tensor, and then apply tensor PCA and its variants to get summary scores of networks (Zhang and Dunson, 2019; Zhang et al., 2019). These scores are treated as brain network surrogates in subsequent analyses, e.g., relating brain networks to human traits (Zhang et al., 2019). Tensor PCA is relatively efficient computationally while providing a simple low-dimensional summary of an individual's brain structure, but it is linear, limiting the ability to represent brain networks parsimoniously. Other matrix-based approaches, including spatial independent component analysis (ICA), non-negative matrix factorization (NMF), and spatial sparse coding algorithms, are also widely used in the analysis of functional brain connectomes; see Beckmann et al. (2005), Xie et al. (2017), and references therein. An alternative is graph representations based on geometric deep learning. Kipf and Welling (2017), Hamilton et al. (2017) proposed graph convolutional networks (GCN) that use structure information in learning a low-dimensional feature representation for each node in a graph. Kawahara et al. (2017), Ktena et al. (2018) applied GCN to functional brain connectomes for classification and similarity ranking. Zhao et al. (2019) proposed a variational autoencoder-based Gaussian mixture model for functional brain connectomes classification. Advanced graph embeddings for structural brain connectomes are still lacking.

A major motivation of this article is to develop a non-linear latent factor modeling approach to (1) provide a characterization of the population distribution of brain graphs and (2) output low-dimensional features that can be used to summarize an individuals graph. Compared with the original high-dimensional adjacency metrics, the low-dimensional features of brain networks can further facilitate visualization, prediction, and inference on relationships between connectomes and human traits. With this motivation, we are particularly intrigued by deep neural networks for non-linear dimension reduction. Generative algorithms, such as Variational Auto-Encoders (VAEs) (Kingma and Welling, 2014; Rezende et al., 2014), have proven successful in representing images via low dimensional latent variables. VAEs model the population distribution of image data through a simple distribution for the latent variables combined with a complex non-linear mapping function. A key to the success of such methods is the use of convolutional operators to encode symmetries often present in images. However, structural brain networks have a fundamentally different geometric structure, and such methods cannot be employed directly.

We develop a model-based variational Graph Auto-Encoder (GATE) for brain connectome analysis. GATE consists of two components. The first component is a generative model that specifies how the latent variables z_i give rise to the observations A_i through a non-linear mapping, parametrized by neural networks. The second component is an inference model that learns the inverse mapping from A_i to z_i . Our main contributions can be summarized as follows.

First, GATE learns the embedding and the population distribution of brain connectomes simultaneously. This is achieved by: (1) a nonlinear latent factor model to obtain a low-dimensional representation z_i of brain network A_i ; and (2) a hierarchical generative model designed to learn the conditional distribution $p(A_i | z_i)$ so that one can accurately reconstruct the brain network from the latent embedding. We model each cell $A_{i[uv]}$ in A_i using a latent space model (Hoff et al., 2002), with the latent coordinates of the regions u and v varying as a nonlinear function of the individual-specific features z_i . This step involves a novel graph convolutional network that relies on the intrinsic locality of the brain networks to propagate node-specific k -nearest neighbor information.

Second, we extend GATE to relate human phenotypes to brain structural connectivity, which we refer to as regression with GATE (reGATE). reGATE is a supervised embedding method that simultaneously learns the population distribution of brain networks, network embeddings, and a predictive model for human traits. Although there has been some work integrating regression models and VAEs, the focus has been on multi-stage approaches; e.g., see Yoo et al. (2017) as an example. reGATE can generate from the population distribution of brain networks conditionally on the value of a human trait. This provides invaluable information

about how traits and brain networks are associated while characterizing variation across individuals. We further draw inference on selected network summary measures of interest, such as network density and average path length, to understand how these properties are distributed depending on human traits.

We apply GATE and reGATE to brain connectomes from ABCD and HCP and find strong relationships between structural connectomes and cognition traits in both datasets. reGATE shows superior performance in predicting the relationship between cognition and brain connectomes, particularly when trained with data from large numbers of individuals. For example, using more than five thousand brain scans in the ABCD study, reGATE improves prediction of cognitive traits by 30% – 40% compared with existing competitors. Through detailed inference based on reGATE, we show that individuals with high cognitive traits tend to have denser connections between hemispheres, higher overall network density, and lower average path length. Such network summary measures have higher variability across the children evaluated in ABCD compared with adults in HCP.

2. Methods

2.1. Brain imaging datasets and structural connectome extraction

We focus on two large datasets in this paper: the Adolescent Brain Cognitive Development (ABCD) dataset and the Human Connectome Project (HCP) dataset.

ABCD dataset: The ABCD study in the United States focuses on tracking brain development from childhood through adolescence to understand biological and environmental factors that can affect the brain's developmental trajectory. The research consortium consists of 21 research sites across the country and invited 11,878 9-10-year-old children to participate. Researchers track their biological and behavioral development through adolescence into young adulthood. The dataset can be downloaded from NIH Data Archive (NDA, <https://nda.nih.gov>). The imaging protocol is harmonized for three types of 3T scanners: Siemens Prisma, General Electric (GE) 750, and Philips. We downloaded the structural T1 MRI and diffusion MRI (dMRI) data for 5252 subjects from the ABCD 2.0 release in NDA. The structural T1 images were acquired with an isotropic resolution of 1 mm³. The diffusion MRI images were obtained based on imaging parameters: 1.7 mm³ resolution, four different b -values ($b = 500, 1000, 2000, 3000$) and 96 diffusion directions. There are 6 directions at $b = 500$, 15 directions at $b = 1000$, 15 directions at $b = 2000$, and 60 directions at $b = 3000$. Multiband factor 3 is used for dMRI acceleration. See Casey et al. (2018) for more details about data acquisition and preprocessing of the ABCD data.

HCP dataset: The HCP aims at characterizing human brain connectivity in about 1,200 healthy adults to enable detailed comparisons between brain circuits, behavior, and genetics at the level of individual subjects (Van Essen et al., 2012). Customized scanners were used to produce high-quality and consistent data to measure brain connectivity. The data containing various traits and MRI data can be easily accessed through <https://db.humanconnectome.org/>.

To obtain structural connectomes, we used a state-of-the-art dMRI data preprocessing framework – population-based structural connectome (PSC) mapping (Zhang et al., 2018a). PSC uses a reproducible probabilistic tractography algorithm (Girard et al., 2014; Maier-Hein et al., 2017) to generate whole-brain tractography. PSC borrows anatomical information from high-resolution T1 images to reduce bias in the reconstruction of tractography. We used the Desikan-Killiany atlas (Desikan et al., 2006) to define the brain regions of interest (ROIs) corresponding to the nodes in the structural connectivity network. The Desikan-Killiany parcellation has 68 cortical surface regions with 34 nodes in each hemisphere. For each pair of ROIs, we extracted the streamlines connecting them. In this process, several procedures were used to increase reproducibility: (1) each gray matter ROI is dilated to include a small portion of white matter region, (2) streamlines connect-

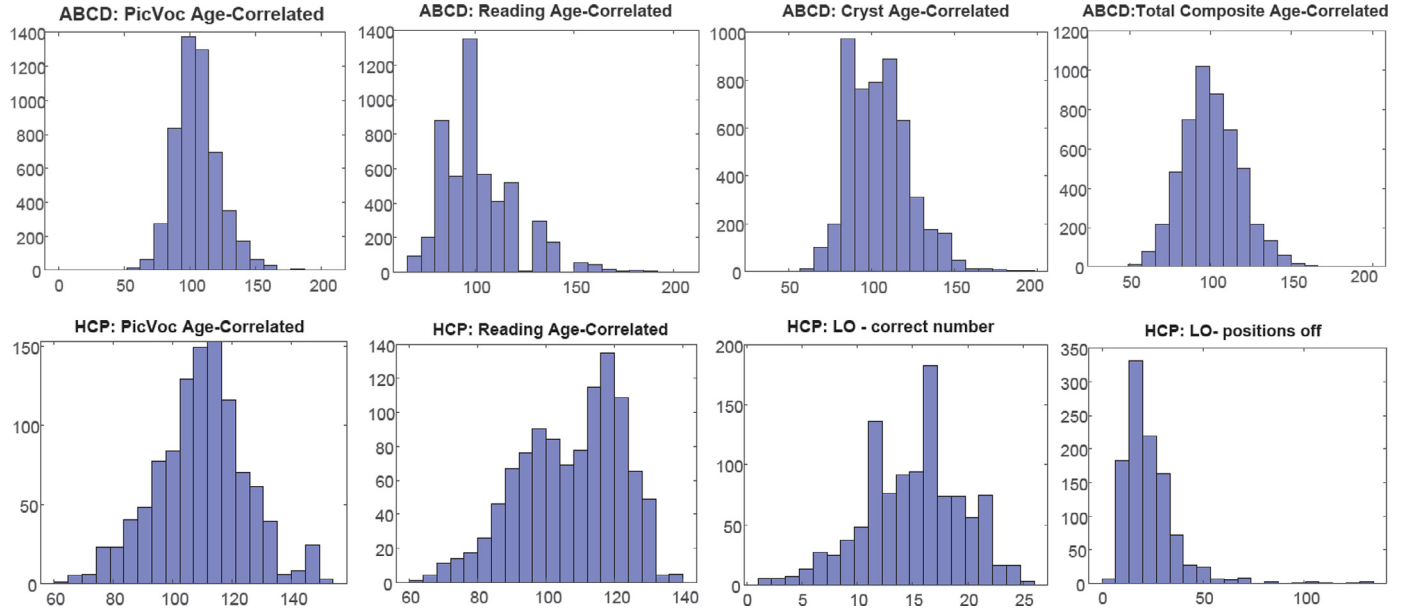


Fig. 1. Histograms of cognitive traits. The first row: ABCD study with 5252 subjects involved; from left to right, the traits include picture vocabulary score, oral reading recognition test score, crystallized composite age-corrected standard score, cognition total composite score. The second row: HCP study with 1065 subjects; from left to right: picture vocabulary test score, oral reading recognition test score, line orientation (LO) - total number correct, and line orientation (LO) - total positions off for all trials.

ing multiple ROIs are cut into pieces so that we can extract the correct and complete pathway and (3) outlier streamlines are removed. We use the number of fibers connecting each pair of ROIs to summarize connectivity in our analyses.

For the ABCD dataset, we processed 5252 subjects using PSC. We focus our analyses on four cognitive traits: (a) picture vocabulary score, (b) oral reading recognition test score, (c) crystallized composite age-corrected standard score, and (d) cognition total composite score. The first row in Fig. 1 demonstrates the distribution of the four traits. Similarly, for the HCP dataset, we preprocessed 1065 subjects using PSC. We focused on the cognitive traits: (a) picture vocabulary test score, (b) oral reading recognition test score, (c) line orientation - total number correct, and (d) line orientation - total positions off for all trials.

2.2. The graph auto-encoder model

The brain connectome for individual i is represented as a $V \times V$ symmetric adjacency matrix A_i , where $A_{i[uv]}$ is the count of the number of fibers connecting regions u and v in the i th individual's brain. Using the Desikan atlas, we choose $V = 68$ ROIs. We let

$$\begin{aligned} L(A_i) &= (A_{i1}, \dots, A_{iV(V-1)/2}) \\ &\equiv (A_{i[21]}, A_{i[31]}, \dots, A_{i[V1]}, A_{i[32]}, \dots, A_{i[V2]}, \dots, A_{i[V(V-1)]}) \end{aligned}$$

denote the lower triangular elements of matrix A_i . We let y_i denote the value of a cognitive trait for individual i . Our goal is to model the population distribution of the A_i 's and learn the relationship between cognitive traits y_i and brain structural connectomes A_i .

2.2.1. Latent space model for brain connectomes

Latent space models (Hoff et al., 2002) provide a probabilistic framework that assumes the edges in the networks are conditionally independent given their corresponding edge probabilities, with these probabilities defined as a function of pairwise distances between the nodes in a latent space. Borrowing the conditional independence idea, we first introduce a general latent space model for brain connectomes. We assume that the number of fibers connecting brain regions are conditionally independent Poisson variables, given individual- and edge-specific rates

$$\lambda_i = \{\lambda_{i1}, \dots, \lambda_{iV(V-1)/2}\}^T,$$

$$A_{i\ell} | \lambda_{i\ell} \sim \text{Poisson}(\lambda_{i\ell}), \quad (1)$$

independently for each pair $\ell = 1, \dots, V(V-1)/2$ and $i = 1, \dots, n$. We assume $\log(\lambda_{i\ell})$ has the following factorization form:

$$\log(\lambda_{i\ell}) = \gamma_\ell + \psi_\ell^{(i)}, \quad (2)$$

$$\psi_\ell^{(i)} = \sum_{r=1}^R \alpha_r X_{ur}^{(i)} X_{vr}^{(i)}, \quad \text{for } \ell = [uv], \quad (3)$$

$$\text{and } X_r^{(i)} = (X_{1r}^{(i)}, \dots, X_{Vr}^{(i)})^T. \quad (4)$$

As shown in (2), $\log(\lambda_{i\ell})$ is decomposed into two parts: a baseline parameter γ_ℓ controlling connection strength between the ℓ th pair of brain regions, representing shared structure across individuals, and an individual deviation $\psi_\ell^{(i)}$. Taking into account symmetry constraints and excluding the diagonal elements, there are $V(V-1)/2$ unknown $\{\lambda_{i\ell}\}$ for each subject, leading to a daunting dimensionality problem. To reduce dimensionality, Durante et al. (2017) proposed an SVD-type latent factorization, as shown in (3), where $r = 1, \dots, R$ indexes the different latent dimensions, $\alpha_r > 0$ is a weight on the importance of dimension r , and $X_{ur}^{(i)}$ is the r th latent factor specific to brain region u and subject i . According to (3)-(4), if $X_{ur}^{(i)}$ and $X_{vr}^{(i)}$ have the same sign and neither are close to zero, we have $X_{ur}^{(i)} X_{vr}^{(i)} > 0$ and there will be a positive increment on $\psi_\ell^{(i)}$ and hence on the expected number of fibers connecting regions u and v for subject i .

Model (1)-(4) can flexibly characterize variability across individuals in brain connectivity, while accommodating the complexity of network structures within each individual. However, we face challenges in learning the latent representations using existing latent space models: (1) Non-linearity: graph data are generally non-Euclidean with complicated structures. Designing a model to efficiently capture the non-linear structure is difficult. (2) Sparsity: brain regions are not fully connected, particularly structural brain networks. (3) Speed: existing latent space approaches often rely on Markov chain Monte Carlo sampling, which is computationally intensive for high dimensional graphs. It is desirable to develop a fast non-linear factorization model. To address these

challenges, we propose an autoencoder-based approach called Graph Autoencoder (GATE), from which we model the latent coordinates $X_u^{(i)}$ of brain regions as a non-linear function of a lower-dimensional vector z_i , which serves as a low-dimensional representation of the i th brain network A_i .

2.2.2. The graph autoencoder (GATE) model

GATE relies on the variational autoencoder (VAE Kingma and Welling, 2014), which is a popular technique for non-linear dimension reduction. Denote $z_i \in \mathbb{R}^K$ as a low-dimensional latent representation of the individual brain connectome A_i . GATE consists of two components. The first component is a generative model that specifies how the latent variables z_i give rise to the observations A_i through a non-linear mapping, parametrized by neural networks. The second component is an inference model that learns the inverse mapping from A_i to z_i . We frame our proposed Graph Autoencoder (GATE) method in the following context.

2.2.3. Generative model

For each subject i , we assume $A_{i\ell}$ for $\ell = 1, \dots, V(V-1)/2$ are conditionally independent given the latent representation $z_i \in \mathbb{R}^K$. Therefore, the likelihood of $L(A_i)$ is

$$p_\theta(L(A_i) = a_i | z_i) = \prod_{\ell=1}^{V(V-1)/2} p_\theta(A_{i\ell} = a_{i\ell} | z_i). \quad (5)$$

$p_\theta(A_{i\ell} | z_i)$ is a generative model for the weighted adjacency matrix A_i given the latent z_i with $z_i \sim p(z)$. We define $p(z) = N(0, I_K)$, representing all connectomes in the same Gaussian latent space.

We learn the mapping from the Gaussian latent space to the complex observation distribution in (5) by a hierarchical model equipped with parameters θ . Specifically, we assume the observations $A_{i\ell}$ arise from the following generative process:

$$\begin{aligned} z_i &\sim N(0, I_K), \\ A_{i\ell} | z_i &\sim \text{Poisson}(\lambda_{i\ell}(z_i)), \end{aligned} \quad (6)$$

where the Poisson rate parameter $\lambda_{i\ell}(z_i)$ is modeled as a nonlinear function of z_i according to:

$$\lambda_{i\ell}(z_i) = \exp(\gamma_\ell + \psi_\ell(z_i)), \quad (7)$$

$$\psi_\ell(z_i) = \sum_{r=1}^R \alpha_r X_{ur}(z_i) X_{vr}(z_i), \quad \text{for } \ell = [uv], \quad (8)$$

$$X_r(z_i) = (X_{1r}(z_i), \dots, X_{Vr}(z_i))^T = g_r(z_i), \quad (9)$$

where $g_r(\cdot) : \mathbb{R}^K \rightarrow \mathbb{R}^V$ is a nonlinear mapping from z_i to the r th latent factor of the brain regions X_r , parameterized by deep neural networks with parameters θ for $r = 1, \dots, R$.

Denote $\mathbf{X}(z_i) = (X_1(z_i), \dots, X_R(z_i)) \in \mathbb{R}^{V \times R}$. The u th row $(X_{u1}(z_i), \dots, X_{uR}(z_i))$ represents the latent features of brain region $u \in \mathcal{V}$ for individual i . A relatively large positive value for the cross product between the u th and v th rows implies a relatively high connection strength between these brain regions. The nonlinear mapping $\{g_r(\cdot)\}$ ($r = 1, \dots, R$) characterizes the latent embedding of brain regions that is determined via the local collaborative patterns among brain regions.

To take into account the intrinsic locality of structural brain networks, we propose a novel graph convolutional network (GCN) to learn each region's representation by propagating node-specific k -nearest neighbor information. Information in nodes that are closer to each other will be pooled together in GCN. The intrinsic locality refers to the relative distance between brain regions measured through the length of white matter fiber tracts connecting them. We extract this information from brain imaging tractography and store it in a matrix $B \in \mathbb{R}^{V \times V}$, where B_{uv} is the averaged length of fiber tracts between region u and

v , $B_{uv} = B_{vu}$, $B_{uu} = 0$, and we set $B_{uv} = \infty$ if there are no fibers between them. For each region u , we define its k -nearest neighbors (k -NN(u)) as the k ROIs closest to u according to our notion of distance, and denote the region itself as its 0-NN. If a region u has less than k direct neighbors, we will include all the regions v satisfying $B_{uv} \in (0, \infty)$ as its neighbors. In practice, we choose the average degree of nodes as the number of neighbors for simplicity since it measures the average number of collaborations among nodes.

To learn the r -latent coordinate $g_r(z_i)$ for subject i , the key idea is to consider each column $X_r(z_i)$ as an "image" with each region as an irregular pixel; we have R such "images" for each individual. Convolutional neural networks (CNN) are highly effective architectures in image and audio recognition tasks (Hinton et al., 2012; Krizhevsky et al., 2012; Sermanet et al., 2012), thanks to their ability to exploit the local translational invariance structures over their domain. Considering the unique features of the brain connectome networks, we generalize the CNN and define appropriate graph convolutions to learn the nonlinear mapping $\{g_r(\cdot)\}$ via exploiting the local collaborative pattern among brain regions. In particular, we define an M -layer GCN as follows:

$$X_r^{(i,1)} = h_1(W^{(r,1)} z_i + b_1), \quad (10)$$

$$X_r^{(i,m)} = h_m(W^{(r,m)} X_r^{(i,m-1)} + b_m) \quad \text{for } 2 \leq m \leq M, \quad (11)$$

where $X_r^{(i,m)}$ denotes the output of the m th layer of the convolutional neural network, $h_m(\cdot)$ is an activation function for the m th layer, and $W^{(r,m)}$ is a weight matrix characterizing the convolutional operator at this layer. We denote the parameters b_m , $W^{(r,m)}$, together with γ_ℓ , α_r in (7)-(8) ($m = 1, \dots, M$, $r = 1, \dots, R$, $\ell = 1, \dots, V(V-1)/2$) as the model parameter θ . The activation functions $\{h_m(\cdot)\}$ can be chosen from the following candidates based on performance: (1) Rectified linear unit (ReLU) function, which is widely used (Goodfellow et al., 2016) in deep neural networks, with the definition as $\text{ReLU}(x) = \max(0, x)$, where the max operation is applied element-wise; (2) Sigmoid function defined as $h_m(x) = \frac{1}{1+e^{-x}} \in (0, 1)$; (3) linear or identity function $h_m(x) = ax$ with $a \neq 0$.

For $m = 1$, $W^{(r,1)} \in \mathbb{R}^{V \times K}$ maps the latent representation $z_i \in \mathbb{R}^K$ to the latent space $X_r^{(i,1)} \in \mathbb{R}^{V \times 1}$; for $m \geq 2$, $W^{(r,m)}$ is a $V \times V$ weight matrix with the u th row $w_u^{(r,m)}$ satisfying $w_{uv}^{(r,m)} > 0$ if $v = u$ or $v \in k_r\text{-NN}(u)$, and $= 0$ otherwise. (11) implies that the embedding feature of each region at the m th layer is determined by the weighted sum of itself and its nearest neighbor regions at the $(m-1)$ th layer, and the related weights aim to characterize the region-specific local connectivity. For $r = 1, \dots, R$, we can choose different values of k to define its k_r -NN to fully explore the possible collaboration pattern among brain regions.

Figure 2 shows how a three-layer GCN learns $X_r(z_i)$ via a 2-NN GCN. First, we initialize the latent feature for each region as $x_{ur}^{(i,1)}$ based on (10). Then, we construct a "graph" based on the fiber length in B : each region is assigned to connect with 2 nearest neighbors at most according to the fiber length to other brain regions. This information is reflected in $W^{(i,m)}$, whose rows contain at most three non-zero elements (one at the diagonal and two off the diagonal). Next, we update the latent feature of each region in the next layer based on a sum of reweighted features from its 2 nearest neighbors and itself.

We collect all the parameters $(\gamma_\ell, \alpha_r, b_m, W^{(r,m)})$ for $\ell = 1, \dots, V(V-1)/2$, $r = 1, \dots, R$, $m = 1, \dots, M$ as θ . In the following Section 2.2.4, we show how to use variational inference to learn θ .

2.2.4. Variational inference and GATE learning

To train and evaluate the deep generative model in (6), we need to estimate θ , the parameters characterizing the mapping from z_i to A_{ii} and $p_\theta(z_i | A_i)$, the posterior distribution of the latent variable. By applying Bayes' rule, we have the posterior as

$$p_\theta(z_i | A_i) = \frac{p_\theta(A_i | z_i) p(z_i)}{p_\theta(A_i)}.$$

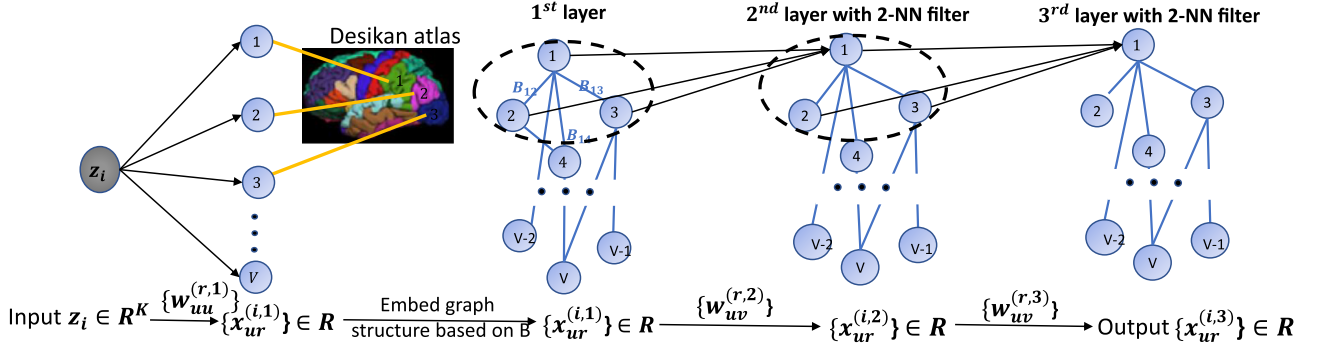


Fig. 2. Illustrative example of three-layer GCN architecture with 2-NN filters. For example, to learn the r th latent coordinate for node 1, the 2-NN is node 2 and node 3. After input z_i , the first layer embedding is $x_{1r}^{(i,1)} = h_1(u_{11}^{(r,1)} z_i)$, then the second layer embedding is $x_{1r}^{(i,2)} = h_2(w_{12}^{(r,2)} x_{2r}^{(i,1)} + w_{13}^{(r,2)} x_{3r}^{(i,1)})$, and the third layer embedding is $x_{1r}^{(i,3)} = h_3(w_{11}^{(r,3)} x_{1r}^{(i,2)} + w_{12}^{(r,3)} x_{2r}^{(i,2)} + w_{13}^{(r,3)} x_{3r}^{(i,2)})$, with the output $x_{1r}^{(i,3)}$ as the r th latent coordinate for node 1.

Since the likelihood function $p_\theta(A_i|z_i)$ is parameterized via the neural network with non-linear transformations, both the marginal distribution $p_\theta(A_i)$ and the posterior probability distribution $p_\theta(z_i|A_i)$ are intractable. Hence, we resort to variational inference (VI) (Jordan et al., 1999; Hoffman et al., 2013), a widely-used tool for approximating intractable posterior distributions. VI seeks a simple distribution $q_\phi(z_i|A_i)$ parameterized by ϕ from a variational family, e.g., a Gaussian distribution family, that best approximates $p_\theta(z_i|A_i)$. We call such $q_\phi(z_i|A_i)$ as the probabilistic encoder, which maps the input A_i to a low dimensional latent representation z_i . The approximated posterior $q_\phi(z_i|A_i)$ should be close to $p_\theta(z_i|A_i)$. We use Kullback-Leibler (KL) divergence to quantify the separation between these two distributions, which is defined as $D_{KL}(Q||P) = E_{z \sim Q} \log \frac{Q(z)}{P(z)}$, measuring how much information is lost if the distribution Q is used to represent P . We choose

$$q_\phi(z_i|A_i) \sim N(\mu_\phi(A_i), \text{diag}\{\sigma_\phi^2(A_i)\}), \quad (12)$$

i.e., a fully factorized (diagonal covariance) Gaussian distribution, to facilitate computation. We design deep neural networks to learn the parameters in μ_ϕ and σ_ϕ^2 , and denote the parameters involved in deep neural networks as ϕ . The details are in Supplementary S.1.

Our objective is to maximize the observed data log-likelihood $\log p_\theta(A_i)$, and also minimize the difference between the true posterior $p_\theta(z_i|A_i)$ and approximated posterior distribution $q_\phi(z_i|A_i)$. We express the above objective as

$$\begin{aligned} & \log p_\theta(A_i) - D_{KL}(q_\phi(z_i|A_i)||p_\theta(z_i|A_i)) \\ &= E_{q_\phi(z_i|A_i)}[\log p_\theta(A_i|z_i)] - D_{KL}(q_\phi(z_i|A_i)||p_\theta(z_i)) := -\mathcal{L}(A_i; \theta, \phi), \end{aligned} \quad (13)$$

where detailed calculation of (13) can be found in Supplementary S.2. Since $D_{KL}(q_\phi(z_i|A_i)||p_\theta(z_i|A_i))$ is nonnegative, $-\mathcal{L}(A_i; \theta, \phi)$ can be viewed as a lower bound on the marginal log-likelihood, referred to as the evidence lower bound (ELBO), which is a function of both θ and ϕ . Therefore, the training objective is minimizing the negative of the ELBO, i.e., minimizing

$$\mathcal{L}(A_i; \theta, \phi) = -E_{q_\phi(z_i|A_i)}[\log p_\theta(A_i|z_i)] + D_{KL}(q_\phi(z_i|A_i)||p(z_i)). \quad (14)$$

$\mathcal{L}(A_i; \theta, \phi)$ consists of two parts: the first term is the reconstruction error, measuring how well the model can reconstruct A_i ; while the second term, defined as the KL divergence of the approximate posterior from the prior, is a regularizer that pushes $q_\phi(z_i|A_i)$ to be as close as possible to its prior $N(0, I_K)$.

In practice, the expectation in the ELBO (14) is intractable. To address this, we employ Monte Carlo variational inference (Kingma and Welling, 2014) by approximating the troublesome expectation with samples of the latent variables from the variational distribution $z_i \sim q_\phi(z_i|A_i)$. Particularly, we form the Monte Carlo estimates of the expectation as

$$E_{q_\phi(z_i|A_i)}[\log p_\theta(A_i|z_i)] \simeq \frac{1}{L} \sum_{\ell=1}^L \log p_\theta(A_i|z_i^\ell),$$

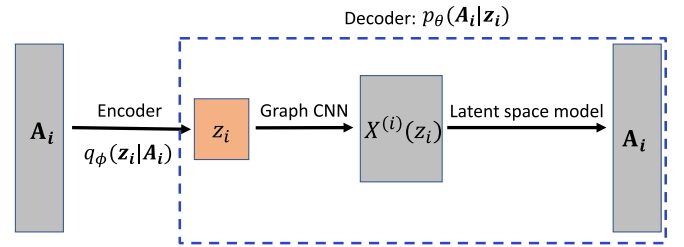


Fig. 3. A taxonomy of GATE. The encoder step is based on Eq. (12) to map A_i to z_i . The decoder step is based on Eq. (6)-(9) to map z_i to A_i through a latent space model with node embedding $X^{(i)}(z_i)$ learned from a Graph CNN.

Table 1

Algorithm 1: Training GATE model using gradients.

Input: $\{A_i\}_{i=1}^n$, $\{z_i\}_{i=1}^n$, geometric matrix B , latent space dimension R .

Randomly initialize θ, ϕ

while not converged **do**

 Sample a batch of $\{A_i\}$ with mini-batch size m , denote as \mathcal{A}_m .

for all $A_i \in \mathcal{A}_m$ **do**

 Sample $\epsilon_i \sim N(0, I_R)$, and compute $z_i = \mu_\phi(A_i) + \epsilon_i \odot \Sigma_\phi(A_i)$.

 Compute the gradients $\nabla_\theta \tilde{\mathcal{L}}(A_i; \theta, \phi)$ and $\nabla_\phi \tilde{\mathcal{L}}(A_i; \theta, \phi)$ with z_i .

 Average the gradients across the batch.

 Update θ, ϕ using gradients of θ, ϕ .

Return θ, ϕ .

where z_i^ℓ is sampled with the reparametrization trick: sampling $\epsilon_i^\ell \sim N(0, I_K)$ and reparametrizing $z_i^\ell = \mu_\phi(A_i) + \epsilon_i^\ell \odot \Sigma_\phi(A_i)$, where $\Sigma_\phi(A_i) = \text{diag}\{\sigma_\phi^2(A_i)\}$. A simple calculation shows that the KL-divergence $D_{KL}(q_\phi(z_i|A_i)||p(z_i)) = \frac{1}{2} \sum_{k=1}^K (\mu_k^2 + \sigma_k^2 - 1 - \log(\sigma_k^2))$, where μ_k and σ_k are the k th element of $\mu_\phi(A_i)$ and $\Sigma_\phi(A_i)$ respectively. Therefore, the ELBO in (14) can be approximated as

$$\begin{aligned} \mathcal{L}(A_i; \theta, \phi) &\simeq \tilde{\mathcal{L}}(A_i; \theta, \phi) \\ &= -\frac{1}{L} \sum_{\ell=1}^L \log p_\theta(A_i|z_i^\ell) + \frac{1}{2} \sum_{k=1}^K (\mu_k^2 + \sigma_k^2 - 1 - \log(\sigma_k^2)), \end{aligned}$$

which is differentiable with respect to θ and ϕ . Then, given n observed networks, we can construct an estimator of the ELBO of the full dataset, based on the minibatches $\frac{n}{m} \sum_{i=1}^m \tilde{\mathcal{L}}(A_{(i)}; \theta, \phi)$, where $\{A_{(i)}\}_{i=1}^m$ is a randomly drawn sample of size m from the full observed data with sample size n . Viewing $\frac{n}{m} \sum_{i=1}^m \tilde{\mathcal{L}}(A_{(i)}; \theta, \phi)$ as the objective, we implement a stochastic variational Bayesian algorithm to optimize θ and ϕ , respectively. Figure 3 shows the graphical diagram of the GATE approach; Algorithm 1 summarizes the GATE training procedure. Once the GATE model is learned, we can do the following: (1) have a low-dimensional representation for each individual network; (2) generate brain networks

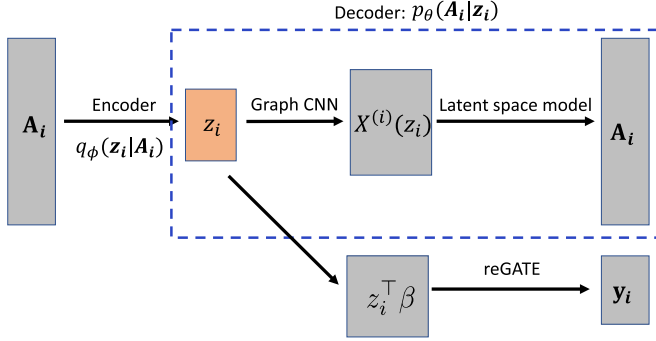


Fig. 4. reGATE to predict human traits. The encoder step is based on Eq. (12) to map A_i to z_i . The decoder step maps z_i to A_i based on Eq. (6)–(9); and a regression of y_i with respect to z_i .

to learn the population distribution of brain connectomes and features of these connectomes.

2.3. Regression with GATE and inference

Relating brain connectomes with traits. In addition to finding low-dimensional representations of brain structure networks, we are also interested in inferring the relationship between brain networks and human traits, such as cognition. With this goal in mind, we develop a supervised version of GATE, referred to as regression GATE (reGATE). Let y_i be a trait of the i th subject. We first express the joint log likelihood of (A_i, y_i) as

$$\log p_\theta(A_i, y_i) = -\mathcal{L}(A_i, y_i; \theta, \phi) + D_{KL}(q_\phi(z_i | A_i) || p_\theta(z_i | y_i, A_i)), \quad (15)$$

where

$$-\mathcal{L}(A_i, y_i; \theta, \phi) = \mathbb{E}_{q_\phi(z_i | A_i)} \log p_\theta(y_i | z_i) + \mathbb{E}_{q_\phi(z_i | A_i)} \log p_\theta(A_i | z_i) - D_{KL}(q_\phi(z_i | A_i) || p_\theta(z_i))$$

is called the ELBO of $\log p_\theta(A_i, y_i)$. We assume the human trait y_i and the brain connectivity A_i are conditionally independent given the latent representation z_i for the i th subject, and show the derivation of (15) in Supplementary S.3. In (15), we divide the log-likelihood of (A_i, y_i) into two parts: the ELBO denoted as $-\mathcal{L}(A_i, y_i; \theta, \phi)$, and the non-negative KL-divergence between $q_\phi(z_i | A_i)$ and $p_\theta(z_i | y_i, A_i)$. Different from the unsupervised ELBO in (13), (15) can be considered as a supervised ELBO with an extra term $p_\theta(y_i | z_i)$ that essentially formulates a regression of y_i with respect to z_i . Here we consider y_i as a continuous random variable, and set $p_\theta(y_i | z_i)$ as a univariate Gaussian, i.e., $p_\theta(y_i | z_i) \sim N(z_i^\top \beta + b, \sigma^2)$, where $\beta, \sigma^2 \in \theta$ are parameters to be learned. Figure 4 shows the flowchart of the reGATE architecture.

Similarly to Section 2.2.4, we form the Monte Carlo estimate of $\mathcal{L}(A_i, y_i; \theta, \phi)$ and estimate θ, ϕ, β, b following the stochastic variational Bayesian Algorithm 1 by replacing $\mathcal{L}(A_i; \theta, \phi)$ with $\mathcal{L}(A_i, y_i; \theta, \phi)$. We show the detailed sampling steps in Supplementary S.4. For the trained reGATE model, we have: (1) a low-dimensional representation for each individual network; (2) human trait prediction for each individual network; (3) ability to generate brain networks for inference on how features of the networks vary across individuals and with traits.

Conditional generative model: We are interested in inferring how brain networks vary across levels of a trait. For example, if y_i measures a person’s memory ability, we would like to study differences in the distribution of brain networks between people with good and bad memory skills. To address questions of this type, we generate samples from the posterior distribution of A_i given particular y_i values using Gibbs sampling: sample z_i from $p_\theta(z_i | y_i)$, then sample A_i from $p_\theta(A_i | z_i)$. The conditional $p_\theta(A_i | z_i)$ is learned while implementing reGATE. The posterior distribution of z_i given y_i can be expressed as $p_\theta(z_i | y_i) \propto p_\theta(y_i | z_i) p_\theta(z_i)$,

where $p_\theta(z_i) \sim N(0, I_K)$ and $p_\theta(y_i | z_i) \sim N(z_i^\top \beta + b, \sigma^2)$. Therefore, we have $p_\theta(z_i | y_i) \sim N(\mu_z(y_i), \Sigma_z(y_i))$, where

$$\mu_z(y_i) = (I_K + \beta \beta^\top / \sigma^2)^{-1} \beta (y_i - b) / \sigma^2, \quad \text{and} \quad \Sigma_z(y_i) = (I_K + \beta \beta^\top / \sigma^2)^{-1}; \quad (16)$$

the derivation of (16) is in Supplementary S.5.

The latent representation z_i is unidentifiable in VAE since the log-likelihood and ELBO are rotationally invariant for z_i . For example, letting $\tilde{z}_i = U^\top z_i$, then $P_{U, \theta}(A_i) = P_\theta(A_i)$ and

$$D_{KL}(q_{U, \phi}(z_i | A_i) || p_{U, \theta}(z_i | A_i)) = D_{KL}(q_\phi(z_i | A_i) || p_\theta(z_i | A_i)),$$

where U is an orthogonal matrix, $q_{U, \phi}(\cdot)$ and $p_{U, \theta}(\cdot)$ are defined by replacing z_i with \tilde{z}_i in $q_\phi(\cdot)$ and $p_\theta(\cdot)$. Rotational invariance can be solved by post-processing to rotationally align the z_i ’s. However, this would only be necessary if one is attempting to compare z_i ’s from different datasets or analyses of a given dataset. Within an analysis, the main focus is on inference on the relative values of z_i ’s, and these relative values are well defined. In addition, when the focus is on relating brain structure to human traits or in predicting traits based on brain structure or vice versa, the non-identifiability issue does not present a problem.

3. Simulation study

We conduct a simulation study to evaluate the performance of GATE and reGATE on a broad application in graph-value data. We simulate different types of random graphs using the Python package *NetworkX*. Particularly, we consider four network structures: sparse networks according to the model in Johnson (1977), community structures under the model of Nowicki and Snijders (2001), small-worldness from the model in Watts and Strogatz (1998), and scale-free property from the model in Barabási and Albert (1999). We simulate 100 networks with $V = 68$ nodes for each type by sampling their edges from conditional independent Bernoulli random variables given their corresponding structure-specific edge probability. Each structure-specific edge probability vector is carefully constructed to assign a high probability to a subset of network configurations characterized by a specific property. Figure 5 displays some example networks we generated with the four different network structures.

We first generate y_i according to $y_i = \alpha^\top A_i \alpha + \epsilon_i$, where $\alpha = (1, \dots, 1, 0, \dots, 0)^\top \in \mathbb{R}^{68}$, and $\epsilon_i \sim N(0, 1)$. We then standardize y_i , so that

it ranges from -1.5 to 2.0 . These settings aim to generate separable y_i ’s according to the topological structures of A_i . The histograms in Fig. 7 clearly demonstrate how y_i varies for different network structures. Our goals in this simulation study include (1) learning the latent representation under both GATE and reGATE; (2) inferring how the network connectivity structure varies with y_i ; (3) assessing the predictive performance of the reGATE model.

We train GATE and reGATE to obtain low-dimensional representations z_i ’s. Specifically, for the n simulated networks, we have $\bar{z}_1 = (\mathbb{E}(p_\theta(z_1 | A_1)), \dots, \bar{z}_n = \mathbb{E}(p_\theta(z_n | A_n)))$ as the posterior means after training the GATE model. We then conduct principal component analysis (PCA) analysis on the posterior means $\{\bar{z}_1, \dots, \bar{z}_n\}$ and plot each \bar{z}_i using their first two PC scores in \mathbb{R}^2 colored according to the corresponding y value. We can clearly observe the separation between these four types of networks in the low-dimensional representation space inferred using both GATE and reGATE, with reGATE yielding greater separation across the groups.

To infer how brain networks A_i vary according to y_i , we simulate networks from the posterior distribution of $A_i | y_i$ according to the conditional generative model in Section 2.3. Specifically, we first sample $z_i | y_i$ based on Eq. (16) with the parameters obtained from the previously trained reGATE and y_i ranging from -1.5 to 2 , then generate networks via $p_\theta(A_i | z_i)$. Figure 7 shows the generated networks with structure varying for different y_i ’s. We can clearly observe that the network

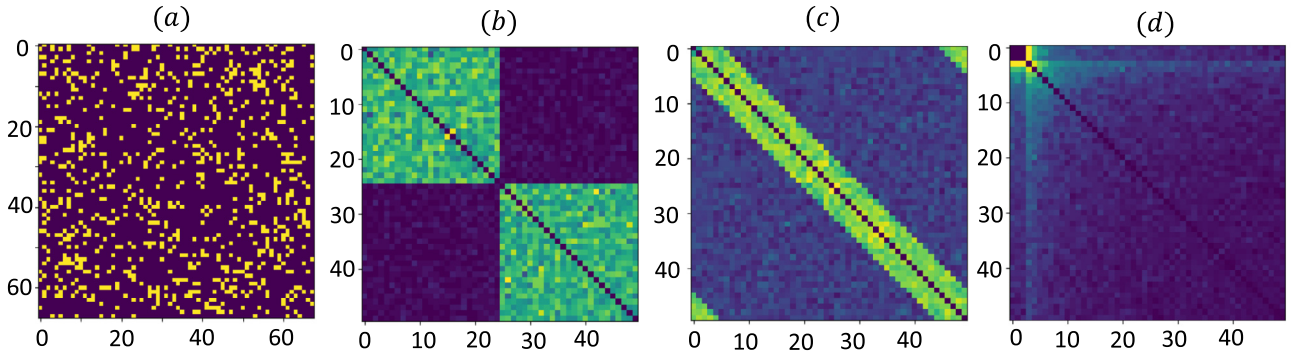


Fig. 5. The edge probability vectors (rearranged in matrix form) from four network structures: (a) sparse graph, (b) community structure, (c) small world, (d) scale free.

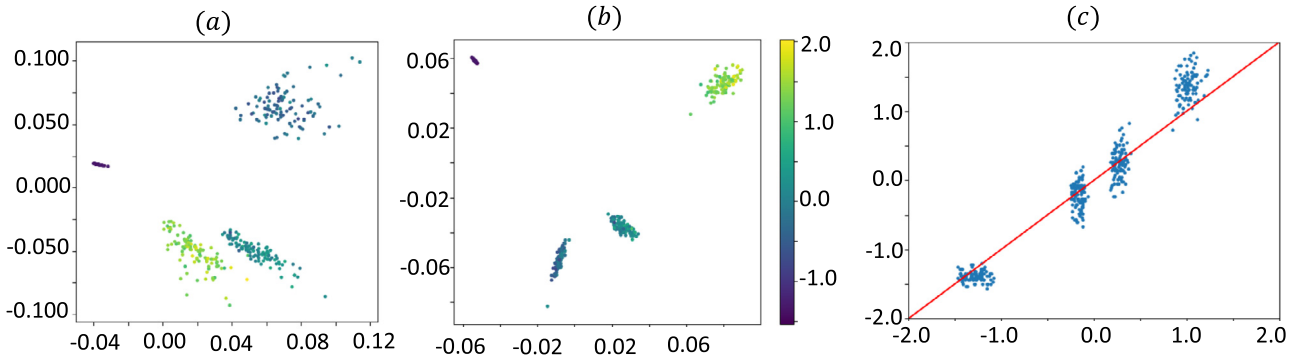


Fig. 6. Plot of the first 2 PC scores of the posterior mean of $z_i|A_i$ with corresponding y_i . Different colors refer to various values of y . (a) GATE, (b) reGATE, (c) Predicted y v.s. true y in reGATE, where x -axis is the true value and y -axis is the predicted value.

Table 2

The first row and the second row are the MSE under different methods. All numbers are calculated based on the mean of 100 replicated simulations.

	reGATE	LR-TNPCA	LR-PCA	CPR
MSE: case 1	0.0252	0.0271	0.0340	0.0388
MSE: case 2	0.0505	0.0593	0.0700	0.1078
Time (mins)	17.5	197.6	5.3	253.7

shows a sparse structure when $y_i = -1.5$, a mixture of community and small-world structure when $y_i = -0.1$, and a clear scale-free structure when $y_i = 2$. These generated structures are consistent with the ground truth in our simulation settings.

To address our third goal, we evaluate the predictive accuracy of reGATE. We consider two cases: (1) $y_i = \alpha^T A_i \alpha + \epsilon_i$; (2) $y_i = (\alpha^T A_i \alpha)^2 + (\alpha^T A_i \alpha)^3 + \epsilon_i$. We also compare reGATE with a few popular methods in the literature for predicting human traits using network data. The first method is a regular linear regression based on tensor network principal component analysis (LR-TNPCA) in Zhang et al. (2019). The second method is linear regression based on a regular PCA applied to the vectorized networks. The third method is tensor regression proposed in Zhou et al. (2013), denoted as CPR here. The mean square error (MSE) from five-fold cross-validation was used to compare different approaches. Figure 6(c) shows the association between the predicted value and the true value under reGATE in case 1. The first and second rows in Table 2 show the MSE under different methods. We can see reGATE outperforms other methods in predictive accuracy. The third row in Table 2 reports the computing time with 100 replicated simulations. reGATE is both fast and accurate based on these results.

We summarize the computing details used in the simulation study here. We run stochastic gradient descent with momentum (the Adam

algorithm in Kingma and Ba (2015)) on GATE and reGATE with the learning rate 0.001 on one NVIDIA Titan-V GPU. In GATE, we use a batch size of 128, sampled uniformly at random at each epoch and repeated for 1000 epochs. In reGATE, we use 5-fold CV to calculate the MSE and run 200 epochs with batch size of 128 for each training dataset. The experimental details and network architectures for the inference model and generative model training are summarized in Table 3. The latent dimension K is chosen as the smallest value that achieves the minimal training loss, provided that the network architectures are fixed as in Table 3. Figure 8(a) shows the log-training loss with latent dimensions K varying from 5 to 100, and the training loss achieves the best performance when K is around 45. The sensitivity of the training loss with respect to the random initialization is also explored. As shown in Fig. 8(b), under different initializations, the paths of the training loss converge to the same level as the number of epochs increases. All code used to produce the results and figures is available online via GitHub (<https://github.com/meimeiliu/GATE>).

4. Applications to the ABCD and HCP data

We apply our method to both the ABCD and HCP datasets described in Section 2.1 to examine the relationship between structural brain networks and cognition for adolescents and young adults. The ABCD study uses a reliable and well-validated battery of measures that assess a wide range of human functions, including cognition. The core of this battery is comprised of the tools and methods developed by the NIH Toolbox for assessment of neurological and behavioral function (Gershon et al., 2013). The Toolbox includes measures of cognitive, emotional, motor, and sensory processes. Since we are particularly interested in cognition, we extract four cognition related measures as y from ABCD, including

- (a) Picture vocabulary test: the picture vocabulary test uses an audio recording of words, presented with four photographic images on

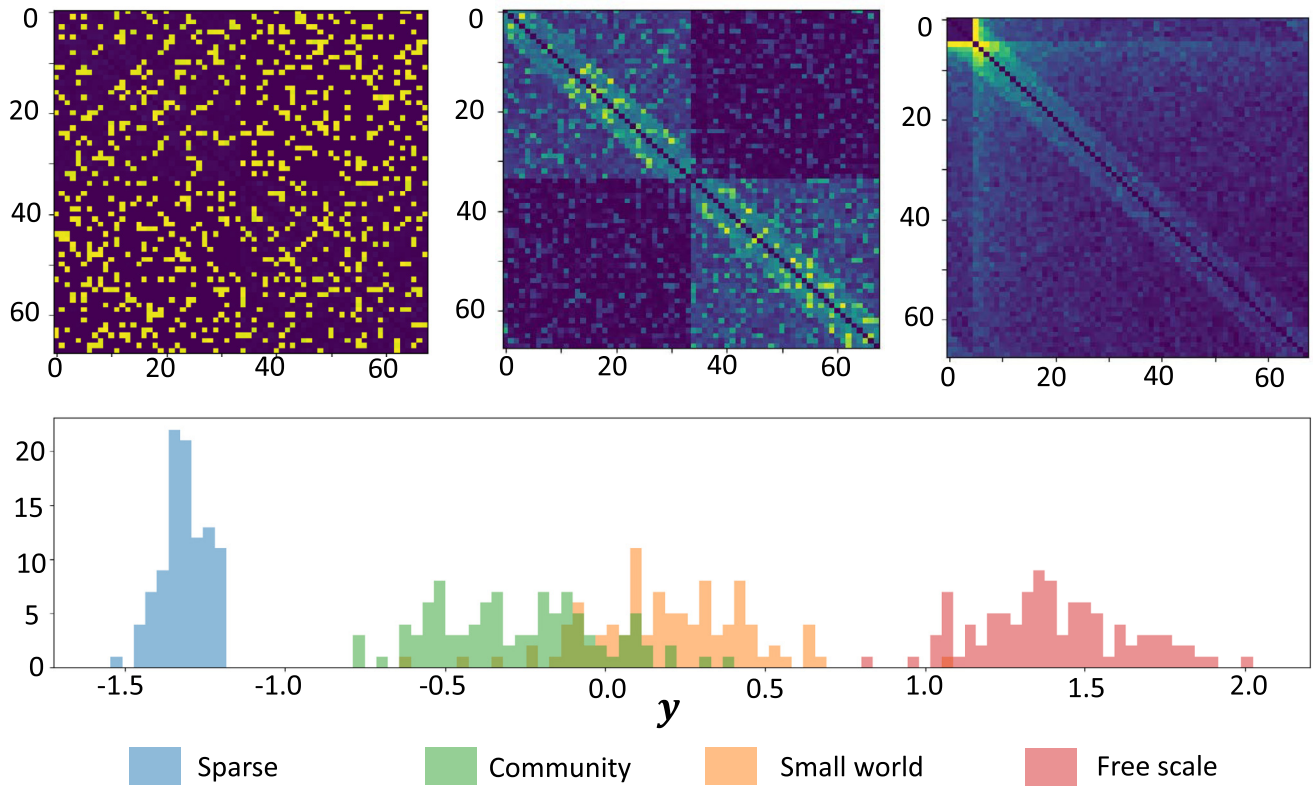


Fig. 7. First row: generated networks conditional on the specific y_i using reGATE, corresponding to $y_i = -1.5, -0.1, 0.2$ respectively from left to right. Second row: histogram of y_i with respect to the network structure; the x -axis is the value of y_i , the y -axis is the frequency that belongs to a specific structure in the training data.

Table 3

Experimental details and network architectures. K is the dimension of z_i , N is the number of layers in the inference network, M is the number of layers in GCN, and R is the dimension of $X^{(i)}$.

	Inference model (μ_ϕ, σ_ϕ)		Generative model	
	μ_ϕ ($N = 2$)	σ_ϕ ($N = 2$)	setting	activation
GATE/reGATE ($K = 45$)	$W_{1,\mu} : 45 * 400$ $W_{2,\mu} : 400 * 45$ $b_{1,\mu} : 400 * 1$ $b_{2,\mu} : 45 * 1$ $\phi_{1,\mu} = \text{ReLU}$ $\phi_{2,\mu} = \text{Linear}$	$W_{1,\sigma} : 45 * 400$ $W_{2,\sigma} : 400 * 45$ $b_{1,\sigma} : 400 * 1$ $b_{2,\sigma} : 45 * 1$ $\phi_{1,\sigma} = \text{ReLU}$ $\phi_{2,\sigma} = \text{Linear}$	k -NN: 16 $M = 2$ $R = 5$	h_1 : Sigmoid h_2 : Sigmoid

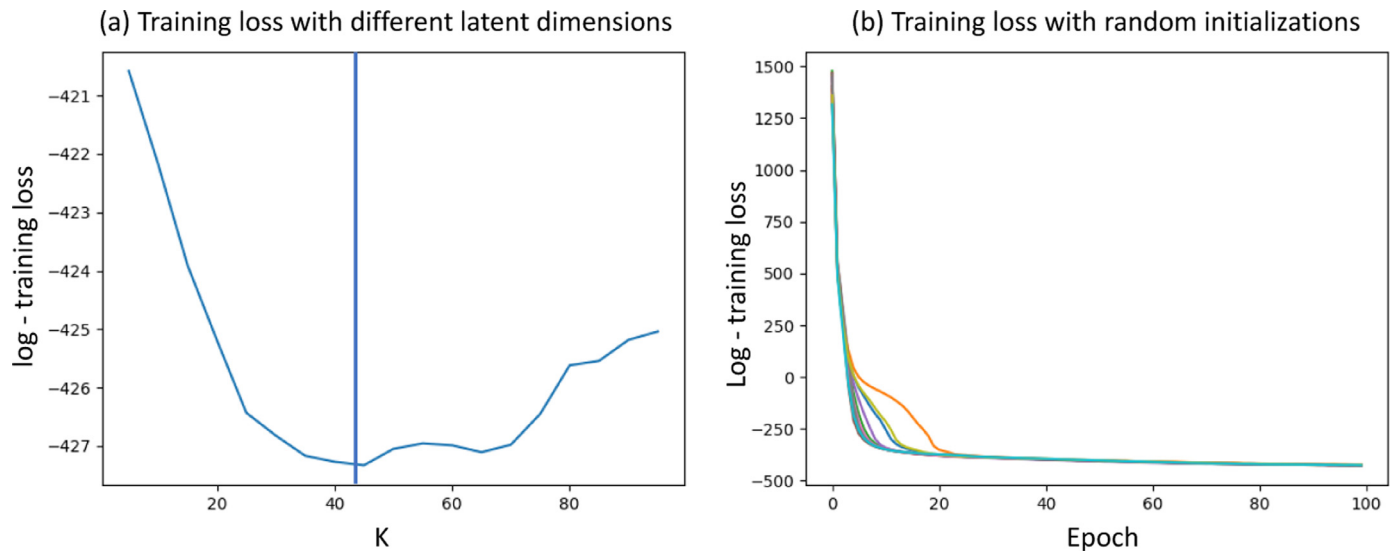


Fig. 8. (a) Log-training loss with latent dimension K varies; (b) the paths of the log-training loss with random initializations.

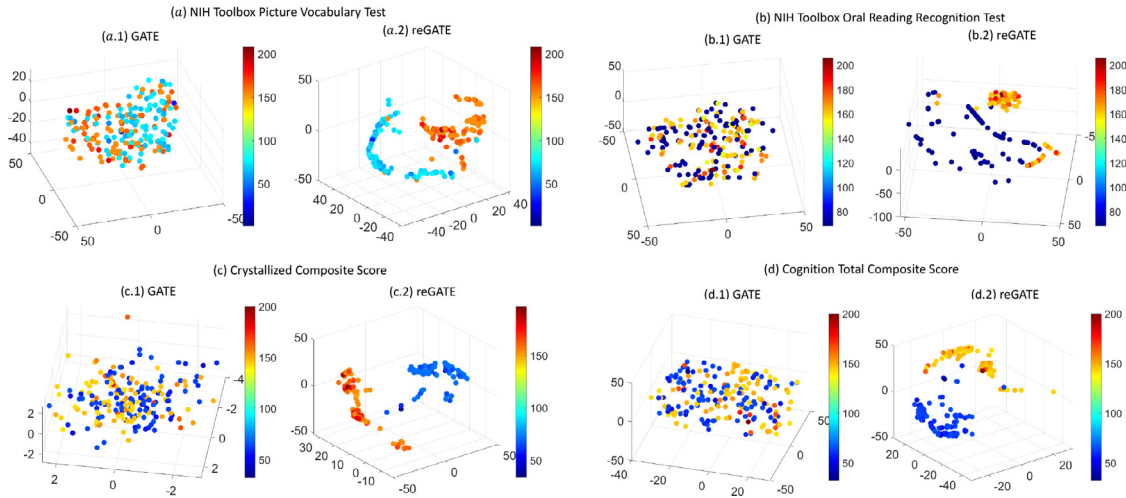


Fig. 9. Visualization of low dimensional embedding z_i learned from GATE and reGATE applied to the ABCD study. We display 100 subjects with the lowest trait scores and 100 subjects with the highest scores for each trait. Colors represent the trait scores.

the computer screen. Participants are asked to select the picture that best matches the meaning of the word.

- (b) Oral reading recognition test: participants on this test are asked to read and pronounce letters and words as accurately as possible.
- (c) Crystallized composite score: crystallized cognition composite can be interpreted as a global assessment of verbal reasoning. We use the age-corrected standard score.
- (d) The cognition total composite score: this composite score measures overall cognition and is obtained from a factor analysis (Heaton et al., 2014).

We extracted two matched cognitive measures from HCP: picture vocabulary test score and oral reading recognition test. In addition, we add two more cognitive measures in our data analysis: total number correct answers and total positions off in a line orientation test. A more detailed description of these traits can be found in Gershon et al. (2013).

4.1. Visualization: show network data in low-dimensional space

Both GATE and reGATE output low-dimensional representations of the brain networks. We can visualize the latent features of each individual's connectome and examine the relationship between structural connectivity and the four traits via the latent features. We train GATE on 5252 brain networks extracted from the ABCD dataset to obtain low-dimensional representations z_i for $i = 1, \dots, 5252$. We then plot the posterior mean of $z_i | A_i$ using t-SNE (Van der Maaten and Hinton, 2008) colored with its corresponding trait score in \mathbb{R}^3 . We show 200 subjects' data for each cognition trait, with the first 100 subjects having the lowest trait scores and the second 100 subjects having the highest scores. As shown in Fig. 9(a)–(d), under both GATE and reGATE, we obtain a large separation between the two groups of subjects, indicating that brain connection patterns are different for these groups. reGATE has better performance since we incorporate the trait information in learning the z_i s. A similar analysis is conducted using the HCP data, and the result is shown in Figure 16 in the supplement.

4.2. Prediction: predict traits with reGATE

In this section, we demonstrate the predictive ability of reGATE and compare it with several competitors: LR-TNPCA (Zhang et al., 2019), LR-PCA, CPR (Zhou et al., 2013), and BLR (Wang et al., 2019). BLR is supervised bi-linear regression (BLR) with emphasis on signal sub-network selection. To verify the role played by the latent space model, we further compare our approach with a simplified reGATE without considering the

node-level latent space structure. Specifically, we simplify the generative model as $z_i \sim N(0, I_K)$, and $A_{i \ell} | z_i \sim \text{Poisson}(\lambda_{i \ell}(z_i))$, where $\lambda_{i \ell}(z_i)$ is directly learned via a fully connected neural network without the latent space setting in Eqs. (7)–(9). We denote this approach as S-reGATE.

MSE from five-fold cross-validation is used to assess performance. Table 4 shows results for ABCD and HCP datasets. reGATE significantly outperforms other methods in both datasets for all traits. S-reGATE has a better prediction performance in ABCD data but cannot compete with BLR in the HCP dataset. Hence, the proposed latent space model does indeed improve prediction; interpretation is also improved relative to S-reGATE, which does not capture some aspects of variation in the brain networks.

With the bigger sample size ($n = 5252$) in the ABCD data, the improvements of reGATE are more significant than those for the HCP data. It is well known that deep neural networks tend to perform exceptionally well for large training sample sizes.

We also calculate the percentage of MSE improvement upon a baseline model, the sample means \bar{y} , for the prediction result from the ABCD and HCP data. As shown in Fig. 10(a.1)–(a.4), most methods in ABCD study demonstrate an improved MSE compared with the sample mean \bar{y} , indicating that there is a detectable relationship between the structural connectome and cognitive traits. However, reGATE achieves stable and significant prediction improvements ranging from 30% to 40%, while the competitors' performance improvements fluctuate from -5% to 5% . From Fig. 10(b.1)–(b.4), we see reGATE still has the best performance among other competitors.

Besides MSE, we evaluate the correlation between the predicted value \hat{y}_i and the observed value y_i via five-fold cross-validation for different approaches. Correlation is reported in the parenthesis in Table 4. We can see that reGATE significantly improves the prediction by increasing the correlation from around 0.2 to 0.4.

4.3. Inference: generate brain connectomes conditioned on traits

An appealing characteristic of reGATE is the ability to generate brain networks for new individuals conditionally on their trait value. This facilitates inference on how brain networks, and their topological properties, change across different levels of a trait and how this dependence varies across individuals. In our first experiment, we assess the performance of GATE in characterizing the observed brain network data via posterior predictive checks (Gelman et al., 2013) for relevant network topological properties, including network density, mean eigencentrality, average path length, and average degree. Denote $\eta_k = g_k\{L(\mathcal{A})\}$ as the random variable associated with the k th network summary measure,

Table 4

Comparison of prediction results of different methods in both the ABCD and HCP datasets. Each cell shows the MSE and correlation (in the parenthesis) between the observed and predicted measures via five-fold CV.

	reGATE	S-reGATE	LR-TNPCA	LR-PCA	CPR	BLR
ABCD ($n = 5252$)						
Pic Voc	186.8 (0.40)	212.5 (0.27)	280.1 (0.22)	285.2 (0.20)	285.1 (0.23)	200.4 (0.29)
Oral Reading	206.5 (0.41)	224.8 (0.32)	342.3 (0.19)	346.5 (0.14)	357.0 (0.13)	337.1 (0.19)
Cryst Comp	279.4 (0.39)	304.5 (0.31)	315.8 (0.23)	322.9 (0.22)	318.4 (0.19)	321.3 (0.30)
CogTot Comp	248.0 (0.38)	255.6 (0.35)	297.6 (0.17)	308.6 (0.26)	306.7 (0.22)	290.2 (0.32)
HCP ($n = 1065$)						
Pic Voc	209.1 (0.28)	214.4 (0.25)	214.5 (0.25)	219.0 (0.21)	257.2 (0.17)	216.1 (0.23)
Oral Reading	198.9 (0.26)	205.7 (0.22)	202.1 (0.22)	204.2 (0.21)	252.1 (0.19)	200.4 (0.26)
LO: correct number	17.8 (0.27)	20.4 (0.20)	18.5 (0.23)	18.2 (0.22)	21.9 (0.19)	18.8 (0.18)
LO: positions off	195.8 (0.27)	203.5 (0.25)	201.8 (0.22)	202.6 (0.24)	259.0 (0.19)	200.7 (0.27)

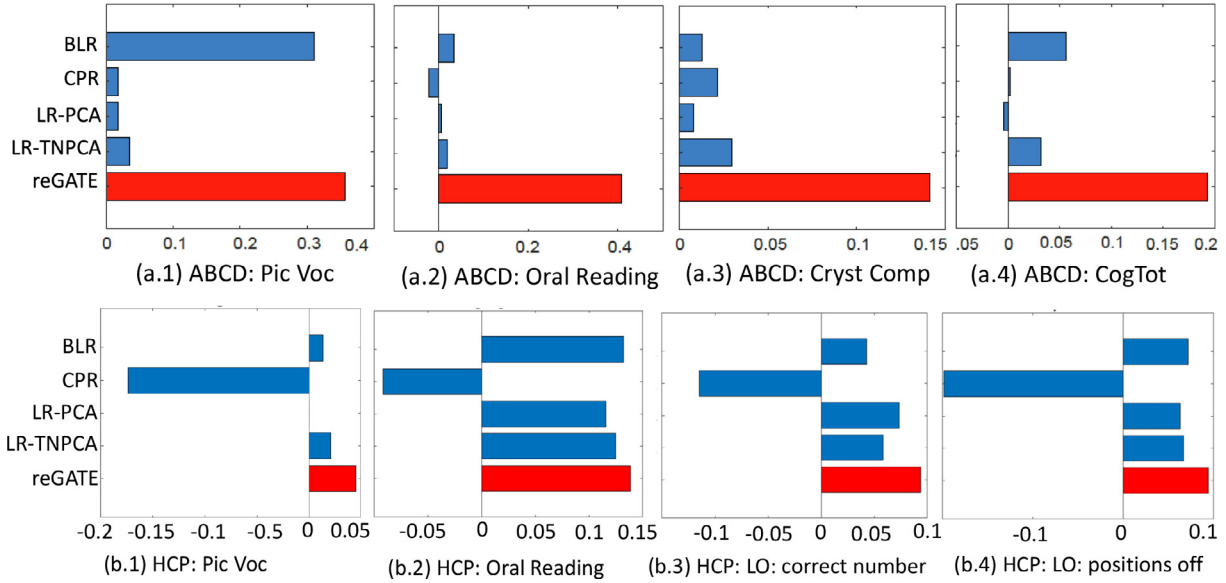


Fig. 10. Percentage of the MSE improvement compared with a baseline method of using \bar{y} as the predicted trait. (a.1)–(a.4): ABCD with picture vocabulary score, oral reading recognition test score, crystallized composite age-corrected standard score, and cognition total composite score. (b.1)–(b.4): HCP with picture vocabulary score, oral reading recognition test score, LO-correct number, LO-positions off. The X-axis is the improved proportion; the y-axis marks different methods.

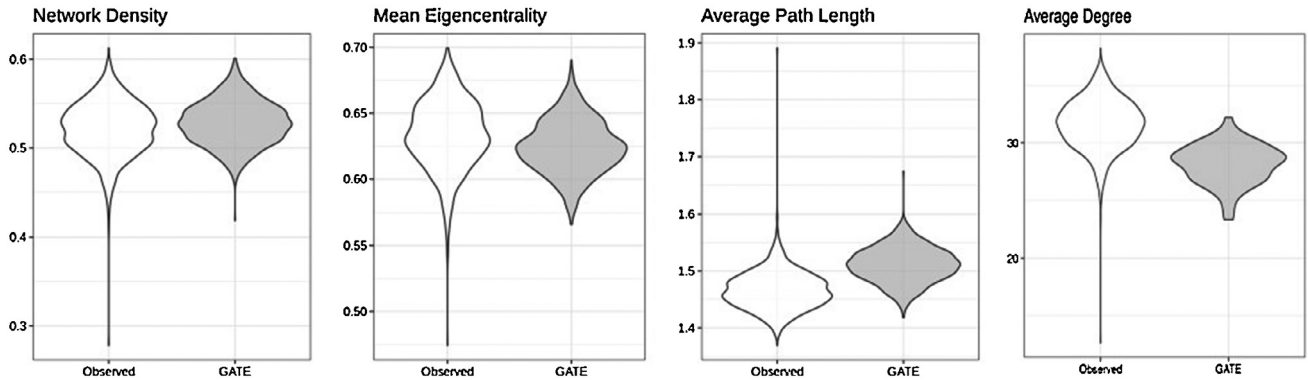


Fig. 11. HCP: Goodness-of-fit assessments for selected network summary measures. The violin plots summarize the distribution in the observed data (white) and the posterior predictive distribution arising from GATE (gray).

for $k = 1, 2, 3, 4$, where \mathcal{A} represents the brain connectomes. Then we calculate the posterior predictive distributions for these summary measures based on the generative model $p_{\theta}(\mathcal{A} | z)$ in Section 2.2.3 as $p_{\eta_k}(\eta | z) = \sum_{a \in \mathcal{A}_v; g_k(a) = \eta} p_{\theta}(\mathcal{A} = a | z)$, for $k = 1, 2, 3, 4$, respectively. Figure 11 compares the network summary measures computed using all 1065 subjects from the HCP data (white-colored) and the generated network data (gray-colored) from GATE. GATE achieves good performance in characterizing the observed network summary measures.

We next consider the conditional predictive distribution of network topological summaries given trait scores, i.e., $\eta_k | y$. We consider network density and average path length for both the ABCD and HCP data and focus on different levels of picture vocabulary test score as y . The distribution of $\eta_k | y$ can be expressed as $p_{\eta_k}(\eta | y) = \sum_{a \in \mathcal{A}_v; g_k(a) = \eta} p_{\theta}(\mathcal{A} = a | y)$, for $k = 1, 2$, respectively, where $p_{\theta}(\mathcal{A} = a | y)$ can be obtained via the conditional generative model in Section 2.3. More specifically, conditional on each particular picture vocabulary test score level, we first

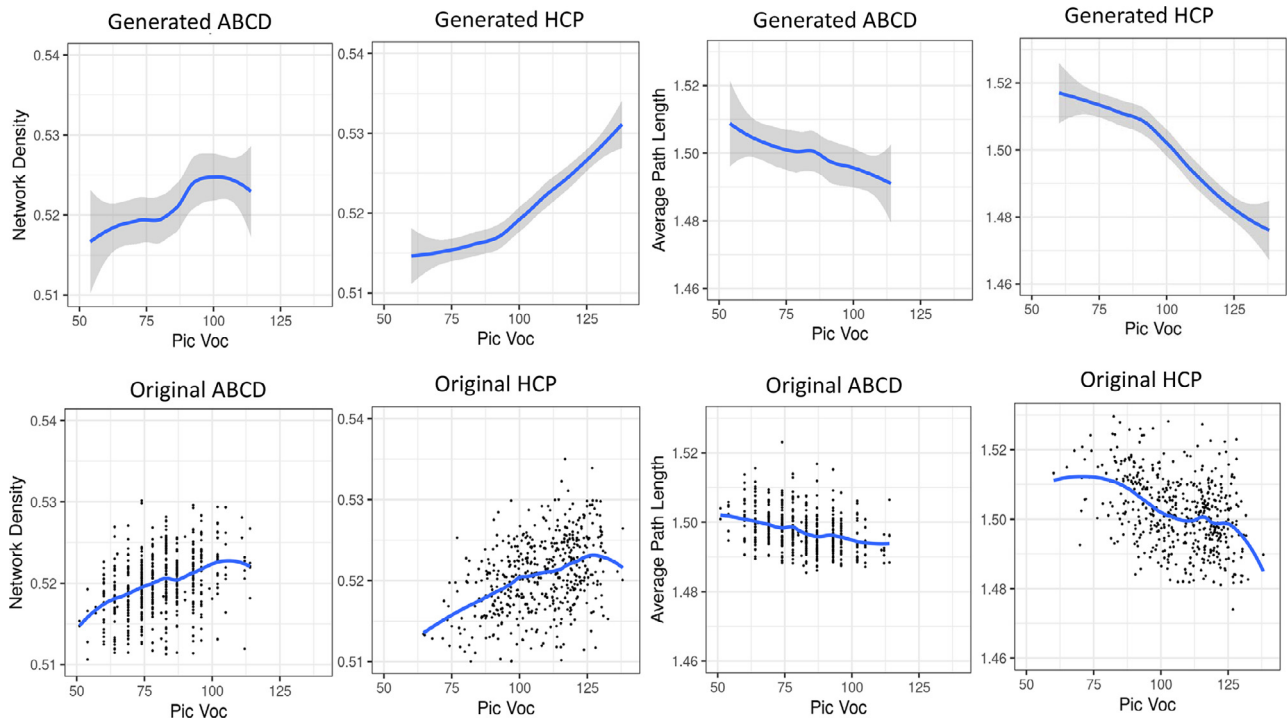


Fig. 12. Comparison between generated brain networks and the observed brain networks for different trait levels in network density and average path length for ABCD and HCP datasets. First row for ABCD and HCP: posterior predictive network summaries with confidence bands for network density and average path length. The x-axis is the picture vocabulary test score; the y-axis corresponds to network density and average path length, respectively. The confidence bands are calculated based on 500 generated network summaries conditional on different picture vocabulary test scores. Second row: scatter plot and the fitted curve for different trait levels in network density and average path length for the original ABCD and HCP datasets.

generate 500 networks based on the well-trained reGATE model. Next, we calculate the network summaries for each generated network, and then calculate the mean and confidence bands based on 500 calculated network summaries. The first row in Fig. 12 is the constructed confidence band of network density and average path length versus y based on the generated networks via reGATE. As shown in the first row in Fig. 12, the network density increases, and the average path length decreases as the picture vocabulary score increases. There is more variability in network topological summaries for the adolescent subjects in ABCD than for the adults in HCP.

It is challenging to conduct a similar inference on the real data (ABCD and HCP data) due to the limited observed brain networks for a specific picture vocabulary test score. Instead, we show the scatter plots of network density and average path length versus y on the real datasets in the second row of Fig. 12. The blue curves are the fitted mean trajectory. It can be seen that the overall trends are similar: the network density is increasing as y increases, while the average path length is decreasing as y increases.

We next explore how the brain network varies across different levels of trait y . We consider three levels that range from the minimum to the maximum for the oral reading recognition score in the HCP and generate multiple networks for each trait value. In particular, we consider the oral reading recognition score $y = 60, 91$ and 138 , and generate brain networks according to the conditional generative procedure in Section 2.3. If we set $y = 60$, then the posterior $p_{\theta}(A | y = 60)$ indicates the distribution of the brain networks for people with oral reading recognition test equal to 60. A mean network is used to summarize the generated network data for a given y , and we further dichotomize the connectome to $\{0, 1\}$ depending on whether a connection exists for better visualization. The results are shown in Fig. 13, where Fig. 13(a.1) shows the histogram of the observed trait scores, and Fig. 13(a.2)–(a.4) shows the generated mean networks (with the first 34 nodes from the left side of the brain and the next 34 nodes from the right side of the

brain) for different y . We further compare the generated mean network with the observed network in HCP. From the observed HCP, we have one subject for $y = 60$, ten subjects with $y = 91$, and two subjects for $y = 138$. There are two issues in empirically estimating the posterior mean: (1) for different y , we have different numbers of observations; (2) we have very little data to estimate a high-dimensional object. These two issues make the estimation directly from observed brain connectomes less credible. Fig. 13(b.1)–(b.3) shows the mean connectome estimated from the observed HCP data. From the result in the first row, more connections between the brain's two hemispheres are correlated with better reading ability. This result is consistent with findings in the literature (Bullmore and Sporns, 2009; Durante and Dunson, 2018). From the result of real data, we also observe such a trend, but not so obvious compared with the first row generated from reGATE.

To enhance the above analysis, we reduce the number of levels and select only the 10%th and 90%th quantiles from the observed y as two representative levels. For each level, we generate 100 conditional networks $A_i | y_i$ and calculate the mean difference between the networks in the high and low trait groups. Figure 14 plots the top 50 connections (based on absolute values) in the mean difference network for the picture vocabulary test for both HCP and ABCD. This procedure is done separately for the two datasets. The 50 connections are further separated into positive and negative connections and are plotted in separate panels in Fig. 14. After ranking the connections by absolute values, these connections are dominated by positive ones, indicating that better vocabulary ability is associated with more connections.

The first row in Fig. 14 shows results from the HCP data, and the second row shows results from the ABCD data. Considering the different populations (young adults vs adolescents) in our data analysis, it is interesting to observe many similar results. For example, we observe denser connections both within the left and right frontal lobes and between them. In particular, we see that brain regions such as $l26, r26$ (rostral middle frontal), $l27, r27$ (superior frontal), and $l3, r3$ (caudal middle

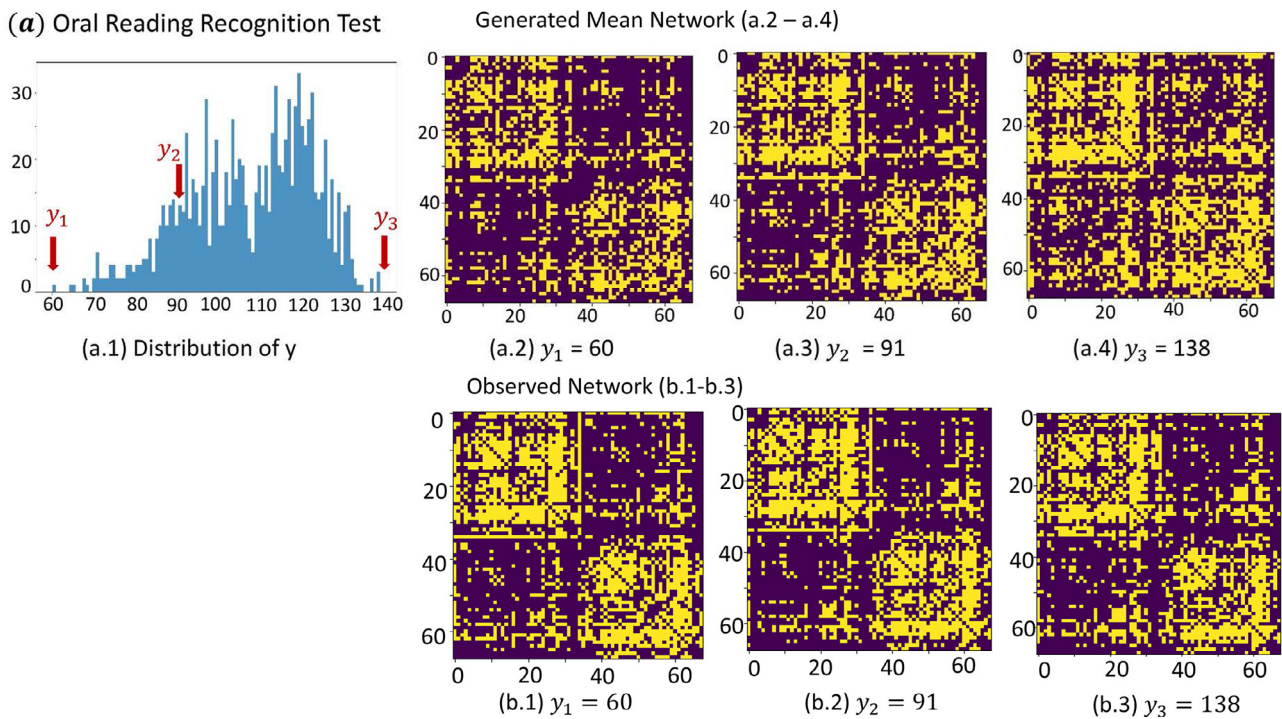


Fig. 13. Examples on generating conditional brain networks $A | y$ for oral reading recognition test in HCP dataset; and comparison between generated brain networks and the observed brain networks for different y in oral reading recognition test for HCP. The first row (a.1) shows histograms of observed trait scores for the 1065 individuals in HCP (the x -axis shows the trait value, and the y -axis shows the frequency). (a.2)-(a.4) show the means of generated brain networks (after dichotomization) conditional on different y under the reGATE model (with the x -axis and the y -axis indexing brain regions; the Supplement contains a table with descriptions of each ROI). (b.1)-(b.3) show the observed brain networks for different y in HCP (with the x -axis and the y -axis indexing brain regions).

frontal) are densely involved (nodes with high degrees) in the top 50 connections. These brain regions are thought to contribute to higher cognition and particularly working memory (Boisgueheneuc et al., 2006), and are important for language-related activities (Binder et al., 1997; Friederici, 2002). We also observe that the four nodes in the occipital lobe (regions 4, 10, 12 and 20) all appear in the top 50 connections. This visual preprocessing center has a few connections (fiber bundles) to the parietal lobe (e.g., regions 7, 24, 28) and then to the frontal lobe. Hence, our results show that richer connections between the visual, sensory, and working memory systems are strongly associated with higher picture vocabulary in adolescents and adults. Some negative connections within each hemisphere are also observed, although their numbers and strengths are smaller than the positive ones. These connections may arise from errors introduced during the connection recovery stage or simply statistical noise.

We further compare the mean difference of the generated data with original brain connectomes in the high and low trait groups. Due to the limited and unbalanced observations for the observed connectome in the particular trait level, we first choose one group of observed networks with the picture vocabulary test score below the 10% quantile and one group of observed networks with the test scores above the 90% quantile, and then calculate the mean difference for the two groups. Figure 15 plots the top 50 connections (based on absolute values) in the mean difference network for the picture vocabulary test from the HCP and ABCD dataset. The 50 connections are further separated into positive and negative connections and are plotted in separate panels in Fig. 15. Comparing with Fig. 14, we can see a similar trend: after ranking the connections by absolute values, these connections are dominated by positive ones, indicating that better vocabulary ability is associated with more connections. Besides the shared dominated connections in both the positive and negative connections, the plots generated from the original weight matrix have more connections that may be caused by the individual's random effects.

4.4. Computing details

We summarize the computing details for our analyses of ABCD and HCP data. We run Adam on GATE and reGATE with the learning rate 0.001 on one NVIDIA Titan-V GPU. In GATE, we used a batch size of 128 sampled uniformly at random at each epoch and ran 200 epochs. In reGATE, we used 5-fold CV to calculate the MSE and ran 100 epochs with a batch size of 128 for each training dataset. The latent dimension K can be chosen as 68 under the same criteria as in Section 3. The computing time in modeling the ABCD dataset is 14.34 mins for GATE and 15.2 mins for reGATE. The computing time in the HCP dataset is 4.74 mins for GATE and 5.18 mins for reGATE.

Table 5 shows the detailed network architectures for the inference model and generative model training.

5. Discussion

We develop a novel nonlinear latent factor model to characterize the population distribution of brain connectomes across individuals and depending on human traits. GATE outputs two layers of low dimensional nonlinear latent representations: one on the individuals that can be used as a summary score for visualization and prediction of human traits of interest; and one on the nodes for characterizing the network structure of each individual based on a latent space model. A supervised model reGATE is proposed to analyze the relationship between human traits and brain connectomes. GATE/reGATE are developed based on a deep neural network framework and implemented via a stochastic variational Bayesian algorithm. The algorithm is computationally efficient and can be applied to massive networks with large number of nodes (brain ROIs).

With applications to the ABCD and HCP data, we used GATE and reGATE to study the relationship between brain structural connectomes and various cognition measures. Using the generative model of reGATE,

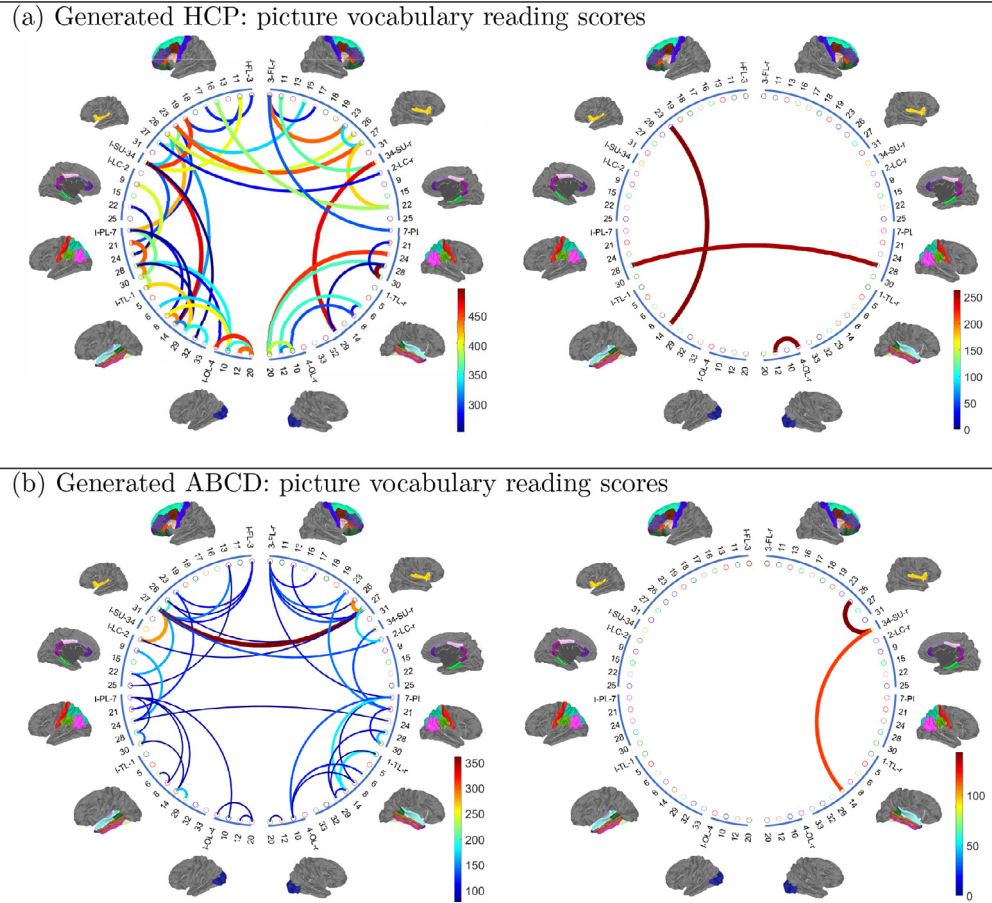


Fig. 14. (a)-(b): Top 50 pairs of brain regions in terms of the mean changes in the generated brain connectivity between two levels of picture vocabulary reading scores (90% and 10% quantiles). The left panel shows the positive connections among the 50 pairs, and the right panel shows negative connections. More details about each node can be found in the Excel spreadsheet in Supplementary Material II.

Table 5

Experimental details and network architectures. K is the dimension of z_i , N is the number of layers in the inference network, M is the number of layers in GCN, R is the dimension of $X^{(i)}$.

	Inference model (μ_ϕ, σ_ϕ)		Generative model	
	μ_ϕ ($N = 2$)	σ_ϕ ($N = 2$)	setting	activation
GATE/reGATE ($K = 68$)	$W_{1,\mu} : 68 * 256$	$W_{1,\sigma} : 68 * 256$	k -NN: 32	h_1 : Sigmoid
	$W_{2,\mu} : 256 * 68$	$W_{2,\sigma} : 256 * 68$	$M = 2$	h_2 : Sigmoid
	$b_{1,\mu} : 256 * 1$	$b_{1,\sigma} : 256 * 1$	$R = 5$	
	$b_{2,\mu} : 68 * 1$	$b_{2,\sigma} : 68 * 1$		
	$\phi_{1,\mu} = \text{ReLU}$	$\phi_{1,\sigma} = \text{ReLU}$		
	$\phi_{2,\mu} = \text{Linear}$	$\phi_{2,\sigma} = \text{Linear}$		

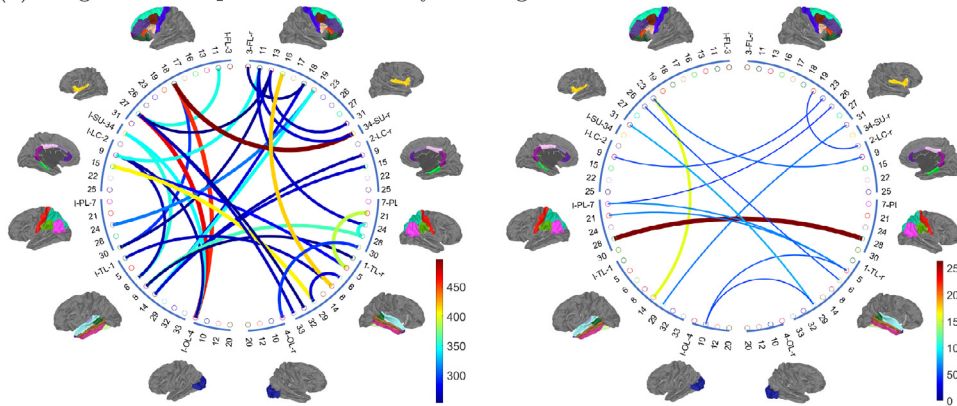
we can simulate brain networks for a given y (e.g., cognition measure) and compare these brain connectomes and related network topological summaries under different cognition levels while allowing variability across individuals. For these cognition traits, we clearly observe that cross hemispheres connections are essential. In the ABCD and HCP datasets, we found that the network density increases while the average path length decreases as the cognition level increases; these network measures for adolescents evaluated in ABCD have a higher variation than adults in HCP. reGATE had superior performance in predicting trait scores from brain networks, with the gain particularly notable in the larger ABCD study.

The generative aspect of GATE/reGATE has a wide range of applications, including data augmentation, outlier network detection, and potentially sensitive data release. As an example in data augmentation, most existing neuroimaging studies contain only a few subjects, and models like GATE that require a large sample size do not work well

for such data. With the help of large datasets, such as the ABCD and UK-Biobank data, we can pre-train a model using the large datasets and then refine the model with the smaller dataset. With our generative model, we can generate more data to mimic the data distribution of the smaller dataset and release confidential datasets generated from this distribution.

In the future, we would like to extent GATE/reGATE in the following directions. First, in brain networks, a refined brain division can give a larger number of nodes, providing a more detailed description of the brain. GATE and reGATE provide a new set of tools for handling high-resolution brain networks, and it becomes interesting to extend the methodology to handle multiresolution data. Moreover, current large studies all collect both functional MRI and diffusion MRI data. It is straightforward to extend GATE/reGATE to jointly embed both functional and structural connectomes, even allowing the strength and nature of the link to vary across individuals and with traits.

(a) Original HCP: picture vocabulary reading scores



(b) Original ABCD: picture vocabulary reading scores

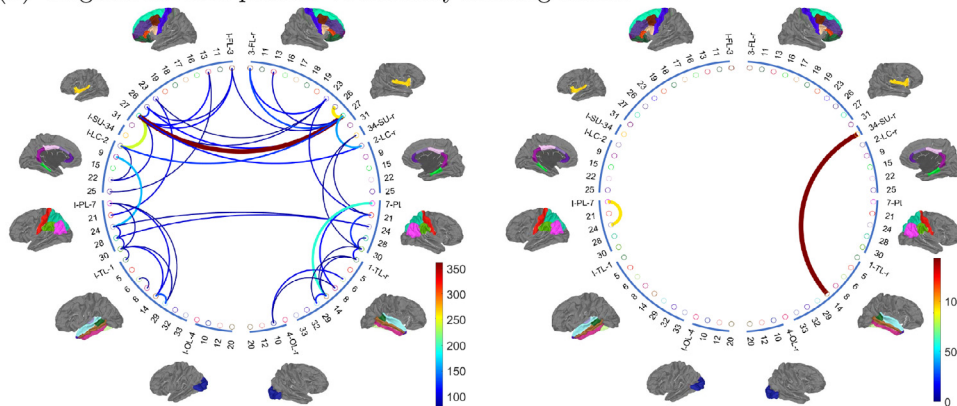


Fig. 15. (a)-(b): Top 50 pairs of brain regions in terms of the mean changes in the original brain connectivity dataset HCP and ABCD between two levels of picture vocabulary reading scores (above 90% quantile and below 10% quantile). The left panel shows the positive connections among the 50 pairs, and the right panel shows negative connections. More details about each node can be found in the Excel spreadsheet in Supplementary Material II.

Credit authorship contribution statement

Meimei Liu: Conceptualization, Methodology, Software, Validation, Formal analysis, Writing – original draft, Writing – review & editing. **Zhengwu Zhang:** Data curation, Formal analysis, Writing – review & editing, Visualization, Funding acquisition. **David B. Dunson:** Conceptualization, Methodology, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Acknowledgements

Data used in the preparation of this article were obtained from the Human Connectome Project (<https://www.humanconnectome.org/>) and Adolescent Brain Cognitive Development (ABCD) Study (<https://abcdstudy.org>). The HCP WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) were funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University. The ABCD study is supported by the National Institutes of Health and additional federal partners under award numbers U01DA041022, U01DA041028, U01DA041048, U01DA041089, U01DA041106, U01DA041117, U01DA041120, U01DA041134, U01DA041148, U01DA041156, U01DA041174, U24DA041123, U24DA041147, U01DA041093, and U01DA041025. A full list of supporters is available at <https://abcdstudy.org/federal-partners.html>. A listing of participating sites and a complete listing of the study investigators can be found at <https://abcdstudy.org/scientists/workgroups/>. ABCD consortium investigators designed and implemented the study and/or provided data but did not necessarily participate in the analysis or writing of this

report. This paper reflects the views of the authors and may not reflect the opinions or views of the NIH or ABCD consortium investigators.

Dunson and Zhang want to acknowledge support from the National Institutes of Health (NIH) of the United States under award number MH118927. We also want to acknowledge the support from the NVIDIA Corporation with their donation of the Titan-V GPU.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:[10.1016/j.neuroimage.2021.118750](https://doi.org/10.1016/j.neuroimage.2021.118750).

References

- Barabási, A.-L., Albert, R., 1999. Emergence of scaling in random networks. *Science* 286 (5439), 509–512.
- Beckmann, C.F., DeLuca, M., Devlin, J.T., Smith, S.M., 2005. Investigations into resting-state connectivity using independent component analysis. *Philos. Trans. R. Soc. B* 360 (1457), 1001–1013.
- Binder, J.R., Frost, J.A., Hammeke, T.A., Cox, R.W., Rao, S.M., Prieto, T., 1997. Human brain language areas identified by functional magnetic resonance imaging. *J. Neurosci.* 17 (1), 353–362.
- Boisgueheneuc, F.d., Levy, R., Volle, E., Seassau, M., Duffau, H., Kinkingnehun, S., Samson, Y., Zhang, S., Dubois, B., 2006. Functions of the left superior frontal gyrus in humans: a lesion study. *Brain* 129 (12), 3315–3328.
- Bullmore, E., Sporns, O., 2009. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* 10 (3), 186.
- Casey, B., Cannonier, T., Conley, M.I., Cohen, A.O., Barch, D.M., Heitzeg, M.M., Soules, M.E., Teslovich, T., Dellarco, D.V., Garavan, H., Orr, C., 2018. The adolescent brain cognitive development (ABCD) study: imaging acquisition across 21 sites. *Dev. Cogn. Neurosci.*
- Craddock, R.C., Jabdi, S., Yan, C.-G., Vogelstein, J.T., Castellanos, F.X., Di Martino, A., Kelly, C., Heberlein, K., Colcombe, S., Milham, M.P., 2013. Imaging human connectomes at the macroscale. *Nat. Methods* 10 (6), 524–539.
- Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., Albert, M.S., Killiany, R.J., 2006. An

- automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* 31 (3), 968–980.
- Durante, D., Dunson, D.B., 2018. Bayesian inference and testing of group differences in brain networks. *Bayesian Anal.* 13 (1), 29–58.
- Durante, D., Dunson, D.B., Vogelstein, J.T., 2017. Nonparametric Bayes modeling of populations of networks. *J. Am. Stat. Assoc.* 112 (520), 1516–1530.
- Fornito, A., Zalesky, A., Breakspear, M., 2013. Graph analysis of the human connectome: promise, progress, and pitfalls. *Neuroimage* 80, 426–444.
- Friederici, A.D., 2002. Towards a neural basis of auditory sentence processing. *Trends Cogn. Sci.* 6 (2), 78–84.
- Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., Rubin, D.B., 2013. *Bayesian Data Analysis*. Chapman and Hall/CRC.
- Gershon, R.C., Wagster, M.V., Hendrie, H.C., Fox, N.A., Cook, K.F., Nowinski, C.J., 2013. NIH toolbox for assessment of neurological and behavioral function. *Neurology* 80 (11), S2–S6.
- Girard, G., Whittingstall, K., Deriche, R., Descoteaux, M., 2014. Towards quantitative connectivity analysis: reducing tractography biases. *Neuroimage* 98, 266–278.
- Glasser, M.F., Smith, S.M., Marcus, D.S., Andersson, J.L., Auerbach, E.J., Behrens, T.E., Coalson, T.S., Harms, M.P., Jenkinson, M., Moeller, S., Robinson, E.C., Sotiropoulos, S.N., Xu, J., Yacoub, E., Ugurbil, K., Van Essen, D.C., 2016. The human connectome project's neuroimaging approach. *Nat. Neurosci.* 19 (9), 1175–1187.
- Goodfellow, I., Bengio, Y., Courville, A., 2016. *Deep Learning*. MIT press.
- Hamilton, W., Ying, Z., Leskovec, J., 2017. Inductive representation learning on large graphs. *Adv. Neural Inf. Process. Syst.*
- Heaton, R.K., Akshoomoff, N., Tulsy, D., Mungas, D., Weintraub, S., Dikmen, S., Beaumont, J., Casaleto, K.B., Conway, K., Slotkin, J., Gershon, R., 2014. Reliability and validity of composite scores from the NIH toolbox cognition battery in adults. *J. Int. Neuropsychol. Soc.* 20 (6), 588.
- Hinton, G., Deng, L., Yu, D., Dahl, G., Mohamed, A.-r., Jaitly, N., Senior, A., Vanhoucke, V., Nguyen, P., Kingsbury, B., Kingsbury, B., 2012. Deep neural networks for acoustic modeling in speech recognition. *IEEE Signal Process. Mag.* 29.
- Hoff, P.D., Raftery, A.E., Handcock, M.S., 2002. Latent space approaches to social network analysis. *J. Am. Stat. Assoc.* 97 (460), 1090–1098.
- Hoffman, M.D., Blei, D.M., Wang, C., Paisley, J., 2013. Stochastic variational inference. *J. Mach. Learn. Res.* 14 (1), 1303–1347.
- Johnson, D.B., 1977. Efficient algorithms for shortest paths in sparse networks. *J. ACM* 24 (1), 1–13.
- Jones, D.K., Knosche, T.R., Turner, R., 2013. White matter integrity, fiber count, and other fallacies: the do's and don'ts of diffusion MRI. *Neuroimage* 73, 239–254.
- Jordan, M.I., Ghahramani, Z., Jaakkola, T.S., Saul, L.K., 1999. An introduction to variational methods for graphical models. *Mach. Learn.* 37 (2), 183–233.
- Kawahara, J., Brown, C.J., Miller, S.P., Booth, B.G., Chau, V., Grunau, R.E., Zwicker, J.G., Hamarneh, G., 2017. BrainNetCNN: convolutional neural networks for brain networks; towards predicting neurodevelopment. *Neuroimage* 146, 1038–1049.
- Kingma, D.P., Ba, J., 2015. Adam: a method for stochastic optimization. In: *International Conference on Learning Representations (ICLR)*.
- Kingma, D.P., Welling, M., 2014. Auto-encoding variational Bayes. In: *International Conference on Learning Representations (ICLR)*.
- Kipf, T.N., Welling, M., 2017. Semi-supervised classification with graph convolutional networks. In: *International Conference on Learning Representations*.
- Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. ImageNet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.*
- Ktena, S.I., Parisot, S., Ferrante, E., Rajchl, M., Lee, M., Glocker, B., Rueckert, D., 2018. Metric learning with spectral graph convolutions on brain connectivity networks. *Neuroimage* 169, 431–442.
- Van der Maaten, L., Hinton, G., 2008. Visualizing data using t-SNE. *J. Mach. Learn. Res.* 9 (11).
- Maier-Hein, K.H., Neher, P.F., Houde, J.-C., Côté, M.-A., Garyfallidis, E., Zhong, J., Chamberland, M., Yeh, F.-C., Lin, Y.-C., Ji, Q., Reddick, W., 2017. The challenge of mapping the human connectome based on diffusion tractography. *Nat. Commun.* 8 (1), 1349.
- Miller, K.L., Alfaro-Almagro, F., Bangerter, N.K., Thomas, D.L., Yacoub, E., Xu, J., Bartsch, A.J., Jbabdi, S., Sotiropoulos, S.N., Andersson, J.L., Griffanti, L., 2016. Multi-modal population brain imaging in the UK Biobank prospective epidemiological study. *Nat. Neurosci.* 19 (11), 1523.
- Nowicki, K., Snijders, T.A.B., 2001. Estimation and prediction for stochastic blockstructures. *J. Am. Stat. Assoc.* 96 (455), 1077–1087.
- Park, H.-J., Friston, K., 2013. Structural and functional brain networks: from connections to cognition. *Science* 342 (6158), 1238411.
- Rezende, D.J., Mohamed, S., Wierstra, D., 2014. Stochastic backpropagation and approximate inference in deep generative models. In: *International Conference on Machine Learning*, pp. 1278–1286.
- Sermanet, P., Chintala, S., LeCun, Y., 2012. Convolutional neural networks applied to house numbers digit classification. In: *International Conference on Pattern Recognition (ICPR 2012)*.
- Smith, R.E., Tournier, J.D., Calamante, F., Connelly, A., 2012. Anatomically-constrained tractography: improved diffusion MRI streamlines tractography through effective use of anatomical information. *Neuroimage* 62 (3), 1924–1938.
- Tuch, D.S., 2004. Q-ball imaging. *Magn. Reson. Med.* 52 (6), 1358–1372.
- Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E., Yacoub, E., Ugurbil, K., Consortium, W.-M.H., 2013. The WU-Minn human connectome project: an overview. *Neuroimage* 80, 62–79.
- Van Essen, D.C., Ugurbil, K., Auerbach, E., Barch, D., Behrens, T., Bucholz, R., Chang, A., Chen, L., Corbetta, M., Curtiss, S.W., Della Penna, S., 2012. The human connectome project: a data acquisition perspective. *Neuroimage* 62 (4), 2222–2231.
- Wang, L., Zhang, Z., Dunson, D.B., 2019. Symmetric bilinear regression for signal subgraph estimation. *IEEE Trans. Signal Process.* 67 (7), 1929–1940.
- Watts, D.J., Strogatz, S.H., 1998. Collective dynamics of 'small-world' networks. *Nature* 393 (6684), 440.
- Xie, J., Douglas, P.K., Wu, Y.N., Brody, A.L., Anderson, A.E., 2017. Decoding the encoding of functional brain networks: an fMRI classification comparison of non-negative matrix factorization (nmf), independent component analysis (ica), and sparse coding algorithms. *J. Neurosci. Methods* 282, 81–94.
- Yoo, Y., Yun, S., Jin Chang, H., Demiris, Y., Young Choi, J., 2017. Variational auto-encoded regression: high dimensional regression of visual data on complex manifold. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3674–3683.
- Zhang, J., Sun, W. W., Li, L., 2018a. Network response regression for modeling population of networks with covariates. *arXiv preprint arXiv:1810.03192*.
- Zhang, L.W.Z., Dunson, D., 2019. Common and individual structure of multiple networks. *Ann. Appl. Stat.* 13 (1), 85–112.
- Zhang, Z., Allen, G.I., Zhu, H., Dunson, D., 2019. Tensor network factorizations: relationships between brain structural connectomes and traits. *Neuroimage* 197, 330–343.
- Zhang, Z., Descoteaux, M., Zhang, J., Girard, G., Chamberland, M., Dunson, D., Srivastava, A., Zhu, H., 2018. Mapping population-based structural connectomes. *Neuroimage* 172, 130–145.
- Zhao, Q., Honnorat, N., Adeli, E., Pfefferbaum, A., Sullivan, E.V., Pohl, K.M., 2019. Variational autoencoder with truncated mixture of Gaussians for functional connectivity analysis. In: *International Conference on Information Processing in Medical Imaging*. Springer, pp. 867–879.
- Zhou, H., Li, L., Zhu, H., 2013. Tensor regression with applications in neuroimaging data analysis. *J. Am. Stat. Assoc.* 108 (502), 540–552.