Integrative Bayesian Analysis of Brain Functional Networks Incorporating Anatomical Knowledge

Ixavier A. Higgins[a], **Suprateek Kundu**[a], and **Ying Guo**[a]

[a]Department of Biostatistics and Bioinformatics, Emory University, Atlanta, GA 30322, USA

Abstract

Recently, there has been increased interest in fusing multimodal imaging to better understand brain organization by integrating information on both brain structure and function. In particular, incorporating anatomical knowledge leads to desirable outcomes such as increased accuracy in brain network estimates and greater reproducibility of topological features across scanning sessions. Despite the clear advantages, major challenges persist in integrative analyses including an incomplete understanding of the structure-function relationship and inaccuracies in mapping anatomical structures due to inherent deficiencies in existing imaging technology. This calls for the development of advanced network modeling tools that appropriately incorporate anatomical structure in constructing brain functional networks. We propose a hierarchical Bayesian Gaussian graphical modeling approach which models the brain functional networks via sparse precision matrices whose degree of edge specific shrinkage is a random variable that is modeled using both anatomical structure and an independent baseline component. The proposed approach adaptively shrinks functional connections and flexibly identifies functional connections supported by structural connectivity knowledge. This enables robust brain network estimation even in the presence of misspecified anatomical knowledge, while accommodating heterogeneity in the structure-function relationship. We implement the approach via an efficient optimization algorithm which yields maximum a posteriori estimates. Extensive numerical studies involving multiple functional network structures reveal the clear advantages of the proposed approach over competing methods in accurately estimating brain functional connectivity, even when the anatomical knowledge is misspecified up to a certain degree. An application of the approach to data from the Philadelphia Neurodevelopmental Cohort (PNC) study reveals gender based connectivity differences across multiple age groups, and higher reproducibility in the estimation of network metrics compared to alternative methods.

Keywords

Corresponding author: Suprateek Kundu, suprateek.kundu@emory.edu, 1518 Clifton Rd., NE , Atlanta , GA 30322-4201.
Conflicts of interest
Declarations of interest: none

1 Introduction

The human brain is an extremely complex organ responsible for all thought and bodily function. Various approaches have sought to explain the brain's functionality as a result of neurotransmissions between individual neurons, reflected as the co-activation between voxels in brain images. Recently, there has been a rapid increase in research on brain connectome analysis focused on linking interregional dependencies to brain function. Methods for both structural connectivity (SC) and functional connectivity (FC) have seen increasing developments with the emergence of non-invasive technologies such as diffusion tensor imaging (DTI) and functional magnetic resonance imaging (fMRI). FC measures the temporal coherence in brain activity across two distinct brain regions, while SC approaches based on DTI reconstruct white matter pathways in the brain by measuring the diffusivity of water molecules in brain tissues. These two types of connectivity offer complementary and interdependent information about brain structure and function.

Despite strong evidence regarding the role of white matter fiber tracts in regulating FC (Damoiseaux and Greicius, 2009; Honey et al., 2010; Sporns, 2013) and considerable progress in separately estimating FC and SC, there have been comparatively limited advances in FC approaches which are guided by underlying anatomical knowledge. Incorporating anatomical knowledge in estimating FC is clearly desirable since it is expected to produce more accurate estimates of the network, which translates to greater reproducibility of the findings as illustrated via our fMRI data analysis. However, several considerations need to be taken into account, such as the complexity of the structure-function relationship (Hermundstad et al., 2013), heterogeneity in FC for a given SC strength, which presumably is attributed to the fact that FC is only partially dependent on SC (Damoiseaux and Greicius, 2009; Messé et al., 2014) and regulated by unobservable dynamics in underlying neuronal activity (Bressler and Tognoli, 2006).

Recently, Venkataraman et al. (2012) and Xue et al. (2015) proposed approaches to jointly model the probability of co-activation based on fMRI data while incorporating direct structural connections. They provide measures of functional co-activation deviating from standard measures of FC such as Pearson or partial correlations. Hinne et al. (2014) proposed a Bayesian approach which uses fMRI data to model the distribution of partial correlations for edges determined by the given SC information. The assumption that FC only exists between anatomically connected regions ignores the contributions of indirect anatomical pathways and does not capture the complexity of the relationship between brain structure and function (Honey et al., 2009, 2010; Messé et al., 2014). Moreover, the above approaches use multi-subject data which requires registration of images to a shared template under the assumption that the volumes are similar and can be matched. Unfortunately, this assumption has limitations for human brain images considering the substantial variability in cortical anatomy and function (Zhu et al., 2012). This variability is especially pronounced during the developmental phases of childhood and adolescence as is the case with our motivating Philadelphia Neurodevelopmental Cohort (PNC) study. Ng et al. (2012) and Pineda et al. (2014) proposed approaches for estimating sparse functional networks for individual subjects via an adaptive graphical lasso. In both models, edge specific shrinkage parameters are deterministic functions of the SC information. Under these approaches, FC

with less anatomical support are more heavily penalized, and vice-versa. However, the parametric form of the shrinkage parameters may not adequately capture the complex underlying structure-function relationships and does not account for heterogeneity in FC for a given SC strength resulting from non-anatomical sources of variation such as BOLD signal sensitivity to proximity to blood vessels (Zhang et al., 2016) as well as cardiac, acquisition and preprocessing effects (Skudlarski et al., 2008). Moreover, such a parametric relationship may lead to network estimates which are not robust and more sensitive to the misspecification of anatomical knowledge. Figure 1 illustrates a representative FC-SC relationship in the PNC study (refer to section 3.2.1 for details on the calculation of SC strengths).

The above discussions highlight a serious need for integrative modeling approaches which adaptively estimate FC by incorporating structural knowledge in an appropriate manner. In designing such an approach, our primary goals for the method are that it will (a) correctly identify true interregional functional connections and identify which connections are absent in the true network; (b) lead to reproducible estimates of FC such that the estimated brain network can be replicated across multiple scanning sessions, which is a topic of great importance in current literature (Varoquaux et al., 2010); and (c) specify a flexible structure-function relationship which is robust to misspecification of SC information (arising from limitations in existing image acquisition technology) and can accommodate heterogeneity in FC for a given SC. We propose a novel hierarchical Bayesian Gaussian graphical modeling approach for estimating FC based on single subject fMRI data which incorporates given SC information in a manner that addresses the aforementioned objectives. The FC is computed via sparse precision matrices whose elements are estimated under Laplace type priors having edge specific shrinkage parameters that are random variables modeled using SC information and an independent baseline component. The prior encourages stronger FC given a large SC (and vice-versa), but also accounts for edge specific variations in FC unrelated to the brain anatomy, via the baseline component. Thus, the approach is flexible in accounting for anatomical knowledge with the FC being guided by, but not completely determined by, the SC information. Our method is motivated by the variable selection approach in Chang et al. (2016), which incorporates prior graph knowledge in a linear regression setting, but is distinct in addressing graphical model selection and precision matrix estimation, as well as the manner in which the prior knowledge is incorporated. Under certain choices of model parameters, the proposed approach reduces to an adaptive shrinkage approach specifying a parametric relationship between the shrinkage parameters and the anatomical information, similar to Ng et al. (2012) and Pineda et al. (2014).

While Markov chain Monte Carlo (MCMC) can be used to implement the proposed approach, it is not scalable to large networks needed for whole brain connectome analysis as in our application. Moreover, MCMC samples cannot take exact zeroes under a Laplace prior, and an additional thresholding step is often needed for model selection. We propose an optimization algorithm to obtain the *maximum a posteriori* (MAP) estimate, which is computationally efficient, scales to a large dimensions, and does not require post-hoc thresholding. Under various simulation studies, we observe superior performance of the proposed method as compared to alternative approaches with or without SC information. The advantages of our approach become more evident as the degree of misspecification of

anatomical knowledge increases, and/or the number of nodes grows larger which is particularly relevant for whole brain connectome analysis.

Our efforts are motivated by data from the PNC study, a large-scale, NIMH funded initiative to understand the developmental trajectory of the brain from childhood to adolescence (Satterthwaite et al., 2014). The PNC data contains both DTI and resting state fMRI measurements from boys and girls ages 8–21 years, with suggestive but unclear SC-FC relationships (Figure 1). We fuse multimodal brain imaging data to examine gender differences in brain networks across three age brackets—pre-teens (ages 8–12), teens (ages 13–17), and young adults (ages 18–21)—and discover several gender based differences in FC within and across the age brackets. We also assess the reliability of computed network metrics across scanning sessions and find that the proposed approach yields strong reproducibility in the estimation of network metrics, which is almost always higher than alternative approaches. While other studies have separately examined the reproducibility of functional and structural brain networks (see Welton et al. (2015) for a review), ours is one of the first to examine the reproducibility of anatomically informed functional networks to the best of our knowledge.

Section 2 describes the proposed methodology and the optimization routine for estimating networks, while sections 3 presents numerical studies and application of our method to PNC data. We conclude with a brief discussion in Section 4.

2 Materials and methods

2.1 Gaussian graphical model for brain networks

While early work on brain network estimation utilized Pearson correlation to measure undirected interregional dependencies, recent literature has focused on Gaussian Graphical Models (GGMs) which are parametrized by the inverse covariance matrix and measure functional connectivity via partial correlations. The precision matrix has often been reported to result in more accurate and robust estimation of underlying brain network structures as compared to other methods such as thresholding of the covariance matrix (Smith et al., 2011). Compared to the covariance approach, the precision matrix can distinguish a true, direct functional connection between two regions from those that exist because of confounding with other nodes in the network.

GGMs assume observations are normally distributed and that zeros in the inverse covariance matrix correspond to absent edges in the network, $\mathscr{G}$. A standard GGM specifies $y_t \sim N_p(0, \Omega^{-1})$ where y_t is a vector containing the fMRI signal at each of the p ROIs for the tth time-index, where t=1,..,T and T is the total number of image volumes acquired during the scanning session. In everything that follows, $\mathbf{Y}_{T \times p} = \{y_1, \ldots, y_T\}$ is the data matrix where each row contains the fMRI signals across all brain regions at time index t. $\mathbf{\Omega}$ is the inverse covariance matrix (or precision matrix) which belongs to the cone of $p \times p$ symmetric positive definite matrices, denoted by M_p^+. Under this framework, estimating $\mathscr{G}$ is equivalent to estimating structural zeros in the positive definite precision matrix Ω.

Due to tradeoffs between the cost and efficiency of information transfer, it is typically assumed that the brain seeks efficient organization favoring a sparse set of active connections at any point in time (Eavani et al., 2015). The GGM approach is well equipped to handle such sparse networks by imposing penalties that shrink sufficiently weak functional connections to zero, where the L_1 penalty under the graphical lasso (Friedman et al., 2008) is a popular choice (Ng et. al, 2012; Monti et. al, 2014 ; Pineda-Pardo et. al, 2014). The graphical lasso can be thought of as an extension of the Lasso approach in regression settings and penalizes the full data likelihood to estimate the inverse covariance matrix as

$$\hat{\Omega} = \arg \max_{\Omega \in M_p^+} \log \det(\Omega) - \mathrm{tr}(S\Omega) - \lambda \sum_{j \leq k} \left| w_{jk} \right| \quad 1$$

Where $s = \sum_{t=1}^{T} (y_t' y_t)/T$ is the sample covariance matrix, $|x|$ denotes absolute value of x, det(.) is the determinant operator, tr(.) is the matrix trace operator, and $\lambda > 0$ is the penalty parameter controlling overall network sparsity. If $\lambda = 0$, one obtains the maximum likelihood estimate, while large values of λ shrinks an increasing number of off-diagonal elements to zero. The typical graphical lasso approach fits a series of graphs under various choices of the tuning parameter λ and chooses the optimal network as the one minimizing some goodness of fit criteria (Yuan and Lin, 2006).

2.2   Structurally informed Bayesian Gaussian graphical model

Bayesian GGM approaches have been successfully used for estimating brain networks (see Mumford et. al (2014) for a review). One such approach is the Bayesian graphical lasso (Wang, 2012), which has similarities with the penalized graphical lasso approach in the sense that the *maximum a posteriori* (MAP) estimator is equivalent to the estimate under (1). This approach assumes that the *p* dimensional fMRI signal at time index t is distributed as $y_t \sim N_p(0, \Omega-1)$, $t = 1, .., T$, with the prior on the inverse covariance as

$$\pi(\Omega | \lambda) = C_\lambda^{-1} \prod_{k=1}^{p} \mathrm{Exp}(\omega_{kk}; \lambda) \prod_{j<k} \mathrm{DE}(\omega_{jk}; \lambda) \mathrm{I}(\Omega \in M_p^+) \quad 2$$

where $\pi(.)$ represents the prior distribution and I(x) is an indicator function that takes the value one when condition x is true. The diagonal element ω_{kk} is modeled under an exponential prior distribution $Exp(\lambda)$, the off-diagonal element ω_{jk} is modeled with double exponential or Laplace prior distribution $DE(\lambda)$, and λ is the shrinkage parameter. In a fully Bayesian paradigm, λ is typically assigned a prior distribution, and is thus learned from the data, resulting in an adaptive shrinkage of the elements in Ω.

In order to incorporate anatomical knowledge in functional connectivity estimation, we propose a hierarchical Bayesian structurally informed Gaussian graphical model (siGGM). It is based on the generic Bayesian GGM in (2), but involves edge specific shrinkage parameters which are modeled using anatomical knowledge. Throughout this article, we will

denote the structural connectivity metric as p_{jk} for edge (j, k), where a larger value denotes a stronger anatomical connection and vice-versa. For example, in our data application, pjk corresponds to the probability of SC obtained via probabilistic tractography (please see section 3.2.1 for details). The proposed approach to estimating the brain functional network incorporating anatomical knowledge is defined as follows

$$\pi(\Omega|\lambda) = C_{\lambda,\nu}^{-1} \prod_{k=1}^{p} \mathrm{Exp}(\omega_{kk}; \tfrac{\nu}{2}) \prod_{j<k} \mathrm{DE}(\omega_{jk}; \nu\lambda_{jk}) I(\Omega \in M_p^+),$$

$$\pi(\lambda|\mu,\eta) = C_{\lambda,\nu} \prod_{j<k} \mathrm{LN}(\mu_{jk} - \eta p_{jk}, \sigma_{\lambda}^2) \quad 3$$

where (i) $\lambda = \left\{ \lambda_{jk,\ j\ <\ k}, j, k\ =\ 1,\ \ldots,\ p \right\}$ denotes the collection of edge specific shrinkage parameters having a log-normal (LN) type distribution which restricts the shrinkage parameters to non-negative values; (ii) ν refers to the tuning parameter controlling the network's overall sparsity and also corresponds to the scale parameter for the exponential prior on the diagonals; (iii) η is a positive random variable which controls the average effect of SC on FC; (iv) μ_{jk} denotes the random edge specific baseline component representative of non-anatomical sources of variations regulating FC; and (v) $C_{\lambda,\nu}$ is the intractable normalizing constant for the prior on the precision matrix depending on λ and ν. This constant assures a proper prior distribution on Ω; however, it is not possible to analytically evaluate it due to the constraint that $\Omega \in M_p+$, as described in Wang (2012). We utilize the trick introduced in that paper where the intractable constant is included in the prior on λ such that it cancels with the term in $\pi(\Omega/\lambda)$, leading to a closed form full posterior that facilitates computation. We note that in the extreme case when $\log(\lambda_{jk}) = \mu_{jk} - \eta p_{jk}$, the model specifies aparametric relationship which has similarities with Ng et. al (2012) and Pineda-Pardo et. al (2014).

The anatomically informed prior on the shrinkage parameters in (3) specifies a probabilistic relationship between the edge specific shrinkage parameters and the given SC knowledge via η. In particular, increasing positive values of η implies an increasing dependence on the given SC, potentially resulting in a functional connection even for small SC weights. Figure 2 illustrates that for large η and increasing SC, the Laplace prior has heavier tails and less mass around zero, which is interpreted as increased probability of strong FC. In contrast, small values of η do not result in a noticeable change in the prior distribution under varying SC strengths, implying a negligible relationship between SC and FC. Moreover, the shrinkage parameters are stochastically monotonically decreasing with respect to the SC strength, under the restriction $\eta > 0$. This implies that as the SC strength (pjk) for the edge (j, k) is increased, the corresponding shrinkage parameter λjk will take smaller values in probability, resulting in values of ωjk which are away from zero. Hence, the presence of FC at edge (j, k) shrinkage for ωjk resulting in the absence of FC at edge (j, k).

Additionally, the baseline effect, μ_{jk}, corresponds to variations in underlying neuronal activity that are independent of the brain anatomy. This formulation enables (a) more flexibility in the FC-SC relationship by allowing the possibility of strong FC when an anatomical connection is not obvious, and vice-versa; and (b) heterogeneity in FC across edges which possesses similar SC strength that is often encountered in practice. Overall, increasing(decreasing) absolute values of the baseline effect discourages(encourages) the presence of an edge in a manner that is independent of the anatomical information. Although one could also accommodate variations in FC for a given SC strength via σ_λ while keeping μ fixed, our formulation incorporating an edge specific baseline effect confers several advantages. First, it permits the accommodation of edge specific tuning of the shrinkage parameters while $\sigma\lambda$ controls global variation in FC. Second, μ_{jk} can be used to differentiate edge specific variation in FC that is not attributed to direct SC, and can be used to characterize edges with enhanced or trivial SC influences. The hyperparameters μ and η are unknown and are learned in a data-adaptive manner under the following priors

$$\pi(\mu_{jk}) = \text{N}(\mu_0, \sigma_\mu^2) \quad \text{for} \quad j<k, j,k=1,\ldots,p \quad 4$$
$$\pi(\eta) = Ga(a_\eta, b_\eta)$$

where (μ_0, σ_μ^2), and (a_η, b_η), are typically pre-specified. In equation (4), Ga(a_η, b_η) is the gamma distribution with scale parameter a_η, rate parameter b_η, and expected value a_η/b_η. We select these distributions for μ_{jk} and η because they are conjugate and provide adequate performance under the proposed method. Hyperparameters governing these priors are discussed in Appendix A. We note that the scale parameter ν controls the overall network sparsity and is treated as a tuning parameter, enabling the estimation of a series of networks with varying sparsity levels. The optimal network is chosen as the point estimate corresponding to the value of ν minimizing the Bayesian Information Criteria.

2.3 Model Estimation

Although the proposed model can be implemented using MCMC, it is not scalable to high dimensional settings involving a large number of nodes. Moreover, MCMC samples require a post hoc thresholding step to select important edges since estimates cannot take exact zeros under a Laplace prior. We bypass these limitations by computing a MAP estimate for the parameters of interest. Our iterative optimization approach employs an existing graphical lasso algorithm to sample the precision matrix given all other parameters, coupled with additional optimization steps to sample the shrinkage parameters and associated hyperparameters inherent in the Bayesian specification (3). In order to fit the proposed model, we estimate $\Theta = (\Omega, a, \eta, \mu)$ by maximizing the log-posterior distribution in (6), where $a = \log(\lambda) = \{\log(\lambda_{jk}), j < k\}$ and $\mu = \{\mu_{jk}, j < k\}$ denotes the vector of edge specific log-shrinkage parameters and baseline effects in (3), respectively. Note that a is normally distributed due to the log-normal prior placed on λ. The posterior distribution can be written as

$$P(\Theta|y_1,...,y_T)\alpha P(\Theta)P(y_1,...,y_T|\Theta)$$

$$= P(\Omega,\alpha,\mu,\eta)\prod_{t=1}^{T}P(y_t|\Theta)$$

$$= P(\Omega|\alpha,\mu,\eta)P(\alpha|\mu,\eta)P(\mu,\eta)\prod_{t=1}^{T}P(y_t|\Omega^{-1})$$

$$= P(\Omega|\lambda)P(\alpha|\mu,\eta)P(\mu)P(\eta)\prod_{t=1}^{T}P(y_t|\Omega^{-1})$$

$$= P(\eta)\prod_{t=1}^{T}P(y_t|\Omega^{-1})\prod_{j,k}P(\Omega|\lambda)P(\alpha|\mu,\eta)P(\mu)$$

$$\mathrm{Ga}(\eta;a_\eta,b_\eta)\prod_{t=1}^{T}\mathcal{N}_p(y_t;0,\textstyle\sum)\prod_{k=1}^{p}\mathrm{Exp}(w_{kk};v/2)$$

$$\prod_{j<K}\mathrm{DE}(\omega_{jk};v\lambda_{jk})\mathcal{N}(\alpha_{jk};\mu_{jk}-\eta p_{jk},\sigma_\lambda^2)\mathcal{N}(\mu_{jk};\mu_0,\sigma_\mu^2)\quad 5$$

We find the MAP solution for the model parameters by maximizing over the the posterior log-likelihood as $\Theta\hat{} = \arg\max l(\Theta)$, where

$$i(\Theta) = -\frac{T}{2}\log|\Omega| + \frac{1}{2}\mathrm{tr}(S|\Omega|) + v\sum_{j<k}e^{\alpha_{jk}}|\omega_{jk}| + \sum_{j<k}\frac{(\alpha_{jk}-(\mu_{jk}-\eta p_{jk}))^2}{2\sigma_\lambda^2}$$

$$-(a_\eta-1)\log(\eta) + b_\eta\eta + \sum_{j<k}\frac{(\mu_{jk}-\mu_0)^2}{2\sigma_\mu^2} - p\log(\tfrac{1}{2}v) + \frac{1}{2}v\sum_{k=1}^{p}\omega_{kk}.$$

All parameters in the posterior distribution are updated iteratively until convergence. The precision matrix is updated given other parameters using the existing graphical lasso algorithm, whereas μ_{jk} and η are updated via closed form expressions and α is updated via a

Newton-Raphson step since a closed form solution does not exist. The iterative updates continue until $|l(\Theta^{(m)}) - l(\Theta^{(m+1)})|/ < \epsilon |l(\Theta^{(m+1)})|$ for $\epsilon = 10^{-4}$. At convergence, $\hat{\Theta} = (\hat{\Omega}, \hat{\alpha}, \hat{\eta}, \hat{\mu})$ is the solution, where $\hat{\Omega}$ is the anatomically informed functional brain network based on single subject data. In general, the method makes its largest improvements within the first three iterations (see Figure 6) and converges rapidly. We note that one could alternatively treat μ and η as tuning parameters and compute a range of networks over a grid of (μ, η) values, and then select the optimal network as the one minimizing some goodness of fit criteria. However, this strategy did not result in adequate numerical performance, highlighting the advantages of specifying suitable priors on hyperparameters in order to estimate them in a data-adaptive manner. The computational steps for updating the model parameters are detailed in Appendix A. The method was developed in R version 3.3.0 and is available on github (https://github.com/IxavierHiggins/siGGMrepo).

3   Results and Discussion

3.1   Simulations

3.1.1   Simulation Setting—

We conduct numerical studies to assess the performance of siGGM relative to SC naive and SC informed competitors. SC naive approaches are representative of methods that do not incorporate auxiliary information, and includes the graphical lasso or *Glasso* (Friedman et al., 2008), the partial correlation approach *Space* proposed by Peng et al. (2009), and the proposed approach in (3) without structural information obtained by setting $\eta = 0$, denoted *siGGM* ($\eta = 0$). SC informed approaches incorporate anatomical information into the estimation routine. We consider the Bayesian G-Wishart approach by Hinne et. al (2014) which treats FC as completely determined by SC and is denoted by G-Wishart as well as the adaptive graphical lasso approach by Pineda-Pardo et. al (2014) which specifies a parametric relationship between the shrinkage parameters and SC, and is denoted by aGlasso. All of the above approaches, except *Space* which estimates partial correlations, calculate sparse inverse precision matrices, where a zero off-diagonal entry implies the absence of an edge. The *Glasso* and *Space* approaches are implemented via the R packages *glasso* and *space*, respectively. We estimate the precision matrix under the G-Wishart approach using Matlab scripts available on the author's website and incorporate adaptive weights in the *glasso* algorithm to implement aGlasso.

Data Generation: In order to assess the performance of our approach, we simulate data under three assumed network structures and consider various relationships between SC and FC. The network structures are (a) Erdos-Renyi (ER) networks consisting of edges randomly generated with probability 0.15; (b) small-world (SM) networks generated under the Watts-Strogatz model (Watts and Strogatz, 1998) in which most nodes may not be directly connected, but can reach other nodes via a small number of steps, and (c) scale-free (SF) networks generated using the preferential attachment model (Barabási and Albert, 1999), in which nodes are more likely to link to a highly connected node than to a node with few connections, resulting in a hub network. For each network, we consider varying the number of nodes corresponding to $p = 100, 200$. The data was generated using a Gaussian graphical model $\mathbf{y}_t \sim N_p(0, \Omega^{-1})$ with $T = 200$ time points ($t = 1, \ldots, 200$), where $\Omega_{\mathscr{G}}$ is the precision

matrix with zero off-diagonal elements corresponding to absent edges in the binary network $\mathcal{G}$, and non-zero off-diagonal elements otherwise. Conditional on $\mathcal{G}$, $\Omega_{\mathcal{G}}$ is constructed as follows: the non-zero off-diagonal elements corresponding to important edges are generated from a Uniform($-1,1$) distribution, and the diagonal elements were fixed to be one. In order to ensure positive definiteness, we subtracted the minimum of the eigenvalues from each diagonal element of the generated precision matrix.

We also assess the methods' performance for non-Gaussian data which is generated using realistic fMRI time series under spatiotemporal separability assumptions of independent component analysis (ICA), using the framework in the SimTB (Allen et. al, 2011) Matlab toolbox. We generate $Y_{T \times p} = A_{T \times q}S_{q \times p} + E_{T \times p}$ where the columns of A represent the temporal dynamics of functional networks and contain realistic BOLD-type resting state fMRI time courses sampled using the neuRosim R package (Welvaert et al., 2011); S is the source map representing the spatial distribution of functional subnetworks constructed under small-world networks replicating properties of realistic brain networks; and E is a matrix of zero mean white noise. The true precision matrix, Ω_0, can be computed theoretically as $(S'\mathrm{cov}(A')S + \epsilon^2 I)^{-1}$, where $\epsilon = .5$ and I is the $p \times p$ identity matrix. Although this simulation approach produces realistic time series data, it results in dense precision matrices with no off-diagonal zeros. Given strong support in the literature of the sparsity of brain networks (Eavani et al., 2015; Kim et al., 2015; Power et al., 2013), we calculate the true network as having 10% density by retaining only those edges for which the corresponding partial correlation was in the top 10 percentile. However, we note that the results under the proposed method are stable over varying true network densities.

Prior SC Knowledge: Conditional on $\Omega_{\mathcal{G}}$, we construct several types of SC information according to the following schemes, where the SC strengths were generated randomly between (0, 1) and FC is measured by partial correlations. For scenario *MI*, we specify that 50% of those edges with strong FC (partial correlation $> .06$) also have strong SC(> 0.7), while 25% of those edges with strong FC have moderate SC($0.3 - 0.7$), and the remaining 25% have weak SC($0 - 0.3$). For scenario *MII*, the proportion of edges that have strong FC, coupled with strong SC, moderate SC, and weak SC, are 30%, 35% and 35% respectively. For each of the scenarios *MI* and *MII*, we also consider two levels of misspecification of the SC information, which are denoted as *MI*(*a*), *MI*(*b*), and *MII*(*a*), *MII*(*b*), respectively. For *MI*(*a*)/*MII*(*a*), we specify that 10% of those edges with zero FC have non-zero SC, while for *MI*(*b*)/*MII*(*b*), we fix 20% of those edges with zero FC to have non-zero SC. All the other edges with zero or weak FC are assumed to have small SC, while remaining edges with moderate FC have non-zero SC strengths. We note that edges having zero FC but non-zero SC represent potential misspecification of anatomical knowledge, based on the notion that strong SC typically underlies robust non-zero FC (Shen et al., 2015; Kemmer et al., 2017). For the non-Gaussian data based on the ICA model, we generate the SC information as in scenario MI, and consider varying levels of misspecification.

Comparison Metrics: To assess the performance under different approaches, we compute the area under the curve (AUC), which is a measure of the estimated sensitivity versus specificity over different network sparsity levels. Sensitivity is computed as $TP/(TP + FN)$,

while specificity is defined as $TN/(TN + FP)$, where *TP, TN, FP, FN* denote the number of edges that are true positives, true negatives, false positives and false negatives, respectively. These measures are derived by comparing the true binary network structure, $\mathscr{G}$, to $\hat{G}$ Which contains edges corresponding to non-zero elements in the estimate $\hat{\Omega}$. To evaluate the point estimate of the network obtained under the Bayesian information criteria (BIC), we compute the Matthews Correlation Coefficient (MCC), which is a scalar measure combining sensitivity and specificity and is defined as MCC=

$$\frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$ (Matthews, 1975; Wang, 2012). We compute the relative L_1 norm error, $(\|\hat{\Omega} - \Omega\|_1)/\|\Omega\|_1$, to assess accuracy in estimating the precision matrix encapsulating the FC strengths. Since brain networks are also often evaluated in terms of summary statistics reflecting network organization, we evaluate the accuracy in estimating the global efficiency which is a commonly used measure for global integration of brain connectivity.

Since the true precision matrix is dense for the ICA generated data, we consider two alternate measures of performance under this scenario—the inverse error (Padmanabhan et al., 2016) and Kullback-Leibler divergence (Hinne et al., 2014). The inverse error captures the accuracy of the estimated precision matrix, and is defined as $\|\Omega_{-1}\hat{\Omega} - I\|_F$ where $\hat{\Omega}$ is the estimated precision matrix, I is a $p \times p$ identity matrix, and $\|.\|_F$ is the frobenius norm. On the other hand, the Kullback-Leibler divergence is defined as $1/\{2(\log(2))\}[\log(\det(\Omega 0)/\det(\hat{\Omega})) + \text{tr}(\hat{\Omega}\Omega - 1) - p]$, with a larger divergence corresponding to a poor fit.

3.1.2 Results—The results under siGGM and SC naive approaches under *MI*(*a*) are presented in Table B.1, and Table 1 displays results for SC informed approaches under various levels of misspecification. Table B.1 illustrates that either the proposed siGGM approach, or the variant of the proposed approach with no prior knowledge ($\eta = 0$), have the lowest bias in estimating the global efficiency (Eglob) across all network sizes (p=100, 200). Moreover, the proposed approach always has higher MCC and AUC values, and lower L_1 error norm compared to all other SC naive approaches. These results demonstrate the advantages of using structural knowledge to guide network estimation.

When the misspecification levels are varied, Table 1 illustrates that the proposed method has a consistently lower bias in estimating the global efficiency for both 10% and 20% misspecification levels, compared to alternative SC informed approaches. Moreover, while the G-Wishart approach may have a higher MCC for $p = 100$ when the misspecification level is 10% (cases *MI*(*a*) and *MII*(*a*)), the proposed method has a comparable or higher MCC for 20% misspecification levels (cases *MI*(*b*) and *MII*(*b*)). Moreover under $p = 200$, the MCC under the proposed approach is the highest for small-world and scale-free networks, and comparable to the G-Wishart method for the Erdos-Renyi network. We also note that while aGlasso often has the lowest MCC values, it may sometimes yield a higher AUC under small-world and scale-free networks for $p = 100$. However, the proposed approach is shown to have the highest AUC for $p = 200$ for all scenarios, highlighting the advantages of incorporating prior knowledge in a flexible manner in higher dimensions. Finally, siGGM consistently has the lowest L_1 error in estimating the precision matrix across all networks and dimensions. The above results illustrate a robust performance of the proposed method

for $p = 100$ and a superior performance for $p = 200$ under the small-world and scale-free networks, which closely resemble brain networks encountered in practical applications.

Although Table 1 provides some idea about the relative performance under misspecification, it is of interest to look at the effects of misspecification in more detail. Hence, we examined the AUC and $L1$ error values as the misspecification level was gradually increased from 4% to 50%, under different networks for $p = 100$. The results, presented in Figure 4, illustrate that the proposed method has a signifi tly higher AUC under the Erdos-Renyi network across all misspecification levels, and registers a significantly higher AUC for larger misspecification levels under the small-world network. The differences in AUC between siGGM and aGlasso are not significant for $p = 100$, but we note that siGGM has a higher AUC for larger dimensions ($p = 200$) as in Table 1. For all networks, the proposed method is seen to have a significantly lower L_1 error across all misspecification levels, while the error under the G-Wishart increases sharply as the misspecification level is increased.

The superior performance of siGGM relative to SC informed and SC naive approaches is also observed in the ICA based simulated data exhibiting the spatiotemporal dynamics of the BOLD signals. In Figure 5, siGGM exhibits the largest or comparable AUC across all levels of SC misspecification, with significant improvements in AUC for $p = 100$. Our method also has significantly lower inverse error and Kullback-Leibler divergence scores across misspecification levels. On the other hand, the AUC for aGlasso drops sharply for $p = 100$ for higher misspecification levels. Moreover, G-Wishart's strict adherence to the SC information contributes to poor detection and estimation of edges, with the AUC declining sharply and the inverse error and Kullback-Leibler divergence increasing steeply as the misspecification proportion increases. The above discussions clearly illustrate the ability of siGGM to recover the true network of connection strengths via flexible incorporation of anatomical information under both Gaussian and non-Gaussian settings, with a robust performance under varying SC misspecification levels.

Finally, we note that the siGGM can be implemented fairly quickly. On a 2.5Gz Intel Core i5 processor, the procedure estimates the optimal graph structure in approximately three seconds for p=40, twenty seconds for p=100, and approximately four minutes for p=200. While these computation times are slightly slower compared to generic graphical modeling approaches naive to anatomical knowledge, the overall computation is sufficiently quick and feasible for practical implementation in whole brain connectome analysis. Moreover, the siGGM approach converges fairly quickly over a wide range of simulation scenarios, as illustrated in Figure 6.

3.2 PNC Data Application

Existing literature has examined various neural substrates for age related changes using structural and functional neuroimaging (Gur et al., 2012; Shaw et al., 2008; Raznahan et al., 2011). Moreover, gender differences have been extensively documented in behavioral measures (Halpern et al., 2007; Hines, 2010), structural neuroimaging (Lenroot et al., 2007), and functional imaging measures (Lenroot and Giedd, 2010). However, gender related differences in the developmental trajectory of the brain functional network from childhood to adolescence are still not understood well (Gur et al., 2012), and further, limited attempts

have been made to investigate such differences by fusing functional and structural neuroimaging data. We use resting state fMRI and DTI data from the Philadelphia Neurodevelopment Cohort (PNC) study to obtain preliminary answers to these questions. After estimating brain functional connectivity based on SC knowledge, we examine FC differences between boys and girls across different age groups.

We perform the analysis separately for each gender within the three age groups 8–12 (preteen), 13–17 (teen), 18–21 (young adult), where each age group contains approximately 9 to 12 individuals, and is constructed as in Ingalhalikar et al. (2014). All subjects are right-handed, physically, and mentally healthy, enabling a fair comparison between the groups. In addition to assessing gender based network differences, we also perform a secondary analysis to assess our method's ability to reliably estimate functional networks. For this analysis, we split each subjects' resting state fMRI time series into two equally sized scanning sessions (60 scans each) and calculate the intraclass correlation coefficient or ICC (refer to equation (A.4.1)) for seven network metrics which are widely used to summarize brain networks. The network metrics include clustering coefficient, characteristic path length, local efficiency, global efficiency, modularity, hierarchy, and degree, and they were calculated with the Matlab toolboxes Brain Connectivity Toolbox (Rubinov and Sporns, 2010a) and GRETNA (Wang et al., 2015). Mathematical definitions for each metric are presented in Appendix C. We note that the ICC is a commonly used measure designed to assess the similarity of network estimates across scanning sessions (Braun et al., 2012; Choe et al., 2017; Niu et al., 2013), and it is usually derived by calculating the proportion of the total variation attributed to variability across scanning sessions. Thus, small variation across sessions relative to variation between individuals produces high ICC values, indicating strong reproducibility.

3.2.1   Data preprocessing—Resting-state fMRI scans were acquired on a single-shot, interleaved multi-slice, gradient-echo, echo planar imaging (GE-EPI) sequence (Satterthwaite et al., 2014). Nominal voxel size is 3×3×3mm with full brain coverage achieved with the following parameters: TR/TE=3000/32 ms, flip=90°, FOV=200 $\times$ 220 mm, matrix= 64 $\times$ 64, 46 slices, slice thickness/gap=3 mm/0 mm for a total of 6.2 minutes. Participants were instructed to remain awake, motionless, and fixated on a crosshair throughout the duration of the data acquisition. Several standard preprocessing steps were applied to the rs-fMRI data, including despiking, slice timing correction, motion correction, registration to MNI 2mm standard space, normalization to percent signal change, removal of linear trend, regressing out CSF, WM, and 6 movement parameters, bandpass filtering (0.009 to 0.08), and spatial smoothing with a 6mm FWHM Gaussian kernel. Subsequent voxel level data is aggregated into 90 regions of interest (ROI) based on the Automated Anatomical Labelling atlas (Tzourio-Mazoyer et al., 2002). For each ROI, the average time series of all constituent voxels represents the region's temporal BOLD signal.

Diffusion weighted images permit us to localize and orient white matter fiber bundles via the diffusion of water in the brain. Images were acquired on a twice-refocused spin-echo (TRSE) single-shot EPI sequence for a total of 64 diffusion-weighted directions with b=1000 s/mm^2 and 7 scans with b=0 s/mm^2 (Satterthwaite et al., 2014). Acquisition parameters were TR/TE=8100/82ms, matrix=128$\times$128, FOV=240mm, slice thickness=2mm,

GRAPPA factor=3. Due to gradient induced vibrations disturbing image quality, DWI images were acquired in two imaging runs to reduce the continuous duration in which subjects tolerate the scan. Standard pre-processing procedures, such as eddy current correction and bias-field correction are applied to the diffusion weighted data. Subsequently, we use the FSL functions bedpostx and probtracx2 to estimate the distribution of fiber tensors at each voxel and the count of white matter fibers tracts connecting all pairs of brain regions, respectively. In order to obtain the SC scores, we compute $p_{jk} = $ average$\{N_{jk}/N_{j*}, N_{kj}/N_{k*}\} \in (0, 1)$, where N_{jk} equals the number of permissible tracts initiated at region j that pass through region k and N_{j*} refers to the total number of permissible tracts initiated at region j. Fiber tracks passing through gray matter or cerebrospinal fluid are discarded. These SC scores can be interpreted as the probability of structural connectivity between regions j and k, which we often refer to as the strength of SC.

3.2.2   Results—In Figure 7, we see that for males (top left panel) and females (top right panel), the association between FC and SC along structurally connected regions is largest for aGlasso, indicating close adherence with the anatomical information. However, the literature suggests that FC is not fully explained by direct structural connections and thus a large association is not realistic. This strong dependence on SC contributes to aGlasso's inferior correlation with the empirical FC (*bottom row* in Figure 7). On the other hand, siGGM's flexible incorporation of SC while accounting for non-anatomical sources of variation produces desirable results since it adheres to the SC information while maintaining an association with the empirical FC observed in SC-naive approaches. Additionally, we also discover that siGGM leads to larger shrinkage and smaller variance for conditional dependencies between anatomically isolated brain regions compared to the generic graphical lasso without prior knowledge. This yields a smaller number of functional connections between anatomically disconnected ROIs.

For network analysis, we classify each ROI into one of eight functional modules corresponding to resting state networks as defined in Smith et. al (2009). These functional modules include a medial visual network, an occipital pole and lateral visual network ("VIS", 18 nodes), the default mode network ("DMN", 8 nodes), a sensorimotor network ("SM", 9 nodes), an auditory network ("AUD", 10 nodes), an executive control network ("EC", 19 nodes), right and left frontoparietal modules ("FPR" and "FPL", 11 and 10 nodes, respectively) and an unknown module containing unassigned nodes ("UNK", 5 nodes). Figure 8 shows that males and females have similar connectivity patterns with primarily positive connections within functional modules. Further comparisons of male and female brain networks within each age group reveals that consistent connections across age groups persist within module while inconsistent connections mainly exist between modules. After standardizing by the number of nodes in each module, the SM and AUD were found to be the two most highly connected functional modules in males and females across all age groups. Figure 9 (A) illustrates the similarity in network architecture for males and females with all metrics having non-significant differences across genders (except for local efficiency for teens and young adults), which implies shared patterns in brain organization across gender and age groups. Figure 9 (B) illustrates that males exhibit greater (but non-significant) between module but smaller within-module connectivity differences in teens and

young adults. These findings are supported by previous work and has been linked to variations in emotional identification and spatial cognitive tasks (Satterthwaite et al., 2016).

As a second level of the analysis, we are also interested in the distribution of differentially weighted edges between males and females within each of the eight functional modules. Differentially weighted edges were identified as connections for which the FC strength was significantly different between genders under a permutation test. To evaluate if the number of differentially weighted edges within and between modules occur more often than allowed by chance, we define a goodness of fit measure (equation (A.3.1) in Appendix A) which represents the deviation between observed and expected numbers of differentially weighted edges for each module block, standardized by the expected number. This measure captures whether a given module block has unusually high or low occurrence of differentially weighted edges and enables us to identify modules with the most pronounced differences across gender. From the results presented in Table 2, we discover statistically significant differences in the number of differentially weighted edges occurring in the executive control (EC) module in pre-teens and young adults, which is supported by previous results on gender related differences in the EC (Hyde, 1981; Mansouri et al., 2016). Table 2 also suggests that gender based differences attenuate with development, with the largest number of differentially weighted edges in the pre-teen group (377) and the smallest in the young adult group (272). We also find the differentially weighted edge between the cingulum_ant_L in the EC and parietal_inf_L in the DMN exists in pre-teens, teens, and young adults, which suggests consistent gender based differences during the developmental phase. These regions are known to have brain volume differences between males and females which may point to subtle cognitive variations (Frederikse et al., 1999; Ruigrok et al., 2014).

A major challenge in resting state connectivity studies is to ensure reproducibility of the findings (Griffanti et al., 2016). We demonstrate that appropriately incorporating anatomical connectivity information leads to stable topological features of estimated networks across scanning sessions. Figure 10 displays the ICC of seven network metrics under different approaches, where the details for computing the ICC are outlined in equation (A.4.1) in Appendix A. It is clear that the proposed siGGM produces estimates that have notably larger ICC measures for all the network metrics compared to all the other approaches considered. The reproducibility under the proposed approach is substantial for the clustering coefficient, global efficiency, and degree, and is moderate for all the other metrics. Moreover, it is reassuring to see that these three metrics with the highest ICC values under the proposed approach have been shown to be the most reproducible network metrics in independent studies (Niu et al., 2013; Telesford et al., 2010; Wang et al., 2011). In contrast, reproducibility is barely moderate under aGlasso for most metrics and weak under SC naive approaches. We note that siGGM has significantly higher ICC (p<.001) than Glasso and Space with respect to all network metrics, and a significantly higher ICC compared to aGlasso for all network metrics (p<.004) except local efficiency and degree (p>.05). These findings highlight the benefits of incorporating anatomical information in a flexible manner. Although not presented, we note that the G-Wishart approach leads to an unrealistic ICC value of one in all cases, which is starkly different than the reliability values reported in previous studies (Welton et al., 2015). The perfect reliability is due to the fact that G-Wishart

relies entirely on the SC information for specifying the functional network structure, resulting in the exact same network for both the sessions. Hence the reproducibility results under G-Wishart are not comparable.

4 Conclusion

In this paper, we introduce a novel Bayesian approach and an associated optimization algorithm for fusing structural and functional imaging data in order to estimate brain functional networks. We propose a flexible method for incorporating *apriori* known anatomical connectivity information in order to estimate the functional brain networks, bypassing the limitations of existing approaches by accommodating complex structure-function relationships while allowing unknown sources of variation independent of underlying anatomical structure. The proposed model is biologically more realistic compared to existing methods and often outperforms alternative approaches with or without anatomical knowledge as illustrated via extensive numerical studies. In particular, the advantages under the proposed method become more evident as the misspecification levels for the anatomical knowledge and/or the number of nodes is increased, which has important practical implications. An analysis of the PNC data yields brain networks which have strongly reliable network metrics under the proposed approach, whereas the reproducibility under other approaches are moderate at best.

Although our focus is on direct SC, the literature suggests that indirect connections have a non-trivial role in mediating cortical activity. Unfortunately, there is no consensus as to how it should be measured or incorporated into a model. Skudlarski et al. (2008) found that the maximal correlation between SC and FC occurs when there are at most three intermediary brain regions structurally connecting the two regions of interest, but these results may need to be replicated before wider use in the neuroimaging community. There are several possible approaches for incorporating indirect SC into the model estimation. First, one could craft a scalar score measuring structural connection strength inclusive of direct and higher ordered SC information. A second approach modifies the mean of $\pi(\lambda_{jk}|\mu_{jk}, \eta)$ to include iterated powers of SC, which would require slight modifications of the proposed siGGM approach. In future work, we intend to explore various direct and indirect measures of SC so as to assess which anatomical measures yield the most meaningful and reproducible FC results. This is a topic of open research requiring further statistical inquiry with the guidance of neuroscience experts.

While the incorporation of SC information clearly improves the detection of edges, we see a slightly deteriorated performance under non-Gaussian data under all GGM based approaches. However, the siGGM approach performs well regardless of the Gaussianity of the observed data, and often has increased gains over competing methods under non-Gaussian data over varying misspecification levels, underlining it's robustness to model misspecification. In future work, we plan to incorporate the anatomical information into nonparametric network estimation approaches such as the nonparanormal graphical model. Moreover, we note that the estimated FC under siGGM has a weaker correspondence to SC compared to alternate SC informed approaches, suggesting a less stringent adherence to the

given SC information that is attributed to the incorporation of non-anatomical sources of variation prevalent in real data.

Acknowledgements

Appendix A

A.1 Parameter Updates

We optimize (6) by iteratively updating model parameters as follows.

Update Ω: Given the data, $\mathbf{Y} = \{\mathbf{y}1, \ldots, \mathbf{y}_T\}$, and current estimates for all other model parameters, we solve

$$\hat{\Omega}^{(m+1)} = \arg\min_{\Omega} -\log\det(\Omega) + \text{tr}(S\Omega) + \frac{v}{2}\sum_{j<k} e^{\alpha_{jk}^{(m+1)}}\left|\omega_{jk}\right| + \frac{1}{2}v\sum_{j=k}\left|\omega_{kk}\right| \quad \text{A.1.1}$$

for $\hat{\Omega}$. This resembles the penalized likelihood framework of the traditional Gaussian graphical model. Define $\delta_{ij}^{(k+1)} = \frac{1}{2}\exp(\alpha_{ij}^{(m+1)})$ for $j \neq k$ and $\delta_{ij}^{(m+1)} = \frac{1}{2}$ for $j = k$. We can re-express (A.1.1) as

$$\hat{\Omega}^{(m+1)} = \arg\min_{\Omega} -\log\det(\Omega) + \text{tr}(S\Omega) + v\sum_{j<k}\delta_{jk}^{(m+1)}\left|\omega_{jk}\right|,$$

where we update Ω using a approximation solver, *QUIC* (Hsieh et al., 2011), available in R.

Update μ_{jk}: Given $\mathbf{Y}$, $\eta^{(m+1)}$, and $\alpha_{jk}^{(m)}$, , we update μ_{jk} via the closed form equation

$$\mu_{jk}^{(m+1)} = \frac{\sigma_\mu^2(\alpha_{jk}^{(m)} + \eta^{(m+1)}p_{jk}) + \sigma_\lambda^2\mu_0}{\sigma_\mu^2 + \sigma_\lambda^2}$$

Update η: Given $\mathbf{Y}$ and $\boldsymbol{\alpha}^{(m)}$, we can update η via closed form equation

$$\hat{\eta}^{(m+1)} = \frac{-\beta^{(m)} + \sqrt{(\beta^{(m)})^2 - 4\gamma\rho}}{2\gamma}$$

where $\beta^{(m)} = b_\eta + \dfrac{\sum_{j\,<\,k} \alpha_{jk}^{(m)} p_{jk}}{\sigma_\lambda^2} - \dfrac{1}{\sigma_\lambda^2} \sum_{j\,<\,k} \mu_{jk}^{(m)} p_{jk},\; \gamma = \dfrac{\sum_{j\,<\,k} p_{jk}^2}{\sigma_\lambda^2},\; \text{and } \rho = -\dfrac{1}{\sigma_\lambda^2}(a_\eta - 1)$

Update $\boldsymbol{\alpha}$: Given $\mathbf{Y}$, $\boldsymbol{\Omega}(m)$, $\boldsymbol{\mu}(m{+}1)$, and $\eta(m{+}1)$, we can estimate $\alpha_{jk}^{(m+1)}$ for $1 \le j < k \le p$ by solving

$$\hat{\alpha}^{(m+1)} = \arg\min_{\alpha} v \sum_{j\,<\,k} e^{\alpha_{jk}} \left|\omega_{jk}^{(m)}\right| + \sum_{j\,<\,k} \frac{(\alpha_{jk} - (\mu_{jk}^{(m+1)} - \eta^{(m+1)} p_{jk}))^2}{2\sigma_\lambda}$$

A closed form solution doesn't exist, so we implement a Newton Raphson solver to find the optimal choice of $\boldsymbol{\alpha}$. Re-expressing this problem, we have

$$\arg\min_{\alpha} exp(\alpha)' \left|\omega^{(m+1)}\right| - \frac{1}{2\sigma_\lambda^2}(\alpha - (\mu_{jk}^{(m+1)} - \eta^{(m+1)}\tilde{p}))'(\alpha - (\mu_{jk}^{(m+1)} - \eta^{(m+1)}\tilde{p}))$$

where $\boldsymbol{\alpha} = \{a_{12}, a_{13}, \ldots, a_{(p-1)p}\}$, P' denotes the upper diagonal elements of the structural connectivity matrix P, $e^{\boldsymbol{\alpha}}$ is the element wise exponential for each component of $\boldsymbol{\alpha}$, and $\mathbf{1}$ as a vector of 1's of length $\dfrac{p(p-1)}{2}$. Since $\boldsymbol{\Omega}$ is symmetric and we do not shrink diagonal elements, we simplify our estimation of $\boldsymbol{\alpha}$ by only focusing upon the upper diagonal elements.

The Newton Raphson updating equation based on step size γ is $\boldsymbol{\alpha}^{m+1} = \boldsymbol{\alpha}^m - \gamma\, g(\boldsymbol{\alpha}^m) H(\boldsymbol{\alpha}^m)^{-1}$, where

$g(\alpha) = v\sigma_\lambda^2 D_{|w(m)|} e^\alpha + [\alpha - (\mu^{(m+1)} - \eta^{(m+1)}\tilde{P})]$ and $H(\alpha) = v\sigma_\lambda^2 D_{|w(m)|}\, D|e\alpha| + I$, $D_{|w(m)|}$ is a $\dfrac{p(p-1)}{2} \times \dfrac{p(p-1)}{2}$ diagonal matrix with elements as the upper triangular elements of $\boldsymbol{\Omega}$, and similarly for $D_{|e\alpha|}$, and I is an identity matrix. Since H is a diagonal matrix, it is easily inverted and serves as an appropriate Hessian matrix. We search for the step size ( ) using back tracking line search for each update of $\boldsymbol{\alpha}$ as in Chang et. al (2017).

A.2 Hyperparameter Choice and Initial Values

The proposed siGGM approach iteratively solves for the MAP estimator and works best when reasonable starting values are provided. We first find an initial estimate for the graph structure and the sparse inverse precision matrix ($\boldsymbol{\Omega}_0$), using the graphical lasso. We initialize all edge specific penalty parameters as λ_0, which is the global tuning parameter corresponding to $\boldsymbol{\Omega}_0$. We set $\sigma_\mu^2 = 5$, corresponding to an uninformative prior which reflects our lack of knowledge regarding the baseline effects and choose $\mu_0 = 0$ as a default setting. We randomly generate the edge specific baseline effects μ_{jk} from the prior distribution $N(\mu_0, \sigma_\mu^2)$ and use these as initial values. The initial value of η is chosen by

averaging $-\frac{\sum_{l < k} (\exp(\lambda_0) - \mu_{jk})\,/\,p_{jk}}{p(p-1)/2}$ which is the average of all possible η values under the

relationship $\exp(\lambda_0) = \mu_{jk} - \eta p_{jk}$, $j < k$ corresponding to $\sigma_\lambda = 0$. We choose

$$\sigma_\lambda^2 = \frac{1}{p(p-1)/2} \sum_{j < k} \sum_{j,k=1}^{p} (\exp(\lambda_0) - \mu_{jk} - \eta p_{jk})^2.$$

Finally, we found that choosing a_η and b_η to attain $\mathrm{E}[\eta] \approx 6$ and $\mathrm{Var}[\eta] \approx 1$ incorporates structural information in a flexible manner. However, larger first moments for the prior on η may lead to increased false positives as our method places more weight on smaller structural connections, and similarly, smaller first moment may decrease the overall impact of structural information. For example, when $a_\eta > 1$ and $b_\eta \to \infty$, we have $\mathrm{E}[\eta] \to 0$, which makes the siGGM indistinguishable from SC naive methods. In simulations, we found values selected for the set of hyperparameters—$(a\eta\,,\,b\eta\,,\,\sigma\mu,\,\sigma\lambda)$—led to fast estimation of model parameters and accurate results across a wide array of settings.

A.3 Measure for computing between module diff ences

We define the goodness of fit measure

$$X_{g_1,g_2}^2 = \frac{(Q_{(g_1,g_2)} - E_{(g_1,g_2)})^2}{E_{(g_1,g_2)}}, \quad \text{A.3.1}$$

where $g1$, $g2 \in \{1,\ldots,G\}$ are the indices corresponding to one of the $\mathscr{G}$ functional modules, $Q_{g1,g2}$ represents the observed number of differentially weighted edges in the (g_1, g_2) block, $E_{g1,g2}$ represents the expected number of differentially weighted edges in the (g_1, g_2) block when edges distribute randomly across the module blocks. $X_{g1,g2}^2$ measures the goodness of fit for each within-module block ($g_1 = g_2$) or between-module block ($g_1 \neq g2$). In equation (A.3.1), the expected value can be derived in a straightforward manner as $E_{g1,g2} = 0.5p*\{|g_1| (|g_2| - 1)\}$ for within module blocks ($g_1 = g_2$) and $E_{g1,g2} = p*|g_1||g_2|$ for between-module blocks ($g1 \neq g2$), where $|g|$ represents the total number of nodes within the gth module, and $p*$ represents the proportion of differentially weighted edges among all the edges across the network. Using 5000 permutations of group labels at each edge, the differentially weighted edges are identified as those connections with significant FDR-adjusted p-values.

A.4 Calculation of ICC

The intraclass correlation coefficient is a widely used reliability metric for assessing test-retest reliability of brain network topology in neuroimaging applications. Using ICC(3,1), (two-way mixed single measures testing for consistency) we investigate the reliability of graph metrics across two scanning session (Guo et al., 2012; Telesford et al., 2010). The quantity is calculated as

$$\mathrm{ICC}(3,1) = \frac{\mathrm{BMS} - \mathrm{EMS}}{\mathrm{BMS} + (k-1)\,\mathrm{EMS}}, \quad \text{A.4.1}$$

where k is the number of scanning sessions per participant, *BMS* is the between mean square and *EMS* is the mean residual sum of squares. BMS captures the variability between subjects while EMS measures unexplained within-subject variation in functional connectivity across scanning sessions (see Shrout and Fleiss (1979)). This metric is commonly used to measure test-retest network stability in brain networks (Braun et al., 2012) with agreement scale $0 < ICC \quad .2$ (slight), $.2 < ICC \quad .4$ (fair), $.4 < ICC \quad .6$ (moderate), $.6 < ICC \quad 8$ (strong), and $.8 < ICC \quad 1$ (near perfect) as suggested by Telesford et al. (2010).

Appendix B

Appendix B presents the result for siGGM and SC naïve approaches on stimulated data.

Appendix C

Network metrics quantify properties of graphs, such as the local connectedness of regions and global connectivity across the network. These features distill complex organizing principles into scalar values that have led to critical insights into the brain's functionality in health and disease (Lord et al., 2012). In the following section, we provide interpretations and mathematical formulas for the seven network metrics investigated in this work. All functions assume the brain network is a $p \times p$ binary adjacency matrix, A, where a_{ij} is equal to one if there is an edge between regions i and j or zero otherwise. We first describe two metrics that are building blocks for our network features of interest. The first is the mnumber of triangles around node $i, t_i = \frac{1}{2} \sum_{j,m=1}^{p} a_{ij} a_{im} a_{jm}$, which measures connectivity among regions adjacent to region i. The second is the shortest path length between regions i and j which is defined as $d_{ij} = > : a_{uv} \in g_{i \leftrightarrow j} a_{uv}$, where $g_{i \leftrightarrow j}$ is the set of edges comprising the shortest geodesic distance between the regions. This metric is a basis for quantifying global integration across the brain (Rubinov and Sporns, 2010*b*). For additional details, please see https://sites.google.com/site/bctnet/measures.

1. **Degree**

   The degree of region i is the number of edges incident to it. Mathematically, it is defined as $k_i = \sum_{j \neq i}^{p} a_{ij}$ Regions i and j are considered neighbors (or adjacent) if $a_{ij} = 1$.

2. **Characteristic Path Length**

   The characteristic path length is the average shortest path length in the network and is defined as $L = \frac{1}{p} \sum_i L_i = \frac{1}{p} \sum_i \frac{\sum_{j \neq i} d_{ij}}{p-1}$ where d_{ij} is the shortest path length and L_i is $p \frac{ip}{ip-1}$ the characteristic path length of region i.

3. **Clustering coefficient**

   The clustering coefficient measures the number of closed triangles containing a ROI. It is equivalent to the number of neighbors of region i that are also

neighbors of each other (Watts and Strogatz, 1998). It is calculated as

$$C = \frac{1}{p}\sum_i C_i = \frac{1}{p}\sum_i \frac{2t_i}{k_i(k_i-1)},$$ where Ci is the clustering coefficient of region i.

4. Local Efficiency

Braun et al. (2012) define local efficiency as "the efficiency of the local subgraph of a node that only contains direct neighbors of the region" and is a measure of local connectedness. It is defined as

$$E_{loc} = \frac{1}{p}\sum_i E_{loc,i} = \frac{1}{p}\sum_i \frac{\sum_{j,m=1,\neq i} a_{ij}a_{im}\left[d_{jm}(N_i)^{-1}\right]}{k_i(k_i-1)},$$ where ki(ki−1)*ii* Eloc,i is

the local efficiency of region i and $d_{jm}(N_i)$ is the length of the shortest path between regions j and m that contains only neighbors of i.

5. Global Efficiency

The global efficiency is the average inverse shortest path length in the network and quantifies integration across spatially distant regions in the brain. It is

calculated as $E = \frac{1}{p}\sum_i Ei = \frac{1}{p}\sum_i \frac{\sum_{j=1,j\neq i}^{P} d^{-1}_{ij}}{p-1}$ where E_i is the global

efficiency of region i.

6. Modularity

This metric measures how well the network decomposes into non-overlapping clusters of connected regions and is calculated as

$$Q = \sum_{u\in M}\left[euu - \left(\sum_{v\in M} e_{uv}\right)^2\right],$$ where M constitutes the sets of non-

overlapping modules, euv is the proportion of all links that connect regions in module u with regions in module v.

7. Hierarchy

Network hierarchy is an ordering of regions such that high degree regions tend to have lower local clustering coefficient than low degree regions. Mathematically, this ratio is quantified as $C \sim k_{-\beta}$, where C is the clustering coefficient, k is the region's degree, and β is the hierarchy coefficient that is large for strongly hierarchical structured networks. β is estimated as the slope coefficient from a linear regression of log(C) on log(k) (Braun et al., 2012).

References

Barabási A-L and Albert R (1999), 'Emergence of scaling in random networks', science 286(5439), 509–512. [PubMed: 10521342]

Braun U, Plichta MM, Esslinger C, Sauer C, Haddad L, Grimm O, Mier D, Mohnke S, Heinz A, Erk S et al. (2012), 'Test–retest reliability of resting-state connectivity network characteristics using fmri and graph theoretical measures', Neuroimage 59(2), 1404–1412. [PubMed: 21888983]

Bressler SL and Tognoli E (2006), 'Operational principles of neurocognitive networks', International Journal of Psychophysiology 60(2), 139–148. [PubMed: 16490271]

Chang C, Kundu S and Long Q (2016), 'Scalable bayesian variable selection for structured high-dimensional data', arXiv preprint arXiv:1604.07264.

Choe AS, Nebel MB, Barber AD, Cohen JR, Xu Y, Pekar JJ, Caffo B and Lindquist MA (2017), 'Comparing test-retest reliability of dynamic functional connectivity methods', Neuroimage 158, 155–175. [PubMed: 28687517]

Damoiseaux JS and Greicius MD (2009), 'Greater than the sum of its parts: a review of studies combining structural connectivity and resting-state functional connectivity', Brain Structure and Function 213(6), 525–533. [PubMed: 19565262]

Eavani H, Satterthwaite TD, Filipovych R, Gur RE, Gur RC and Davatzikos C (2015), 'Identifying sparse connectivity patterns in the brain using resting-state fmri', Neuroimage 105, 286–299. [PubMed: 25284301]

Frederikse ME, Lu A, Aylward E, Barta P and Pearlson G (1999), 'Sex differences in the inferior parietal lobule', Cerebral Cortex 9(8), 896–901. [PubMed: 10601007]

Friedman J, Hastie T and Tibshirani R (2008), 'Sparse inverse covariance estimation with the graphical lasso', Biostatistics 9(3), 432–441. [PubMed: 18079126]

Griffanti L, Rolinski M, Szewczyk-Krolikowski K, Menke RA, Filippini N, Zamboni G, Jenkinson M, Hu MT and Mackay CE (2016), 'Challenges in the reproducibility of clinical studies with resting state fmri: An example in early parkinson's disease', Neuroimage 124, 704–713. [PubMed: 26386348]

Guo CC, Kurth F, Zhou J, Mayer EA, Eickhoff SB, Kramer JH and Seeley WW (2012), 'One-year test–retest reliability of intrinsic connectivity network fmri in older adults', Neuroimage 61(4), 1471–1483. [PubMed: 22446491]

Gur RC, Richard J, Calkins ME, Chiavacci R, Hansen JA, Bilker WB, Loughead J, Connolly JJ, Qiu H, Mentch FD et al. (2012), 'Age group and sex differences in performance on a computerized neurocognitive battery in children age 8– 21.', Neuropsychology 26(2), 251. [PubMed: 22251308]

Halpern DF, Benbow CP, Geary DC, Gur RC, Hyde JS and Gernsbacher MA (2007), 'The science of sex differences in science and mathematics', Psychological science in the public interest 8(1), 1–51. [PubMed: 25530726]

Hermundstad AM, Bassett DS, Brown KS, Aminoff EM, Clewett D, Freeman S, Frithsen A, Johnson A, Tipper CM, Miller MB et al. (2013), 'Structural foundations of resting-state and task-based functional connectivity in the human brain', Proceedings of the National Academy of Sciences 110(15), 6169–6174.

Hines M (2010), 'Sex-related variation in human behavior and the brain', Trends in cognitive sciences 14(10), 448–456. [PubMed: 20724210]

Hinne M, Ambrogioni L, Janssen RJ, Heskes T and van Gerven MA (2014), 'Structurally-informed bayesian functional connectivity analysis', NeuroImage 86, 294–305. [PubMed: 24121202]

Honey CJ, Thivierge J-P and Sporns O (2010), 'Can structure predict function in the human brain?', Neuroimage 52(3), 766–776. [PubMed: 20116438]

Honey C, Sporns O, Cammoun L, Gigandet X, Thiran J-P, Meuli R and Hagmann P (2009), 'Predicting human resting-state functional connectivity from structural connectivity', Proceedings of the National Academy of Sciences 106(6), 2035–2040.

Hsieh C-J, Dhillon IS, Ravikumar PK and Sustik MA (2011), Sparse inverse covariance matrix estimation using quadratic approximation, in 'Advances in neural information processing systems', pp. 2330–2338.

Hyde JS (1981), 'How large are cognitive gender differences? a meta-analysis using! w^2 and d..', American Psychologist 36(8), 892.

Ingalhalikar M, Smith A, Parker D, Satterthwaite TD, Elliott MA, Ruparel K, Hakonarson H, Gur RE, Gur RC and Verma R (2014), 'Sex differences in the structural connectome of the human brain', Proceedings of the National Academy of Sciences 111(2), 823–828.

Kemmer PB, Bowman FD, Mayberg H and Guo Y (2017), 'Quantifying the strength of structural connectivity underlying functional brain networks', arXiv preprint arXiv:1703.04056.

Kim J, Pan W, Initiative ADN et al. (2015), 'Highly adaptive tests for group differences in brain functional connectivity', NeuroImage: Clinical 9, 625–639. [PubMed: 26740916]

Lenroot RK and Giedd JN (2010), 'Sex differences in the adolescent brain', Brain and cognition 72(1), 46–55. [PubMed: 19913969]

Lenroot RK, Gogtay N, Greenstein DK, Wells EM, Wallace GL, Clasen LS, Blumenthal JD, Lerch J, Zijdenbos AP, Evans AC et al. (2007), 'Sexual dimorphism of brain developmental trajectories during childhood and adolescence', Neuroimage 36(4), 1065–1073. [PubMed: 17513132]

Lord A, Horn D, Breakspear M and Walter M (2012), 'Changes in community structure of resting state functional connectivity in unipolar depression', PloS one 7(8), e41282. [PubMed: 22916105]

Mansouri FA, Fehring DJ, Gaillard A, Jaberzadeh S and Parkington H (2016), 'Sex dependency of inhibitory control functions', Biology of sex diff ences 7(1), 11.

Matthews BW (1975), 'Comparison of the predicted and observed secondary structure of t4 phage lysozyme', Biochimica et Biophysica Acta (BBA)-Protein Structure 405(2), 442–451.

Messé A, Rudrauf D, Benali H and Marrelec G (2014), 'Relating structure and function in the human brain: relative contributions of anatomy, stationary dynamics, and non-stationarities', PLoS computational biology 10(3), e1003530. [PubMed: 24651524]

Monti RP, Hellyer P, Sharp D, Leech R, Anagnostopoulos C and Montana G (2014), 'Estimating time-varying brain connectivity networks from functional mri time series', NeuroImage 103, 427–443. [PubMed: 25107854]

Mumford JA and Ramsey JD (2014), 'Bayesian networks for fmri: a primer', Neuroimage 86, 573–582. [PubMed: 24140939]

Ng B, Varoquaux G, Poline J-B and Thirion B (2012), 'A novel sparse graphical approach for multimodal brain connectivity inference', Medical Image Computing and Computer-Assisted Intervention–MICCAI 2012 pp. 707–714.

Niu H, Li Z, Liao X, Wang J, Zhao T, Shu N, Zhao X and He Y (2013), 'Test-retest reliability of graph metrics in functional brain networks: a resting-state fnirs study', PLoS One 8(9), e72425. [PubMed: 24039763]

Padmanabhan N, White M, Zhou HH and O'Connell R (2016), 'Estimating sparse precision matrices', Monthly Notices of the Royal Astronomical Society 460(2), 1567–1576.

Peng J, Wang P, Zhou N and Zhu J (2009), 'Partial correlation estimation by joint sparse regression models', Journal of the American Statistical Association 104(486), 735–746. [PubMed: 19881892]

Pineda-Pardo JA, Bruña R, Woolrich M, Marcos A, Nobre AC, Maestú F and Vidaurre D (2014), 'Guiding functional connectivity estimation by structural connectivity in meg: an application to discrimination of conditions of mild cognitive impairment', Neuroimage 101, 765–777. [PubMed: 25111472]

Power JD, Schlaggar BL, Lessov-Schlaggar CN and Petersen SE (2013), 'Evidence for hubs in human functional brain networks', Neuron 79(4), 798–813. [PubMed: 23972601]

Raznahan A, Lerch JP, Lee N, Greenstein D, Wallace GL, Stockman M, Clasen L, Shaw PW and Giedd JN (2011), 'Patterns of coordinated anatomical change in human cortical development: a longitudinal neuroimaging study of maturational coupling', Neuron 72(5), 873–884. [PubMed: 22153381]

Rubinov M and Sporns O (2010a), 'Complex network measures of brain connectivity: uses and interpretations', Neuroimage 52(3), 1059–1069. [PubMed: 19819337]

Rubinov M and Sporns O (2010b), 'Complex network measures of brain connectivity: uses and interpretations', Neuroimage 52(3), 1059–1069. [PubMed: 19819337]

Ruigrok AN, Salimi-Khorshidi G, Lai M-C, Baron-Cohen S, Lombardo MV, Tait RJ and Suckling J (2014), 'A meta-analysis of sex differences in human brain structure', Neuroscience & Biobehavioral Reviews 39, 34–50. [PubMed: 24374381]

Satterthwaite TD, Connolly JJ, Ruparel K, Calkins ME, Jackson C, Elliott MA, Roalf DR, Hopson R, Prabhakaran K, Behr M et al. (2016), 'The philadelphia neurodevelopmental cohort: a publicly available resource for the study of normal and abnormal brain development in youth', Neuroimage 124, 1115–1119.

Satterthwaite TD, Elliott MA, Ruparel K, Loughead J, Prabhakaran K, Calkins ME, Hopson R, Jackson C, Keefe J, Riley M et al. (2014), 'Neuroimaging of the philadelphia neurodevelopmental cohort', Neuroimage 86, 544–553. [PubMed: 23921101]

Shaw P, Kabani NJ, Lerch JP, Eckstrand K, Lenroot R, Gogtay N, Greenstein D, Clasen L, Evans A, Rapoport JL et al. (2008), 'Neurodevelopmental trajectories of the human cerebral cortex', Journal of Neuroscience 28(14), 3586–3594. [PubMed: 18385317]

Shen K, Miši B, Cipollini BN, Bezgin G, Buschkuehl M, Hutchison RM, Jaeggi SM, Kross E, Peltier SJ, Everling S et al. (2015), 'Stable long-range interhemispheric coordination is supported by direct anatomical projections', Proceedings of the National Academy of Sciences 112(20), 6473–6478.

Shrout PE and Fleiss JL (1979), 'Intraclass correlations: uses in assessing rater reliability.', Psychological bulletin 86(2), 420. [PubMed: 18839484]

Skudlarski P, Jagannathan K, Calhoun VD, Hampson M, Skudlarska BA and Pearlson G (2008), 'Measuring brain connectivity: diffusion tensor imaging validates resting state temporal correlations', Neuroimage 43(3), 554–561. [PubMed: 18771736]

Smith SM, Fox PT, Miller KL, Glahn DC, Fox PM, Mackay CE, Filippini N, Watkins KE, Toro R, Laird AR et al. (2009), 'Correspondence of the brain's functional architecture during activation and rest', Proceedings of the National Academy of Sciences 106(31), 13040–13045.

Smith SM, Miller KL, Salimi-Khorshidi G, Webster M, Beckmann CF, Nichols TE, Ramsey JD and Woolrich MW (2011), 'Network modelling methods for fmri', Neuroimage 54(2), 875–891. [PubMed: 20817103]

Sporns O (2013), 'Structure and function of complex brain networks', Dialogues in clinical neuroscience 15(3), 247. [PubMed: 24174898]

Telesford QK, Morgan AR, Hayasaka S, Simpson SL, Barret W, Kraft RA, Mozolic JL and Laurienti PJ (2010), 'Reproducibility of graph metrics in fmri networks', Frontiers in neuroinformatics 4.

Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B and Joliot M (2002), 'Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain', Neuroimage 15(1), 273–289. [PubMed: 11771995]

Varoquaux G, Gramfort A, Poline J-B and Thirion B (2010), Brain covariance selection: better individual functional connectivity models using population prior, in 'Advances in neural information processing systems', pp. 2334–2342.

Venkataraman A, Rathi Y, Kubicki M, Westin C-F and Golland P (2012), 'Joint modeling of anatomical and functional connectivity for population studies', IEEE transactions on medical imaging 31(2), 164–182. [PubMed: 21878411]

Wang H (2012), 'Bayesian graphical lasso models and efficient posterior computation', Bayesian Analysis 7(4), 867–886.

Wang J-H, Zuo X-N, Gohel S, Milham MP, Biswal BB and He Y (2011), 'Graph theoretical analysis of functional brain networks: test-retest evaluation on short-and long-term resting-state functional mri data', PloS one 6(7), e21976. [PubMed: 21818285]

Wang J, Wang X, Xia M, Liao X, Evans A and He Y (2015), 'Gretna: a graph theoretical network analysis toolbox for imaging connectomics', Frontiers in human neuroscience **9**.

Watts DJ and Strogatz SH (1998), 'Collective dynamics of "small-world" networks', nature 393(6684), 440–442. [PubMed: 9623998]

Welton T, Kent DA, Auer DP and Dineen RA (2015), 'Reproducibility of graph-theoretic brain network metrics: a systematic review', Brain connectivity 5(4), 193–202. [PubMed: 25490902]

Welvaert M, Durnez J, Moerkerke B, Verdoolaege G and Rosseel Y (2011), 'neurosim: An r package for generating fmri data', Journal of Statistical Software 44(10), 1–18.

Xue W, Bowman FD, Pileggi AV and Mayer AR (2015), 'A multimodal approach for determining brain networks by jointly modeling functional and structural connectivity', Frontiers in computational neuroscience **9**.

Yuan M and Lin Y (2006), 'Model selection and estimation in regression with grouped variables', Journal of the Royal Statistical Society: Series B (Statistical Methodology) 68(1), 49–67.

Zhang Y, Wang Z, Cai Z, Lin Q and Hu Z (2016), 'Nonlinear estimation of bold signals with the aid of cerebral blood volume imaging', Biomedical engineering online 15(1), 22. [PubMed: 26897355]

Zhu D, Li K, Guo L, Jiang X, Zhang T, Zhang D, Chen H, Deng F, Faraco C, Jin C et al. (2012), 'Dicccol: dense individualized and common connectivity-based cortical landmarks', Cerebral cortex 23(4), 786–800. [PubMed: 22490548]
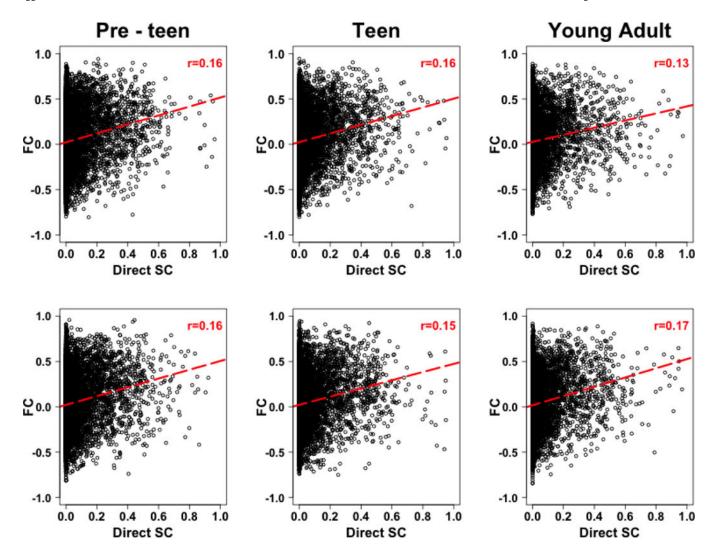
Figure 1:
Plots of the associations between partial correlation (FC) and direct structural connectivity (SC) for all males (*top row*) and females (*bottom row*) in our study. While mild positive correlations—approximately .16—are observed between SC and FC, FC exists between regions with little to no direct structural connections. We also observe large variation in FC for a given SC level (median standard deviation $\approx .275$). The red dashed line is the line of best fit and *r* is the Spearman correlation coefficient between direct SC and FC.
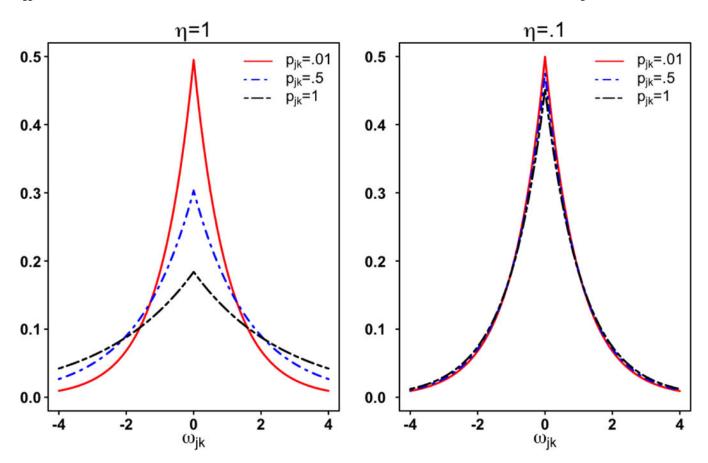
Figure 2:
Prior distribution of ω_{jk} when $\lambda_{jk} = \mu_{jk} - \eta p_{jk}$. μ_{jk} is fixed at zero with varying values of η and SC. Solid, dashed and dotted lines correspond to $p_{jk} = 0.01, 0.5$, and 1, respectively. Left panel (a): for large η values ($\eta = 1$), the prior places increasing mass at the tails, which encourage stronger functional connectivity; right panel (b): for small values ($\eta = 0.1$), the prior on ω does not change noticeably with the change in SC information.
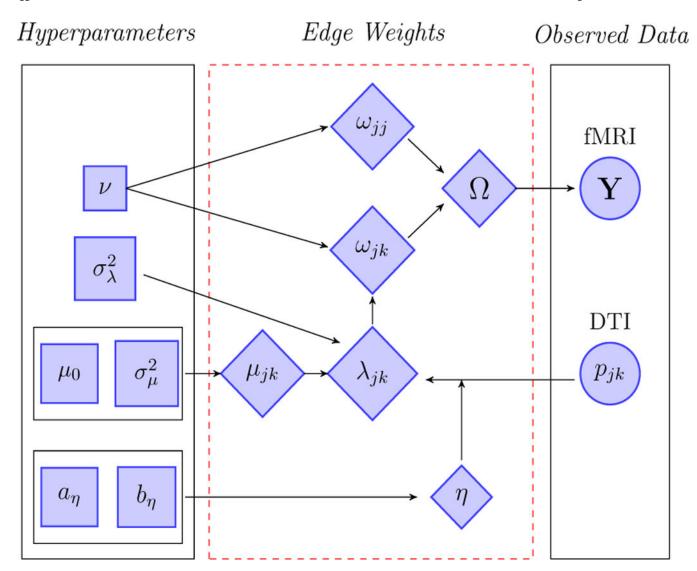
Figure 3:
Graphical illustration of model parameters and their contribution to estimation of anatomically informed functional connectivity based on resting state fMRI data **Y**. Circles represent observed data, diamonds represent parameters to be updated, and squares represent fixed values. Hyperparameters contained in rectangles jointly inform the distribution of the associated edge weight parameter. Parameters in the red, dashed edge weights pane are estimated outputs of the siGGM.
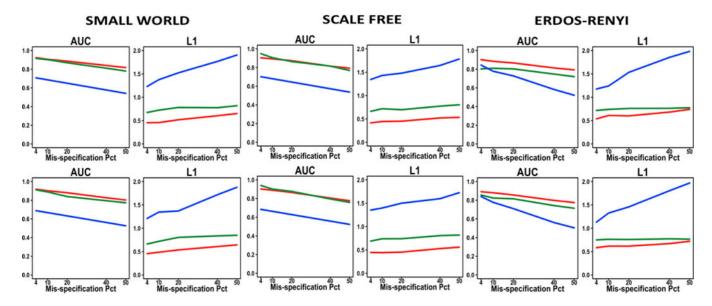
Figure 4:
Comparison of siGGM (**red**), G-Wishart (**blue**), and aGlasso (**green**) simulation results for different network structures with p=100 regions under scenario MI (top row) and MII (bottom row). Each panel displays the AUC or L1 relative error as the percentage of conditionally independent edges with non-zero anatomical connectivity increases. The AUC is significantly higher under siGGM for higher misspecification levels under the small-world network and for all misspecification levels for the Erdos-Renyi network. The $L1$ error is significantly lower under siGGM for non-trivial misspecification levels and all networks.
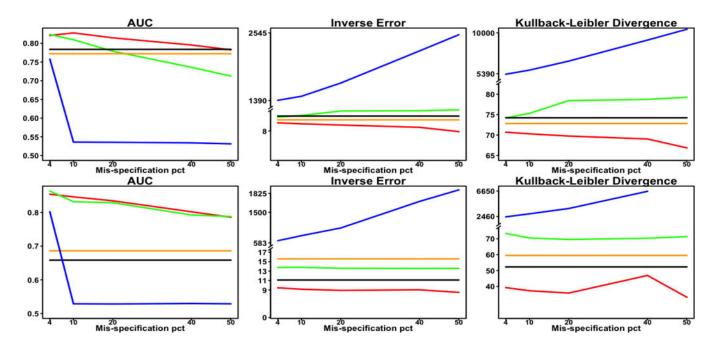
Figure 5:
Results from the simulation studies for non-Gaussian ICA data for 100 (*top row*) and 200 (*bottom row*) brain regions. We compare the performance of siGGM(**red**), aGlasso (**green**), G-Wishart (**blue**), glasso(**black**), and Space (**orange**) with respect to AUC, inverse error, and Kullback-Leibler divergence. The AUC under siGGM is significantly higher for most misspecification levels under $p = 100$, while the inverse error and Kullback-Leibler divergence is significantly lower under siGGM under all misspecification levels, for $p = 100$ and $p = 200$.

Figure 6:
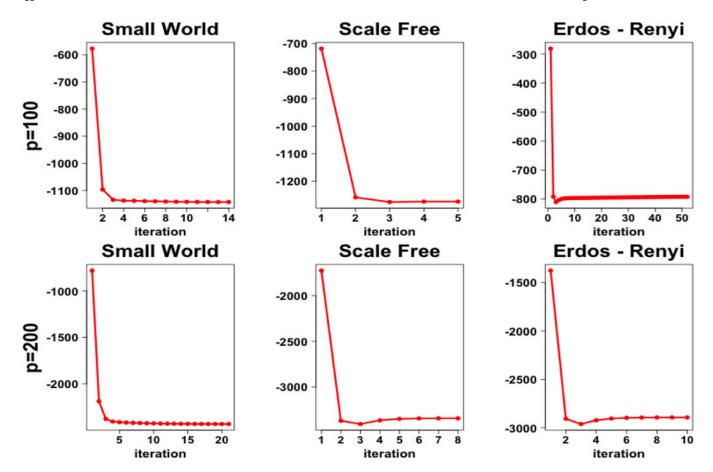Convergence of the log posterior likelihood for the structurally informed Gaussian graphical model (siGGM) under the three network structures (small-world, scale-free, Erdos-Renyi) and sizes (p=100, 200) investigated. The method typically makes the most substantial improvements within the first three iterations.
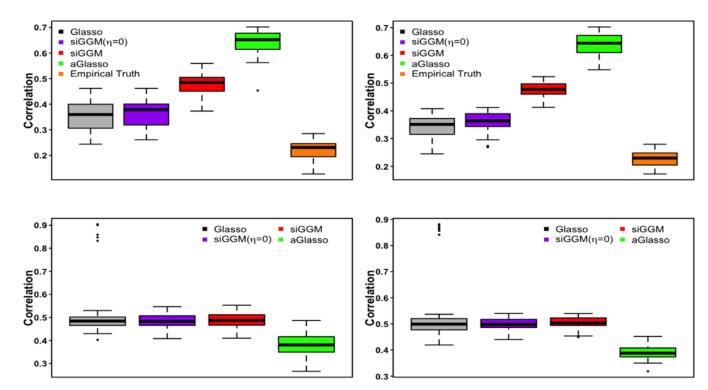
Figure 7:
Correlation between structural and functional connectivity males(left column) and females (right column). The top row displays the correlation between FC estimates and SC for regions structurally connected, and the bottom row displays the correlation between the estimated FC and the empirical FC for structurally connected regions. We note that aGlasso closely adheres to the SC information due to its shrinkage parameter specification.

(A) Female

(B) Male
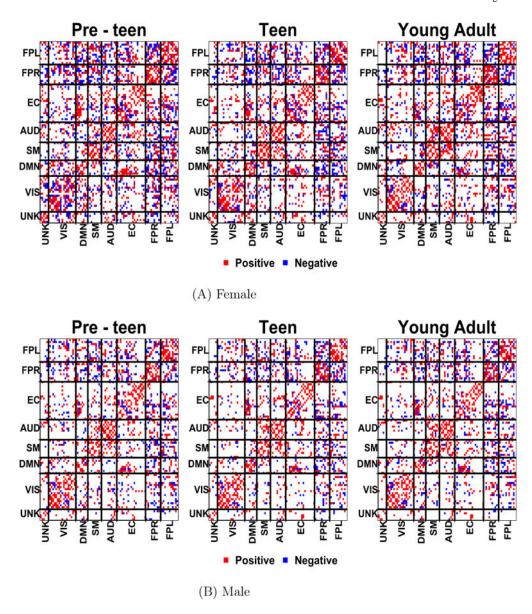
Figure 8:
Network estimates for females (*top row*) and males (*bottom row*) in each of the age ranges, illustrating those connections corresponding to absolute partial correlations > 0.005. While both sexes have similar network structures across the three age groups and have network densities close to 13%, female networks exhibit slightly increased connectivity relative to the networks of males.

(A)

(B)
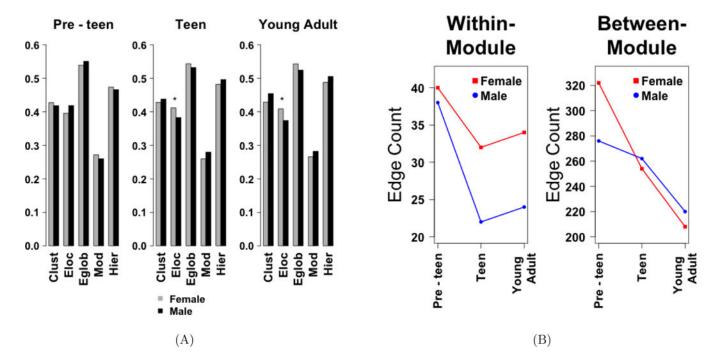
Figure 9:
Topological features of estimated networks in males and females across the three age groups. (A) displays five network properties—Clust (clustering coefficient), Eloc (local efficiency), Eglob (global efficiency), Mod (modularity), Hier (hierarchy)—averaged over the respective gender and age group; (B) displays differentially weighted edges within- and between- module stratified by gender. In teens and young adults, females have more within module connections and fewer between module connections than males. In (A), the local efficiency is statistically significantly different between males and females in teens and young adults ($p < .05$). In (B), there are no significant results at the .05 level of significance.
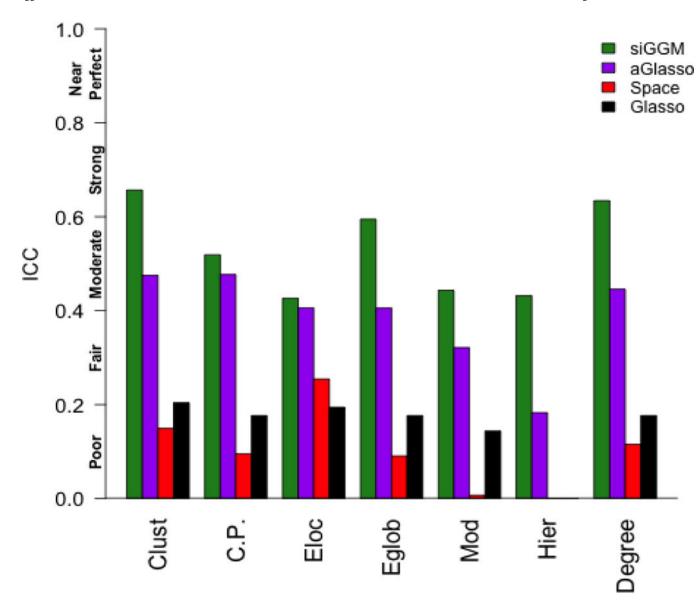
Figure 10:

Reliability of network metrics across scanning sessions for siGGM, aGlasso, *Space*, and *glasso*. We estimate seven network attributes (clustering coefficient (Clus.), characteristic path length (C.P.), local efficiency (Eloc), global efficiency (Eglob), modularity (Mod), hierarchy (Hier), and degree (Degree)) and report ICC(3,1) for all subjects. ICC values are classifi according to the agreement scale $0 < ICC \quad .2$ (poor), $.2 < ICC \quad .4$ (fair), $.4 < ICC \quad .6$ (moderate), $.6 < ICC \quad 8$ (strong), and $.8 < ICC \quad 1$ (near perfect) as suggested by Telesford et. al (2010). siGGM has significantly higher ICC (p<.001) than Glasso and Space with respect to all network metrics, and a significantly higher ICC compared to aGlasso for all network metrics (p<.004) except local efficiency and degree (p>.05). Generally, anatomically informed FC estimates produce more reliable networks than SC naive methods.

Table 1:

Performance of SC informed methods on simulated network data with p=100 and 200 nodes. Eglob is the bias in global efficiency.

	p=100				p=200			
Small World	Eglob	MCC	AUC	L1	Eglob	MCC	AUC	L1
G-Wishart MI(a)	0.120	0.592	0.698	1.345	0.175	0.468	0.865	1.510
G-Wishart MI(b)	0.173	0.447	0.574	1.498	0.224	0.337	0.819	1.801
G-Wishart MII(a)	0.117	0.567	0.676	1.319	0.174	0.447	0.842	1.506
G-Wishart MII(b)	0.171	0.424	0.650	1.492	0.224	0.318	0.797	1.788
aGlasso MI(a)	−0.145	0.522	0.903	0.724	−0.234	0.454	0.800	0.773
aGlasso MI(b)	−0.240	0.407	0.869	0.782	−0.258	0.396	0.788	0.801
aGlasso MII(a)	0.183	0.477	0.889	0.720	−0.221	0.439	0.735	0.770
aGlasso MII(b)	−0.279	0.364	0.840	0.805	−0.272	0.376	0.775	0.814
siGGM MI(a)	0.078	0.590	0.889	0.478	0.121	0.526	0.906	0.532
siGGM MI(b)	0.112	0.500	0.880	0.531	0.166	0.419	0.875	0.603
siGGM MII(a)	0.075	0.576	0.879	0.486	0.125	0.514	0.901	0.563
siGGM MII(b)	0.122	0.490	0.846	0.547	0.169	0.406	0.870	0.633
Scale Free								
G-Wishart MI(a)	0.104	0.590	0.695	1.410	0.146	0.471	0.864	1.503
G-Wishart MI(b)	0.159	0.446	0.671	1.483	0.196	0.339	0.822	1.583
G-Wishart MII(a)	0.102	0.568	0.675	1.397	0.145	0.452	0.846	1.442
G-Wishart MII(b)	0.156	0.424	0.650	1.481	0.196	0.323	0.801	1.604
aGlasso MI(a)	0.223	0.509	0.905	0.719	−0.311	0.375	0.701	0.730
aGlasso MI(b)	−0.195	0.460	0.864	0.697	−0.312	0.336	0.685	0.736
aGlasso MII(a)	−0.256	0.442	0.901	0.739	−0.333	0.351	0.690	0.739
aGlasso MII(b)	−0.252	0.404	0.877	0.740	−0.303	0.320	0.658	0.746
siGGM MI(a)	0.054	0.562	0.853	0.428	−0.075	0.473	0.868	0.442
siGGM MI(b)	0.093	0.467	0.822	0.457	0.131	0.359	0.843	0.492
siGGM MII(a)	0.061	0.552	0.845	0.447	0.078	0.459	0.865	0.457
siGGM MII(b)	0.099	0.451	0.812	0.469	0.132	0.346	0.839	0.523
Erdos-Renyi								
G-Wishart MI(a)	0.174	0.505	0.821	1.300	0.140	0.519	0.860	1.572
G-Wishart MI(b)	0.239	0.368	0.765	1.491	0.187	0.380	0.821	1.976
G-Wishart MII(a)	0.171	0.483	0.807	1.277	0.139	0.501	0.838	1.560
G-Wishart MII(b)	0.237	0.349	0.747	1.464	0.186	0.365	0.805	1.956
aGlasso MI(a)	−0.335	0.404	0.810	0.741	−0.424	0.171	0.596	0.712
aGlasso MI(b)	−0.339	0.353	0.803	0.761	−0.423	0.176	0.600	0.708
aGlasso MII(a)	−0.362	0.333	0.824	0.764	−0.424	0.162	0.631	0.707
aGlasso MII(b)	−0.360	0.318	0.815	0.759	−0.424	0.147	0.648	0.709
siGGM MI(a)	0.124	0.442	0.861	0.624	0.049	0.514	0.862	0.689
siGGM MI(b)	0.186	0.363	0.838	0.646	0.110	0.380	0.826	0.697
siGGM MII(a)	0.122	0.421	0.852	0.638	0.050	0.499	0.846	0.690

	p=100				p=200		
siGGM MII(b)	0.171	0.344	0.825	0.665	0.106	0.367	0.810	0.706

Table 2:

Within- and between- module differences in functional connectivity between males and females. Bolded values with an asterisk indicate statistically significant modules at the .05 level of significance (FDR correction for multiplicity) and values within parenthesis are the number of differentially weighted edges where the average edge weight is larger in males than females. The total number of differentially weighted edges decrease across the age groups (pre-teen 377 DWE, teen 312 DWE, young adult 272 DWE).

	Unknown	Visual	DMN	SM	Aud	EC	FP left	FP right
Pre–Teen								
Unknown	0(0)							
Visual	8(3)	24(12)						
DMN	7(3)	15(7)	0					
SM	3(1)	14(7)	3(1)	2(2)				
Aud	2(1)	17(8)	13(8)	5(2)	4(0)			
EC	6(5)	32(17)	9(4)	20(9)	21(13)	**42∗(26)**		
FP left	2(1)	14(7)	8(5)	4(4)	10(5)	13(6)	0(0)	
FP right	5(2)	18(9)	9(6)	11(8)	11(8)	16(10)	3(1)	6(0)
Teen								
Unknown	2(2)							
Visual	6(3)	22(14)						
DMN	3(2)	8(3)	2(0)					
SM	1(1)	11(8)	9(4)	4(2)				
Aud	3(0)	7(3)	7(4)	13(6)	4(2)			
EC	8(5)	22(11)	13(6)	12(8)	12(5)	18(10)		
FP left	3(1)	16(9)	7(3)	6(4)	9(4)	13(3)	0(0)	
FP right	2(1)	12(6)	9(4)	5(5)	15(5)	18(10)	8(3)	2(2)
Young Adult								
Unknown	2(0)							
Visual	1(1)	10(4)						
DMN	1(1)	14(4)	2(0)					
SM	1(0)	10(3)	9(4)	4(4)				
Aud	1(1)	15(5)	5(2)	5(3)	8(4)			
EC	5(4)	16(8)	10(3)	12(4)	12(6)	**26∗(16)**		
FP left	1(1)	7(4)	5(3)	7(4)	11(8)	9(5)	6(6)	
FP right	1(1)	8(4)	7(2)	9(5)	7(6)	17(7)	8(5)	0(0)

Table B.1:

Performance of siGGM and SC naive approaches on simulated network data with p=100 and 200 nodes. Eglob is the bias in global efficiency.

	p=100				p=200			
	Eglob	MCC	AUC	L1	Eglob	MCC	AUC	L1
Small World								
Glasso	0.177	0.327	0.827	0.575	0.128	0.333	0.757	0.668
Space	−0.206	0.585	0.839	0.407	−0.374	0.597	0.763	0.430
siGGM(η=0)	0.061	0.538	0.847	0.509	−0.019	0.506	0.843	0.587
siGGM	0.078	0.590	0.884	0.478	0.121	0.526	0.906	0.532
Scale Free								
Glasso	0.117	0.365	0.798	0.560	0.038	0.324	0.657	0.605
Space	−0.219	0.495	0.772	0.491	−0.403	0.358	0.664	0.555
siGGM(η=0)	0.005	0.509	0.808	0.528	−0.100	0.411	0.769	0.573
siGGM	0.054	0.562	0.853	0.428	−0.075	0.469	0.868	0.442
Erdos-Renyi								
Glasso	0.245	0.247	0.789	0.859	0.065	0.182	0.659	0.837
Space	−0.125	0.580	0.824	0.465	−0.415	0.253	0.638	0.577
siGGM(η=0)	0.020	0.363	0.792	0.679	−0.204	0.208	0.661	0.700
siGGM	0.124	0.442	0.861	0.624	0.049	0.514	0.862	0.689